សន្និសីទវិទ្យាសាស្ត្រ លើកទី១៤

ការប្រើប្រាស់សក្តានុពលនៃការស្រាវជ្រាវ និងអភិវឌ្ឍន៍ដើម្បីនវានុវត្តន៍ និងកំណើន



PROCEEDINGS OF THE 14TH SCIENTIFIC DAY

ITC

Leveraging R&D for Innovation and Growth

Institute of Technology of Cambodia
Phnom Penh, Cambodia
June 05-06, 2025

ORGANIZED BY

SUPPORTED BY

COOPERATED WITH

PLATINUM SPONSOR

Panasonic

PREMIUM SPONSOR

TABLE OF CONTENTS

FORWARD MESSAGE I
ORGANIZING COMMITTEE..... II
ABOUT THE 14TH SCIENTIFIC DAY 1
EVENT SUMMARY 4
EVENT AGENDA.....7
ORAL SESSION..... 15
POSTER SESSION..... 593
ACKNOWLEDGEMENTS.....671
PROMOTIONAL PAGE 673

FORWARD MESSAGE

The Institute of Technology of Cambodia (ITC), a Cambodian Higher Education Institution, was founded in 1964. Currently, there are three campuses that are operating, Phnom Penh, Tbong Khmum Province, and Kep province. The fourth campus is currently being constructed in Kampong Cham province, while the fifth campus, situated near the win-win monument in Phnom Penh, is also in the construction phase. ITC is expanding to produce human resources for supporting the economic and social development of Cambodia. In today's dynamic global landscape, ITC plays a vital role in Cambodian society while staying at the forefront of progress. Our goals are to cultivate talented and skilled graduates in science, technology and engineering and driven applied research aims at fostering start-up and technology transfer. We equip students with technical expertise and critical analysis skills, enabling seamless integration and growth in the evolving job market. To achieve this vision, we prioritize a synergistic approach to academic pursuits and international scientific research.

Research and Innovation Center (RIC), established in 2015, comprises five research units: Energy Technology and Management (ETM), Food Technology and Nutrition (FTN), Mechatronics and Information Technology (MIT), Materials Science and Structure (MSS), and Water and Environment (WAE). With about 100 researchers, including senior, lecturer, and full-time researchers, RIC supports sustainable development and reduces social disparities through internal practices and international connections. It has forged partnerships with universities in Europe, Asai, and local, improving educational quality, creating new programs, and fostering research collaboration. ITC and RIC also maintain strong ties with Cambodian and multinational companies.

To strengthen the academic and research performance, the 14th Scientific Day under the theme of "*Leveraging R&D for Innovation and Growth*" was organized on 05-06 June 2025. Supported by the Ministry of Education, Youth and Sports and various partners, the event aimed to bring together policymakers, academic and research institutions, and private sector stakeholders from local, regional, and international networks to share and discuss new developments, ideas, breakthroughs, and cutting-edge technologies in engineering and science towards contribution to development of digital society and economy. Finally, on behalf of ITC, I would like to thank the Ministry of Education, Youth and Sports (MoEYS); French Embassy in Cambodia; Japan International Cooperation Agency (JICA); the DX.sea project; Panasonic; Platform for Research and Training on Power System of the EDC-AFD-EU Project; Institut de Recherche pour le Développement (IRD); Food Technology, Research and Innovation Platform (FTRIP); TEM Trading M&E Product Co., Ltd.; and all partners and sponsors.

H.E. Prof. Dr. PO Kimtho
Director General of ITC

ORGANIZING COMMITTEE

Advisory Board

1. H.E. Prof. Dr. PO Kimtho
2. Mr. SOY Ty
3. Assoc. Prof. Dr. BUN Kim Ngun
4. Asst. Prof. Dr. NGUON Kollika
5. Dr. BUN Long
6. Assoc. Prof. Dr. OR Chanmoly

Organizing Committee

1. Asst. Prof. Dr. TAN Reasmey
2. Asst. Prof. Dr. YOS Phanny
3. Asst. Prof. Dr. PHAT Chanvorleak
4. Asst. Prof. Dr. PENG Chanthol
5. Asst. Prof. Dr. VALY Dona
6. Dr. DOUNG Piseth
7. Dr. KRET Kakda
8. Dr. HENG Sounean
9. Ms. KAING Vinhteang
10. Dr. BUN Polyka
11. Ms. LY Soda
12. Mr. PLACK Sokhit

Exhibition, Sound, and Order Committee

1. Asst. Prof. Dr. PENG Chanthol
2. Mr. PENH San
3. Mr. KEO Chhomsethy

Communication Team

1. Asst. Prof. Dr. YOS Phanny
2. Dr. CHHUNHIENG Thavarith
3. Asst. Prof. Dr. SIM Tepmony
4. Asst. Prof. Dr. IN Sokneang
5. Asst. Prof. Dr. CHHUON Kong
6. Asst. Prof. Dr. ENG Chandoeun
7. Asst. Prof. Dr. HAN Virak
8. Mr. LAY Heng
9. Asst. Prof. Dr. PHUN Vengkheang
10. Dr. CHRIN Phok
11. Asst. Prof. Dr. SRENG Sokchenda
12. Asst. Prof. Dr. CHAN Sarin
13. Mr. LY Hav
14. Mr. HASH Chanly
15. Asst. Prof. Dr. LIN Mongkolsery
16. Mrs. SREY Malis
17. Mrs. KHEMTRAN Krasel
18. Mr. SO Phea
19. Asst. Prof. Dr. PHAUK Sokkhey
20. Asst. Prof. Dr. SRANG Sarot
21. Asst. Prof. Dr. YIN Molika
22. Mrs. HANG Vichothy
23. Mr. SIEANG Phen
24. Mrs. KOY Sophary
25. Mr. SOK Kimheng
26. Asst. Prof. Dr. BUN Saret
27. Asst. Prof. Dr. HOUNG Peany
28. Ms. LY Soda

Students, High School Students, and Laboratory Management Committee

1. Mr. NUTH Sothan
2. Asst. Prof. Dr. LIN Mongkolsery
3. Mr. KHUON Rithymean
4. Mrs. SREY Malis
5. Mr. LONG Sovann
6. Mr. ANN Buneng
7. Mr. LAY Khem
8. Mr. THAB Sophatt
9. Ms. MEAS Sreymom
10. Ms. NEANG Vorleak
11. Mr. POV Hengkry
12. Mr. POS Sothea
13. Mr. HENG Chheng
14. Mr. MENG Channy
15. Mr. MOA Chanvattanak
16. Mr. HENG Theany
17. Mr. BEUK Bunpanha
18. Mr. SUN Sopheak
19. Mr. TONG Bunthoeun
20. Mr. CHEAM Ang
21. Mr. SOY Sokhom
22. Ms. LENG Rasie
23. Mr. POENG Runsovann
24. Mr. CHEA Chanry
25. Ms. KHUT Pichraksa
26. Ms. TAN Chandara
27. Mr. PHAUK Ponna
28. Mr. PEN Chentra
29. Mrs. KVAI Sopheap

Scientific, Publication, and IT Committee

1. Dr. HENG Sounean
2. Asst. Prof. Dr. VALY Dona
3. Mr. KHIEV Samnang
4. Ms. KAING Vinhteang
5. Mr. SEANG Chamroeun
6. Mr. LUN Chanmaranden
7. Asst. Prof. Dr. SUONG Malyna
8. Dr. SANG Davin

Finance Committee

1. Dr. KRET Kakda
2. Ms. LY Soda

ABOUT THE 14TH SCIENTIFIC DAY

Leveraging R&D for Innovation and Growth

I. Background

The Scientific Day event is an annual conference organized by the Institute of Technology of Cambodia (ITC). Originally a local forum, it expanded into an international platform in 2022, bringing together academics, researchers, engineers, and industry professionals to share insights and advancements in Science, Technology, Engineering, and Mathematics (STEM), with a focus on sustainable development.

The previous event, the 13th Scientific Day, held on June 06–07, 2024, was a resounding success, attracting approximately 1,500 participants from over 70 organizations. The event featured:

- 4 panelists in a high-level panel discussion
- 15 guest speakers across 9 parallel sessions
- 72 oral presentations and 37 poster presentations

Additionally, students from various high schools across Cambodia attended and participated in the **Open House**, further fostering engagement with STEM disciplines.

The **14th Scientific Day** is scheduled for **June 05–06, 2025**, at **ITC in Phnom Penh, Cambodia**. Under the theme **"Leveraging R&D for Innovation and Growth"**, this year's event aims to facilitate deeper discussions and collaborations in research and development. All submitted papers will undergo **peer review**, and accepted full papers will be published in the **Scientific Day Proceedings**. Furthermore, **exceptional research papers** with significant findings will be recommended for further revision and potential publication in the **Techno-Science Research Journal (TSRJ)**.

We invite you to participate in the **14th Scientific Day 2025** and contribute to the success of this dynamic and impactful conference. Your research and insights will play a crucial role in advancing innovation and knowledge in STEM.

II. Objective

- To promote the dissemination of scientific research outputs and enhance cooperation at local, regional, and international levels;
- To raise awareness and prepare human capital through research and industry linkages that foster innovation;
- To encourage multidisciplinary collaboration and joint research initiatives among ITC, public, and private sectors;
- To strengthen research capacity and innovation for students in Science and Engineering disciplines;
- To disseminate research findings through theoretical demonstrations, simulations, proof-of-concept studies, and prototypes in relevant research fields.

III. Topics of Interest

The conference includes, but is not limited to, the broad areas of:

1. Energy and Technology Management
2. Food Technology and Nutrition
3. Mechatronics and Information Technology

4. Materials Science and Structure
5. Water and Environment

Important Dates:

1. Abstract submission deadline: 31 March 2025
2. Full paper submission deadline: 02 May 2025
3. Notification of acceptance: 19 May 2025

Note: Abstracts and full papers were reviewed by Techno-Science Research Journal Editorial members and researchers

IV. Abstract and Paper Submission

Full papers must be written in English, with a maximum of 6 pages. The full paper template will be sent after abstract submission. Registration, abstract, and full paper submission can be accessed by scanning the QR code below:

Registration

V. Date and Venue

Date of Conference: 05-06 June 2025
Venue: Institute of Technology of Cambodia (ITC)
Russian Federation Blvd, P.O. Box 86, Phnom Penh, Cambodia

VI. Main Events and Tied-in Activities

The Scientific Day of ITC is a dynamic event featuring a range of activities designed to foster knowledge exchange, innovation, and collaboration. It serves as a platform for students, researchers, engineers, scientists, the academic community, and industry representatives to come together to connect, share experiences, and explore future opportunities in science, technology, and innovation. The event promotes learning and dialogue through the presentation of research findings, practical demonstrations, and discussions that reflect diverse perspectives and visions for individual and societal development. Key event of the Scientific Day include:

- **Research Presentations** – showcasing recent scientific and technical findings from students and researchers
- **Poster Exhibition** – visual displays of academic and applied research projects
- **Booth Exhibition** – interactive booths from academic programs, research units, and partner institutions
- **Networking Opportunities** – connecting students, academics, and industry professionals
- **ITC Laboratory Visits** – guided tours to explore ITC's laboratories and facilities.

VII. Sponsorship and Exhibition Opportunities

The Scientific Day event is organized by ITC and aims to promote STEM research and education in Cambodia. For sponsorship opportunities or information on obtaining a booth to display products, please contact us using the information provided below.

Contact info

For detail information you can contact us by:
Phone: (+855) 95 353 112
Email: scientificday@itc.edu.kh
Join our telegram group for updated information

Telegram group

EVENT SUMMARY

The 14th Scientific Day was held at ITC as an onsite event on 5–6 June 2025, under the theme "*Leveraging R&D for Innovation and Growth*". The event aimed to bring together policymakers, academic and research institutions, and private sector stakeholders from local, regional, and international networks to share and discuss new developments, ideas, breakthroughs, and cutting-edge technologies in engineering and science towards contribution to development of digital society and economy. This scientific event serves as an important platform for sharing knowledge, experiences, and research findings between academician, researchers, and practitioners.

This event was organized by the Research and Innovation Center and various departments of the Institute of Technology of Cambodia, with support from the Ministry of Education, Youth and Sports (MoEYS), French Embassy in Cambodia, Japan International Cooperation Agency (JICA) through the INACON project, and the DX.sea project, and with collaboration of Food Technology, Research, and Innovation Platform (FTRIP). The event was financially sponsor by *Platinum Sponsor*—Panasonic; *Premium Sponsors*—Platform for Research and Training on Power System of the EDC-AFD-EU Project; Institut de Recherche pour le Développement (IRD); TEM Trading M&E Product Co., Ltd.; and Food Technology, Research and Innovation Platform (FTRIP); *Gold Sponsors*—Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD); DKK-TOA Corporation & Grande Solid Co., Ltd.; Grande Solid Co., Ltd. & Kyoto Electronics Manufacturing Co., Ltd.; and LHSBI Insurance; *Silver Sponsors*—The Capacity for Cambodia Energy Efficiency (CapCEE); DENSO (Cambodia) Co., Ltd.; Dynamic Scientific; SNP-PT International Co., Ltd., Thai-YAZAKI Electric Wire Co., Ltd., and Yonemori Construction Co., Ltd.; *Internet Sponsors*—Today Internet; and NTC NeocomISP Limited; and *In-kind sponsor*—V.N.C CORPERATION CO., LTD.

The opening ceremony of the 14th Scientific Day was conducted in the morning of June 5, 2025, under the high chairmanship of **H.E Dr. HANG Chuon Naron**, Deputy Prime Minister and Minister of Ministry of Education, Youth and Sports. The event began with welcoming remarks from **H.E. Prof. Dr. PO Kimtho**, Director General of ITC, followed by addresses from **Mr. Kazumasa SANUI**, Chief Representative of JICA Cambodia; **Mr. Emmanuel DOLLFUS**, Deputy Director of AFD Cambodia; **Mr. Pierre Vincent**, Counsellor for Cooperation and Cultural Action, and Director of the French Institute of Cambodia, French Embassy; and **Mr. Bryan FORNARI**, Head of Cooperation at the EU Delegation to Cambodia. The ceremony was officially opened by **H.E. Dr. HANG Chuon Naron**, Deputy Prime Minister and Minister of Education, Youth and Sports. Following the opening, a high-level panel discussion was held on the topic "*Leveraging R&D for Innovation and Growth*" featured a discussion forum with **four panelists** from private and public institutes and moderated by **Assoc. Prof. Dr. Or Chanmoly**, Director of Research and Innovation Center of ITC. Prior to the discussion forum, insightful presentations were delivered by:

- **Prof. HENG Kreng**, Head of Management Committee of Department of Research and Innovation, and Deputy Secretary General of Secreteriate General of MoEYS, on topic: Research in Cambodian Higher Education
- **Dr. BUN Long**, Deputy Director of ITC, on topic: Strengthening Research and Innovation Ecosystems at ITC: Pathways to Impact
- **Mr. Hisakazu MAEDA**, Chief Representative of Branch of Panasonic Asia Pacific Pte. Ltd, on topic: Leveraging R&D to Build a Sustainable and Wellness Society

- **Dr. KHUN Kimang**, Chief Technology Officer, AI Farm Robotics Factory, on topic: Next-Gen Robotics: How AI is Transforming Human-Machine Interaction.

Forthermore, the opening ceremony were celebrated with the present of our **Honorable Guests** including Secretary of State, Under Secretary of State, and Directors of Department from the MoEYS; development and cooperation partners from European Union, French Embassy, AFD Cambodia, JICA Cambodia, IRD, AUF Cambodia, INACON project, and Panasonic company; our partner university and institutes in Cambodia such as National University of Management, University of Health Sciences, Battambang Teacher Education College, Royal University of Phnom Penh, Royal University Of Law And Economics, Royal University of Fine Arts, National Meanchey University. The ceremony was also participated by hundreds of ITC students as the witness and for gaining knowledge from the event.

The afternoon session was followed by **123 oral presentations** by national and international researchers and students. These presentations were organized into 11 parallel sessions, covering the following themes:

- Energy Tech for a Sustainable, Efficient, and Resilient Future
- Special Session for Food Technology, Research and Innovation Platform
- The Future of Food Production: Novel Technologies and Alternative Food Sources
- Building Resilient and Sustainable Food Supply Chains for the Future
- AI-Driven Digital Transformation
- Mechatronics Innovations and Control Systems
- Materials Science and Engineering
- Structure and Infrastructure Engineering
- Special Session for Project for Establishment of Risk Management Platform for Air Pollution
- Advanced Water Quality and Pollution Assessment
- Hydrology and Climate: Driving Innovation for Resilient Water Systems

Throughout the two-day event, a total of **42 exhibition booths** were showcased by local and international companies, program sponsors, universities, research and innovation centers, postgraduate schools, and international programs of the Institute of Technology of Cambodia. In addition, a Poster Presentation Session featuring 68 posters highlighted the work of researchers and students and was exhibited on both days. The event also welcomed approximately **1,500 high school students** from Phnom Penh and provincial schools, who participated in guided visits to various laboratories, along with distinguished guests and delegates.

On the second day, the organizers formally presented certificates of sponsorship to all sponsors and certificates of appreciation to both oral and poster presenters in recognition of their valuable contributions. Awards were also presented to five Best Oral Presenters and three Best Poster Presenters, all of whom were students. Additionally, certificates were awarded to all event sponsors. The event concluded with closing remarks delivered by **Mr. SOY Ty**, Deputy Director of the Institute of Technology of Cambodia. The detail schedule of the event is listed in the Event Agenda.

EVENT AGENDA



Agenda for the 14th Scientific Day
Theme: Leveraging R&D for Innovation and Growth
Venue: Institute of Technology of Cambodia
June 5-6, 2025

Day 1: Thursday, June 5, 2025

Time	Activities	Speakers/Presenters	Affiliation
Welcoming Remarks and Opening Ceremony MC: Mr. SENG Mean and Ms. CHOEUN Chakrya			
07:30 - 08:30	Registration	All Participants	
08:30 - 08:40	Welcome Remarks by the Institute of Technology of Cambodia (ITC)	**H.E. Prof. Dr. PO Kimtho**	Director General of ITC
08:40 - 08:50	Welcome Remarks by JICA Cambodia	**Mr. Kazumasa SANUI**	Chief Representative of JICA Cambodia
08:50 - 09:00	Welcome Remarks by AFD	**Mr. Emmanuel DOLLFUS**	AFD Cambodia Deputy Director
09:00 - 09:10	Welcome Remarks by the French Embassy	**Mr. Pierre VINCENT**	Counsellor for Cooperation and Cultural Action, and Director of the French Institute of Cambodia, French Embassy
09:10 - 09:20	Welcome Remarks by the Delegation of the European Union	**Mr. Bryan FORNARI**	Head of Cooperation
09:20 - 09:40	Opening Remarks by the Ministry of Education, Youth and Sport (MoEYS) of Cambodia	**H.E. Dr. HANGCHUON Naron**	Deputy Prime Minister and Minister of MoEYS
09:40 - 09:50	Group Photos	All Participants	
09:50 - 10:10	Coffee Break	All Participants	
Panel Discussion (4 Panelists) on Leveraging R&D for Innovation and Growth Moderator: Assoc. Prof. Dr. OR Chanmoly Venue: Auditorium of Samdec Akka Moha Sena Padei Techo HUN SEN			
10:10 - 11:40 (10 min each)	- Presentation on "Research in Cambodian Higher Education" by Prof. HENG Kreng, Head of Management Committee of Department of Research and Innovation, and Deputy Secretary General of Secreteriate General of MoEYS - Presentation on "Strengthening Research and Innovation Ecosystems at ITC: Pathways to Impact" by Dr. BUN Long, Deputy Director of ITC - Presentation on "Leveraging R&D to Build a Sustainable and Wellness Society" by Mr. Hisakazu MAEDA, Chief Representative of Branch of Panasonic Asia Pacific Pte. Ltd - Presentation on "Next-Gen Robotics: How AI is Transforming Human-Machine Interaction" by Dr. KHUN Kimang, Chief Technology Officer, AI Farm Robotics Factory - Discussion		
11:40 - 12:50	Lunch break		
Energy Technology Management (ETM) 1 Sub-parallel Session Energy Tech for a Sustainable, Efficient, and Resilient Future Chair: Dr. KRET Kakda, Co-Chair: Dr. SEANG Sirisokha Venue: A-417 (A Builing)			
13:00 -13:25	Guest Speaker: Oil and Gas Industry Toward Energy Security and Net-Zero Carbon	Dr. NEAK Kimhak	Chief of Joint Development Office, General Department of Petroleum
13:25 - 13:40	Energy Consumption Simulation of a Three-Story Villa Using DesignBuilder	Mr. BUN Panha	Master's Student, ITC
13:40 - 13:55	Calibration of DEM Parameters Using the Angle of Repose for Biomass Materials	Mr. CHEA Chantola	Master's Student, ITC
13:55 - 14:10	Maximizing PV Hosting Capacity in Low-Voltage Distribution Networks: A Case Study in Cambodia	Mrs. ENG Samphors	PhD Student, ITC
14:10 - 14:25	Heat Stress and Construction Labor Productivity: Identifying Key Factors Affecting Cambodian Construction Workers	Mr. HEANG Latin	PhD Student, ITC
14:25 - 14:40	3D Modeling of a Continuous Mixing Machine for Biomass Briquette Systems	Mr. HOEUNG Hollin	Master's Student, ITC
14:40 - 14:55	A Review of Environmental Factors Affecting Students' Cognitive Performance	Mr. KOR Pheak	Lab Assistance, ITC
14:55 - 15:10	Break		
15:10 - 15:25	Optimal Number of EV Charger Station in Distribution System and Estimates 24-hour EV load profiles	Ms. OENG Kechki	Master's Student, ITC
15:25 - 15:40	Coordinated Reactive Power Control Using Genetic Algorithm for Loss Minimization in PV-Integrated Distribution Networks	Mr. SREY Sak	Master's Student, ITC
15:40 - 15:55	A Frequency Stability Analysis for BESS Placement under High Penetration of IBRs (Solar PV) Considering Load and Solar PV Location	Mr. SREYNATTHA Pagnasatyavorn	Master's Student, ITC
15:55 - 16:10	Inverter Control Strategies for Active and Reactive Power in Photovoltaic Production for Grid Integration	Ms. TY Tangly	Engineering Student, ITC
16:10 -16:25	Electric Field Analysis on Distribution Suspension Insulator with Different Outer Diameter	Mr. VUTH Sireyroth	Engineering Student, ITC
16:25 -16:40	Economic Analysis of Solar Rooftop at Toyota Cambodia Office in Phnom Penh by Using PVsyst Software: EAC OLD vs New Regulation	Mr. PROM Pisith	Master's Student, ITC
Food Technology and Nutrition (FTN) 3 Sub-parallel Sessions			
FTN-1: Special Session for Food Technology, Research and Innovation Platform Chair: Asst. Prof. Dr. IN Sokneang, Co-Chair: Dr. MORM Elen Venue: A-109 (A Building)			
13:00 -13:20	Opening and Welcome Remark by Capfish Project	Dr. THOMBATHU Seetharama Shetty	Chief Technical Advisor, CAPFish-UNIDO
13:20 - 13:40	Guest Speaker: The Important of University and Industry Cooperation (UIC)	Dr. LY Sokny	Deputy General Director, MISTI
13:40-14:00	Development of Oyster Sauce Using Japanese Oysters, Green Mussels, and Clams (RP013-PTCP-07)	Dr. SAY Manit	Lecturer, ITC

14:00-14:20	Application of Locally Fabricated Hybrid Solar Dryer for Improving Dried Fish Quality and Increasing Value Added for Smallholder Fish Processors (RP01-PTCP-01)	Dr. HIN Lyhour	Lecturer, RUA
14:20-14:40	Development of Instant Fish Soups for Commercialization (RP06-PTCP 02)	Ms. NET Marinich	Lecturer, ITC
14:40-15:00	The Development of Snack Products from Low-Value Fish through Dehydration Methods (RP07-PTCP03)	Mr. TONG Socheath	Lecturer, RUA
15:00-15:20	Break		
15:20-15:40	Development of Nutrient-dense Wafer Roll for Children by Incorporating Water Fish Powder (RP08-PTCP04)	Dr. EK Pickmony	Lecturer, ITC
15:40-16:00	Effect of Steaming and Drying on Physicochemical and Microbiological Properties of Green Mussels (RP-	Ms. HOEUN Seanghai	Lecturer, ITC
16:00-16:20	Development of Fish Sausage Product using Pangasius (Pangasius bocourti) Species (RP011-PTCP06)	Dr. CHHE Chinda	Chef of Research and Development STIC Office, NISTI
16:20-16:40	Development of Fish Hydrolysate Fertilizers from Fish Wastes and Discards using Different Methods, and their Effects on Nutrients Contents and Foodborne Microbial Contaminations (RP022-PTCP10)	Dr. THO Kim Eang	Lecturer, RUA
16:40-17:00	Closing Remark	Dr. THOMBATHU Seetharama Shetty	Chief Technical Advisor, CAPFish-UNIDO
FTN-2: The Future of Food Production: Novel Technologies and Alternative Food Sources Chair: Assist. Prof. Dr. MITH Hasika, Co-Chair: Dr. PHUONG Hengsim Venue: A-110 (A Building)			
13:00 -13:25	Guest Speaker: Converting MeeChiet Instant Noodle Product Prototype Into Feasibility Study	Mr. BUN Than	Manager, Business Operations of One More Manufacturing Co., Ltd
13:25 - 13:40	Isolation and Characterization of Salmonella Bacteriophages from Chicken Breast Samples in Phnom Penh Markets as a Biocontrol Agent	Mr. VUTH Sireyvathanak	Lecturer, RUPP
13:40 - 13:55	Assessment of Antimicrobial Resistance Patterns in Salmonella Isolates from Different Food Matrices in Cambodia	Ms. SOVANN Chakriya	Master's Student, ITC
13:55 - 14:10	Study on Consumer Preferences and Development of the Sourdough Bread-making Process with Added Flavors	Mr. CHEA Sokly	Engineering Student, ITC
14:10 - 14:25	Effects of Monoglyceride and Guar Gum on the Quality and Sensory Properties of Mango Frozen Yogurt	Ms. HUOR Sodaneth	Engineering Student, ITC
14:25 - 14:40	Nutritional Quality of Farmed Striped Snakehead (C. striata): Fatty Acid Profile and Proximate Composition Analysis	Mr. LAY Sovannmony	PhD Student, ITC
14:40 - 15:10	Break		
15:10 - 15:25	Optimisation of Method Liquid Extraction of Aroma Compound (2-acetyl-1-pyrroline) from Phka Rumdoul Rice Variety	Ms. SARY Sokharoath	Master's Student, ITC
15:25 - 15:40	Comparing the Quality of Fermented Soy Sauces Made by Using Whole Soybeans and Defatted Soybeans	Mr. LY Luka	PhD Student, ITC
15:40 - 15:55	Comparing the Physicochemical Properties of Soybean Oil Blended with Different Oils	Ms. MICH Monika	Lecturer, ITC
FTN-3: Building Resilient and Sustainable Food Supply Chains for the Future Chair: Assist. Prof. Dr. KHOEURN Kimleang, Co-Chair: Dr. NGET Sovannmony Venue: A-221 (A Building)			
13:00 - 13:25	Guest Speaker 1: Importance of IISO/IEC 17025-Accredited Laboratories in Ensuring Food Safety Compliance	Dr. SRENG Navin	Head of the Laboratory of Environment and Food Safety, IPC
13:25 - 13:50	Guest Speaker 2: Cropping Practices and Use of Pesticides in Rice	Dr. Sester MATHILDE	Senior Researcher, CIRAD. UPR AIDA. ITC
13:50 - 14:05	Identification of Cuticle-binding Bacteria from Meloidogyne graminicola as Potential Biocontrol Agent Affecting the Pest Mobility	Ms. YAN Sreynich	Master's Student, ITC
14:05 - 14:20	Evaluation of Nutrient Profiles of Single and Mixed Cover Crops Species Collected from Rovieng, Preah Vihear Province	Mr. HAN Menghuy	Engineering Student, ITC
14:20 - 14:35	Evaluation of Rice Seed Quality of Cambodian Varieties Produced under Conventional and Agroecology Practices	Ms. HO Singy	Master's Student, ITC
14:35 - 14:50	Influence of the Soil Physicochemical Properties on Phka Rumdoul Rice Quality and Yield in Cambodia	Ms. VICHETH Vichny	Master's Student, ITC
14:55 - 15:10	Break		
15:10 - 15:25	Analysis of Pesticide Residues in Rice Grains using QuEChERS Method Followed by GC-MS	Ms. HENG Aykim	Engineering Student, ITC
15:25 - 15:40	Assessment of Pesticide Residues in Rice and Water from Kanghot Rice Fields in Battambang	Mr. BUN Leanghok	Research Assistant, ITC
15:40 - 15:55	Assessment of Pesticide Residues in Rice and Irrigation Canal: A Study of Agrochemical-Free Farming in Rovieng District, Preah Vihear Province	Mr. LY Rethireach	Engineering Student, ITC
15:55 - 16:10	Seasonal Pesticide Residues in Surface Waters of Banteay Meanchey Province, Cambodia	Mr. TIT Vivath	Engineering Student, ITC
16:10 - 16:25	Development of Dishwashing Liquid Through Utilization of Various Active Ingredients	Mr. VAPHAT Kosal	Engineering Student, ITC
Mechatronic and Information Technology (MIT) 2 Sub-parallel Sessions			
MIT-1: AI-Driven Digital Transformation Chair: Assist. Prof. Dr. VALY Dona, Co-Chair: Assist. Prof. Dr. PHAUK Sokkhey Venue: J-701 (J Building)			
13:00 - 13:25	Guest Speaker 1: AI Meets Public Data: A Prompt-Based Agent for Intuitive Analysis of Government Datasets	Dr. BUOY Rina	Applied AI Researcher, Techo Startup Center
13:25 - 13:50	Guest Speaker 2: Digital Transformation in Education	Ms. SEAK Leng	Deputy Head of GIC, ITC
13:50 - 14:05	Uncovering Factors Contributing to Cambodian High School Students' Mathematics Achievement through Machine Learning Algorithms	Mr. LY Tong	Student, Royal Academy of Cambodia
14:05 - 14:20	Harnessing Predictive Analytic for Detecting Students'Learning and Behavioral Engagement in ITC Moodle Platform	Mr. PEN Chentra	Lecturer, ITC
14:20 - 14:35	Applying Deep Learning Approaches to Khmer Speech Recognition and Text-to-Speech in Conversational Domains	Ms. SOENG Prakmlis	Master's Student, ITC
14:35 - 14:50	AI-Based Skill Trend Analysis: Enhancing Job Market Insights and Curriculum Development in Cambodia	Ms. SONGEAM Sela	Master's Student, ITC
14:55 - 15:10	Break		
15:10 - 15:25	Lip-Synchronized Talking Avatar Generation for Khmer Language Using Audio-to-Landmark Modeling	Mr. YANN Sovanvichea	Master's Student, ITC
15:25 - 15:40	KHR/USD Daily Exchange Rate Short Term Prediction using Deep Learning	Mr. HORT Hourpheng	Master's Student, ITC
15:40 - 15:55	A Multilingual Deep Learning Approach Fine-tuning to Khmer News Summarization	Mr. KOUM Soknan	Engineering Student, ITC
15:55 - 16:10	Handwritten Khmer Digit Recognition Using Sequential RNN Models	Mr. PHOEURN Kimhor	Engineering Student, ITC
16:10 - 16:25	Analysis on Machine Learning Models for Imbalanced Data Problem in Payment Fraud Detection	Ms. SEUN Siuphing	Master's Student, ITC
16:25 - 16:40	Interacting Neural Networks: A Comparative Study of Standard and Distributed Architectures on Image Classification	Ms. UANN Sreyvi	Master's Student, ITC
16:40 - 16:55	Enhancing Khmer Language AI: A Privacy-Focused Chatbot with Retrieval-Augmented Generation	Ms. LI Nita	Master's Student, ITC

MIT-2: Mechatronics Innovations and Control Systems Chair: Assist. Prof. Dr. SRANG Sarot, Co-Chair: Assist. Prof. Dr. PEC Rothna Venue: J-702 (J Building)			
13:00 - 13:20	Guest Speaker 1: How to Make Yourself be the Source of Innovation in Informatics Research	Prof. Moriyama TSUYOSHI	Tokyo Polytechnic University, Japan
13:20 - 13:40	Guest Speaker 2: Support for Primary Industries Using Life Mechatronics Technology	Prof. Toda MASAHI	Kumamoto University, Japan
13:40 - 14:00	Guest Speaker 3: AlphaZero for Ouk Chaktrang: leveraging self-play for Khmer Chess	Dr. KHUN Kimang	AI Farm Robotics Factory
14:00 - 14:15	Balloon-observed Gravity Wave Momentum Flux Reconstruction using Transfer Learning Bayesian Neural Network	Dr. HAS Sothea	Researcher, ITC
14:15 - 14:30	Comparing Time Series with Machine Learning and Deep Learning Models for Paddy Rice Price Forecasting in Battambang Province, Cambodia	Ms. LENG Phirum	Master's Student, ITC
14:30 - 14:45	The Forecasting for Cambodia's Rice Production Using Multiple Linear Regression and Tree-Based Ensemble Learning Methods	Mr. THONG Sithach	Master's Student, ITC
14:45 - 15:00	Model-Based PMSM Control in the dq Frame Using Full-State Feedback with Integral Action and EKF for State Estimation under Stochastic Dynamics	Mr. CHHEAN Rotanak	Master's Student, ITC
15:00 - 15:10	Break		
15:10 - 15:25	Reconstructing Higher-Order Interaction in Algal Communities	Mr. CHHORN Danut	Master's Student, ITC
15:25 - 15:40	Continuously Uploading ECG Waveform Data to Google Sheet using ESP32	Ms. LEANG Phalla I Ning	Engineering student, ITC
15:40 - 15:55	Toward a Case Study on Event Detection in Non-Intrusive Load Monitoring Using Geometric Features of the Cumulative Sum	Mr. SOK Song	PhD Student, ITC
15:55 - 16:10	Investigating Efficiency of PMSM Using Full State Feedback Control with Integral Action and Dynamic Compensation	Mr. TAING Bunsrung	Master's Student, ITC
16:10 - 16:25	Predictive Modeling of Garment Trade Flows in Cambodia: A Data-Driven Approach to Forecasting Export Trends	Mr. EM Dara	Master's Student, ITC
Materials Science and Structure (MSS) 2 Sub-parallel Sessions			
MSS-1: Materials Science and Engineering Chair: Dr. LIM Sovanvichet, Co-chair: Dr. OENG Thaileng Venue: J-703 (J Building)			
13:00 - 13:15	Analysis of Iron Ores and Slags from Archeological Sites by Using Energy Dispersive X-ray Spectroscopy and X-ray Diffraction Technique	Ms. DUM Sreynoch	Student, Royal University of Phnom Penh
13:15 - 13:30	Investigation on Mechanical Properties and Antimicrobial Activity of Natural Rubber Latex Examination Gloves Incorporated with Pomelo Peel Powder	Ms. HENG Likea	Engineering Student, ITC
13:30 - 13:55	Guest Speaker 1: Recycling Construction and Demolition Wastes: Opportunities for Innovation	Prof. Luc COURARD	University of Liege, Belgium
13:55 - 14:20	Guest Speaker 2: Glass from Making to Breaking: Experiment vs. Simulation	Dr. TO Theany	Université de Rennes, France
14:20 - 14:35	Degradation Study of Few-Layer SnP2X6 (X=S, Se) in Ambient Conditions Using OM, AFM, and SEM	Mr. HOR Sok Heng	PhD Student, Royal University of Phnom Penh
14:35 - 14:50	Lithology, Geochemistry, Hydrothermal Alteration and Mineralization of Cu-Mo Porphyry Deposits at Canada Wall Prospect, Andong Meas District, Ratanakiri Province, Cambodia.	Mr. IEM Mengly	Engineering Student, ITC
14:50 - 15:10	Break		
15:10 - 15:25	Design and Modelling of Cost-effective Manufacturing of Glass Tempering Furnace Used in Industrial Applications.	Mr. KEAT Rayuth	PhD Student, ITC
15:25 - 15:40	Investigation of Utilization of Natural Oils as Processing Aids and Activators Alternative to Naphthenic Oils in Rubber Composites	Ms. KONG Sreynich	Engineering Student, ITC
15:40 - 15:55	Mineralogical and Physicochemical Characterizations of Silica Sand by Using Microbeam, XRD, and XRF	Ms. MEUN Muylin	Student, Royal University of Phnom Penh
15:55 - 16:10	Determination of Crystallite Size of Transition Metal Oxides (NiO, Mn_3O_4, and α-Fe_2O_3) Using the Scherrer Formula and Williamson–Hall Plot	Ms. SREANG Sokly	Student, Royal University of Phnom Penh
16:10 - 16:25	Utilization of Fly Ash as an Eco-friendly Filler in Styrene Butadiene Rubber Compound for Tire Tread Application	Mr. YIEM Kakda	Engineering Student, ITC
16:25 - 16:40	Development of Low-carbon Concrete using Biomass Burning Ash for Port and Harbor Facilities	Mr. CHEA Sovannarith	Graduate of Kagoshima University/ Yonemori Construction Co., LTD.
MSS-2: Structure and Infrastructure Engineering Chair: Dr. DOUNG Piseth, Co-chair: Dr. CHHUN Kean Thai Venue: J-704 (J Building)			
13:00 - 13:15	Resilient Modulus of Lime-treated Materials for Pavement Design in the Cambodian Context	Ms. BUOY Chanbo	Master's Student, ITC
13:15 - 13:30	Deploying ANPR Technology to Detect Traffic Violation at Intersections in Phnom Penh	Mr. KEO Sambath	Master's Student, ITC
13:30 - 13:55	Guest Speaker 1: Strategic Approaches to Infrastructure Maintenance and Seismic Design in Japan	Prof. Toshinobu YAMAGUCHI	Kagoshima University, Japan
13:55 - 14:20	Guest Speaker 2: Frontier in Architectural Engineering	Prof. Junne KIKATA	Kagoshima University, Japan
14:20 - 14:35	3D Road Surface Reconstruction: A Case Study of Indoor Corridor Using RPLiDAR 360° A2M8	Mr. MAO Phearun	Master's Student, ITC
14:35 - 14:50	Assessing the Impact of Motorcycle Emissions on Urban Air Quality: A Case Study at the Institute of Technology of Cambodia	Mr. PHINY Lihong	Master's Student, ITC
14:50 - 15:10	Break		
15:10 - 15:35	Guest Speaker 3: Data-driven Urban Research toward Better Life Space Designing	Assoc. Prof. Hideo SHIRAISHI	Kanazawa University, Japan
15:35 - 15:50	Correlation Between Dynamic Cone Penetrometer (DCP) and California Bearing Ratio(CBR) for Subgrade Soil Materials in Cambodia	Mr. PHO Bopheaktra	Master's Student, ITC
15:50 - 16:05	Geotechnical Properties of Clay Soil with River Sand for Sustainable Road Embankment Applications	Mr. SOK Samnang	Student, Techo Sen Institute of Public Works and Transport
16:05 - 16:20	Machine Learning-Based Prediction of Public Bus Ridership: A Case Study in Phnom Penh City	Mr. SRORN Piseth	Master's Student, ITC
16:20 - 16:35	Evaluation of Modified Asphalt Binders Incorporating Recycled Plastic Waste for Enhanced Pavement Performance	Mr. TENG Chy	Master's Student, ITC
16:35 - 16:50	Improving the Durability and Strength of Expansive Soils Stabilized with Emulsion Asphalt and Cement	Mr. YOEUNTHY Vanntheng	Master's Student, ITC
Water and Environment (WAE) 3 Sub-parallel Sessions			
WAE-1: Special Session for Project for Establishment of Risk Management Platform for Air Pollution Chair: Prof. Furuuchi MASAMI; Co-chair: Assist. Prof. PENG Chanthol Venue: I-106 (I Building)			
13:00 - 13:25	Guest Speaker 1: Source Analysis of High PM2.5 Concentration Events During the 2025 Dry Season in Phnom Penh, Cambodia, Using Organic Tracer Compounds.	Assoc. Prof. Fumikazu IKEMORI	Graduate School of Integrated Science and Technology, Nagasaki University, Japan
13:25 - 13:50	Guest Speaker 2: A recent study of Ultrafine Particles (PM0.1) toward Low-carbon Societies in Upper Southeast Asia.	Dr. Worradorn Phairuang	Department of Geography, Faculty of Social Sciences, Chiang Mai University, Thailand
13:50 - 14:15	Guest Speaker 3: Air Quality Management and Policy	Mr. YIM Raksmey	Deputy Director of Air Quality, Noise, and Vibration Management, Ministry of Environment

14:15 - 14:30	Preliminary Assessment of Air Quality in the Workplace in Phnom Penh	Dr. YOEUN Sereyvath	Researcher, ITC
14:30 - 14:45	Variations in Particulate Matter (PM2.5) Concentration in Phnom Penh: The Role of Wind Speed and Wind Direction During Daytime and Nighttime	Dr. SAO Sochan	Researcher, ITC
14:45 - 15:00	Gravimetric Analysis of PM2.5 Mass Concentration: Daily Behavior in Early Dry season in Phnom Penh, Cambodia	Mr. SARY Vahet	Engineering Student, ITC
15:00 - 15:15	Driving Factors of Particulate Matter (PM2.5 and PM10) Dynamics During Heavy Air Pollution in Phnom Penh, Cambodia	Ms. SRONG Sokkeav	Engineering Student, ITC
15:15 - 15:30	Break		
15:30 - 15:45	Preliminary Assessment of Indoor Air Quality in Residences in Phnom Penh	Dr. THANH Channmuny	Researcher, ITC
15:45 - 16:00	Daytime and Nighttime Monitoring of Ultrafine Particulate Matter (PM0.1) during Dry Season in Phnom Penh, Cambodia	Mr. VORNG Chanrat	Engineering Student, ITC
16:00 - 16:15	Variability of PM2.5 Distribution in Relation to Gaseous Pollutants Concentrations in Phnom Penh, Cambodia	Ms. YOU Titsothearna	Engineering Student, ITC
16:15 - 16:30	Characterization of Bacterial Communities in Aquatic Environments of Phnom Penh as Potential Bioaerosol Sources Using Nanopore Sequencing	Ms. SAMBATH Saksoleakhena	Researcher, ITC
WAE-2: Advanced Water Quality and Pollution Assessment Chair: Dr. MUON Ratha; Co-chair: Dr. SANG Davin Venue: F-S1 (F Building)			
13:00 - 13:25	Guest Speaker: Macroplastics in Rivers: Present Issues, Challenges and Solutions in Cambodia	Dr. KET Pinnara	Vice-Dean, Faculty of Hydrology and Water Resources Engineering, ITC
13:25 - 13:50	Assessment of Water Quality along Bassac River and Its Tributaries in Agricultural Communities	Ms. CHANTO Monychot Tepy	PhD Student, ITC
13:50 - 14:05	Chemical Characterization of Acid Mine Drainage from Artisanal Mine Site in Chong Plash Commune, Keo Seima District, Mondulkiri Province, Cambodia	Ms. CHEA Srey Noch	Engineering Student, ITC
14:05 - 14:20	Impact of City Activities on Macroplastic Accumulation in Two Tonle Sap Lake Tributaries	Mr. CHHENG Por Y	Engineering Student, ITC
14:20 - 14:35	Monitoring and Evaluation of Macroplastic Pollution In Siem Reap River	Ms. HENG Chanmonida	Engineering Student, ITC
14:35 - 14:50	Photoinduced Reactive Intermediates Formation from Microplastics and Leachate Dissolved Organic Matter	Ms. KAING Vinhteang	PhD Student, Institute of Science Tokyo
14:55 - 15:10	Break		
15:10 - 15:25	Water Quality Assessment of Groundwater in Mondolkiri Province, Cambodia: Implications for Agriculture	Ms. CHAN Solita	Engineering Student, ITC
15:25 - 15:40	Assessing Macroplastic Pollution in Kampong Khleang River: Urban vs. Rural Pollution and Seasonal Variability	Ms. MENH Leanggek	Researcher, ITC
15:40 - 15:55	Monitoring Macroplastic Accumulation in the Sangker River: Distribution Patterns and Pollution Hotspots	Ms. PHEAP Phoungsy	Engineering Student, ITC
15:55 - 16:10	Pesticide Contamination in Groundwater of Agriculturally Intensive Regions of Cambodia: A Case Study of Koh Thum and Kanghot	Mr. RUOS Vuthy	Engineering Student, ITC
16:10 - 16:25	Development of Coconut Filter to Remove Microplastics at Sea Salt Farm Kampot Province, Cambodia	Ms. SAM Sonarith	Engineering Student, ITC
WAE-3: Hydrology and Climate: Driving Innovation for Resilient Water Systems Chair: Dr. THENG Vouchlay; Co-chair: Mr. SOK Kimhuy Venue: F-S2 (F Building)			
13:00 - 13:25	Spatio-temporal Distribution of Electric Conductivity, Stable Isotopes of Water and Water Level along Tonle Sap River	Assist. Prof. Dr. DOUNG Ratha	Researcher, ITC
13:25 - 13:50	Daily Tidal Dynamics at Chaktomuk Confluence by a 2D Hydrodynamic Model	Dr. PEN Sytharith	Researcher, ITC
13:50 - 14:05	Dynamic Seasonal Variations in River Mixing Patterns at a Major Confluence-Diffluence Site (Chaktomuk Junction, Cambodia) Revealed by Hydrochemical Tracing	Mr. CHORK Vuthy	Researcher, ITC
14:05 - 14:20	Historical Changes In Extreme Daily Temperature Indices In The Last 3 Decades In Cambodia	Assist. Prof. Dr. SOK Ty	Researcher, ITC
14:20 - 14:35	Contemporary Changes in Precipitation Patterns and Extreme Rainfall Indices in Cambodia	Mr. HOUT Meng Hour	Researcher, ITC
14:35 - 14:50	Assessment of Inundation at Prek System of Koh Thum Inducing by Bassac River and Regional Precipitation Pattern	Mr. LY Kim Hort	Engineering Student, ITC
14:55 - 15:10	Break		
15:10 - 15:25	Review on Historical Tropical Hurricanes, Typhoons, and Cyclones Frequency, Intensity, and Trajectory	Mr. CHHENG Por Y	Engineering Student, ITC
15:25 - 15:40	Evaluating Water-Saving Irrigation Methods for Rice Cultivation, Case Study in Cambodia	Ms. PHOL Mengheak	PhD Student, ITC
15:40 - 15:55	The Dependency of the Planting Date on Rainfall in Paddy Fields in the Pursat River Basin	Ms. SIENG Kimhor	Engineering Student, ITC
15:55 - 16:10	Hydraulic Propagation of Pesticides at the Watershed Scale: Tonle Sap River Case Study	Ms. CHIN Ech	Master's Student, ITC
Day 2: Friday, June 6, 2025 Venue: Auditorium of Samdec Akka Moha Sena Padei Techo HUN SEN			
10:30 - 11:15	Delivering Best Presetnation and Best Poster Awards, and Delivering Certificate of Appreciation to Sponsors Representative: **Assoc. Prof. Dr. OR Chanmoly**, Director of Research and Innovation Center, ITC Venue: Auditorium of Samdec Akka Moha Sena Padei Techo HUN SEN		
11:15 - 11:30	Closing Remarks **Dr. BUN Long**, Deputy Director of ITC Venue: Auditorium of Samdec Akka Moha Sena Padei Techo HUN SEN		
Parallel Activities			
Day 1: Thursday, June 5, 2025 Venue: Auditorium of Samdec Akka Moha Sena Padei Techo HUN SEN			
09:00 - 10:30	**Poster Exhibition** Venue: Ground floor, A-Building		
09:00 - 17:00	**Booth Exhibitions** Venue: Ground floor, G-Building		
Lab Tours			

08:00 - 10:00	**1st Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopy Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors
10:00 -12:00	**2nd Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopy Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors
14:00 - 16:00	**3rd Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopy Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors
Day 2: Friday, June 6, 2025		
09:00 - 10:30	**Poster Exhibition** Venue: Ground floor, A-Building	
09:00 - 17:00	**Booth Exhibitions** Venue: Ground floor, G-Building	
Lab Tours		
08:00 - 10:00	**4th Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopy Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors
10:00 -12:00	**5th Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopy Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors

14:00 - 16:00	**6th Lab Tours** 1. Construction Materials Lab (101A) 2. Food Processing Lab (201B) 3. Library (Ground Floor-I) 4. Mechanical Workshop 2 (113J) 5. Microscopic Lab (116B) 6. E-learning Center (221B) 7. Electrical Power Supply System (407J) 8. Water and Environmental Lab (105D) 9. GAR Workshop (101K) 10. Transport and Infrastructure (402K) 11. Food Processing Technology Hall (313J) 12. Network and Security Lab (507J) 13. Geotechnical lab (303H)	- High School Students - Visitors

ORAL SESSION

Energy Technology Management (ETM)

1 Sub-parallel Session

"Energy Tech for a Sustainable, Efficient, and Resilient Future"

Chair: Dr. KRET Kakda, Co-Chair: Dr. SEANG Sirisokha

No.	Topic
1	Energy Consumption Simulation of a Three-Story Villa Using Design Builder *Authored by: Panha BUN, Kinnaleth VONGCHANH, Sarin CHAN*
2	Calibration of DEM Parameters Using the Angle of Repose for Biomass Materials *Authored by: Chantola CHEA, Kinnaleth VONGCHANH, Hollin HOEUNG, Latin HEANG, Sarin CHAN*
3	Maximizing PV Hosting Capacity in Low-Voltage Distribution Networks: A Case Study in Cambodia *Authored by: Samphors ENG, Julien ALLARD, Vannak VAI, François VALLEE*
4	Heat Stress and Construction Labor Productivity: Identifying Key Factors Affecting Cambodian Construction Workers *Authored by: Latin HEANG, Kinnaleth VONGCHANH, Sarin CHAN*
5	3D Modeling of a Continuous Mixing Machine for Biomass Briquette Systems *Authored by: Hollin HOEUNG, Kinnaleth VONGCHANH, Chantola CHEA, Latin HEANG, Sarin CHAN*
6	A Review of Environmental Factors Affecting Students' Cognitive Performance *Authored by: Pheak KOR, Kinnaleth VONGCHANH*
7	Optimal Number of EV Charger Station in Distribution System and Estimates 24-hour EV load profiles *Authored by: Kechki OENG, Vannak VAI, Chhith CHHLONH*
8	Coordinated Reactive Power Control Using Genetic Algorithm for Loss Minimization in PV-Integrated Distribution Networks *Authored by: Sak SREY, Vannak VAI, Kimsrornn KHON*
9	A Frequency Stability Analysis for BESS Placement under High Penetration of IBRs (Solar PV) Considering Load and Solar PV Location *Authored by: Pagnasatyavorn SREYNATTHA, Vannak VAI, Kimsrornn KHON*
10	Inverter Control Strategies for Active and Reactive Power in Photovoltaic Production for Grid Integration *Authored by: Tangly TY, Menghorng BUN, Vannak VAI, Sievlong SUK*
11	Electric Field Analysis on Distribution Suspension Insulator with Different Outer Diameter *Authored by: Sireyroth VUTH, Chansideth VUTHY, Sreylen VA, Ryda VANN, Lyhour YOU*
12	Economic Analysis of Solar Rooftop at Toyota Cambodia Office in Phnom Penh by Using PVsyst Software: EAC OLD vs New Regulation *Authored by: Piseth PROM, Vannak VAI, Kimsrornn KHON*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Energy Consumption Simulation of a Three-Story Villa Using DesignBuilder

Bun Panha [1], Kinnaleth Vongchanh [1,2,*], Chan Sarin [1]

[1] *Thermal Lab, Department of Industrial and Mechanical Engineering, Faculty of Electrical Engineer, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: kinnalethv@itc.edu.kh*

Abstract: *Cambodia's rapid economic growth has positioned the building sector as one of the primary contributors to overall energy demand. The country's tropical climate, combined with this rapid development, has substantially increased energy consumption and raised concerns about long-term sustainability. This study employs DesignBuilder, a widely used building performance simulation tool, to model the energy use of a typical three-story villa (approximately 260 m² of floor area) in Phnom Penh and to assess the impact of seasonal variations on consumption patterns. Key parameters such as building dimensions, envelope materials, lighting systems, air-conditioning units, occupant behavior, and electricity bills, were collected to ensure realistic modeling. The household includes multiple energy-consuming appliances such as air conditioners, lighting, fans, a washing machine, and a refrigerator, and has an average monthly energy consumption in 2024 about 380 kWh, which rises significantly to average 510 kWh during the hot season (March to May). By comparing simulated results with measured data, this study validates DesignBuilder's capacity to accurately model residential energy consumption in Cambodia. By focusing on Cambodia's specific climate and occupant behaviors, this research sets itself apart from broader tropical studies and provides practical insights for architects, engineers, and policymakers. The findings highlight how thoughtful design refinements and behavioral changes can foster greater energy efficiency, reduce costs, and enhance thermal comfort in Cambodia's rapidly expanding building sector.*

Keywords: Energy usage intensity; Energy efficiency; Residential building; DesignBuilder.

1. INTRODUCTION

In Cambodia's rapidly growing economy, the building sector has become the largest final energy consumer. According to national assessments, the residential sector alone accounted for approximately 26.7% of total final energy consumption in 2022 [1]. Notably, it represented around 36% of the country's electricity use, totaling 17,313 terajoules [2]. In response, Cambodia's energy policy has set ambitious targets: the National Energy Efficiency Policy (2022–2030) calls for a 19% reduction in total energy use by 2030 (relative to business-as-usual), including a 34% cut in residential consumption [2,3]. Likewise, the 2022 National Cooling Action Plan (NCAP) emphasizes that "significant and growing energy consumption" must be addressed by controlling cooling loads through efficient design and technology [5]. These policies underscore the importance of curbing residential electricity growth, especially for cooling.

Cambodia's challenge mirrors a wider regional trend. Across Southeast Asia, space-cooling electricity use has risen steeply—from ~12 TWh in 1990 to ~80 TWh in 2020—and is projected to approach 300 TWh by 2040 if current policies remain unchanged [6,7]. By then, air-conditioning could account for nearly one-fifth of all electricity consumed in ASEAN, illustrating the scale of the efficiency opportunity. [7].

Since large-scale field experiments are costly and time-consuming. Tools such as DesignBuilder, which is powered by the EnergyPlus engine, are widely used to model building energy consumption under local climate conditions. These simulations enable the analysis of envelope materials, shading, HVAC systems, and occupant behavior. However, the accuracy of simulation outputs depends heavily on the quality and accuracy of input data. Arbulu *et al.* (2024) showed that a calibrated model of a residential building

[1]* Corresponding author: Kinnaleth Vongchanh
E-mail: kinnalethv@itc.edu.kh*; Tel: +855-99 351 199*

reduced prediction errors dramatically – improving hourly indoor-temperature accuracy by about 67% and heating-energy predictions by ~16% compared to a standard model [8]. Likewise, field studies have demonstrated that incorporating measured utility and usage data into a simulation model can yield simulated consumption figures that closely match real bills. Choun *et al.* (2024) modeled a Phnom Penh apartment using EnergyPlus (via the OpenStudio SketchUp plug-in). The model was calibrated against actual monthly electricity bills. The simulated annual energy use matched the billed consumption to within just 1.78% [9]. These calibration efforts are crucial in humid climates, where occupant behavior (window use, thermostat settings, etc.) greatly affects energy use.

Building on that evidence, this study evaluates the energy performance of a three-storey villa in Phnom Penh. We first develop a high-resolution DesignBuilder model of the dwelling's geometry, envelope and mechanical systems. Measured utility bills, sub-metered air-conditioner data and on-site occupancy surveys are then used to calibrate key input parameters—schedules, equipment efficiencies, shading factors—until the model reproduces observed monthly electricity within accepted ASHRAE Guideline 14 error bounds. Ultimately, the objective is to establish a reliable methodology for predicting and analyzing residential energy use in Cambodia, thereby informing designers and policymakers on how to improve efficiency in the nation's rapidly expanding building sector.

2. METHODOLOGY

2.1 Preliminary Data Collection

The case study focuses on a three-story residential villa in Phnom Penh, Cambodia. Monthly electricity consumption data were collected over a full year (January–December 2023) from utility bills to serve as baseline performance data. Choun *et al.* (2024) similarly used 12 monthly bills to calibrate an EnergyPlus model of a Phnom Penh flat [9]. The homeowner also provided detailed architectural drawings in AutoCAD (DWG) format. These inputs are foundational for both model creation and calibration.

- Electricity bills: Twelve monthly utility bills of 2024 were gathered and tabulated.
- Architectural plans: Floor plans and elevations in AutoCAD DWG format were obtained from the building's design team.

These data allow the simulation model to be tied to actual performance and ensure the geometry matches the real building.

2.2 Geometry Modeling

The DWG drawings were first simplified to include only the essential floor plan outlines (walls, doors, windows) and then exported to DXF format for compatibility with DesignBuilder. In DesignBuilder, each 2D floor plan was imported and extruded into 3D space to form a complete building model. Interior partitions, floor slabs, and roof levels were defined according to the plans. All exterior walls and roofs were classified as thermal zones.

- The converted DXF was imported into DesignBuilder to create the 3D mass model.
- Each floor's plan view was stacked with appropriate inter-floor spacing to match the actual building height.
- Openings were added like doors and windows with the same size from the plans.

Using AutoCAD in the DXF geometry, the window-to-wall ratio (WWR) was calculated for each façade by measuring the total window area and dividing by the wall area. The measured WWR values were then input into DesignBuilder to ensure the model accurately reflects glazing proportion.

$$\text{WWR (\%)} = \frac{\text{Total Wall Area}}{\text{Total Window Area}} \times 100\% \qquad \text{(Eq. 1)}$$

2.3 Schedule Assignment

Internal schedules for occupancy, lighting, and equipment were specified using typical residential patterns. DesignBuilder's built-in templates for "residential" usage schedules were adapted. For example, weekday occupancy was set high during morning and evening hours and low during the daytime, reflecting family living patterns. Lighting and plug-load schedules were similarly defined based on typical hours of use for a home (e.g. high lighting in the evening, low at night) with a constant monthly profile. HVAC operation schedules were set according to the occupant use their air condition system.

2.4 Weather Data

The simulation used a local weather file for Phnom Penh. Specifically, a Typical Meteorological Year (TMY) EPW file for Phnom Penh was obtained and loaded into DesignBuilder. This file contains hourly values of temperature, humidity, solar radiation, and other climate variables for the region.

Using real weather data ensures that seasonal cooling loads are accurately captured.

2.5 Simulation and Calibration

With geometry, schedules, envelope, internal gains, and weather defined, a year-long energy simulation was performed in DesignBuilder with EnergyPlus engine. The outputs of interest were hourly and monthly electricity consumption (kWh) for the entire villa. These simulated monthly totals were then compared to the actual utility bills collected.

The model calibration proceeded iteratively. In each iteration, uncertain parameters (notably internal gains levels, minor material properties and especially occupancy schedules such as the operation of HVAC system, and the other device.) were adjusted so that the simulated monthly consumption more closely matched the measured data.

By integrating real billing data, detailed CAD plans, local material practices, and iterative adjustment of usage assumptions, the methodology ensures that the DesignBuilder model reliably reproduces the villa's actual energy consumption. This calibrated model can then be used confidently to evaluate efficiency measures or design alternatives.

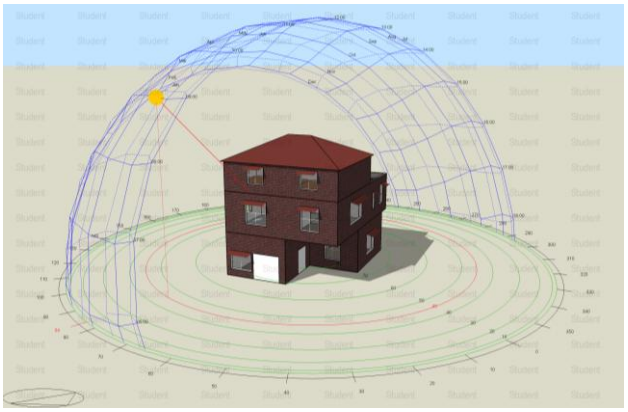
3. RESULTS AND DISCUSSION

The simulation model from DesignBuilder for the three-story residential villa in Phnom Penh model shown in Figure. 1., help us to see the actual building conditions with the sun path throughout the year shown in Figure. 1 with the household characteristics in Table 1.

Table 1. Household characteristics.

Space/Zone	Floor
1 Bedroom with bathroom attech	1st
2 Bedroom with bathroom attach	2nd
1 Master Bedroom with bathroom attach	2nd
Extra bathroom	1st
Kitchen	1st
Living room	1st
Family Loung	2nd
Storage	2nd
Roof top Terrance	3rd

Fig. 1. The building model in Designbuilder in each time of the year.



The sun-path overlays generated in DesignBuilder in Figure 1 show that the dwelling's long axis is aligned approximately north–south, so that the larger façades face the cardinal north and south quadrants, whereas the narrower façades address the east and west. Such an arrangement is widely recognised as a sound passive-design strategy for low-latitude locations (≈ 11° N, Phnom Penh). Field and simulation studies conducted in comparable tropical climates have demonstrated that rotating a building from an east–west exposure to a north–south exposure can reduce annual cooling loads by around 8.5–11.5 % [10] and attenuate peak afternoon gains that are otherwise difficult to shade. Because the sun remains at high altitude for most of the year, north façades receive only diffuse sky radiation, while south façades receive high-angle midday sun that can be intercepted by modest horizontal overhangs; in contrast, east and west façades are subjected to low-angle solar incidence that penetrates deeply indoors and is costly to block. By minimizing the glazed area on those latter orientations, the present design limits incident solar energy during the morning and late-afternoon periods when outdoor temperatures already approach their daily maxima. The approach therefore reduces sensible cooling demand without sacrificing daylight availability.

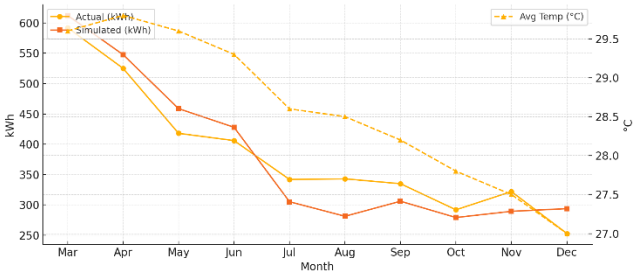
The simulation result for the designbuilder is shown on Table 2. Although we could do the simulation data for multiple year since we are lacking of the actual energy consumption to compared to we are chosen to compare the last ten month of the actual data with the simulation data started from March to December 2024.

Table 2. Simulation result and actual data.

Month	Actual (kWh)	Simulated (kWh)	Difference	% Error
Mar	592	612.23	+20.23	3.4 %

Month	Actual (kWh)	Simulated (kWh)	Difference	% Error
Apr	525	547.60	+22.60	4.3 %
May	418	458.64	+40.64	9.7 %
Jun	406	427.92	+21.92	5.4 %
Jul	342	305.29	−36.71	10.7 %
Aug	343	281.34	−61.66	18.0 %
Sep	335	306.11	−28.89	8.6 %
Oct	292	279.28	−12.72	4.4 %
Nov	322	289.49	−32.51	10.1 %
Dec	253	293.93	+40.93	16.2 %

Fig 2. The simulation result of the total energy consumption of 2024.



While Figure 2., shows the model reproduces the seasonal *shape* of electricity demand—falling with temperature, including both the actual, based on the electrical bill from EDC (Electricite du Cambodge), and simulation of the energy consumption with the curve showing the temperature data of each month average temperature. We can see that

Key quantitative results are as follows. As Table 2. and Figure 2. illustrates the simulated monthly energy consumption of a three-story residential villa located in Phnom Penh, utilizing the DesignBuilder software. The simulation presents the monthly breakdown for total electrical consumption, specifically disaggregated into cooling (air conditioning) and lighting energy consumption from January to December 2024. Cooling energy represents a significant fraction of the total simulated electricity consumption, highlighting the climatic influence and occupant reliance on mechanical cooling during hotter months. For instance, peak cooling loads are observed from March to May, with values reaching approximately 592 kWh, 525 kWh, and 418 kWh respectively. This pattern aligns with typical climatic conditions experienced in Phnom Penh, characterized by elevated temperatures and increased solar radiation during the dry season. Conversely, the lowest cooling demands occur during the transitional months (October), where values drop to approximately 292 kWh, indicating reduced thermal loads due to milder weather. With the non-AC appliances and lighting assumption to the same operation hour we could see the simulation result is under the actual use. In the baseline simulation every plug load, kitchen appliance and lighting circuit was assigned a constant monthly profile. While this worked reasonably well for the hot season, it ignores seasonal behavior such as longer evening lighting in the wet season, increased cooking during holidays, and eat less refrigeration food. As a result, the model over-predicts spring (Mar–Jun) consumption and under-predicts mid-wet-season use (Jul–Aug), when occupants actually spend more time indoors and operate fans, refrigerators and electronics for longer hours. In December the simulation still assumes “normal” occupancy, whereas the homeowners typically travel, unplugging some appliances and switching off lights for extended periods. The model therefore over-estimates December electricity by 16 %.

Table 3. Calibration of the simulation to actual energy consumption.

Statistic	Value
MAPE	9.08%
CV (RMSE)	9.02%
RMSE	-2.62kWh
MAE (Mean Absolute Error)	34.54 kWh

The initial model satisfies, with a global MAPE of 9.08% and CVRMSE of 9.02% well under the Ashrae 14 standard guideline the CVRMSE for the monthly evaluation within 15% [11]. The mean bias error (RMSE) is only –2.62 kWh, indicating that the simulation is essentially unbiased. The MAE of 34.54 kWh shows that a “typical” month deviates from measured consumption by roughly one day’s operation of a single split-type air-conditioner, a modest miss for residential-scale modelling.

Most of the residual variance is concentrated in three months—July (-36.7 kWh, 10.7 %), August (-61.7 kWh, 18.0 %) and December (+40.9 kWh, 16.2 %)—which together contribute more than half of the total squared error. This pattern confirms that the building fabric, HVAC performance curves and baseline internal gains are well captured for the bulk of the year, while behaviour-driven loads (reduced AC use but steady plug loads in mid-wet season, extended holiday vacancy in December) are not yet fully represented. Fine-tuning occupancy-dependent plug/lighting schedules and adding explicit holiday set-backs therefore offers the greatest potential for further error reduction.

CONCLUSIONS

This study demonstrates that a rigorously calibrated DesignBuilder model can reproduce the monthly electricity use of a typical three-story villa in Phnom Penh with acceptable accuracy and thereby serve as a reliable decision-support tool for Cambodia's rapidly expanding residential sector. After successive refinement of envelope properties, internal gains and HVAC schedules, the final model achieved a mean absolute percentage error (MAPE) of 9.08% and a coefficient of variation of the root-mean-square error (CVRMSE) of 9.02% for the March–December 2024 comparison.

The calibrated results confirm that space-cooling accounts for roughly 70–80 % of the household's annual electricity load, with peak demands occurring in the late dry-season months (March–May). The building's north–south orientation and judicious limitation of glazing on the east and west façades help suppress incident solar gains—a design choice that the simulation and supporting literature indicate can trim annual cooling requirements by about 8.5–11.5 % in low-latitude climates. Lighting and other plug loads remain comparatively stable year-round because there were set to a average monthly profile which lead to under-predicts mid-wet-season use (Jul–Aug). However, the AC unit still shared a large amount of energy consumptions, suggest that cooling-focused interventions (better scheduling, higher-efficiency AC units, or passive measures such as shading and natural-ventilation enhancements) offer the greatest potential for cost-effective savings.

Future work should couple the present model with sub-hourly monitoring, incorporate more detail occupant-behavior algorithms, the data from the indoor temperature, thermal comfort, and the measurement power consumption of each device and there working hour.

ACKNOWLEDGMENTS

We express our gratitude to the United Nations Development Programme (UNDP) for the training that facilitated the skilled utilization of the simulation software, and to the United Nations Economic and Social Commission for Asia and the Pacific (UN ESCAP) for supplying the DesignBuilder license to the Institute of Technology of Cambodia. We also acknowledge the homeowner's assistance in providing equipment inventories, two-dimensional blueprints, and utility bills that facilitated our model validation.

REFRENCES

[1] OECD, "Energy Statistics Data Browser," 2023. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=KHM&fuel=Energy consumption&indicator=TFCbySector (accessed May 02, 2025)

[2] OECD, "Energy Statistics Data Browser," 2023. https://www.iea.org/data-and-statistics/data-tools/energy-statistics-data-browser?country=KHM&fuel=Energy consumption&indicator=ElecConsBySector (accessed May 02, 2025)

[3] E. Asia, "World Bank Group," World Bank Gr. A to Z 2016, pp. 189a – 190, 2015, doi: 10.1596/978-1-4648-0484-7_world_bank_group.

[4] Kingdom of Cambodia, "Cambodia's National Cooling Action Plan," no. August, 2022, [Online]. Available: https://epa.moe.gov.kh/pdf/post/bf8229696f7a3bb4700cfddef19fa23f.pdf

[5] C. J. Mugnier, "Kingdom of Cambodia," Coord. Syst. World, pp. 135–138, 2023, doi: 10.1201/9781003307785-34.

[6] I. E. Agency, "Roadmap Towards Sustainable and Energy - Efficient Space Cooling in ASEAN INTERNATIONAL ENERGY".

[7] International Energy Agency (IEA), "The Future of Cooling in Southeast Asia," Futur. Cool. Southeast Asia, no. October, 2019, [Online]. Available: https://www.iea.org/reports/the-future-of-cooling-in-southeast-asia

[8] M. Arbulu, S. Perez-Bezos, A. Figueroa-Lopez, and X. Oregi, "Opportunities and Barriers of Calibrating Residential Building Performance Simulation Models Using Monitored and Survey-Based Occupant Behavioural Data: A Case Study in Northern Spain," Buildings, vol. 14, no. 7, 2024, doi: 10.3390/buildings14071911.

[9] P. Choun, V. Heang, K. Vongchanh, and S. Chan, "Simulation of Energy Consumption for Different Types of Walls and Colour for A Residential Building, Case Study: Phnom Penh City, Cambodia," Malaysian J. Sci. Adv. Technol., pp. 448–455, 2024.

[10] N. H. Wong and S. Li, "A study of the effectiveness of passive climate control in naturally ventilated

residential buildings in Singapore," Build. Environ., vol. 42, no. 3, pp. 1395–1405, 2007, doi: 10.1016/j.buildenv.2005.11.032.

[11] ANSI/ASHRAE, "ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings," Ashrae, vol. 8400, p. 170, 2002.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Calibration of DEM Parameters Using the Angle of Repose for Biomass Materials

CHEA Chantola [1], Kinnaleth VONGCHANH [1,2*], HOEUNG Hollin [1], HEANG Latin [1,2], CHAN Sarin [1]

[1] *Thermal lab., Department of Industrial and Mechanical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: kinnatlethv@itc.edu.kh*

Abstract: *The Discrete Element Method (DEM) is widely used to simulate biomass material handling, but accurate simulation requires parameter calibration. This study calibrates DEM parameters for fibrous and irregular biomass materials, which exhibit complex flow behaviors unlike traditional granular substances. The parameters including static friction, rolling friction, restitution coefficient, and surface energy, using angle of repose (AoR) experiments on sawdust and tree leaves. A fixed funnel method was employed to measure static AoR with five repetitions were conducted per material, and the final AoR was reported as mean ± standard deviation. For sawdust, AoR was 39.29° ± 0.865°, and for tree leaves, 41.71° ± 1.802°. These values were used as benchmarks in DEM simulations using the Hertz-Mindlin contact model with JKR cohesion. Initial parameter sets were derived from literature and refined through simulations to minimize deviation from experimental AoR. The best-fit parameters were Poisson's ratio (0.3), shear modulus (6 MPa for sawdust, 10 MPa for tree leaves), restitution (0.1 for sawdust, 0.002 for tree leaves), static friction (0.65 for sawdust, 0.6 for tree leaves), rolling friction (0.06 for sawdust, 0.07 for tree leaves), and JKR surface energy (0.1 J/m²). The final simulations showed relative AoR errors of 0.10% for sawdust and 1.58% for tree leaves. These calibrated parameters support improved accuracy in DEM modeling of biomass processing systems.*

Keywords: DEM; Angle of Repose (AoR); sawdust; tree leaves

1. INTRODUCTION

Biomass is a CO_2-neutral alternative to fossil fuels, absorbing carbon during growth and releasing it during combustion [1]. Briquetting biomass waste like sawdust and tree leaves enhances solid fuel utility and waste management [2–4]. However, biomass flow behavior is often hindered by irregular shapes, cohesion, and moisture variability [5–7]. While mechanical preprocessing improves flow, it also raises energy demands [8].

To better predict bulk behavior, the Discrete Element Method (DEM) has been widely adopted since its introduction by Cundall and Strack [9]. DEM can simulate particle interactions, energy dissipation, and cohesive forces [10–13], making it valuable for optimizing biomass processing equipment [14]. Accurate simulation relies on calibrated input parameters (e.g., friction, restitution, stiffness), which significantly impact model reliability [15]. While comprehensive calibration frameworks exist [16], simplified methods like the angle of repose (AoR) test offer practical alternatives [17,18]. AoR-based calibration has proven effective across many materials using fixed funnel or stacking methods [19–21], with rolling friction, restitution, and surface energy being key influencing factors [22,23]. Nonetheless, calibrating irregular, cohesive biomass such as sawdust and tree leaves remains challenging.

This study aims to measure and simulate AoR using the fixed funnel method to calibrate DEM parameters for future mixing simulations, contributing to the limited research on fibrous, irregular biomass materials with complex flows.

2. METHODOLOGY

2.1 Angle of Repose Measurement

The angle of repose (AoR) is a crucial parameter for assessing granular material flow, but measurement methods

[1]* Corresponding author: Kinnaleth VONGCHANH
E-mail: kinnatlethv@itc.edu.kh; Tel: +855 99 351 199

are not standardized [24]. Common AoR methods include tilting boxes, fixed funnels, hollow cylinders, revolving drums, internal flow funnels, and tilting boxes [19,24–31]. Each method offers unique insights, but some have limitations like setup complexity and low repeatability. The fixed funnel method, which involves pouring material from a fixed height to form a conical pile, was chosen for this study due to its simplicity and applicability [24].

2.2 Biomass Materials

Biomass materials such as sawdust (SD) and tree leaves (TL), processed with particle sizes less than 8 mm, were used based on prior briquetting research [4], as illustrated in **Fig. 1**. Bulk densities were 230 kg/m³ (SD) and 97 kg/m³ (TL). To maintain equal volume, 60 g of SD and 30 g of TL were used per test.



a. Sawdust b. Tree leaves

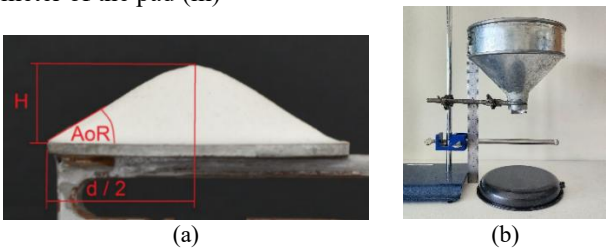
Fig. 1. Biomass Materials [4]

2.3 Angle of Repose (AoR) Experiment

The experiment followed ISO 4324:1997 [32] procedures; a 45 mm funnel was used at a 150 mm drop height same as pad diameter. The AoR was calculated using Eq. 1 based on the pile's height as shown in **Fig.2**.

$$\phi = \tan^{-1} H/(d/2) \quad \text{(Eq. 1)}$$

where: ϕ: angle of repose (°); H: height of the cone (m); d: diameter of the pad (m)



(a) (b)

Fig. 2. AoR (a) measurement theory [33] (b) experimental setup

Parallelly, images of the heap were captured and analyzed using ImageJ [28,34,35] that images were binarized, and pile slopes were measured using the Angle Tool. Multiple measurements per image were averaged to determine the final AoR. Then each sample was experimentally conducted 5-6 times to ensure consistency.

2.4 DEM Simulation

The Hertz-Mindlin contact model with JKR bonding was used to simulate the inter-particle behavior of cohesive biomass materials like SD and TL [36,37]. This model is suitable for materials with static electricity, fibrous structure, or moisture content [46–48]. Key parameters influencing flow behavior include static and rolling friction, restitution coefficient, Young's modulus, and surface energy [40–44]. Adhesion and cohesion modeled via the JKR framework are essential for capturing van der Waals and capillary forces in moist or cohesive biomaterials [46]. DEM parameters were selected through a trial-and-error approach, guided by literature [21,23,50] and the EDEM GEMM database [50]. Parameter ranges explored include: Poisson's ratio (0.1–0.5 SD, 0.1–0.4 TL), shear modulus (1–10 MPa), restitution coefficient (0.1–0.8 SD, 0.002–0.005 TL), static friction (0.01–0.8 SD, 0.2–0.9 TL), rolling friction (0.01–0.8 SD, 0.02–0.07 TL), and JKR surface energy (0.05–0.5 SD, 0.1–3.5 J/m² TL). These were organized into parameter series A–E (TL) and refined as series F (SD). Simulations used a virtual box (510 mm × 316 mm) with particles released into a blocked funnel, using a fixed time step of 1.31388e-05 s. Particle counts were 1946 for SD and 2307 for TL. Post-processing of AoR data was performed with EDEMPy.

3. RESULTS AND DISCUSSION

3.1 Angle of Repose

This study measured the AoR for TL and SD. For TL, three samples (T1, T2, T3) were used to capture variability in their irregular shapes, while SD samples (S1, S2) represented more uniform characteristics. T1J, T2J, and T3J represent AoR values measured from experimental images using the ImageJ method, whereas T1F, T2F, and T3F are the corresponding AoR values calculated by formula.

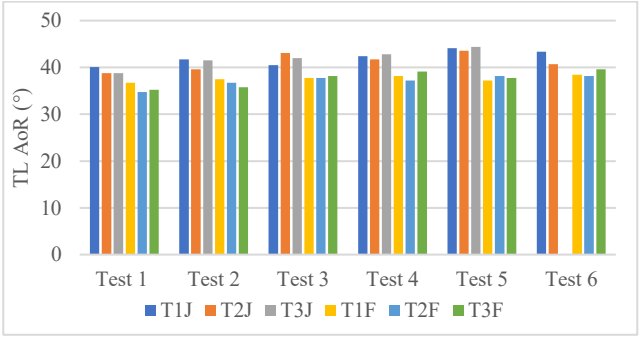


Fig. 3. Comparison of AoR measurements for tree leaves

Fig. 3, compares AoR results from six tests on three TL samples (T1, T2, and T3). The ImageJ method consistently yielded slightly higher values than the formula-based method. Thus, in Test 5, T1J measured 44.133° with ImageJ versus 38.190° with the formula, suggesting that ImageJ captures finer surface angles that manual calculations may miss.

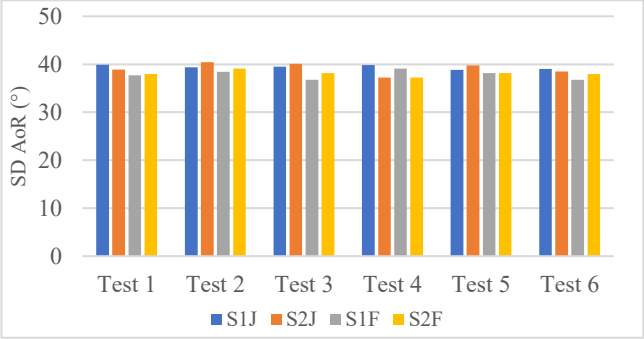


Fig.4. Comparison of AoR measurements for sawdust

Fig. 4, compares the AoR of SD (S1J and S2J) using ImageJ and the formula method (S1F and S2F). ImageJ consistently gave slightly higher values due to more precise slope detection. For instance, S1J in Test 4 measured 39.95° with ImageJ versus 37.80° with the formula. Overall, differences stayed within 2–3°, showing good agreement between methods.

Table 1, summarizes the average AoR values for SD and TL using both ImageJ and formula-based methods. The ImageJ results showed higher averages (41.71° for tree leaves and 39.38° for sawdust) compared to the formula method. These values better reflect the actual pile shape observed during experiments. Due to its higher precision and consistency, the ImageJ method is selected for DEM parameter calibration.

Table 1. Summary of Average AoR and Standard Deviation (StDev) for SD and TL Using ImageJ and Formula Methods

Material	Method	Average AoR (°)	StDev (°)
SD (S1J, S2J)	ImageJ	39.38	0.865
SD (S1F, S2F)	Formula	37.85	0.880
TL (T1J, T2J, T3J)	ImageJ	41.71	1.802
TL(T1F,T2F,T3F)	Formula	37.45	1.257

3.2 Simulation Parameter Calibration Results

DEM simulations were carried out to calibrate material parameters for TL and SD by matching the experimentally measured AoR. For TL, five simulation series: A, B, C, D, and E, were conducted to progressively refine the parameter space (**Table 2**), with 8, 13, 6, 3, and 12 sets, respectively.

Table 3, shows that Series A produced A7, which had an AoR (41.05°) closely matching the experimental value. Series B identified B7 with the closest AoR (41.99°) and the lowest relative error (0.666%). In contrast, Series C and D resulted in higher AoR or greater variability, while Series E showed inconsistent results with high coefficient of variation (CoV). Among all runs in the simulation, A7 was selected as the best match to the experimental AoR (41.71°), due to its low variability across trials (StDev = 5.62°) and acceptable error (1.58%) showing that its more stable simulation behavior. While B7 also showed strong accuracy, its larger standard deviation (8.590°) made A7 the more reliable choice.

Table 2. AoR results from initial parameter screening – tree leaves

Series	Set Count	AoR Range (°)	Closest Match (°)	Notes
Serie A	8	37–43	A7 (41.05)	Best baseline for refinement
Series B	13	41–47	B7 (41.99)	Close, but high StDev
Series C	6	43–48	C2 (45.81)	High AoR, not selected
Series D	3	46–48	D2 (46.30)	Too high AoR
Series E	12	38–51	E4 (39.74)	Unstable results

For sawdust, a single parameter series (F) consisting of five sets (F1–F5) were tested (**Table 4**). F5 yielded an AoR of 39.24°, closely matching the experimental value (39.28°) with a 0.10% relative error and a standard deviation of 3.84°. These results demonstrate the reliability of the setup and the robustness of F5, which was selected for all further sawdust DEM simulations.

Table 3. Summary of Simulated AoR Results – Tree leaves

Simulation	AoR (°)	StDev (°)	Relative Error (%)
B11	42.44	7.544	1.755
B2	42.22	8.688	1.230
B7	41.99	8.590	0.666
A7	41.05	5.623	1.584

Table 4. Summary of Simulated AoR Results – Sawdust

Simulation	AoR (°)	StDev (°)	Relative Error (%)
F1	40.47	7.201	3.027
F2	38.27	5.754	2.581
F3	40.27	2.364	2.521
F4	42.65	7.855	8.569
F5	39.24	3.839	0.101

4. CONCLUSIONS

The angle of repose (AoR) was experimentally measured using the fixed funnel method, averaging 41.71° for tree

leaves (three repetitions) and 39.28° for sawdust (two repetitions). These values guided parameter calibration in DEM simulations using the Hertz-Mindlin with JKR model.

For tree leaves, the best match came from A7, which produced an AoR of 41.05°, resulting in a small error of 1.58% and low variation between tests. In the case of sawdust, the F5 test gave an AoR of 39.24°, with an even smaller error of just 0.10%. The optimal parameters were:

Tree leaves: Poisson's ratio = 0.3, shear modulus = 10 MPa, restitution = 0.002, static friction = 0.6, rolling friction = 0.07, JKR surface energy = 0.1 J/m².

Sawdust: Poisson's ratio = 0.3, shear modulus = 6 MPa, restitution = 0.1, static friction = 0.65, rolling friction = 0.06, JKR surface energy = 0.1 J/m².

A limitation of this study is that the DEM model does not fully capture cohesion variability due to moisture or compaction. Scale effects between the experimental setup and industrial systems may also introduce deviations. Nevertheless, these calibrated parameters provide a strong basis for future DEM simulations of biomass mixing, supporting improved flow prediction.

ACKNOWLEDGMENTS

The author sincerely thanks Dr. KAN Kuchvichea of the Road Materials Laboratory, Department of Civil Engineering, ITC, for providing the supporting structure used in the Angle of Repose experiment. His contribution was vital to the success of this study.

REFERENCES

[1] N. T. and A. B. Swapan Suman, Anand Mohan Yadav, "Combustion characteristics and behaviour of agricultural biomass: a short review," *Intech*, vol. 11, no. tourism, p. 13, 2020, [Online]. Available: https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics

[2] N. HODŽIĆ, K. KADIĆ, and A. Kazagić, "Primjenom otpadne drvne biomase i Miscanthusa do održivog sektora termoenergetike u vremenu energetske tranzicije / Sustainable Energy Sector Using Waste Woody Biomass and Miscanthus in the Energy Transition," *Energ. Ekon. Ekol.*, vol. XXV, no. 1, pp. 1–9, 2023, doi: 10.46793/eee23-1.01h.

[3] G. L. C. and Carlos Ujados Lorenzo , José Pablo Paredes Sánchez and J. X. Bernat, "Bioenergy as an Environmental Alternative in Energy Systems: A Case Study," p. 1481, 2018, doi: 10.3390/proceedings2231481.

[4] Latin Heang, Porchaing Choeng, Kinnaleth Vongchanh, and Sarin Chan, "Experimental Investigation on Sawdust-Tree Leaves Briquette Using Fish Oil and Waste Oil As Binder," no. June 2023, 2020.

[5] J. Klinger *et al.*, "Multiscale Shear Properties and Flow Performance of Milled Woody Biomass," *Front. Energy Res.*, vol. 10, no. July, pp. 1–15, 2022, doi: 10.3389/fenrg.2022.855289.

[6] H. Yi, C. J. Lanning, J. H. Dooley, and V. M. Puri, "Finite element modeling of biomass hopper flow," *Front. Energy Res.*, vol. 11, no. May, pp. 1–12, 2023, doi: 10.3389/fenrg.2023.1162627.

[7] T. L. W. and D. S. Hartley, "Biomass Handling and Feeding," *Intech*, vol. 11, no. tourism, p. 21, 2018, [Online]. Available: https://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics

[8] A. Hamed, Y. Xia, N. Saha, J. Klinger, D. N. Lanning, and J. Dooley, "Flowability of Crumbler Rotary Shear Size-Reduced Granular Biomass: An Experiment-Informed Modeling Study on the Angle of Repose," *Front. Energy Res.*, vol. 10, no. May, pp. 1–20, 2022, doi: 10.3389/fenrg.2022.859248.

[9] P. A. Cundall and O. D. L. Strack, "A discrete numerical model for granular assemblies," *Geotechnique*, vol. 30, no. 3, pp. 331–336, 1979, doi: 10.1680/geot.1980.30.3.331.

[10] T. De, A. Das, and J. Kumar, "On the prediction of particle collision behavior in coarse-grained and resolved systems," *Part. Sci. Technol.*, vol. 41, no. 8, pp. 1170–1184, 2023.

[11] S. Rotter, M. Dosta, and A. Düster, "Discrete element simulation of the breakage behavior of porous granules utilizing bond models," *Comput. Part. Mech.*, vol. 11, no. 1, pp. 89–103, 2024, doi: 10.1007/s40571-023-00610-0.

[12] M. Bagheri, S. Roy, and T. Pöschel, "Discrete Element Simulations of particles interacting via capillary forces using MercuryDPM," *SoftwareX*, vol. 29, no. November 2024, p. 101987, 2025, doi: 10.1016/j.softx.2024.101987.

[13] A. Sac-Morane, M. Veveakis, and H. Rattez, "A Phase-Field Discrete Element Method to study

chemo-mechanical coupling in granular materials," *Comput. Methods Appl. Mech. Eng.*, vol. 424, 2024, doi: 10.1016/j.cma.2024.116900.

[14] J. Quist and M. Evertsson, "Framework for DEM Model Calibration and Validation," *Proc. 14th Eur. Symp. Comminution Classif.*, no. September, pp. 103–108, 2015.

[15] Y. Tan, Y. Yu, J. Fottner, and S. Kessler, "Automated measurement of the numerical angle of repose (aMAoR) of biomass particles in EDEM with a novel algorithm," *Powder Technol.*, vol. 388, pp. 462–473, 2021, doi: 10.1016/j.powtec.2021.04.062.

[16] M. Hofmann, "On the Complexity of Parameter Calibration in Simulation Models," *J. Def. Model. Simul. Appl. Methodol. Technol.*, vol. 2, no. 4, pp. 217–226, 2005, doi: 10.1177/154851290500200405.

[17] C. Zheng, M. A. Behjani, J. Hu, and C. Y. Wu, *Discrete Element Modelling of Pharmaceutical Powder Handling Processes*, no. January. 2023. doi: 10.1002/9783527835935.ch7.

[18] R. Hajivand Dastgerdi and A. A. Malinowska, "A Simplified Calibration Procedure for DEM Simulations of Granular Material Flow," *Materials (Basel).*, vol. 17, no. 19, 2024, doi: 10.3390/ma17194833.

[19] H. M. Beakawi Al-Hashemi and O. S. Baghabra Al-Amoudi, "A review on the angle of repose of granular materials," *Powder Technol.*, vol. 330, pp. 397–417, 2018, doi: 10.1016/j.powtec.2018.02.003.

[20] G. Ma, Z. Sun, H. Ma, and P. Li, "Calibration of Contact Parameters for Moist Bulk of Shotcrete Based on EDEM," *Adv. Mater. Sci. Eng.*, vol. 2022, 2022, doi: 10.1155/2022/6072303.

[21] G. Xun, B. Xuewei, H. Haibo, Z. Fengyu, G. Yuanjuan, and W. Desheng, "DEM PARAMETERS CALIBRATION OF MIXED BIOMASS SAWDUST MODEL WITH MULTI-RESPONSE INDICATORS," vol. 65, no. 3, pp. 183–192, 2022.

[22] X. Wang, Z. Yu, D. Zhang, Y. Chen, and D. Ni, "Calibration of Simulation Parameters for Roasted Green Tea Based on Discrete Element Method," *Nongye Jixie Xuebao/Transactions Chinese Soc. Agric. Mach.*, vol. 55, no. 8, pp. 418–427, 2024, doi: 10.6041/j.issn.1000-1298.2024.08.039.

[23] J. Ren *et al.*, "Discrete Element Simulation Modeling Method and Parameters Calibration of Sugarcane Leaves," *Agronomy*, vol. 12, no. 8, pp. 1–21, 2022, doi: 10.3390/agronomy12081796.

[24] M. Rackl, F. E. Grötsch, M. Rusch, and J. Fottner, "Qualitative and quantitative assessment of 3D-scanned bulk solid heap data," *Powder Technol.*, vol. 321, pp. 105–118, 2017, doi: 10.1016/j.powtec.2017.08.009.

[25] L. Zhichao, "Measuring the Angle of Repose of Granular Systems," *Can. Geotech. J.*, p. 143, 2011, [Online]. Available: http://d-scholarship.pitt.edu/6401/1/MSThesisZhichaoLiu03042011.pdf

[26] T. Q. Huynh, T. T. Nguyen, and B. Indraratna, "Evaluating cohesive models in discrete element simulation through drawdown test with new assessment perspectives," *Powder Technol.*, vol. 452, p. 120542, 2024, doi: 10.1016/j.powtec.2024.120542.

[27] D. McGlinchey, *Characterisation of Bulk Solids*. 2004.

[28] J. Zhang, C. Li, Q. Niu, P. Wang, L. Wang, and H. Li, "Characterization of green peppers based on dynamic repose angle," *LWT*, vol. 180, no. November 2022, p. 114703, 2023, doi: 10.1016/j.lwt.2023.114703.

[29] M. G. Kleinhans, H. Markies, S. J. De Vet, A. C. In't Veld, and F. N. Postema, "Static and dynamic angles of repose in loose granular materials under reduced gravity," *J. Geophys. Res. Planets*, vol. 116, no. 11, pp. 1–13, 2011, doi: 10.1029/2011JE003865.

[30] H. N. Pitanga, J. P. Gourc, and O. M. Vilar, "Interface shear strength of geosynthetics: Evaluation and analysis of inclined plane tests," *Geotext. Geomembranes*, vol. 27, no. 6, pp. 435–446, 2009, doi: 10.1016/j.geotexmem.2009.05.003.

[31] Z. Chik and L. E. Vallejo, "Characterization of the angle of repose of binary granular materials," *Can. Geotech. J.*, vol. 42, no. 2, pp. 683–692, 2005, doi: 10.1139/t04-118.

[32] International Standard ISO 4324, "Surface active agents - Powders and granules - Measurement of the angle of repose," vol. 1977, 1977.

[33] O. Macho *et al.*, “Analysis of static angle of repose with respect to powder material properties,” *Acta Polytech.*, vol. 60, no. 1, pp. 73–80, 2020, doi: 10.14311/AP.2020.60.0073.

[34] I. S. B. Ferreira, R. S. Peruchi, N. J. Fernandes, and P. Rotella Junior, “Measurement system analysis in angle of repose of fertilizers with distinct granulometries,” *Meas. J. Int. Meas. Confed.*, vol. 170, no. May, 2021, doi: 10.1016/j.measurement.2020.108681.

[35] P. Datta and S. A. Faroughi, “Angle of repose for superquadric particles: Investigating the effects of shape parameters,” *Comput. Geotech.*, vol. 165, 2024, doi: 10.1016/j.compgeo.2023.105918.

[36] W. Fang, X. Wang, D. Han, and X. Chen, “Review of Material Parameter Calibration Method,” *Agric.*, vol. 12, no. 5, 2022, doi: 10.3390/agriculture12050706.

[37] M. Paulick, M. Morgeneyer, and A. Kwade, “Review on the influence of elastic particle properties on DEM simulation results,” *Powder Technol.*, vol. 283, pp. 66–76, 2015, doi: 10.1016/j.powtec.2015.03.040.

[38] P. Li, M. Ucgul, S. H. Lee, and C. Saunders, “A new approach for the automatic measurement of the angle of repose of granular materials with maximal least square using digital image processing,” *Comput. Electron. Agric.*, vol. 172, no. November 2019, p. 105356, 2020, doi: 10.1016/j.compag.2020.105356.

[39] N. N. Mohsenin, “Physical properties of plant and animial materials. Vol. 1. Structure, physical characterisitics and mechanical properties.,” 1970.

[40] M. J. Jiang, H. S. Yu, and D. Harris, “A novel discrete model for granular material incorporating rolling resistance,” *Comput. Geotech.*, vol. 32, no. 5, pp. 340–357, 2005, doi: 10.1016/j.compgeo.2005.05.001.

[41] W. R. Ketterhagen, “Modeling the motion and orientation of various pharmaceutical tablet shapes in a film coating pan using DEM,” *Int. J. Pharm.*, vol. 409, no. 1–2, pp. 137–149, 2011, doi: 10.1016/j.ijpharm.2011.02.045.

[42] A. Katterfeld *et al.*, “Calibration of DEM Parameters for Cohesionless Bulk Materials under Rapid Flow Conditions and Low Consolidation,” no. July, 2019, doi: DOI: 10.13140/RG.2.2.26318.31048/1.

[43] C. Mangwandi, Y. S. Cheong, M. J. Adams, M. J. Hounslow, and A. D. Salman, “The coefficient of restitution of different representative types of granules,” *Chem. Eng. Sci.*, vol. 62, no. 1–2, pp. 437–450, 2007, doi: 10.1016/j.ces.2006.08.063.

[44] Teffo V.B. and Naudé N., “Determination of the coefficients of restitution, static and rolling friction,” *J. South. African Inst. Min. Metall.*, vol. 113, no. 4, pp. 351–356, 2013.

[45] Z. Yan, S. K. Wilkinson, E. H. Stitt, and M. Marigo, “Discrete element modelling (DEM) input parameters: understanding their impact on model predictions using statistical analysis,” *Comput. Part. Mech.*, vol. 2, no. 3, pp. 283–299, 2015, doi: 10.1007/s40571-015-0056-5.

[46] C. Coetzee, “A Johnson-Kendall-Roberts (JKR) Contact Model, Implementation in PFC,” no. 03 December, 2020, doi: 10.13140/RG.2.2.21772.54404.

[47] J. Hærvig *et al.*, “On the adhesive JKR contact and rolling models for reduced particle stiffness discrete element simulations,” *Powder Technol.*, vol. 319, pp. 472–482, 2017, doi: 10.1016/j.powtec.2017.07.006.

[48] M. J. Carr, W. Chen, K. Williams, and A. Katterfeld, “Comparative investigation on modelling wet and sticky material behaviours with a simplified JKR cohesion model and liquid bridging cohesion model in DEM,” *ICBMH 2016 - 12th Int. Conf. Bulk Mater. Storage, Handl. Transp. Proc.*, pp. 40–49, 2016.

[49] M. A. Madrid, J. M. Fuentes, F. Ayuga, and E. Gallego, “Determination of the Angle of Repose and Coefficient of Rolling Friction for Wood Pellets,” *Agronomy*, vol. 12, no. 2, 2022, doi: 10.3390/agronomy12020424.

[50] Y. S. J Deli, C Xuegeng, YAN Limin, MO Yisong, “Design and experiment on spiral impurity cleaning device for profile modeling residual plastic film collector,” 2019.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Maximizing PV Hosting Capacity in Low-Voltage Distribution Networks: A Case Study in Cambodia

Samphors ENG [1*], Julien ALLARD [2], Vannak VAI [1], François VALLEE [2]

[1] *Faculty of Electrical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Electrical Power Engineering Unit, University of Mons, 31 Boulevard Dolez, Mons 7000, Belgium*

[*] *Corresponding author: samphors@itc.edu.kh*

Abstract: *The integration of rooftop photovoltaic (PV) systems into low-voltage (LV) distribution networks is a key strategy for advancing Cambodia's decarbonization goals. Accurately assessing the maximum PV hosting capacity (HC) is crucial for distribution operators to ensure optimal PV deployment while maintaining grid stability and operational feasibility. However, optimally determining PV hosting capacity is challenging due to the nonlinear and non-convex nature of the power flow (PF) equations, which can lead to a large computational burden and no global optimality guarantee when included as such in an optimization problem. This study proposes an optimization-based approach for HC assessment, leveraging the DistFlow model for PF equations combined with second-order cone programming (SOCP) relaxation. Backward injections from LV to medium voltage (MV) grid can negatively affect the upstream grid (e.g., voltage increase), which may be difficult for grid operators to handle. Therefore, regulators aim to prevent the reverse power flow that occurs during high PV penetration. In this context, a two-step analysis is performed. The first step involves evaluating the siting of the PV to identify which constraints limit the HC. The second step examines operational scenarios based on established rules of thumb for HC: (1) maximizing HC with allowances for reverse power flow into the main grid, (2) maximizing HC without allowing reverse power flow, and (3) maximizing HC when PV capacity is limited to 50% of the rated MV/LV transformer capacity. The comparison is held on a case study of the Tboung Khmum LV distribution network. The results provide valuable insights into optimal PV integration levels, the influence of network constraints, and the impacts of reverse power flow, ultimately contributing to the sustainable expansion of PV systems in Cambodia's distribution networks.*

Keywords: PV hosting capacity; Optimal power flow; DistFlow model; SOC relaxation; Reverse power flow

1. INTRODUCTION

The Master Plan for Cambodia Power Development (PDP) 2022-2040 [1], led by the Ministry of Mines and Energy, with the support of the Asian Development Bank, describing the country's roadmap for the development of the electricity industry, emphasizing energy security, affordability, and clean energy goals aligned with Cambodia's Nationally Determined Contributions (NDC) and Long-Term Strategy for Carbon Neutrality (LTS4CN) [2]. PDP identified five planning scenarios, which combined different mixtures of coal, hydroelectricity, LNG, renewable energy, and energy efficiency. Integrating photovoltaic systems (PV) on the roof in the low-voltage (LV) distribution network plays a significant role in clean energy strategies. However, the exact assessment of the maximum PV hosting capacity (HC) residence is essential to ensure reliable and safe operation. An estimated HC can cause overload of transformer violations and stress. Based on international Distribution System Operator (DSO) practices and their rules of thumb for Distributed Generator (DG) interconnections and preliminary HC assessments [3], Canada has documented that reverse power flow due to DG integration should not exceed 60% of the transformer rating at the main substation. In Italy, the total DG capacity is recommended to be less than 65% of the MV/LV transformer rating, while in Spain, it should be limited to less than 50% of the transformer rating. In [4], geometric and restrictive models are proposed to provide a constructive approach for determining the optimal solution simultaneously with hosting capacity (HC) under non-convex conditions. In [5], an energy management system for Italy's LV microgrid integrates the costs of distributed

[1*] Corresponding author: Samphors Eng
E-mail: samphors@itc.edu.kh; Tel: +855-99 443 463

energy resources (DERs), current flow via optimal power flow (OPF), and optimal inverter limits. Ref. [6] a multi-perspective approach using a modified Grey Wolf Optimization Algorithm (MOAGWO) to simultaneously maximize the HC and minimize power losses across various grid configurations and penetration levels. Ref. [7] presents a benchmark study comparing deterministic and stochastic methods for assessing HC estimation. A probabilistic framework for HC assessment under uncertainty is presented in [8], [9], [10]. Regarding the OPF-based approach, [11] introduces linearized AC power flow models to enable more efficient HC calculations. In [12], an OPF-based optimization method using Python and OpenDSS is applied to evaluate the maximum hosting capacity (MHC) of the EPRI PCT5 system.

This study suggests an OPF model based on Convex DistFlow that effectively estimates the maximum HC and minimizes the power losses in LV distribution networks by using SOC relaxation. Along with HC estimation, the paper examines how PV sitting and generation uncertainty affect operational limits, including voltage and current constraints, highlighting important limitations that prevent further HC expansion. It also examines how well the system performs in situations where reverse power flow is present and absent to determine how it affects HC. Cambodia PV installations are included in the study, and the applicability of the results is directed towards local energy plans and directives.

2. METHODOLOGY

To determine the maximum HC that can be integrated into an LV distribution system over one year, the HC evaluation is structured as a long-term planning problem. Representative time intervals are employed to capture typical fluctuations in both solar generation and load demand, thereby reflecting the system's operational behavior under normal conditions. This planning approach ensures safe and efficient PV integration by accounting for technical constraints and regulatory guidelines relevant to the Cambodian context. The case study is based on the Tboung Khmum LV distribution network has four different load profiles based on the types of consumers in Cambodia. The same daily load profile is used for the whole year.

To compute the optimal HC, the Convex DistFlow model is applied, as it provides a convex optimization framework that guarantees convergence to a global optimum. Unlike traditional methods, which often rely on time-consuming simulations or non-optimal heuristics, this approach improves computational efficiency and guarantees feasible, reliable solutions.

2.1 Test System

This study presents a planning model based on a single-phase LV distribution system, focusing on rooftop solar panels at households. The conductor used has a cross-sectional area of 75 mm², and the system includes 24 poles, with household loads assigned to the nearest pole using a shortest-path algorithm. The system has 68 buses, with a 75 kVA rated MV/LV transformer located at Bus 1. Three feeders distribute power throughout the network. Fig. 2 shows four different load profiles, which classify household electricity use based on the monthly consumption of residents [13] by the Electricity Authority of Cambodia (EAC), such as Profile 1: low consumption, less than 10 kWh, Profile 2: medium consumption, 10–50 kWh, Profile 3: high consumption, 50–200 kWh, Profile 4: very high consumption, more than 200 kWh per month.

Fig. 3 shows the solar irradiation profile of the case study topology, based on one year of data from January 1 to December 31, 2023, obtained from the NASA database [14]. The "all-sky" solar irradiation data was used as the system input, while temperature data was also considered for calculating the photovoltaic (PV) generation factor. On 8th March at 1:00 PM, the solar irradiance hit the highest point of 1011.72 W/m^2.

2.2 DistFlow Model

Distflow stands for "distribution flow," which represents the current model of a distribution system that records active/reactive performance, voltage drop, line loss, and power injection at each bus. OPF problems are usually prescribed either in a bus injection model (BIM) that focuses on power injection or in a branch flow model (BFM) that emphasizes branch current flow and voltages. In this study, BFM is used as the current model to accurately present the LV distribution network and facilitate system analysis.

- Non-Convex DistFlow Model: The non-linear Branch Flow Model (BFM), as formulated in [15], is presented in Equations (2) to (7). These equations represent the power flow and network constraints, which are inherently non-convex. The objective function is non-linear, defined in Equation (1), aims to maximize the hosting capacity while minimizing power losses. To achieve an optimal HC while ensuring efficient system operation, power losses must also be considered.

Objective Function:

$$\max\left(\gamma_1 \sum_{i\in\mathcal{N}} P_{\mathrm{PV},i} - \gamma_2 \sum_{(i,j)\in\mathcal{E},\, t\in\mathcal{T}} R_{ij} I_{ij,t}^2\right) \quad \text{(Eq. 1)}$$

Subject to power flow constraints:

$$P_{ij,t} = P_{j,t} + R_{ij} I_{ij,t}^2 + \sum_{k\in C_j} P_{jk,t} \quad \text{(Eq. 2)}$$

$$Q_{ij,t} = Q_{j,t} + X_{ij} I_{ij,t}^2 + \sum_{k\in C_j} Q_{jk,t} \quad \text{(Eq. 3)}$$

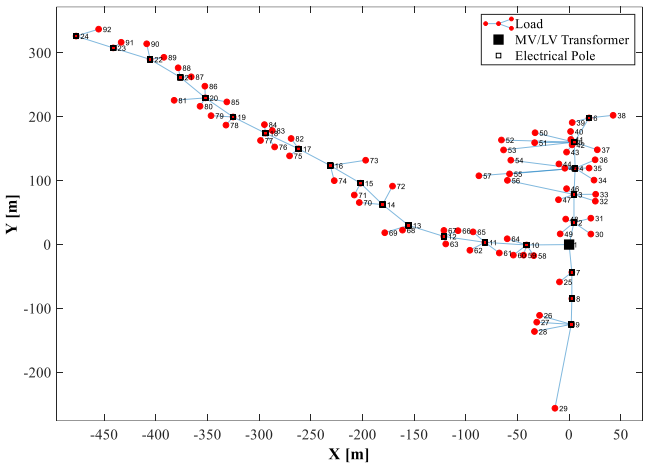


Fig. 1. Tboung Khmum planning topology in Toul Tompong village 10°56'31.4″N,106°02'08.5"E

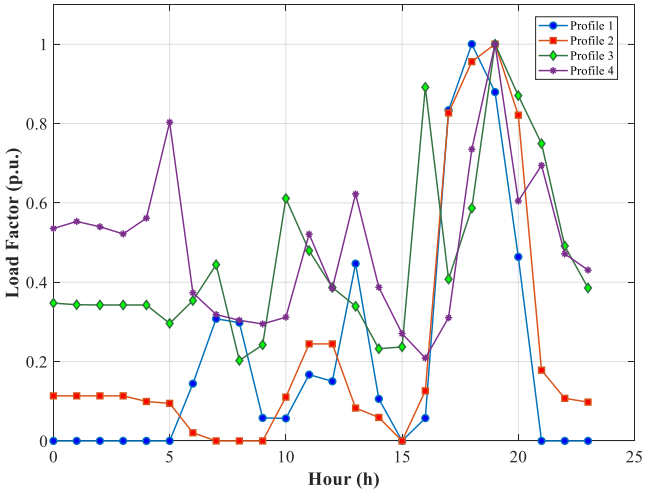


Fig. 2. Daily load profiles for four categories of consumers

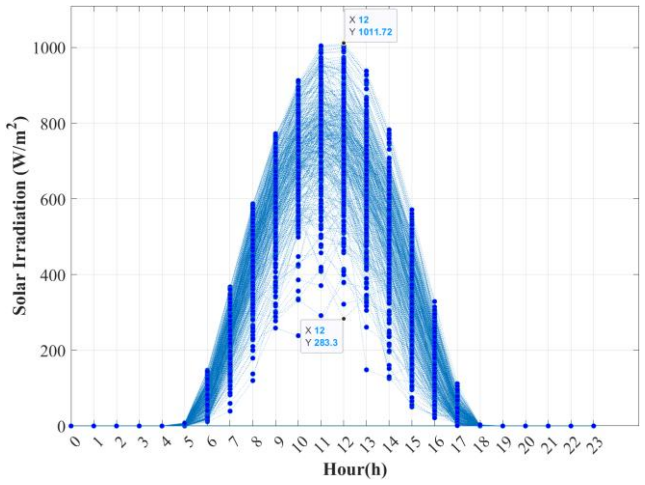


Fig. 3. Solar irradiation at 10°56'31.4″N,106°02'08.5"E of one year Data from NASA

$$V_{j,t}^2 = V_{i,t}^2 - 2\left(R_{ij}P_{ij,t} + X_{ij}Q_{ij,t}\right) + \left(R_{ij}^2 + X_{ij}^2\right)I_{ij,t}^2 \quad \text{(Eq. 4)}$$

Network constraints:

$$P_{ij,t}^2 + Q_{ij,t}^2 \le I_{ij,t}^2 V_{i,t}^2 \quad \text{(Eq. 5)}$$

$$V_{i,t}^{2,min} \le V_{i,t}^2 \le V_{i,t}^{2,max} \quad \text{(Eq. 6)}$$

$$I_{ij,t}^2 \le I_{ij,t}^{2,max} \quad \text{(Eq. 7)}$$

- Convex DistFlow model: To address the non-convexity of the power flow problem, the current squared variable $l_{ij,t} = I_{ij,t}^2$ is introduced. This allows the power flow balance equations (Eq. 2) and (Eq. 3) to be reformulated as linear expressions, as shown in (Eq. 9) and (Eq. 10). Similarly, the voltage drop constraint (Eq. 4) is reformulated using the squared voltage variable $w_{i,t} = V_{i,t}^2$, which makes it convex. The power constraint in (Eq. 5) is relaxed and rewritten as a second-order cone (SOC) constraint in (Eq. 15), allowing it to be expressed in a convex form as in (Eq. 14). These transformations enable the application of SOC relaxation while maintaining the physical interpretability of the distribution network model. The reformulated voltage and current constraints are presented in (Eq. 16) and (Eq. 17). The PV installation capacity is limited to a maximum of 10 kWp on an individual residence [16]. The reverse power flow constraint, as shown in (Eq. 19), is not applied in all case studies but is enforced when reverse power flow to the main grid is not permitted.

Objective Function:

$$\max\left(\gamma_1 \sum_{i\in\mathcal{N}} P_{\mathrm{PV},i} - \gamma_2 \sum_{(i,j)\in\mathcal{E},\, t\in\mathcal{T}} R_{ij} l_{ij,t}\right) \quad \text{(Eq. 8)}$$

Subject to power flow constraints:

$$P_{ij,t} = P_{j,t} + R_{ij} l_{ij,t} + \sum_{k\in\mathcal{C}_j} P_{jk,t} \quad \text{(Eq. 9)}$$

$$Q_{ij,t} = Q_{j,t} + X_{ij} l_{ij,t} + \sum_{k\in\mathcal{C}_j} Q_{jk,t} \quad \text{(Eq. 10)}$$

$$w_{j,t} = w_{i,t} - 2\left(R_{ij}P_{ij,t} + X_{ij}Q_{ij,t}\right) + \left(R_{ij}^2 + X_{ij}^2\right) l_{ij,t} \quad \text{(Eq. 11)}$$

$$P_{j,t} = P_{j,t}^C - Pv_t^f P_{\mathrm{PV},i} \quad \text{(Eq. 12)}$$

$$Q_{j,t} = Q_{j,t}^C \quad \text{(Eq. 13)}$$

Network constraints:

$$P_{ij,t}^2 + Q_{ij,t}^2 \le l_{ij,t} w_{i,t} \quad \text{(Eq. 14)}$$

$$\left\| \begin{matrix} 2P_{ij,t} \\ 2Q_{ij,t} \\ l_{ij,t} - w_{i,t} \end{matrix} \right\| \le l_{ij,t} + w_{i,t} \quad \text{(Eq. 15)}$$

$$w_{i,t}^{min} \le w_{i,t} \le w_{i,t}^{max} \quad \text{(Eq. 16)}$$

$$l_{ij,t} \le l_{ij,t}^{max} \quad \text{(Eq. 17)}$$

$$0 \le P_{\mathrm{PV},i} \le P_{\mathrm{PV},i}^{max} \quad \text{(Eq. 18)}$$

$$P_slack, t \ge 0 \quad \text{(Eq. 19)}$$

3. RESULTS AND DISCUSSION

The system under study has a total annual energy consumption of 12.47 MWh and operates on a single-phase line with a nominal voltage of 230 V. The voltage is

constrained within ±10% of the nominal value, corresponding to a range of 0.9 to 1.1 per unit (p.u.). The optimization problem is formulated in the Julia programming language and solved using Gurobi. This study analyzes the results in two main steps. In the first step, the impact of PV sitting on HC is evaluated. Six different sitting scenarios are proposed to ensure that the total residential PV installation at any location does not exceed 10 kW. The second step presents a detailed analysis through the following four case studies:

- **Base Case**: Analysis of system operating conditions before installing the PV on the roof to assess baseline performance.
- **Case 1**: Optimal HC when reverse power flow to the main grid is permitted.
- **Case 2**: Optimal HC when reverse power flow to the main grid is not permitted.
- **Case 3**: HC analysis based on Cambodia's installed capacity rules. This limits PV installations to less than 50% of rated MV/LV transformers.

3.1 HC Under Different Siting Locations

As shown in Table 1, when PV is installed near the slack bus in Scenario 1, the main constraint on hosting capacity (HC) is line current, which reaches the 175 A limit. Due to low voltage rise and proximity to the slack bus, up to 68.61 kWp can be installed, with a maximum voltage of 1.01 p.u. In Scenario 2, both voltage and current constraints limit HC due to low local consumption, allowing only 38.57 kWp. In Scenario 3, current becomes the main constraint, and higher losses are observed despite a higher HC potential. In the feeder's middle section in Scenario 4, both voltage and current jointly limit HC to 41.9 kWp. In the final sections in Scenarios 5 and 6, voltage rise is the dominant constraint, limiting HC to 31.73 kWp and 25.66 kWp, respectively, although losses are lower due to reduced current flow.

3.2 HC Under Different Proposed Case Study

This section presents the second part of the analysis about the four proposed case studies summarized in Table 2. The peak load is at 8:00 PM, when the total load demand reaches 33.4 kW. The asterisk (*) in the table indicates that the current flow is at the peak load period. During this peak period, the maximum current of the feed is 82.6 and is observed in the base case, case 2, and case 3. The maximum voltage and current are without asterisk state for PV peak production time (t = 1597). The annual energy loss in the basic case is 2.03 MWh.

Table 1. Simulation results of the HC in the proposed PV placement

N^0	Bus	Max Voltage (p.u.)	Max Current (A)	Annual Losses (MWh)	HC (kWp)
1	2, 3, 4	1.010	175.00	2.91	68.61
2	7, 8, 9	1.054	116.67	3.96	38.57
3	10, 11, 12	1.042	175.00	4.21	68.13
4	14, 15, 16	1.058	113.65	3.40	41.90
5	17, 18, 19	1.059	84.37	2.90	31.73
6	20, 21, 22	1.060	71.52	2.66	25.66

Table 2. Simulation results of HC in the proposed case studies

Parameter	Base Case	Case 1	Case 2	Case 3
Max Vol (p.u.)	1.00	1.054	1.00	1.01
Max Current (A)	82.6*	175.0	82.6*	82.6*
Annual Energy Losses (MWh)	2.03	7.00	1.84	1.70
HC (kWp)	0.00	168.62	14.80	37.50

* Indicate the current at their minimum/maximum values during peak load hours.

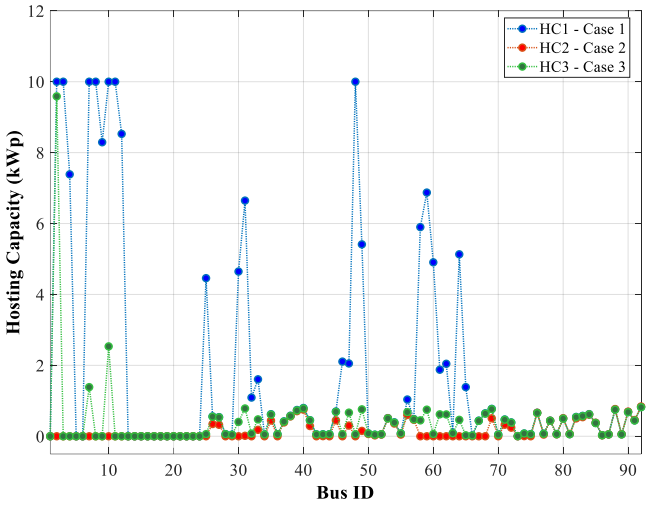


Fig. 4. PV hosting capacity over the three case studies

In case 1, reverse power flow to the main grid is allowed. This allows for a high HC of 168.62 kWp. However, significant currents are reversed in the MV/LV transformer during PV peak generation, leading to a significant increase in energy loss. This is almost 300% more costly than the base case. In Case 2, reverse power flow is prohibited, and the HC is limited to 14.8 kWp to ensure zero reverse power flow to the slack bus. As a result, only a 9 A of current is flowing during peak PV production, as the generated power is primarily used locally. This minimizes the flow of current through the long-distance stick from the MV/LV transformer, reducing the energy loss to 1.84 MWh. In Case 3, Cambodia's PV capacity is limited to 37.5 kWp. PV During PV generation, the maximum current is 42.28 A, with an additional annual energy loss reduced to 1.7 MWh.

Fig. 4 shows the placement and sizing of PV systems. The blue curve represents the HC for Case 1, where PV should be installed near the slack bus to avoid excessive voltage rise. To reduce operational power losses, PV is installed in the middle section of the feeder. To prevent overvoltage, the PV placement and size are adjusted to match the nearby demand,

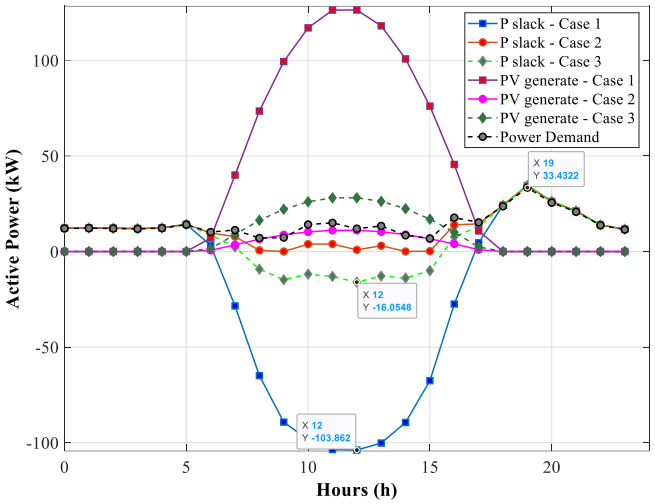


Fig. 5. Import/reverse power flow at the slack bus with corresponding PV output at high PV generation time

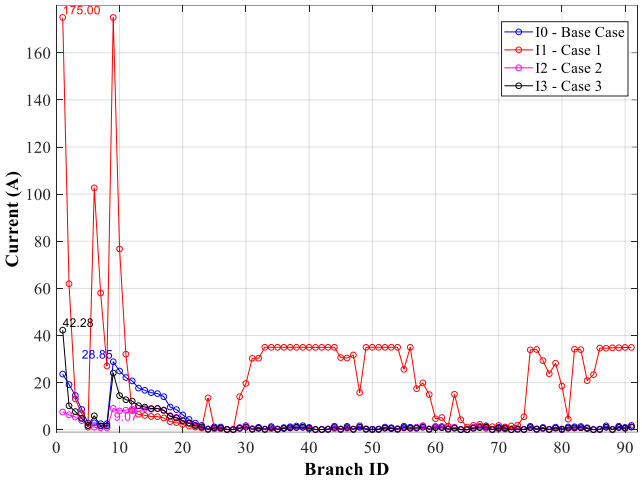


Fig. 6. Branch current flow in each line across the four Cases

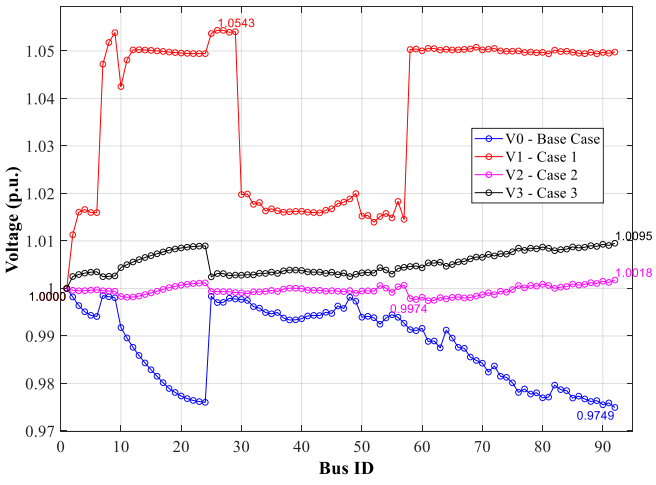


Fig. 7. Bus voltage in each node across the four Cases

as shown at Bus 70. The red curve indicates the HC for Case 2, in which PV is only installed on buses with existing loads, aiming to provide local support and prevent reverse power flow. The green curve represents the HC for Case 3, where PV is installed near the slack bus of each feeder, namely Buses 2, 7, and 10, and other locations close to load centers, similar to Case 2, starting from Bus 25.

Fig. 5 shows the PV generation profile and reverse power flow to the slack bus of the MV/LV transformer for the three cases. The black curve represents the daily load profile, which peaks at night. In Case 1, the maximum reverse power flow (blue curve) reaches 103.86 kW, about 152% of the transformer's nominal capacity. However, this goes beyond the typical first boundary but shows DSO that the system can handle more than 52% of the transformer's capacity for PV hosting. Including energy storage allows the system to better manage this capacity, support higher renewable penetration, and combat rapid demand growth. In case 2, the reverse power is zero (red curve). In Case 3, the PV capacity results in 23.7% of the active power flow to the slack bus. This is within the Canadian guideline, which permits reverse power flow up to 60% of the transformer's rating. As a result, the hosting capacity can be increased to a maximum of 168.62 kWp, up from the initial 37.5 kWp.

Fig. 6 shows the branch currents for the four proposed cases, with the maximum current values indicated on each curve. During peak PV production, the base case has a maximum current of only 28.85 A. In contrast, Case 1 has a maximum of 175 A, Case 2 shows 9.07 A, and Case 3 reaches 42.28 A. These differences highlight how the hosting capacity affects the current flow in the system. Fig. 7 displays the bus voltages at each node for these cases. The voltage is highest in Case 1, due to the larger hosting capacity being proposed. In Case 3, the bus voltage is 1.009 p.u., which corresponds to the maximum voltage at 50% of the transformer's capacity.

4. CONCLUSIONS

In this study, a Convex DistFlow OPF is used to transform complex, nonlinear power flow problems into reliable convex optimizations. This approach helps maximize HC and at the same time minimize energy losses. Reverse power flow is a main issue for DSOs as it determines how many PVs can be installed without any problems. The first step analyzes the PV placement and its effects on voltage and current limit. In the second step, by using an accurate model is used to identify the maximum safe PV capacity and how effectively it can be integrated. The results show that, according to Cambodia's PV installation rule, which assumes that the MV/LV transformer can safely handle up to 50% of its capacity as reverse power flow, the system can support PV injection up to that level. This suggests that the DSO can plan for more PV installations while keeping reverse power within acceptable limits.

ACKNOWLEDGMENTS

This work was funded by Académie de Recherche et d'Enseignement Supérieur (ARES) under Institut de Technologie du Cambodge with Université de Mons collaboration.

NOMENCLATURE

Sets

$\mathcal{N}$: Set of nodes (buses) in the network
$\mathcal{E}$: Set of edges (distribution lines) in the network
$\mathcal{T}$: Set of periods
$\mathcal{C}_j$: Set of child nodes connected to bus *j*

Parameters

R_{ij}, X_{ij}: Resistance/inductance of the distribution line between bus *i* and bus *j*
$P^c_{j,t}, Q^c_{j,t}$: Active/reactive power consumption at bus *j* at time *t*
$P_{j,t}, Q_{j,t}$: Net power consumption at bus *j* at time *t*
Pv^f_t: PV production factor at time *t*
γ_1 : Weighting coefficient for maximizing PV generation
γ_2 : Weighting coefficient for minimizing line losses
$P^{max}_{\mathrm{PV},i}$: Maximum PV capacity at bus *i*
P_slack, t: Power at the slack buses at time *t*

Variables

$P_{ij,t}, Q_{ij,t}$: Active/reactive power flow between bus *i* to *j* at *t*
$P_{PV,i}$: PV installation capacity at bus *i*
$l_{ij,t}$: Current magnitude squared in the line between bus *i* to *j* at time *t*
$w_{i,t}$: Voltage magnitude square at bus *i* at time *t*

REFERENCES

[1] "Regional: Southeast Asia Energy Sector Development, Investment Planning and Capacity Building Facility," Feb. 2024.

[2] "Long-Term Strategy for Carbon Neutrality," Dec. 2021.

[3] Papathanassiou *et al.*, "Capacity of distribution feeders for hosting distributed energy resources," *CIGRE (International Council on Large Electric Systems)*, 2014.

[4] J. Yuan, Y. Weng, and C.-W. Tan, "Determining maximum hosting capacity for PV systems in distribution grids," *International Journal of Electrical Power & Energy Systems*, vol. 135, p. 107342, Feb. 2022.

[5] M. Fresia, L. Bordo, F. Delfino, and S. Bracco, "Optimal day-ahead active and reactive power management for microgrids with high penetration of renewables," *Energy Conversion and Management: X*, vol. 23, p. 100598, Jul. 2024.

[6] B. Ahmadi, O. Ceylan, and A. Ozdemir, "Reinforcement of the distribution grids to improve the hosting capacity of distributed generation: Multi-objective framework," *Electric Power Systems Research*, vol. 217, p. 109120, Apr. 2023.

[7] A. Koirala, T. Van Acker, R. D'hulst, and D. Van Hertem, "Hosting capacity of photovoltaic systems in low voltage distribution systems: A benchmark of deterministic and stochastic approaches," *Renewable and Sustainable Energy Reviews*, vol. 155, p. 111899, Mar. 2022.

[8] M. J. Reno and R. J. Broderick, "Statistical analysis of feeder and locational PV hosting capacity for 216 feeders," in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, IEEE, Jul. 2016, pp. 1–5.

[9] M. Selcen Ayaz, M. Malekpour, R. Azizipanah-Abarghooee, M. Karimi, and V. Terzija, "Probabilistic photovoltaic generation and load demand uncertainties modelling for active distribution networks hosting capacity calculations," *International Journal of Electrical Power & Energy Systems*, vol. 159, p. 110016, Aug. 2024, doi: 10.1016/j.ijepes.2024.110016.

[10] A. Navarro-Espinosa and L. F. Ochoa, "Probabilistic Impact Assessment of Low Carbon Technologies in LV Distribution Systems," *IEEE Transactions on Power Systems*, vol. 31, no. 3, pp. 2192–2203, May 2016.

[11] M. Alturki and A. Khodaei, "Optimal loading capacity in distribution grids," in *2017 North American Power Symposium (NAPS)*, IEEE, Sep. 2017, pp. 1–6.

[12] C. C. Amorim Silva, W. N. Silva, and B. H. Dias, "Photovoltaic Hosting Capacity: An Approach Based on Optimal Power Flow," in *2024 Workshop on Communication Networks and Power Systems (WCNPS)*, IEEE, Dec. 2024, pp. 1–7.

[13] EAC, "Salient Features of Power Development in the Kingdom of Cambodia Until December 2024," Phnom Penh, Dec. 2024.

[14] "NASA Prediction Of Worldwide Energy Resources (POWER) Project." Accessed: Mar. 05, 2025. [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/

[15] M. Farivar and S. H. Low, "Branch Flow Model: Relaxations and Convexification—Part I," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2554–2564, Aug. 2013.

[16] "IEEE Recommended Practice for Utility Interface of Photovoltaic (PV) Systems," Jan. 30, 2000, *IEEE, Piscataway, NJ, USA*.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Heat Stress and Construction Labor Productivity: Identifying Key Factors Affecting Cambodian Construction Workers

Latin Heang[1,2], Kinnaleth Vongchanh[1,2,*], Sarin Chan[2]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.*

[2] *Thermal Lab, Industrial and Mechanical Engineering Department, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: kinnalethv@itc.edu.kh*

Abstract: *Heat stress is a critical factor affecting construction labour productivity, particularly in tropical climates. In Cambodia, labour productivity plays a vital role in the construction sector and the overall economy, influencing efficiency, costs, project timelines, and economic growth. This study aims to analyse key parameters impacting worker productivity, including environmental conditions, physiological responses, and personal attributes such as age, body mass index (BMI), smoking habits, alcohol consumption, work duration, and rest periods. Heat stress was assessed using a heat stress monitor to measure Wet Bulb Globe Temperature (WBGT), while heart rate was tracked via a heart rate sensor. Productivity was observed based on individual work activities. Additionally, demographic information was collected through interviews. Data collection was conducted twice during both the hot and cool months of the year in Phnom Penh to capture seasonal variations. The study involved 74 healthy male construction workers, randomly selected from voluntary participants on-site. The data was analysed using the R programming. The findings indicate significant correlations between labour productivity and factors such as WBGT, age, work intensity, work duration, rest duration, alcohol drinking and smoking habits, BMI, working shift and timeslot. Specifically, heat stress was found to have an adverse correlation with construction labour productivity ($r = -1.156$, $p < 0.001$), highlighting the negative impact of extreme temperatures on worker efficiency. These findings emphasise the key determinants of construction labour productivity, establishing a strong foundation for future model development. These results highlight the urgent need for climate adaptation strategies to protect worker well-being and sustain productivity in the construction sector.*

Keywords: Heat stress; Construction labour productivity; WBGT; Physiological factors; Cambodia

1. INTRODUCTION[1]

Climate change poses significant challenges globally, especially in tropical regions like Cambodia, where average temperatures in Phnom Penh have reached 28.03°C with projections of a rise between 0.7°C and 4.3°C by 2100 [1, 2]. The Ministry of Economics and Finance (MEF) and NCSD indicate that climate change could lead to labour productivity losses in sectors such as agriculture, manufacturing, and construction, which contribute about 10% to Cambodia's GDP [3]. Increased temperatures elevate the risk of heat stress among construction workers, impacting their productivity, safety, and overall well-being [4, 5]. Heat stress, defined as the body's total heat load from both internal metabolic processes and the external environment, can lead to various health issues and productivity declines, with studies showing reductions in work capacity under high temperatures [6]–[8]. Heat stress is a major occupational hazard affecting outdoor workers in sectors such as agriculture, construction, mining, firefighting, transportation, and brick-making [9]. Among those, construction is one of the most vulnerable to heat, which adversely affects construction labour productivity, resulting in delays and cost overruns in building projects [7]. Elements include elevated temperatures and humidity, direct solar exposure, excessive clothing, and insufficient hydration can lead to heat stress [10]. Chinnadurai et al. [7] demonstrated that outdoor workers engaged in moderate to high workloads, exceeding the threshold limit of 28°C, experienced a productivity decline of 18 to 35%. Additionally, Ayessaki & Smallwood [11] found that adverse

[1]* Corresponding author: Kinnaleth Vongchanh
E-mail: kinnalethv@itc.edu.kh; Tel: +855-99351 199

site conditions can adversely affect worker performance by 50%. A worker exerting moderate labour intensity has a 50% reduction in work capacity at temperatures of 33-34°C [8].

Various indices, particularly the Wet Bulb Globe Temperature (WBGT), are used to assess heat stress exposure [12]. Previous research highlights the negative effects of heat stress on construction productivity, yet evidence specific to Cambodia remains limited. This study aims to bridge the gap between occupational health and workforce efficiency in the construction sector by examining key factors that influence labour productivity.

2. METHODOLOGY

2.1 Study design

This cross-sectional study examined the impact of heat stress and other factors on construction labour productivity, using methodologies adapted from Yi and Chan [13]. Ethical approval was obtained from the National Ethics Committee for Health Research in Cambodia (Ref. No. 037 NECHR). Data collection was conducted by trained observers following a pilot study, with each observer assigned to a single worker for consistency. The study included 74 healthy male construction workers from five sites in Phnom Penh during cool months (November 2022 – February 2023) and hot months (March – June 2023) to assess seasonal variations in environmental conditions and physiological responses.

2.2 Data collection

This study employed a structured approach to data collection, focusing on three key components: questioning, measurement, and observation. These methods ensured a comprehensive assessment of factors influencing construction labour productivity, including personal attributes, environmental conditions, physiological responses, and productivity metrics. Personal information was gathered through structured questionnaires, collecting demographic and occupational details such as worker roles, site location, age, education level, and lifestyle habits like smoking and alcohol consumption. Additionally, health screenings were conducted to identify prevalent medical conditions, including hypertension, diabetes, cardiovascular diseases, and neurological disorders, while monitoring for heat stress symptoms such as headaches, dizziness, muscle cramps, and excessive sweating.

Environmental conditions were continuously recorded during working hours using a QuesTemp 36 heat stress monitor, with Wet Bulb Globe Temperature (WBGT) calculated following ISO 7243 standards [14]. Physiological responses were measured using specialised equipment, including a Polar H10 heart rate sensor to monitor heart rate and calculate the percentage of maximum heart rate [16]. Weight was recorded using an OHM Body Composition Scale, and height was measured using a Shinwa Tape Meter, with Body Mass Index (BMI) calculated accordingly.

To assess labour productivity, the study applied the AACE International method, categorising work activities into direct, indirect, and non-productive tasks. Several key indicators, including working time (WT), work duration (WD), average construction labour productivity (CLPavg), work duration average (WDavg), and average resting duration (RDavg), were measured to evaluate workforce efficiency. These metrics were calculated based on the duration spent on specific tasks relative to the total working hours, providing a detailed analysis of productivity trends within the construction sector.

2.3 Data analysis method

The data analysis in this study followed a structured statistical approach to summarise key characteristics of construction workers and examine factors influencing labour productivity. Descriptive statistical methods were first applied to identify trends, patterns, and potential outliers within the dataset. Key variables analysed included age, average Wet Bulb Globe Temperature (WBGTavg), work and rest durations, percentage of maximum heart rate (%HRmax), BMI, working shifts and timeslot.

Data analysis was conducted using R software. For continuous variables, measures of central tendency (mean values) and variability (standard deviation, range) were computed to describe data distribution. Categorical variables, including smoking habits (SH), alcohol consumption (ADH), and educational background, were analysed using frequency and percentage calculations to develop a demographic profile of the workforce.

To examine relationships between key continuous variables, correlation analysis was performed using Pearson correlation coefficients (r). This analysis quantified the relationship between CLPavg (construction labour productivity), WBGTavg, age, work duration (WDavg), rest duration (RDavg), %HRmax, BMI, ADH, SH, working shift, and timeslot, identifying potential predictors that may affect construction labour productivity.

3. RESULTS AND DISCUSSION

3.1 Descriptive statistic

The descriptive statistical result shown in Table 1 analysis examined key factors influencing construction labour productivity (CLP) among 74 workers engaged in rebar (32 participants) and moulding (42 participants) work. The study was conducted across both hot and cold months, with an average heat stress level (WBGT) of 28.75°C, ranging from

20.23℃ to 36.58℃, covering low, moderate, and high heat stress risks.

Participants had an average age of 34 years and a BMI of 22.34 kg/m². Heart rate analysis revealed an average exertion of 51% HRmax, occasionally peaking at 87% HRmax, indicating high physical strain without changes in worker routines. Productivity assessment showed an overall construction labour productivity of 62% ± 48, with workers dedicating 41.73% ± 26.96 of their time to work-related tasks, while 20.85% ± 21.33 was spent resting or engaged in non-productive activities.

Table 1. Descriptive statistics of various variables

Variable	Mean ± SD	Range	Unit
Age	34 ± 11	18 - 57	Years
CLP	62 ± 48	0 - 100	%
WBGT	28.75 ± 2.87	20.23 - 36.58	℃
%HRmax	51.19 ± 8.14	30 - 87.86	%
BMI	22.36 ± 2.92	17.01 - 29.66	kg/m²
WDavg	41.73 ± 26.96	0 - 100	%
RDavg	20.85 ± 21.33	0 - 100	%

3.2 Environmental condition and construction labour productivity

Table 2 showed the hourly distribution of the heat stress risk levels (WBGT) in Cambodian construction sites varies throughout the day. Low-risk conditions (Risk I, WBGT < 27.9°C) occur mainly from 7:00 to 8:00, making mornings the most favourable for work. Moderate-risk conditions (Risk II, WBGT 27.9–32.2°C) peak between 8:00 and 10:00, with a secondary rise around 16:00, requiring caution but remaining manageable. High-risk conditions (Risk III, WBGT ≥ 32.2°C) are most severe from 10:00 to 15:00, significantly impacting productivity due to intense heat stress. To optimise efficiency, construction activities should be strategically scheduled, with essential preventive measures such as hydration, shade, and rest breaks during high-risk periods.

Table 2. The hourly percentage of WBGT risk level and labour productivity activity

Time	WBGT risk level (%)			Productivity (%)		
	I	II	III	D	I	N
7:00-8:00	77	23	0	44	9	46
8:00-9:00	67	27	5	64	12	24
9:00-10:00	59	27	14	64	11	25
10:00-11:00	40	34	17	51	10	39
13:00-14:00	40	36	25	57	14	29
14:00-15:00	47	33	21	61	13	26
15:00-16:00	53	37	11	57	11	32
16:00-17:00	69	29	2	51	9	40

Table 2 illustrates the percentage of time workers dedicate to direct (D), indirect (I), and non-productive activities (N). During the first hour (7:00–8:00), 44% of workers engaged in direct productivity, 9% in indirect tasks, and 46% in non-productive activities. Productivity increased between 9:00, with direct productivity rising to 64%, while indirect productivity remained stable at 12%, and non-productive activities decreased to 24%.

The highest engagement in direct productivity (64%) occurred between 10:00, with 10–12% of workers involved in indirect productivity and 20–25% in non-productive tasks. These shifts highlight the importance of strategic work scheduling to optimise efficiency throughout the day.

3.3 Physiological monitoring

Table 3 presents the hourly distribution of heart rate (HR) and work intensity (%HRmax) among construction workers. During the morning shift (7:00–11:00), the mean heart rate ranged from 92.1 to 96.8 bpm, with standard deviations of 12.1 to 13.5, indicating variability among workers. Corresponding work intensity ranged from 49.7% to 52.5% HRmax, with standard deviations between 7.2% and 8.1%. In the afternoon shift (13:00–17:00), heart rate slightly increased (92.3–98.9 bpm), with standard deviations of 13.6–14.6, while work intensity ranged from 49.66% to 53.44% HRmax, with deviations from 8.06% to 8.69%.

Workers showed slightly higher heart rates and work intensity in the afternoon shift, which may contribute to fatigue and reduced concentration. Monitoring heart rate variations helps assess physiological responses to work conditions, enabling targeted interventions to improve worker health and performance [17].

Table 3. Descriptive statistics of heart rate at different time

Time	Heart Rate (bpm)		% Heart Rate Max (%)	
	Mean ± SD	Range	Mean ± SD	Range
7:00-8:00	96.8 ± 13.2	58 - 158	52.4 ± 8.1	31.7-84.4
8:00-9:00	95.6 ± 12.1	55 - 156	51.5 ± 7.2	30.4-78.6
9:00-10:00	92.7 ± 12.7	55 - 147	49.9 ± 7.4	30.0-79.5
10:00-11:00	92.1 ± 13.4	57 - 158	49.6 ± 7.8	30.0-80.4
13:00-14:00	98.9 ± 13.6	57 - 152	53.4 ± 8.4	30.5-87.9
14:00-15:00	97.3 ± 14.6	57 - 160	52.5 ± 8.6	30.4-86.5
15:00-16:00	94.0 ± 14.5	56 - 157	50.6 ± 8.5	30.0-79.8
16:00-17:00	92.2 ± 13.6	58 - 162	49.6 ± 8.0	30.0-80.2

3.4 Correlation analysis

The results, presented in Table 4, indicate significant correlations between multiple factors. CLPavg demonstrated positive associations with work duration, rest duration, age, alcohol consumption, and smoking habit with $P < 0.001$, suggesting that extended working hours, sufficient rest, and certain behavioural traits contribute to enhanced productivity. Conversely, negative correlations were observed between CLPavg and WBGT ($p < 0.001$), %HRmax ($p < 0.01$), BMI ($p < 0.05$), working shifts and timeslot with $p < 0.001$, highlighting potential adverse effects of environmental and physiological strain on worker productivity. The significant

correlations between productivity and variables such as WBGT, work duration, rest periods, age, BMI, alcohol consumption, smoking habits, working shift, and timeslot highlight their predictive potential. Future model development can leverage these insights to optimise workforce management, improve heat stress mitigation strategies, and enhance productivity forecasting. By applying statistical or machine learning approaches, this framework could support data-driven decision-making to improve worker efficiency and resilience in Cambodia's construction sector.

Table 4. Correlative correlation of each variable

Variables	CLP avg.	WBGTavg.
WBGTavg.	-1.156***	
WDavg.	0.985***	-0.074***
RDavg.	0.418***	0.014
%HRmax	-0.059**	0.155***
BMI	-0.042*	-0.381***
ADH	0.219***	0.080***
SH	0.177***	0.276***
Shift	-0.819***	0.382***
Timeslot	-0.576***	0.410***
Age	0.086***	-0.352***

*Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1*

4. CONCLUSIONS

Rising temperatures are more likely to influence construction labour productivity in Cambodia's tropical climate, where extreme heat poses significant challenges to workers.

The results showed that heat stress, age, work intensity, work duration, rest duration, alcohol consumption, and smoking habits are likely productivity predictors.

Subsequent investigations ought to concentrate on the advancement of predictive modelling specifically designed for Cambodia's labour force, with an emphasis on rebar and moulding workers. Advanced statistical and machine learning methodologies can facilitate the quantification of productivity patterns, evaluate environmental and physiological effects, and enhance workforce management strategies.

ACKNOWLEDGMENTS

We thank Cambodia Climate Change Alliance Phase 3 for financial support, Dr. Jackie Yang, for guidance on data analysis, the NUS team and HeatSafe project for measurement methodology assistance, Professor Albert Chan of The Hong Kong Polytechnic University for support, and the construction company for data collection access. We thank the Thermal Lab members CHEA Chantola, HOENG Hollin, BUN Panha, and LY Senglim for their invaluable data input and Ken Pisal for his data collecting.

REFERENCES

[1] ClimateCHIP, "Your Area: Today (switch to: Tomorrow) | Climate CHIP," *ClimateCHIP*, 2023. http://climatechip.org/your-area-climate-data (accessed Jul. 21, 2023).

[2] S. Aguilar-Gomez, S. Baquie, and P. J. Robyn, "Environmental Hazards, Climate, and Health in Cambodia," 2024.

[3] CCCA2, "Annual Progress Report 2018," Phnom Penh, 2018. [Online]. Available: https://info.undp.org/docs/pdc/Documents/KHM/CCCA%20Anual%20Progress%20Report%202018_Final.pdf

[4] V. S. Fuentes, F. Bühler, and W. B. Markussen, "Use of Industrial Excess Heat to Produce District Cooling in Tropical Countries," 2022, doi: 10.52202/062738-0118.

[5] A. Ojha, S. Shakerian, M. Habibnezhad, H. Jebelli, S. Lee, and F. M. Sadra, "Feasibility of Using Physiological Signals From a Wearable Biosensor to Monitor Dehydration of Construction Workers," 2020, doi: 10.3311/ccc2020-004.

[6] B. L. Jacklitsch, W. J. Williams, K. Musolin, A. Coca, J.-H. Kim, and N. Turner, *Criteria for a recommended standard: occupational exposure to heat and hot environments*. 2016.

[7] J. Chinnadurai, V. Venugopal, P. Kumaravel, and R. Paramesh, "Influence of occupational heat stress on labour productivity–a case study from Chennai, India," *Int. J. Product. Perform. Manag.*, vol. 65, no. 2, pp. 245–255, 2016, doi: 10.1108/IJPPM-08-2014-0121.

[8] T. Kjellstrom, N. Maître, C. Saget, M. Otto, and T. Karimova, "Working on a warmer planet: The impact of heat stress on labour productivity and decent work," *Geneva Int. Labor Organ.*, 2019.

[9] P. Habibi, J. Razmjouei, A. Moradi, F. Mahdavi, S. Fallah-Aliabadi, and A. Heydari, "Climate change and heat stress resilient outdoor workers: findings from systematic literature review," *BMC Public Health*, vol. 24, no. 1, p. 1711, 2024.

[10] NIOSH, *Heat Stress: Criteria for a Recommended Standard*. 2016. [Online]. Available: http://books.google.com/books?id=J7plB39nbDMC

[11] W.-Y. Ayessaki and J. Smallwood, "Influencing workers' performance through health and safety interventions," in *Procedia engineering*, 2017, vol. 182, pp. 42–49. doi: 10.1016/j.proeng.2017.03.111.

[12] D. G. Park, S. C. Shin, S. W. Kang, and Y. T. Kim, "Development of Flexible Self Adhesive Patch for Professional Heat Stress Monitoring Service," 2005, doi: 10.1109/iembs.2005.1617309.

[13] W. Yi and A. P. C. Chan, "Effects of heat stress on

construction labor productivity in Hong Kong: a case study of rebar workers," *Int. J. Environ. Res. Public Health*, vol. 14, no. 9, p. 1055, 2017, doi: 10.3390/ijerph14091055.

[14] K. Parsons, "Heat stress standard ISO 7243 and its global application," *Ind. Health*, vol. 44, no. 3, pp. 368–379, 2006.

[15] G. M. Budd, "Wet-bulb globe temperature (WBGT)-its history and its limitations," *J. Sci. Med. Sport*, vol. 11, no. 1, pp. 20–32, 2008, doi: 10.1016/j.jsams.2007.07.003.

[16] P. Kor, K. Vongchanh, L. Heang, S. Chan, and J. Yang, "Measurement Survey on Construction Labour Productivity under Heat Stress during the Cool, ," *J. Emerg. Technol. Innov. Res.*, vol. 10, no. 10, pp. d909–d916, 2023, [Online]. Available: https://www.jetir.org/view?paper=JETIR2310306

[17] Szymon Siecinski, P. Kostka, and E. J. Tkacz, "Comparison of Heart Rate Variability Indices Based on Seismocardiograms From Healthy Volunteers and Patients With Valvular Heart Diseases," 2022, doi: 10.22489/cinc.2022.014.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

3D Modeling of a Continuous Mixing Machine for Biomass Briquette Systems

Hollin Hoeung [1], Kinnaleth Vongchanh [1,2,*], Chantola Chea [1], Latin Heang [1,2], Sarin Chan [1]

Thermal Lab, Industrial and Mechanical Engineering Department, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.*

[*]*Corresponding author: kinnalethv@itc.edu.kh*

Abstract: *Nowadays, briquettes are widely utilized to replace traditional fuel in different applications for sustainable renewable heat sources to provide in the residential and industrial sector. This study focuses on the design of a continuous mixing machine optimized for biomass briquette production to mix organic waste materials into a uniform blend for high-quality renewable heat sources, such as tree leaves, sawdust, and fish oil as a binder. Continuous mixing is essential in order to achieve consistent briquette quality and quantity to satisfy production requirements. The designed mixing machine features continuous screw and paddle blade mechanisms for the proposed ensuring effective material blending without allowing particle segregation. This machine is able to handle raw materials with varying densities, ensuring homogeneity before feeding into the briquetting process. The design process involves selecting optimal design parameters for a cylindrical mixing tube with a length of 2 meters and a diameter of 0.3 meter, including a mixing capacity of 100 kilograms per hour to mix waste materials to produce high-quality biomass briquettes. It is optimized to handle particle sizes of 10 millimeters or smaller, ensuring uniform mixing and consistent briquette formation, the optimal mixing speed ranges from 12 to 15 RPM, providing effective material blending, the system requires a 0.5 horsepower motor, ensuring efficient power consumption and smooth operation, in addition material selections are mild steel and stainless steel to enhance durability and performance, to creating a 3D model using Solid Edge software providing a detailed visualization of the machine's structure and assembly, and selecting materials will prioritize strength, durability, corrosion resistance, thermal conductivity, and cost-effectiveness. The system consists of several mechanical components, such as a screw conveyor with a diameter of 0.3 m and length of 0.4 m, bearings with an inner diameter of 0.09 m, belts of 0.013 m width, and shaft with a diameter of 0.09 m and length of 2 m to ensure efficient power transfer and minimize energy loss.*

Keywords: 3D Modeling; Tree Leaves; Sawdust; Continuous Mixing Machine; Biomass Briquette system

1. INTRODUCTION

A mixing machine, or mixer, is a mechanical device designed to combine two or more materials, either solids or liquids, into a uniform mixture. In biomass briquette production, a mixer is used to combine waste organic materials from agricultural residues into a well-mixed, uniform feedstock suitable for further processing [1]. It directly influences the quality of the briquettes produced. Biomass briquettes, formed from organic agricultural residues such as tree leaves (TL), sawdust (SD), and binders like fish oil (FO) , serve as an eco-friendly alternative to fossil fuels for heating, cooking, and industrial applications [2,3,4].

There are two general types of mixing methods: batch mixing and continuous mixing. In batch mixing, materials are loaded into a mixer in set amounts and processed in separate cycles. While it offers simplicity, batch mixing often suffers from inconsistencies between batches and longer production times [5].

On the other hand, continuous mixing uninterrupted flow, offering higher efficiency, consistent output, and reduced variation in mixing quality. Continuous mixers also support axial and radial blending, reduce segregation, and are more compact, making them ideal for industrial-scale production. There are several types of continuous mixers such as rotating drum mixers, ribbon mixers, static mixers, and paddle mixers [5][6].

This study focuses on the design and calculation of a continuous mixing machine specifically for biomass briquette production. The machine integrates a screw and paddle blade

[*] Corresponding author: Dr. Kinnaleth Vongchanh
E-mail: kinnalethv@itc.edu.kh; Tel: +855-99351199

mechanism, which allows for the effective transport and mixing of materials with differing densities. The aim is to ensure homogeneous mixing of materials with varying moisture content and particle size before they enter the briquetting stage. A crucial step in the production of high-quality briquettes is the mixing process, which directly affects the uniformity, density, and combustion performance of the final product [2].

2. METHODOLOGY

The design process involves the use of Solid Edge 2023 software to create a 3D model of the mixing system. That consists of few main components such as the mixing tube, screw, paddles, bearings, belts, pulleys, and shaft, providing a detailed visual, fabrication, and assembly. In addition, the design requires the process; mixing machines improve productivity, reduce human error, and support continuous, large-scale operations with the capacity of 100 kg/hr mass flow rate and suitable for TL, SD, and FO.

2.1 System design

The mixture material was described as follows in Fig. 1. The mixing machine has a cylindrical shape and is horizontal. There are screws and paddle blades connected to the main shaft. The continuous raw materials (TL, SD) inlet to the mixing tube by the hopper. At the same time, the main shaft rotated, which accelerated by the motor through the belt transmission to the screw conveyors, and the paddle blade moved along with the main shaft. When the raw materials inside the cylinder tube are pushed by the screw conveyor, push them toward the pump to inject the FO. All the unmixed material will be mixed by the paddle blade and move slowly to uniform the mixture, then the next screw conveyor will continuously be pushing the mixture to the outlet.

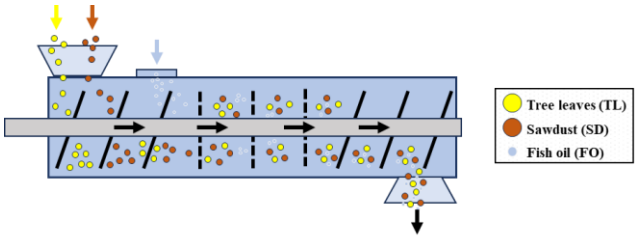


Fig. 1. 2D diagram horizontal continuous mixers

The design analysis is critical to ensuring that the machine operates efficiently and meets the production requirement of a mixing capacity of 100 kg/hr and the density of materials is 100 kg/m^3. It is necessary to calculate key mechanical parameters such as:

The required Mass flow rate for the mixing machine can be determined using the mass flow rate and density:

$$\dot{m}_b \times \rho = \dot{Q}_b \quad \text{(Eq. 1)}$$

where, $\dot{Q}_b$ is the volume flow rate, ρ is the density of bulk materials, and $\dot{m}_b$ is the mass flow rate.

The volume will be used to calculate the shape for the mixing machine:

$$V_{cylinder} = \frac{\pi D^2 L}{4} \quad \text{(Eq. 2)}$$

where, $V_{cylinder}$ is the volume of the cylindrical tube, D is the diameter of the cylindrical tube, and L is the length of the cylindrical tube.

2.2 Screw conveyor and paddle blades

The important step was to study the basics of the continuous mixing machine. It included the machine element. The main components of a biomass mixing machine are the screw and paddle blades. Thus, our focus is more on the research and the study of an appropriate blade design which will serve our purpose. Materials move through an external force during mixing because their movement is not naturally occurring.

The mixing devices can be categorized based on how an external force is applied [7] or made up of various forms of movement [8], such as convective mixing, shear mixing, and diffusive mixing.

The main components of a biomass mixing machine are screws and paddle blades. A screw conveyor, also known as an auger conveyor, is a mechanical device that uses a rotating helical screw blade, called a flighting to move materials, including liquids and granular substances, within a tube or trough. Screw conveyors are widely used in the food industry, building construction and mining companies, and chemical, agricultural and processing industries, mostly for elevating and/or transporting bulk materials over short to medium distances [9][10].

The volume will be used to calculate finding the shape for the mixing machine:

$$V_{sc} = \frac{\pi}{4}(D_{out}^2 - D_{in}^2) \times t_p \quad \text{(Eq. 3)}$$

where, V_{sc} is the volume of the screw, D_{out} is the outer diameter of the screw flight, D_{in} is the inner diameter of the screw flight, and t_p is the thickness of the screw flight.

The force will be used to calculate for the mixing machine:

$$F_{sc} = \frac{mgL(1+tan\phi)}{p} \quad \text{(Eq. 4)}$$

where, F_{sc} is the force that acting on screw, m is the mass of the bulk material, and g is the gravity of earth, p is the length of the screw pitch, and ϕ is the angel between force and surface of the screw flight.

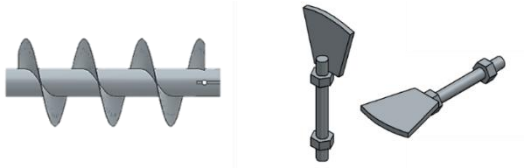


Fig. 2. Screw conveyor and paddle blade

For a continuous mixing machine, a paddle blade or propeller is a blade, typically attached to a shaft, designed to mix liquids, solids, or both. Paddle blades are generally flat and wide, suitable for beating batters and frostings. Propellers, on the other hand, have curved blades, allowing for efficient agitation and liquid circulation in both horizontal and vertical planes. The mixing arm is subjected to the combined effect of tensile and bending stresses. The most vulnerable cross section of blade is at the connection between the mixing arm and the mixing shaft. And when angle between blade and shaft axis 90° [11]. The best quality of the mixture corresponds to the longest processing time. With increasing mixer performance, the quality of the mixture decreases slightly. An increase in the number of paddles significantly improves the quality of the mixture and should be at least 5 [12][13]. The simulation results showed that increasing the impeller speed from 10 RPM to 40 RPM improved mixing quality across all conditions. However, further increasing the speed to 70 RPM negatively affected mixing at a 60% fill level. Higher vessel fills levels (40% to 60%) enhanced mixing at higher speeds, but fill level had little effect at the low speed of 10 RPM. The particle loading arrangement had minimal impact on mixing quality [14].

The volume of paddle will be used:

$$V_{paddle} = 2\pi r^2 h + 3\sqrt{3}s^2 - 2\pi r^2 \quad \text{(Eq. 5)}$$

where, V_{paddle} is the volume of the paddle blade, h is the height of paddle, s is the side length of the hexagonal nut, and r is the radius of the paddle blade.

Force acting on the paddle:

$$F_{pad} = \frac{1}{2}\rho_{pad} A_{eff} C_d v^2 \quad \text{(Eq. 6)}$$

where, F_{pad} is the force that acting on the surface of the blade, ρ_{pad} is the density of the paddle, $A_{eff} = A_{pad} cos\theta$ is the effective area projected, the C_d is drag coefficient of the paddle, v is the velocity of the paddle.

2.3 Main shaft

A transmission shaft is a rotating mechanical component that transfers rotational torque and power from one location to another within a machine. According to Fig. 3. For continuous mixing machine, screw and paddle blades are typically attached to a shaft arranged by 3 zones, with zones one and three being the screw and zone two being the paddle blades. It consisted of 14 paddle blades and 2 parts of screw conveyors. It was observed that the FO, SD, and TL was conveying and mixing roughly.

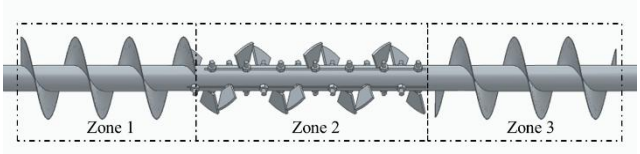


Fig. 3. The assembly of main shaft of mixing machine

Torque that requires to move the mass of materials:

$$T_{sc} = F_{sc} \times \frac{2D_{sc} - p_{sc}}{4} \quad \text{(Eq. 7)}$$

$$T_{pad} = n_{pad} F_{pad} \times A_{pad} \quad \text{(Eq. 8)}$$

where, T_{sc}, T_{pad} are the torque of the screw and paddle blade, ρ_{pad} is the density of the paddle, n_{pad} is the number of the paddle blade, and A_{pad} is area of the paddle blade.

Power consumption that requires for the machine:

$$P = T_{total} \times \omega \quad \text{(Eq. 9)}$$

where, P is the power that require for the mixing machine, T_{total} is the total torque that require, and ω is the angular velocity.

2.5 Frame/Support

In Fig. 5 shows that support structure of the machine. All the machine components are mounted on this support. In order to get the required strength, two plates (Bearing support shaft) are fixed with the help of nut and bolts. The method used for fabrication is arc welding. The machine frame is made of mild steel angle-shaped profile with a size of 2.5 x 1.04 x 1.2 m which is connected through welding process.

2.6 Assembly of 3D model

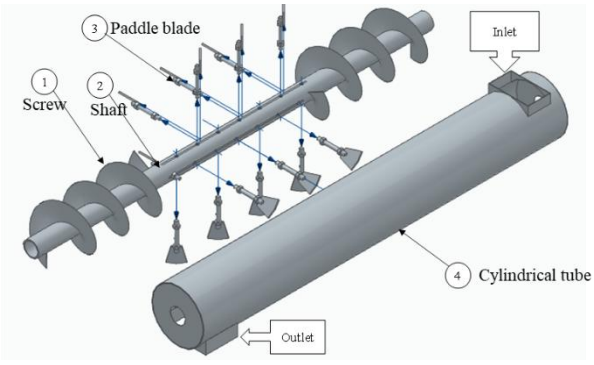


Fig. 4. 3D modeling components of continuous mixing machine

Once the calculation is fixed, the second important thing is the machine design. The design was be done in Solid edge

2023. The different components designed along the blade are shaft, cylindrical tube, and support, etc.

The 3D design of the continuous mixing machine was developed to provide a complete visual representation of the system and its components. The full assembly includes the mixing chamber, screw shaft, paddle blades, bearings, pulley system, belt drive, and support frame. The 3D model serves as a blueprint for fabrication and performance analysis, validating the structural integrity and functional layout of the machine.

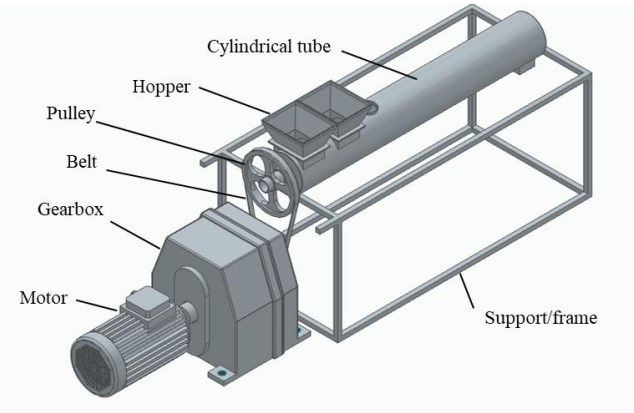


Fig. 5. Isometric view of full assembly of mixing machine

3. RESULTS AND DISCUSSION

As we made the design according to requirements, the necessary calculations were also carried out. In the last stage to checking whether the machine runs properly or not, it was observed that the paper was getting cut into strips but it was returning again with the blades.

The design aims to support a production capacity of 100 kg/hr, processing raw materials with a particle size of 10 mm or smaller, and operates at an optimal speed range of 12 to 20 RPM.

The screw conveyor was assumed to have a thickness of 2 mm, with a pitch defined as two-thirds of its diameter. A clearance of 2 mm between the screw and the mixing cylinder was considered, corresponding to a volume flow rate of 0.0166 m³/min. Over a mixing period of 8 minutes, the system must handle approximately 0.133 m³ of material.

The volume inside the mixer was determined by accounting for the space occupied by the screw, shaft, and paddles, with the paddle geometry derived from catalog data. Each paddle volume was calculated as 0.0037 m³, including the blade and attached nut. For mixing consistency, four paddles per pitch were included over 40% of the total mixer length, and the fill level was set at 80%. This resulted in a required mixing tube length of approximately 2.36 meters.

To calculate the mechanical performance, the torque required to drive both the screw and paddles was calculated by Eq. 7, 8, and 9. The screw torque was found to be 243.31 N·m, while the paddle contributed 3.15 N·m, and the totaling a combined torque of 246.45 N·m.

The rotational speed needed to maintain the desired flow rate was calculated to be 12.5 RPM, corresponding to an angular velocity of 1.308 rad/s. Using these values, the power required to drive the system was determined to be 322.58 W, or approximately 0.439 horsepower, supporting the motor specification of 0.5 HP.

Table 1. Mixing machine specifications

Parameter	Specification	Unit
Scale	2.34 (Length) x 1.2 (Width) x 1.04 (Height)	m
Production Capacity	100	kg/hr
Motor power	≈0.5	HP
Material type	Mild steel / Stainless Steel	
Mixing speed	10-20 [10][14]	RPM
Particle size input	≤10	mm
Suitable Materials	Fish oil (binder), sawdust, tree leaf	
Transmission Components	Shaft, bearings, belt, Pulleys	
Power source	Electric motor	

Based on the design, a continuous mixing machine is able to provide on scale for the screw pressing machine, which fitted to capacity 100kg/hr of production. From the research on experimental investigation on sawdust-tree leaves Briquette [2], the density of briquettes were approximately between 895kg/m^3 to 1180kg/m^3, with the expected the mass of a single briquette with the length of 10cm was 0.25kg to 0.3kg, which deliver consistent around 330 briquettes. When the heating value was between 21MJ/kg to 22MJ/kg, which provide for screw pressing machine produces 6,930MJ in 1hr.

In The design of a continuous mixing machine also presents challenges related to control accuracy, and energy efficiency. Ensuring the system maintains a reliable continuous flow, handles diverse material properties, and operates within acceptable power limits is essential for effective large-scale application. This study aims to overcome these challenges through careful design analysis and engineering validation.

4. CONCLUSIONS

This study successfully designed a continuous mixing machine optimized for biomass briquette production. By integrating a screw conveyor and paddle blade mechanism, the system ensures uniform blending of organic waste

materials such as tree leaves, sawdust, and fish oil binder. The machine effectively processes raw materials with varying densities, minimizing particle segregation and achieving consistent briquette quality. Through calculations and analysis, the optimal design parameters were selected, including a cylindrical mixing tube with a length of 2.34 meters, an operating speed range of 12.5 RPM, and a power requirement is 0.5 HP. The system efficiently mixes up to 100 kg of material per hour, ensuring continuous operation. It is specifically designed to process a mixture of tree leaves, sawdust, and fish oil as a binder, achieving uniform material distribution. The design integrates mechanical engineer principles with 3D modeling to optimize efficiency, reliability, and suitability for integration into a continuous biomass briquette production line.

Nomenclature

A	area	(m^2)
D	diameter	(m)
F	force	(N)
g	gravity of earth	(m/s^2)
L	length	(m)
m	mass	(kg)
$\dot{m}$	mass flow rate	(kg/s)
n	number of pitch	-
N	rotational speed	(rpm)
p	pitch	(m)
P	power	(Watt)
$\dot{Q}$	volume flow rate	(m^3/s)
r	radius	(m)
T	torque	(N.m)
v	velocity	(m/s)
V	volume	(m^3)
ρ	density	(kg/m^3)
ϕ	angel between force and surface of screw	(°)
ω	angular velocity	(rad/s)

REFERENCES

[1] P. Dinesha, S. Kumar, and M. A. Rosen, "Biomass Briquettes as an Alternative Fuel: A Comprehensive Review," *Energy Technol.*, vol. 7, no. 5, 2019, doi: 10.1002/ente.201801011.

[2] S. C. Heang Latin, Porchaing Choeng, Kinnaleth Vongchanh, "Experimental Investigation on Sawdust-Tree Leaves Briquette Using Fish Oil and Waste Oil As Binder," *Techno-Science Res. J. 8*, no. June 2023, pp. 84–93, 2020.

[3] P. Donald, C. Sanchez, M. Me, T. Aspe, and K. N. Sindol, "An Overview on the Production of Bio-briquettes from Agricultural Wastes: Methods, Processes, and Quality," *J. Agric. Food Eng.*, vol. 3, no. 1, pp. 1–17, 2022, doi: 10.37865/jafe.2022.0036.

[4] G. K. Ngusale, Y. Luo, and J. K. Kiplagat, "Briquette making in Kenya: Nairobi and peri-urban areas," *Renew. Sustain. Energy Rev.*, vol. 40, pp. 749–759, 2014, doi: 10.1016/j.rser.2014.07.206.

[5] E. L. Paul, V. A. Atiemo-obeng, and S. M. Kresta, *HANDBOOK OF INDUSTRIAL MIXING HANDBOOK OF INDUSTRIAL MIXING Edited by*. 2004.

[6] C. D. R. P. J. Cullen, Rodolfo J. Romañach, Nicolas Abatzoglou, *Pharmaceutical Blending and Mixing*. 2015.

[7] K. Gotoh, K. Iinoya, and K. Higashitani, *Powder technology handbook*, Forth edit., vol. 10, no. 1. 2020. doi: 10.1080/07373937.2020.1757994.

[8] A. W. N. N Harnby, M F Edwards, *Mixing in the Process Industries*, vol. 2. 1997.

[9] L. Pezo, A. Jovanović, M. Pezo, R. Čolović, and B. Lončar, "Modified screw conveyor-mixers - Discrete element modeling approach," *Adv. Powder Technol.*, vol. 26, no. 5, pp. 1391–1399, 2015, doi: 10.1016/j.apt.2015.07.016.

[10] A. Jovanović, L. Pezo, S. Stanojlović, N. Kosanić, and L. Lević, "Discrete element modelling of screw conveyor-mixers," *Hem. Ind.*, vol. 69, no. 1, pp. 95–101, 2015, doi: 10.2298/HEMIND130412026J.

[11] Z. Yao, R. Yang, W. Yuan, and H. An, "Mechanical analysis and optimal design of mixing paddlesforcsam mixers," *Acad. J. Manuf. Eng.*, vol. 17, no. 4, pp. 80–86, 2019.

[12] A. A. H. Al-Maidi *et al.*, "Modelling the quality of the mixture in a continuous paddle mixer," *Int. J. Agric. Stat. Sci.*, vol. 16, no. 2018, pp. 1769–1774, 2021.

[13] B. Jadidi, M. Ebrahimi, F. Ein-Mozaffari, and A. Lohi, "Effect of the Mixer Design Parameters on the Performance of a Twin Paddle Blender: A DEM Study," *Processes*, vol. 11, no. 3, 2023, doi: 10.3390/pr11030733.

[14] A. Yaraghi, M. Ebrahimi, F. Ein-Mozaffari, and A. Lohi, "Mixing assessment of non-cohesive particles in a paddle mixer through experiments and discrete

element method (DEM)," *Adv. Powder Technol.*, vol. 29, no. 11, pp. 2693–2706, 2018, doi: 10.1016/j.apt.2018.07.019.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

A Review of Environmental Factors Affecting Students' Cognitive Performance

Pheak KOR [1], Kinnaleth VONGCHANH [1,2*]

[1] *Department of Industrial and Mechanical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: kinnalethv@itc.edu.kh*

Abstract: *People usually spend time inside the building, thus a pleasant and healthy environment is essential for their daily lives, work and studies. Many research studies indicate that Indoor Environmental Quality (IEQ) greatly affects individuals' physical and mental well-being, comfort, and, crucially, their academic success. IEQ comprises four key elements: thermal conditions, indoor air quality (IAQ), noise levels, and lighting. Ambient temperature, air velocity, and relative humidity are examples of thermal elements that affect students' comfort and ability to focus. IAQ is assessed through CO_2 levels and Total Volatile Organic Compounds levels (TVOCs), which can affect both cognitive performance and respiratory health. Effective ventilation and air filtration are essential for ensuring good IAQ. Good noise levels assist in minimizing distractions, enabling students to hear and process information effectively. Furthermore, appropriate lighting whether natural or artificial helps relieve eye strain and improve focus. Given the significant role of IEQ and IAQ in shaping students' learning experiences, it is essential to establish reliable methods for measuring their impact on academic performance. In this regard, various assessment techniques are employed to evaluate the extent to which these environmental factors influence students' cognitive abilities. To assess the impact of these environmental factors, neurobehavioral tests are administered to evaluate students' academic and cognitive performance. These typically include 13 assessments: numerical calculations, digit span memory tests (number, letter and symbol recall), visual learning with picture recognition, reading comprehension, letter search tasks, event sequence, graphic abstracting, hand-eye coordination, creative thinking, overlapping, spatial image, conditional reasoning, and visual choice reaction time. Following the review, numerous studies consider reading comprehension (72.22%), digit span memory (61.11%), math (numerical calculation, 55.56%), and visual memory (50%) as indicators of students' cognitive performance across the four educational levels.*

Keywords: Cognitive performance; Education sector; Indoor air quality; Indoor environmental quality; Neurobehavioral tests

1. INTRODUCTION

Individuals spend over 80% of their time indoors, making classroom environments especially important for students' health and academic performance [1–5]. Poor indoor conditions in schools can lead to negative health outcomes and hinder learning, particularly in children who are more sensitive to environmental factors [6–7]. Key components of indoor environmental quality (IEQ) include thermal comfort and indoor air quality (IAQ), both of which are linked to student well-being and concentration [8–11]. Poor ventilation can reduce alertness and memory retention [12], and combined effects of thermal discomfort and low IAQ can impair cognitive performance, as shown in studies by Jaber et al. [13] and Sarbu et al. [14]. Xiong et al. [15] found that optimal learning environments are typically cooler, quiet, and well-lit, although ideal IEQ conditions can vary by task. This paper aims to (1) review how IEQ factors affect student health and academic outcomes, and (2) identify the most commonly used neurobehavioral tests to assess learning performance.

2. METHODOLOGY

This study employs a literature review methodology to summarize research on how Indoor Environmental Quality (IEQ) influences students' cognitive performance. It is based on 18 peer-reviewed studies published between 2009 and

[1*] Corresponding author: Kinnaleth Vongchanh
E-mail: kinnalethv@itc.edu.kh; Tel: +855-99 351 199

2024, covering four educational levels (primary, secondary, high school, and university), and utilizing thirteen neurobehavioral tests through either qualitative or quantitative synthesis. To quantify the use of each test, two calculations were performed: (1) Percentage by educational level—for each test, the number of studies that used it at a specific educational level was divided by the total number of reviewed papers (18 papers); and (2) Overall usage percentage—the total number of studies using each test across all levels was divided by the total number of reviewed papers to determine its overall usage.

3. RESULTS AND DISCUSSION

This section summarizes findings from the reviewed studies on how Indoor Environmental Quality (IEQ) affects students' cognitive performance. Key IEQ components—thermal comfort, air quality, lighting, and acoustics—are shown to significantly influence academic outcomes. Researchers have used a variety of neurobehavioral assessments, particularly focusing on tasks like math calculations, memory recall, and reading comprehension, to evaluate these effects.

3.1 Indoor Environmental Quality (IEQ)

Indoor Environmental Quality (IEQ) refers to the overall conditions within a built environment that influence occupants' health, comfort, and performance. It is typically characterized by four primary dimensions: thermal environment, indoor air quality (IAQ), visual comfort, and acoustic conditions, each contributing to perceived and measurable indoor satisfaction [16-18].

3.1.1 Thermal Conditions

According to ASHRAE-55 [19], thermal comfort refers to a state where individuals feel satisfied with the surrounding temperature. For optimal comfort, it recommends maintaining indoor temperatures between 20°C and 24°C in winter and 23°C to 26°C in summer. Deviations from these ranges can negatively affect occupants; colder conditions may cause cold stress and muscle tension, while excessive heat can lead to fatigue, dehydration, reduced concentration, and decreased cerebral blood flow [20–26].

Maintaining indoor relative humidity is crucial for both comfort and health. Wargocki et al. [27] and Mendell et al. [28] recommend keeping humidity between 40% and 60%, while ASHRAE Standard 55 allows for a broader range of 30% to 60% [19]. Low humidity levels can lead to dry skin, eye irritation, sore throats, and increased risk of airborne infections [20,29]. On the other hand, high humidity promotes mold and bacterial growth, which may cause allergic reactions, infections, and irritation of the skin, eyes, or respiratory system [30,31].

Maintaining appropriate indoor air velocity is also essential for classroom thermal comfort and student learning performance [32]. The recommended range in occupied spaces is 0.15 to 0.25 m/s [33]. When airflow falls below this range, it can create a sense of stagnation, resulting in discomfort, stuffiness, and poor body heat regulation, particularly in warm conditions. Inadequate air movement also hinders pollutant and CO_2 dispersion, potentially causing fatigue, drowsiness, and diminished cognitive function [20,34].

3.1.2 Indoor Air Quality (IAQ)

Indoor Air Quality (IAQ) refers to the air conditions inside buildings and their impact on occupant health and comfort, with pollutants like CO_2 and TVOCs being key indicators [35]. A study across 60 Scottish schools found that higher CO_2 levels correlated with reduced attendance and lower academic performance, even after controlling for external factors [36]. Poor ventilation contributes to elevated CO_2, potentially causing Sick Building Syndrome (SBS), which affects health and learning [37,38]. CO_2 levels above the recommended 1000 ppm threshold can cause symptoms such as fatigue, headaches, and poor concentration [39–42]. Beyond 1500 ppm, cognitive and decision-making functions decline significantly [43], and at extreme levels (5000–100,000 ppm), serious health effects including unconsciousness or death may occur [44,45].

Total Volatile Organic Compounds (TVOCs) are key indicators of indoor air quality, representing the combined levels of volatile substances that can easily evaporate at room temperature [46]. These compounds originate from sources like paints, cleaning agents, furnishings, and construction materials [47,48]. Health effects vary by concentration: levels below 0.2 mg/m³ are considered safe, 0.2–3 mg/m³ may cause mild discomfort, 3–25 mg/m³ can lead to more noticeable symptoms like headaches, and levels above 25 mg/m³ pose serious health risks, including potential neurotoxicity [49,50].

3.1.3 Acoustic factor

Elevated noise levels in educational environments have been shown to negatively affect students' ability to concentrate, communicate, and learn effectively, while also impacting their mental development and overall well-being [15,51,52]. The National Institute for Occupational Safety and Health (NIOSH) warns that continuous exposure to noise above 85 dB may cause permanent hearing loss [53]. According to OSHA [54], sounds under 85 dB are generally safe, but long-term exposure to levels between 85–140 dB can

lead to hearing damage, while sounds exceeding 140 dB may cause immediate, irreversible harm.

The World Health Organization (WHO) emphasizes that environmental noise can cause a wide range of adverse health effects, including discomfort, disturbance, and stress-related symptoms. Prolonged exposure to high noise levels increases the risk of serious conditions such as high blood pressure, blood clotting, sleep disorders, and cardiovascular disease. In extreme cases, sustained exposure may even elevate the risk of mortality [55].

3.1.4 Lighting factor

Lighting in learning environments is crucial for student academic performance, as it directly affects visual comfort and focus. A study by Heschong [58] revealed that elementary students in classrooms with the most daylight showed a 21% improvement in learning rates compared to those with the least daylight. The European and Illuminating Engineering Society standards recommend an illuminance of 300–500 lux for general learning tasks, while levels above 500 lux are suggested for more precise activities [59,60]. Poor lighting can lead to various health and visual issues such as eye strain, headaches, fatigue, and difficulty concentrating [62-65]. Bright or flickering lights can cause ocular fatigue, dryness, and blurred vision, particularly in environments with low humidity [66,67].

3.2 Neurobehavioral Tests

Lan et al. [68,69] employed neurobehavioral tests to assess cognitive performance. Those tests included Math (number calculation), digit span (letter, and symbol memory), visual learning memory or picture recognition, event sequence, reading comprehension, graphic abstracting, hand-eye coordination, creative thinking, overlapping, spatial image (thinking-spatial reasoning, imagination), letter search (perception - visual search), conditional reasoning (thinking - verbal reasoning), visual choice reaction time (executive functions - response speed and accuracy).

Based on **Table 1**, which presents data from 18 reviewed papers, Math tests were conducted in 5 papers (27.78%) at the primary school level, 1 paper (5.56%) at the secondary school level, 1 paper (5.56%) at the high school level, and 3 papers (16.67%) at the university level. In total, Math tests were used in 10 out of 18 studies, representing 55.56% of all reviewed papers.

For Digit Span Memory tests, 4 papers (22.22%) focused on primary school students, 2 papers (11.11%) on secondary school students, none on high school students, and 5 papers (27.78%) on university students. This test was employed in 11 studies overall, accounting for 61.11% of the reviewed literature.

Visual Memory tests appeared in 3 primary-level studies (16.67%), 1 secondary-level study (5.56%), no high school studies, and 5 university-level studies (27.78%). Altogether, 9 studies (50.00%) used visual memory as a cognitive performance indicator.

Table 1. Thirteen types of tests have been used to measure students' learning performance from primary school to university level

Thirteen types of tests	Primary School	Secondary School	High School	University
1. Math (Number Calculation)	[12,32,73-75]	[32]	[76]	[9,14,72]
2. Digit Span, Letter, and Symbol	[12,32,78,79]	[32,79]		[13,70,71,72, 80]
3. Memory Visual Learning Memory or Picture Recognition	[12,78,79]	[79]		[13,14,15,71, 80]
4. Event Sequence	[78,79]	[79]		[13,68,71]
5. Reading Comprehension	[32,73,74,75]	[32,73]	[73,76]	[9,15,68,77, 80]
6. Graphic Abstracting				[15,68]
7. Hand-eye Coordination				[13,68]
8. Creative Thinking				[80]
9. Overlapping	[78,79]	[79]		
10. Spatial Image	[78]			[72,80]
11. Letter Search	[78,79]	[79]		
12. Conditional Reasoning				[80]
13. Visual Choice Reaction Time	[12,32]	[32]		[15]

For Reading Comprehension, the test was used in 4 studies (22.22%) at the primary level, 2 studies (11.11%) at the secondary level, 2 studies (11.11%) at the high school level, and 5 studies (27.78%) at the university level. In total, it was the most widely applied, appearing in 13 of the 18 studies, or 72.22% of their assessment.

These percentages highlight the distribution and popularity of each test type across different educational levels, with reading comprehension tests being the most frequently used across all categories.

4. CONCLUSIONS

Indoor Environmental Quality (IEQ), which includes thermal conditions (ambient temperature, air velocity, and relative humidity), indoor air quality (IAQ: CO_2 and TVOC), noise levels, and illumination or lighting levels, affects students' health, satisfaction, engagement, and learning performance. Therefore, maintaining appropriate levels of IEQ is essential in educational buildings.

After reviewing, many studies identify reading comprehension (72.22%), digit span memory (61.11%), math

(numerical calculation, 55.56%), and visual memory (50%) as key measures of students' cognitive performance.

REFERENCES

[1] J. González-Martín, N. J. R. Kraakman, C. Pérez, R. Lebrero, and R. Muñoz, "A state–of–the-art review on indoor air pollution and strategies for indoor air pollution control," *Chemosphere*, vol. 262, p. 128376, 2021.

[2] H. K. Lai, M. Kendall, H. Ferrier, I. Lindup, S. Alm, O. Hänninen, *et al.*, "Personal exposures and microenvironment concentrations of $PM_{2.5}$, VOC, NO_2 and CO in Oxford, UK," *Atmospheric Environment*, vol. 38, no. 37, pp. 6399–6410, 2004.

[3] S. Vardoulakis and P. Kinney, "Grand challenges in sustainable cities and health," *Frontiers in Sustainable Cities*, vol. 1, p. 7, 2019.

[4] J. A. Leech, W. C. Nelson, R. T. Burnett, S. Aaron, and M. E. Raizenne, "It's about time: a comparison of Canadian and American time–activity patterns," *Journal of Exposure Science & Environmental Epidemiology*, vol. 12, no. 6, pp. 427–432, 2002.

[5] C. Schweizer, R. D. Edwards, L. Bayer-Oglesby, W. J. Gauderman, V. Ilacqua, M. J. Jantunen, *et al.*, "Indoor time–microenvironment–activity patterns in seven regions of Europe," *Journal of Exposure Science & Environmental Epidemiology*, vol. 17, no. 2, pp. 170–181, 2007.

[6] M. J. Mendell and G. A. Heath, "Do indoor pollutants and thermal conditions in schools influence student performance? A critical review of the literature," *Indoor Air*, vol. 15, no. 1, pp. 27–52, 2005.

[7] G. Zhang, J. Spickett, K. Rumchev, A. H. Lee, and S. Stick, "Indoor environmental quality in a 'low allergen' school and three standard primary schools in Western Australia," *Indoor Air*, vol. 16, no. 1, pp. 37–47, 2006.

[8] U. Haverinen-Shaughnessy, R. J. Shaughnessy, E. C. Cole, O. Toyinbo, and D. J. Moschandreas, "An assessment of indoor environmental quality in schools and its association with health and performance," *Building and Environment*, vol. 93, pp. 35–40, 2015.

[9] M. C. Lee, K. W. Mui, L. T. Wong, W. Y. Chan, E. W. M. Lee, and C. T. Cheung, "Student learning performance and indoor environmental quality (IEQ) in air-conditioned university teaching rooms," *Building and Environment*, vol. 49, pp. 238–244, 2012.

[10] A. Merabtine, C. Maalouf, A. A. W. Hawila, N. Martaj, and G. Polidori, "Building energy audit, thermal comfort, and IAQ assessment of a school building: A case study," *Building and Environment*, vol. 145, pp. 62–76, 2018.

[11] J. M. Daisey, W. J. Angell, and M. G. Apte, "Indoor air quality, ventilation and health symptoms in schools: an analysis of existing information," *Indoor Air*, vol. 13, no. 1, pp. 53–64, 2003.

[12] Z. Bakó-Biró, D. J. Clements-Croome, N. Kochhar, H. B. Awbi, and M. J. Williams, "Ventilation rates in schools and pupils' performance," *Building Environ.*, vol. 48, pp. 215–223, 2012.

[13] A. R. Jaber, D. M. Dejan, and M. U. Marcella, "The effect of indoor temperature and CO_2 levels on cognitive performance of adult females in a university building in Saudi Arabia," *Energy Procedia*, vol. 122, pp. 451–456, 2017.

[14] I. Sarbu and C. Pacurar, "Experimental and numerical research to assess indoor environment quality and schoolwork performance in university classrooms," *Building and Environment*, vol. 93, pp. 141–154, 2015.

[15] L. Xiong, X. Huang, J. Li, P. Mao, X. Wang, R. Wang, and M. Tang, "Impact of indoor physical environment on learning efficiency in different types of tasks: A 3 × 4 × 3 full factorial design analysis," *Int. J. Environ. Res. Public Health*, vol. 15, no. 6, p. 1256, 2018

[16] L. T. Wong, K. W. Mui, and P. S. Hui, "A multivariate-logistic model for acceptance of indoor environmental quality (IEQ) in offices," *Building and Environment*, vol. 43, no. 1, pp. 1-6, 2008.

[17] E. Arens, M. A. Humphreys, R. de Dear, and H. Zhang, "Are 'class A' temperature requirements realistic or desirable?," *Building and Environment*, vol. 45, no. 1, pp. 4–10, 2010.

[18] A. Ackley, O. I. Olanrewaju, O. N. Oyefusi, W. I. Enegbuma, T. S. Olaoye, A. E. Ehimatie, *et al.*, "Indoor environmental quality (IEQ) in healthcare facilities: A systematic literature review and gap analysis," *Journal of Building Engineering*, vol. 86, p. 108787, 2024.

[19] ASHRAE, *Thermal Environmental Conditions for Human Occupancy*, ANSI/ASHRAE Standard 55-1992, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, 1992.

[20] ASHRAE, *ANSI/ASHRAE Standard 55-2017: Thermal Environmental Conditions for Human Occupancy*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2017.

[21] L. Lan, P. Wargocki, and Z. Lian, "Thermal comfort, perceived air quality, and cognitive performance at moderate heat stress," *Building and Environment*, vol. 45, no. 1, pp. 22–29, 2011.

[22] O. Seppänen, W. J. Fisk, and Q. H. Lei, *Effect of temperature on task performance in office environment*, Lawrence Berkeley National Laboratory, 2006.

[23] S. Ogoh, K. Sato, K. Okazaki, T. Miyamoto, A. Hirasawa, K. Morimoto, and M. Shibasaki, "Blood flow distribution during heat stress: cerebral and systemic blood flow," *Journal of Cerebral Blood Flow & Metabolism*, vol. 33, no. 12, pp. 1915–1920, 2013.

[24] A. R. Bain, L. Nybo, and P. N. Ainslie, "Cerebral vascular control and metabolism in heat stress," *Comprehensive Physiology*, vol. 5, no. 3, pp. 1345–1380, 2015.

[25] T. McMorris, J. Swain, M. Smith, J. Corbett, S. Delves, C. Sale, *et al*., "Heat stress, plasma concentrations of adrenaline, noradrenaline, 5-hydroxytryptamine and cortisol, mood state and cognitive performance," *Int. J. Psychophysiol.*, vol. 61, no. 2, pp. 204–215, 2006.
[26] L. Nybo, P. Rasmussen, and M. N. Sawka, "Performance in the heat—physiological factors of importance for hyperthermia-induced fatigue," *Compr. Physiol.*, vol. 4, no. 2, pp. 657–689, 2014.
[27] P. Wargocki, D. P. Wyon, and O. Fanger, "Productivity is affected by the air quality in offices," in *Proc. Healthy Build.*, vol. 1, pp. 635–640, 2000.
[28] M. J. Mendell, E. A. Eliseeva, M. M. Davies, M. Spears, A. Lobscheid, W. J. Fisk, and M. G. Apte, "Association of classroom ventilation with reduced illness absence: A prospective study in California elementary schools," *Indoor Air*, vol. 23, pp. 515–528, 2013.
[29] L. M. Reinikainen and J. J. Jaakkola, "Significance of humidity and temperature on indoor air quality in schools," *Archives of Environmental Health*, vol. 56, no. 5, pp. 419–424, 2001.
[30] E. Brągoszewska, A. Mainka, J. S. Pastuszka, K. Lizończyk, and Y. G. Desta, "Assessment of bacterial aerosol in a preschool, primary school and high school in Poland," *Atmosphere*, vol. 9, no. 3, p. 87, 2018.
[31] A. M. Madsen, S. Moslehi-Jenabian, M. Z. Islam, M. Frankel, M. Spilak, and M. W. Frederiksen, "Concentrations of Staphylococcus species in indoor air as associated with other bacteria, season, relative humidity, air change rate, and S. aureus-positive occupants," *Environmental Research*, vol. 160, pp. 282-291, 2018.
[32] M. Puteh, M. H. Ibrahim, M. Adnan, C. N. Che'Ahmad, and N. M. Noh, "Thermal comfort in classroom: constraints and issues," *Procedia - Social and Behavioral Sciences*, vol. 46, pp. 1834-1838, 2012.
[33] ASHRAE, *Thermal Environmental Conditions for Human Occupancy*, ANSI/ASHRAE Standard 55-2004, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, 2004.
[34] A. K. Melikov, "Personalized ventilation," *Indoor Air*, vol. 14, suppl. 7, pp. 157–167, 2004.
[35] U.S. Environmental Protection Agency (EPA), "Introduction to Indoor Air Quality," *EPA.gov*, Feb. 12, 2025.
[36] S. Gaihre, S. Semple, J. Miller, S. Fielding, and S. Turner, "Classroom carbon dioxide concentration, school attendance, and educational attainment," *Journal of School Health*, vol. 84, no. 9, pp. 569–574, 2014.
[37] D. P. Wyon and P. Wargocki, "How indoor environment affects performance," *Thought*, vol. 3, no. 5, p. 6, 2013.
[38] P. Wargocki, J. A. Porras-Salazar, S. Contreras-Espinoza, and W. Bahnfleth, "The relationships between classroom air quality and children's performance in school," *Building Environ.*, vol. 173, p. 106749, 2020.
[39] ASHRAE, *Ventilation for Acceptable Indoor Air Quality*, ANSI/ASHRAE Standard 62-2001, American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, GA, USA, 2001.
[40] ASHRAE, *ANSI/ASHRAE Standard 62.1-2019: Ventilation for Acceptable Indoor Air Quality*. Atlanta, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, 2019.
[41] N. Muscatiello, A. McCarthy, C. Kielb, W. H. Hsu, S. A. Hwang, and S. Lin, "Classroom conditions and CO_2 concentrations and teacher health symptom reporting in 10 New York State Schools," *Indoor Air*, vol. 25, no. 2, pp. 157–167, 2015.
[42] U. Satish, M. J. Mendell, K. Shekhar, T. Hotchi, D. Sullivan, S. Streufert, and W. J. Fisk, "Is CO_2 an indoor pollutant? Direct effects of low-to-moderate CO_2 concentrations on human decision-making performance," *Environmental Health Perspectives*, vol. 120, no. 12, pp. 1671–1677, 2012.
[43] O. Seppänen, W. J. Fisk, and M. J. Mendell, "Association of ventilation rates and CO_2 concentrations with health and other responses in commercial and institutional buildings," *Indoor Air*, vol. 9, no. 4, pp. 226–252, 1999.
[44] Occupational Safety and Health Administration (OSHA), *Carbon Dioxide – Health Hazard Information Bulletin*, U.S. Department of Labor, Publication No. HIB 06-05-1996, Jun. 1996.
[45] X. Zhang, P. Wargocki, and Z. Lian, "Physiological responses during exposure to carbon dioxide and bioeffluents at levels typically occurring indoors," *Indoor Air*, vol. 27, no. 1, pp. 65–77, 2017.
[46] G. Wei, X. Yu, L. Fang, Q. Wang, T. Tanaka, K. Amano, and X. Yang, "A review and comparison of the indoor air quality requirements in selected building standards and certifications," *Building and Environment*, vol. 226, p. 109709, 2022.
[47] M. Hori, "Total Volatile Organic Compound (TVOC) as Index of Indoor Air Quality and Its Measuring and Evaluation," *Journal of the Human-Environment System*, vol. 23, no. 1, pp. 1–11, 2020.
[48] Y. Ding, S. Li, and J. Han, "Pollution characteristics of total volatile organic compounds in indoor air of urban residential buildings in China," *Journal of Environmental and Occupational Medicine*, vol. 39, no. 7, pp. 821–826, 2022.
[49] L. Mølhave, "Volatile organic compounds – indoor air quality and health," *Indoor Air '90*, vol. 5, pp. 447–452, 1990.
[50] L. Mølhave, "The use of TVOC as an indicator in IAQ investigations," in *Proc. INDOOR AIR '96*, vol. 2, pp. 37–46, 1996.
[51] B. M. Shield and J. E. Dockrell, "The effects of noise on children at school: A review," *Building Acoustics*, vol. 10, no. 2, pp. 97–116, 2003.

[52] F. Gheller, G. Spicciarelli, P. Scimemi, and B. Arfé, "The effects of noise on children's cognitive performance: A systematic review," *Environment and Behavior*, vol. 55, no. 8–10, pp. 698–734, 2024.
[53] U.S. Department of Labor, *Occupational Noise Exposure*, NIOSH Publication No. 98-126, June 1998.
[54] Occupational Safety and Health Administration (OSHA), *OSHA Technical Manual (OTM), Section III: Chapter 5 Noise*, 2022.
[55] World Health Organization (WHO), *Burden of Disease from Environmental Noise: Quantification of Healthy Life Years Lost in Europe*, WHO Regional Office for Europe, 2011.
[56] N. Castilla, J. L. Higuera-Trujillo, and C. Llinares, "The effects of illuminance on students' memory: A neuroarchitecture study," *Building and Environment*, vol. 228, p. 109833, 2023.
[57] I. Konstantzos, S. A. Sadeghi, M. Kim, J. Xiong, and A. Tzempelikos, "The effect of lighting environment on task performance in buildings–A review," *Energy and Buildings*, vol. 226, p. 110394, 2020.
[58] L. Heschong, *Daylighting in Schools: An Investigation into the Relationship between Daylighting and Human Performance*, Detailed Report, 1999.
[59] *UNE-EN 12464-1:2012, Light and lighting—Lighting of workplaces—Part 1: Indoor workplaces*, British Standards Institution, London, UK, 2012.
[60] D. L. DiLaura, *The Lighting Handbook: Reference and Application*, Illuminating Engineering Society of North America, New York: IESNA, 2011.
[61] Malaysian standard 1525:2007: Code of Practice on Energy Efficiency and Renewable Energy for Non-Residential Buildings. Malaysia, Department of Standards Malaysia, 2007.
[62] S. A. Rahman, C. M. Shapiro, F. Wang, H. Ainlay, S. Kazmi, T. J. Brown, and R. F. Casper, "Effects of filtering visual short wavelengths during nocturnal shiftwork on sleep and performance," *Chronobiol. Int.*, vol. 30, no. 8, pp. 951–962, 2013.
[63] C. Owsley, "Vision and aging," *Annu. Rev. Vis. Sci.*, vol. 2, no. 1, pp. 255–271, 2016.
[64] J. Kim, H. Oh, W. Kim, S. Choi, W. Son, and S. Lee, "A deep motion sickness predictor induced by visual stimuli in virtual reality," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 554–566, Feb. 2022.
[65] P. R. Boyce, "The impact of light in buildings on human health," *Indoor Built Environ.*, vol. 19, no. 1, pp. 8–20, 2010.
[66] C. Y. Park, "Night light pollution and ocular fatigue," *J. Korean Med. Sci.*, vol. 33, no. 38, p. e257, 2018.
[67] K. H. Lin, C. C. Su, Y. Y. Chen, and P. C. Chu, "The effects of lighting problems on eye symptoms among cleanroom microscope workers," *Int. J. Environ. Res. Public Health*, vol. 16, no. 1, p. 101, 2019.
[68] L. Lan and Z. Lian, "Use of neurobehavioral tests to evaluate the effects of indoor environment quality on productivity," *Building Environ.*, vol. 44, no. 11, pp. 2208–2217, 2009.
[69] L. Lan, Z. Lian, L. Pan, and Q. Ye, "Neurobehavioral approach for evaluation of office workers' productivity: The effects of room temperature," *Building and Environment*, vol. 44, no. 8, pp. 1578–1588, 2009.
[70] W. Cui, G. Cao, J. H. Park, Q. Ouyang, and Y. Zhu, "Influence of indoor air temperature on human thermal comfort, motivation and performance," *Building and Environment*, vol. 68, pp. 114–122, 2013.
[71] W. Liu, J. Yu, S. Li, L. Shi, M. Xie, and Z. Wang, "A method to evaluate cognitive impairment under high temperature based on temple skin temperature," *Energy and Buildings*, vol. 306, p. 113919, 2024.
[72] Y. Geng, W. Ji, B. Lin, and Y. Zhu, "The impact of thermal environment on occupant IEQ perception and productivity," *Building and Environment*, vol. 121, pp. 158–167, 2017.
[73] J. Toftum, B. U. Kjeldsen, P. Wargocki, H. R. Menå, E. M. N. Hansen, and G. Clausen, "Association between classroom ventilation mode and learning outcome in Danish schools," *Building and Environment*, vol. 92, pp. 494–503, 2015.
[74] J. Graff Zivin, S. M. Hsiang, and M. Neidell, "Temperature and Human Capital in the Short and Long Run," *J. Assoc. Environ. Resour. Econ.*, vol. 5, no. 1, pp. 77–105, 2018.
[75] U. Haverinen-Shaughnessy, D. J. Moschandreas, and R. J. Shaughnessy, "Association between substandard classroom ventilation rates and students' academic achievement," *Indoor Air*, vol. 21, no. 2, pp. 121–131, 2010.
[76] H. Cho, "The effects of summer heat on academic achievement: A cohort analysis," *J. Environ. Econ. Manag.*, vol. 83, pp. 185–196, 2017.
[77] G. C. Marchand, N. M. Nardi, D. Reynolds, and S. Pamoukov, "The impact of the classroom built environment on student perceptions and learning," *J. Environ. Psychol.*, vol. 40, pp. 187–197, 2014.
[78] J. Jiang, D. Wang, Y. Liu, Y. Xu, and J. Liu, "A study on pupils' learning performance and thermal comfort of primary schools in China," *Building and Environment*, vol. 134, pp. 102–113, 2018.
[79] D. Wang, C. Song, Y. Wang, Y. Xu, Y. Liu, and J. Liu, "Experimental investigation of the potential influence of indoor air velocity on students' learning performance in summer conditions," *Energy and Buildings*, vol. 219, p. 110015, 2020.
[80] X. Ma, H. Liu, Z. Zhang, and Y. Li, "How does indoor physical environment differentially affect learning performance in various classroom types?" *Building and Environment*, vol. 234, p. 110189, 2023.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Optimal number of EV charger stations in distribution system with 24-hour EV Load Profiles Estimation

Kechki OENG [1*], Vannak VAI [2], Chhith CHHLONH [2]

Department of Energy Technology and Management, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: oeng.kechki@gmail.com*

Abstract: *The rapid growth of electric vehicles (EVs) necessitates robust charging infrastructure planning to ensure grid stability and user satisfaction. This paper proposes a four-stage framework for optimizing number of EV charging station and load profile 24-hour for each charger incorporating time-varying arrival rates, probabilistic state-of-charge distributions, and charger constraints. First, traffic volume data from JICA reports is used to estimate number of EV passing through road. Second, a stochastic model generates EV battery capacity. Third, Random remaining capacity of EV that predict minimum and maximum battery capacity of EV is zero percent to thirty percent. For three of stage provided stochastic analysis to model and predict the probabilistic variations in energy demand based on uncertain factors such as vehicle arrival times, charging durations, and user behavior. Four, Optimal number of chargers number to provide of chargers without queueing and 24-hour load profile. Furthermore, Brute-Force method has given implement in distribution system to determine the optimal number of chargers and display load profile of EV. The result will achievement two parts that we can get the number of the charger and load profile 24-hour for each charger. The framework addresses gaps in existing studies by integrating EV load profiling, and number of chargers optimization. offering a scalable solution for sustainable EV infrastructure.*

Keywords: Electric vehicles (EVs), Battery Capacity (BC), Brute-Force (BF), Stochastic modeling, Charger Station (CS).

1. INTRODUCTION

The rapid adoption of electric vehicles (EVs) has necessitated the development of robust charging infrastructure to meet growing energy demands while ensuring grid stability. Accurate estimation of EV load profiles is critical for planning and optimizing charging stations, particularly in high-demand scenarios such as highway corridors by [1] and [4]. Cambodia's National Road 5 is a critical transport corridor projected to experience significant EV adoption by 2030. Current infrastructure planning lacks integration between traffic patterns, EV charging behaviors, and grid capacity. Traffic volume data for National Road 5, obtained from JICA reports. The several calculations and adjustments to prepare the data for use in subsequent EV charging station simulations.

The modeling required to integrate EVs as loads within the distribution grid is of great importance, since the validity of the results are only as good as the models that are used [18]. The models presented here is based on the physical behavior of EV batteries (EV charging model)[17]. On the voltage pro le of the distribution system in the presence of a charging load is also studied. A stochastic model is developed to estimate hourly EV load from arrival time, departure time and distance traveled.[18]

The work in this paper consist of estimate EV arrival using historical traffic data, mode battery capacity and SOC variations, optimal charger number to minimize abandonment and generate 24-hour load profile for grid planning.

The rest of the paper is arranged as follows. In section I is the introduction. Section II explains the data estimate EV charger station, number of charger and model of capacity.

[1*] Corresponding author: kechki Oeng
E-mail:oeng.kechki@gmail.com ; Tel: +855-96 674 6732

Section III showing the result and discussion of data JICA estimate and actual data estimate of number of EVs, data estimate EV battery capacity distribution and estimate EV arrival state of charge (SOC) distribution and data of 24-hour load profile. Section IV concludes with future work.

2. METHODOLOGY

2.1 Generation of simulation input data

The EV charging simulation proposed in this paper requires 3 types of input data:

1. Hourly arrival rate of EVs
2. Battery capacity in kWh of the arrived EVs
3. State-of-charge of EV upon arrival

These 3 parameters are used for assigning available charger, the charging time of each EV at each State sampling interval. 15 minutes were used as sampling intervals for updating the number of EVs and charging time in the simulation. This interval is chosen as a compromising between simulation speed and accuracy. Smaller interval produces better accuracy but tend to take much longer to simulate.

This paper proposes the following stochastic models for the 3 types of input:

1. Hourly Arrival Rate of EVs: Non-Homogeneous Poisson Process (NHPP)

EV arrivals at charging stations often exhibit time-dependent variability due to commuting patterns, peak demand hours, or seasonal factors. NHPP is ideal for modeling such scenarios because it allows the arrival rate $\lambda(t)$ to vary with time, unlike a standard Poisson process with constant intensity. This aligns with real-world observations where charging demand spikes during evenings or weekends. [6][7]
The number of arrivals in interval $[t, t+s]$ follows:

$$N(t) - N(t+s) \sim Poisson\left(\int_t^{t+s} \lambda(\alpha) d\alpha\right) \quad \text{(Eq. 1)}$$

Where $N(t)$ is nonhomogeneous Poisson process (NHPP), $\lambda(t)$ is the time-dependent arrival rate, s is the sampling interval. [8][9]

2. Battery Capacity (kWh): Normal Distribution

Battery capacities in EVs often cluster around a mean value due to manufacturing standards and market segmentation (e.g., 40 kWh for compact EVs vs. 100 kWh for luxury models). The normal distribution captures this central tendency and variability, reflecting natural dispersion in battery sizes across vehicle types. Normal distributions are empirically validated for technical specifications like battery capacity, as seen in studies analyzing EV fleet data. [6]

The probability density function (PDF) is:

$$f(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}} \quad \text{(Eq. 2)}$$

where μ is the mean capacity and σ^2 the variance. [10]
3. State-of-Charge (SOC) Upon Arrival: Uniform Distribution

SOC at arrival is assumed to be uniformly distributed if drivers' charging behaviors are random within a feasible range (e.g., 10%–30%). This simplifies modeling when no prior data on charging habits is available, treating all SOC levels as equally probable.
The probability density function (PDF) is:

$$f(x) = \frac{1}{b-a} \text{ for } a \le x \le b \quad \text{(Eq. 3)}$$

where a and b define the minimum and maximum SOC. [11]

2.2 Charging Station Load Profile Estimate

The simulation models real-world EV charging behavior without queuing the model displays the number of chargers, representing the total charging points available for EV users, it gets abandoned rather than waiting. It's showing by flowchart below.

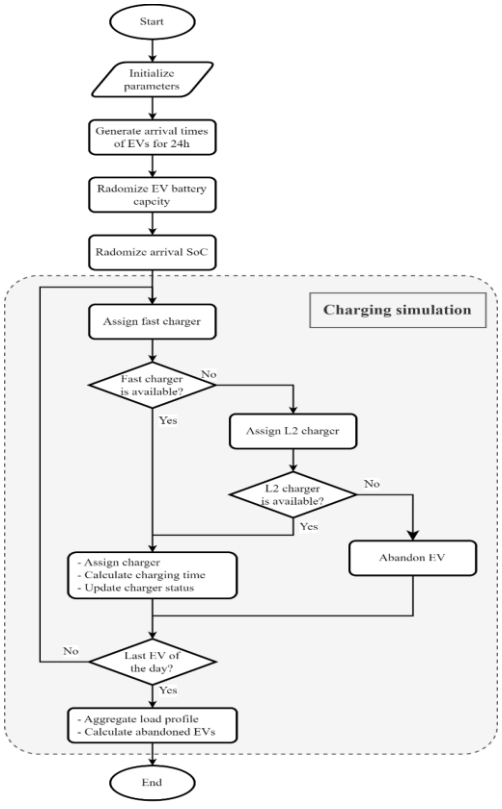


Fig. 1. flowchart of the EV Charger load profiling simulation

Step1: Initialize parameters

- Various parameters related to EV charging, such as charger capacities, arrival rates, and charging durations, are set up.

Step 2: Generate arrival times of EVs for 24h

- The system simulates or records the arrival times of EVs over a 24-hour period.

Step 3: EV Characteristics Assignment

- Randomize EV battery capacity for normal distribution
- Randomize arrival SoC

Step 4: Assign fast charger

- Assign fast charger and then checks availability charger. If available, calculates charging time.
- If not available we are going to "Assign Level 2 charger" and checks availability
- The charger status is updated to reflect its usage. If no charger is available
- The EV is abandoned, meaning it does not get charged.

Step 5: Check if the EV is the last one of the day

- If not, the system returns to process the next arriving EV.
- If it is the last EV, the system proceeds to the final step.

Step 6: Aggregate load profile and calculate abandoned EVs

- The system compiles the overall charging load profile for the day.
- The number of EVs that were abandoned due to unavailability of chargers is recorded.

2.2 Charging Station Load Profile Estimate with optimal

The optimal number of fast and level 2 (L2) chargers needed to serve all EVs without any abandonments by the Brute-Force Charger Optimization. The code implements the actual charging simulation that the brute-force search uses to evaluate each potential charger configuration, tracking which EVs get served and which get abandoned for each tested combination of fast and L2 chargers.

Step1: Initialize parameters

- Various parameters related to EV charging, such as charger capacities, arrival rates, and charging durations, are set up.

Step 2: Generate arrival times of EVs for 24h

- The system simulates or records the arrival times of EVs over a 24-hour period.

Step 3: EV Characteristics Assignment

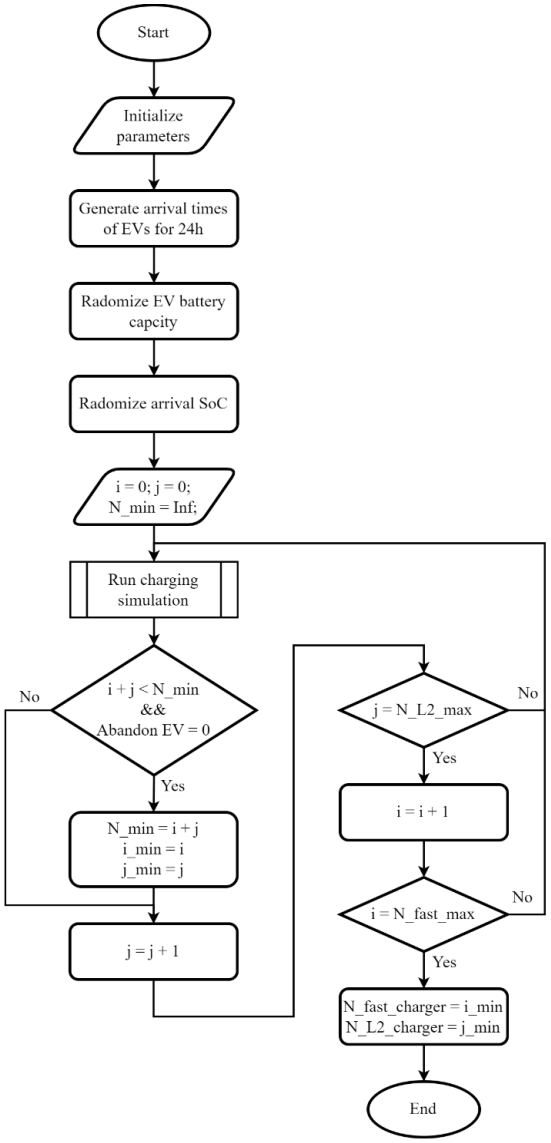


Fig. 2. flowchart of the EV Charger load profiling simulation with optimal

- Randomize EV battery capacity for normal distribution
- Randomize arrival SoC

Step 4: Brute-Force Search Algorithm

- Initializes counters i (fast chargers), j (L2 chargers) and N_min (minimum total chargers)
- Nested loops test all combinations of chargers:
  - Outer loop increments fast chargers (i)
  - Inner loop increments L2 chargers (j)
  - For each combination, runs simulation
  - If combination serves all EVs (Abandon EV = 0) with fewer total chargers (i+j < N_min), updates optimal solution

Step 5: Charging Simulation Core

- Run charging simulation for each arriving EV in current time slot:
  - Calculates energy needed based on battery capacity and SOC difference
  - Attempts to assign highest-power available charger first
  - If assigned, calculates charging_time and updates charger status
  - If no chargers available, increments abandoned_evs counter

Step 6: Optimal Solution Determination

- After testing all combinations, sets:
  - N_fast_charger = i_min
  - N_L2_charger = j_min

3. RESULTS AND DISCUSSION

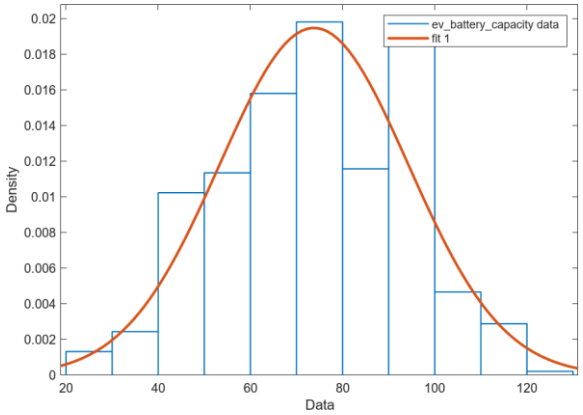


Fig. 3. Distribution fit battery capacity based on global EV data.

Fig. 3. Illustrates the distribution fit of battery capacity obtained from global EV data. The result shows that it follows normal distribution. This distribution is critical in the simulation process, where we need to determine the battery capacity of EV upon arrival at the charging station.

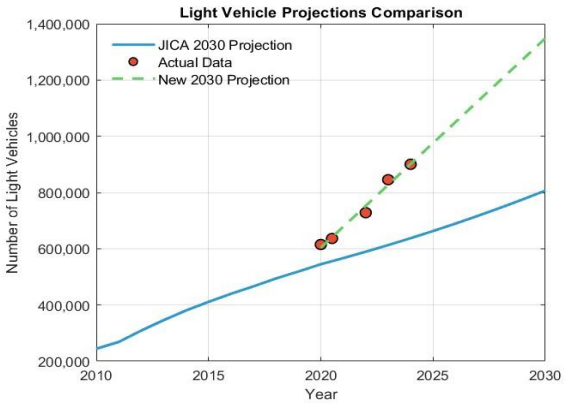


Fig. 4. Result estimate light vehicle of JICA and actual data from Ministry of public works and transport

As seen from **Fig. 4.** and **Table 1**, compares light vehicle projections from JICA's 2030 model with actual data and a revised projection. While the JICA 2030 projection (solid blue line) estimates a gradual increase 806674 EVs in 2030, the actual data (red dots) from 2020 to 2024 consistently exceed this projection. Consequently, a new 2030 projection (green dashed line) predict based on actual data reflects a significantly steeper growth trajectory, suggesting a faster-than-expected increase in the number of light vehicles from JACA estimates.

Table 1. Data of number EV for each year

Year	JICA2030 Projection	Actual Data	New 2030 Projection
2020	544961	636619	
2022	589429	728294	
2023	613006	845626	
2024	632476	900318	
2030	**806674**		**1273680**

Fig.5. presents three distributions related to estimated electric vehicle (EV) arrivals, battery capacity of the arrived EVs and SOC of EV. The three parameters of 24 hours data to input in simulation software to generated numbers of EVs and load profile of each charger that provided in figure 6 and figure 7.

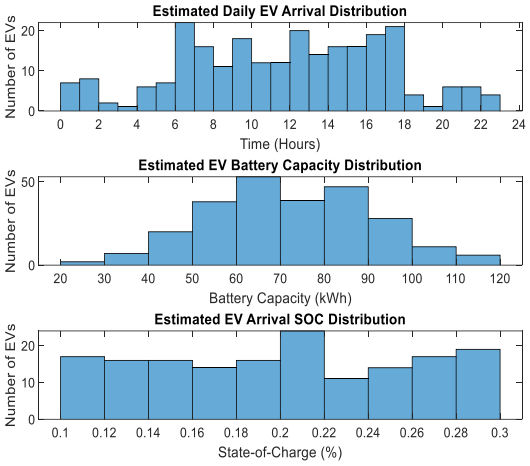


Fig. 5. Result of estimate EV arrival for 24-hour, EV battery capacity and EV arrival SOC distribution

Fig.6 presents a comparative analysis of the 24-hour load profile of power demand using two different time-averaging intervals: 15-minute averages and 1-hour averages. This analysis is critical for understanding short-term fluctuations and overall trends in electricity consumption, which can inform operational planning and peak demand management strategies.

This profile reveals high variability in power usage throughout the day, with numerous short-duration spikes and dips. The maximum power demand recorded in this profile is 1815 kW, indicated by a red dashed line labeled "Peak Load = 1815 kW". The bottom panel shows the same 24-hour period but with power demand averaged over 1-hour intervals. The peak load for this profile is 1515 kW, also denoted by a red dashed line. Although this value is lower than that of the 15-minute profile, it provides a clearer view of broader consumption trends and is useful for long-term planning and energy allocation.

The comparison between the two profiles highlights the impact of data granularity on peak demand estimation. The 15-minute average captures short-term peaks that may be overlooked in longer averaging windows. In contrast, the 1-hour average provides a simplified overview suitable for general planning purposes. 15-minute interval load profile is generated from the simulation. In general, the smaller the sample interval the better the simulation accuracy; however, this will make the simulation slower. On the other hand, 1-hour interval is typically used in load profile planning. Understanding both perspectives is essential for effective load management, infrastructure sizing, and minimizing demand charges.

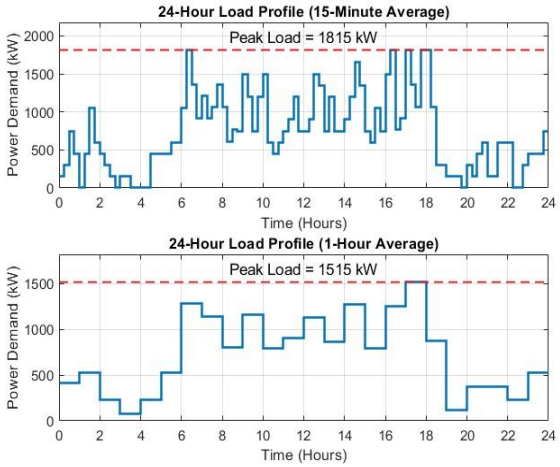


Fig. 6. Result of 24-hour load profile

```
Total number of DC chargers: 12 (fast charger), 1 (L2 charger)
Total EVs: 257
EVs successfully assigned chargers: 257
Abandoned EVs (no charger available): 0
Abandonment rate: 0.00%
```

Fig. 7. Simulation result of the optimal number of chargers

Fig.7. there's the total number of DC chargers: 12 fast chargers and 1 Level 2 (L2) charger. DC fast chargers provide quicker charging compared to L2. The result is provided by MATLAB simulation.

4. CONCLUSIONS

This study presents a comprehensive framework for optimizing the number of EV charging stations and estimating 24-hour load profiles in distribution systems, addressing critical gaps in infrastructure planning. By integrating stochastic modeling of EV arrivals, battery capacity distributions, and state-of-charge (SOC) variations, the proposed four-stage methodology ensures robust demand forecasting and grid stability.

ACKNOWLEDGMENTS

I would like to express my utmost gratitude to the Cambodia Higher Education Improvement Project (**HEIP**) for the financial support this research is successful completion of this research and has greatly facilitated progress. Additionally, I would like to thank the Energy Technology and Management for their guidance and technical support.

REFERENCES

[1] A. Almutairi and Y. Almasri, "Load profile modeling of plug-in EVs.," *IEEE Transactions on Smart Grid, 12(3), 1234-1245. R.*, 2021.

[2] M. Alqahtani *et al.*, "Investigation and Minimization of Power Loss in Radial Distribution Network Using Gray Wolf Optimization," *Energies*, vol. 16, no. 12, pp. 4571–4571, Jun. 2023, doi: https://doi.org/10.3390/en16124571.

[3] RL Institute, "SpiceEV: Simulation program for charging strategies.," *https://github.com/rl-institut/spice_ev.*, 2023.

[4] Springer, "Stochastic model for aggregated EV charging de-mand," *In Proceedings of the International Conference on Smart Grids and Energy Systems (pp. 45-60).* , 2022.

[5] M. B. A. A.-M. Fatma Yaprakdal, "Optimal Operational Scheduling of Reconfigurable Microgrids in Presence of Renewable Energy Sources," *in ResearchGate*, 2019.

[6] J. Brady and M. O'Mahony, "Modelling charging profiles of electric vehicles based on real-world electric vehicle charging data," *in Sustainable Cities and Society, vol. 26, pp. 203-216*, 2016.

[7] Dingtong Yang and undefined, "Dynamic Modeling and Real-time Management of a System of EV Fast-charging Stations," 2020.

[8] "Nonhomogeneous Poisson Processes," *Probabilitycourse.com*, 2025. https://www.probabilitycourse.com/chapter11/11_1_4_nonhomogeneous_poisson_processes.php

[9] "14.6: Non-homogeneous Poisson Processes," *Statistics LibreTexts*, May 05, 2020. https://stats.libretexts.org/Bookshelves/Probability_Theory/Probability_Mathematical_Statistics_and_Stochastic_Processes_(Siegrist)/14%3A_The_Poisson_Process/14.06%3A_Non-homogeneous_Poisson_Processes

[10] A. Cunningham, "Probability Playground: The Normal Distribution," *Buffalo.edu*, 2018. https://www.acsu.buffalo.edu/~adamcunn/probability/normal.html (accessed May 25, 2025).

[11] B. Illowsky and S. Dean, "5.2 The Uniform Distribution - Statistics | OpenStax," *openstax.org*, Mar. 27, 2020. https://openstax.org/books/statistics/pages/5-2-the-uniform-distribution

[12] Ministry of Public Works and Transport , "Cambodia Vehicle Safety and Environmental Protection Ministry of Public Works and Transport General Department of Land Transport." Available: https://www.jasic.org/meeting_docs_admin/contents/uploads/doc/meeting2/4-2%20Slide%20Presentation%20PP%20forum(Cambodia).pdf

[13] Abduraheem Alobaidi, Hunter DesRoches, and Mahdi, Mehrtash, "Impact of Vehicle to Grid Technology on Distribu-tion Grid with Two Power Line Filter Approaches," *IEEE Green Technologies Conference (GreenTech)*, 2021. USA

[14] "Cambodia's vehicle population grows by 11.3% - Khmer Times," Mar. 08, 2023. https://www.khmertimeskh.com/501251641/cambodias-vehicle-population-grows-by-11-3/

[15] Ministry of Public Works and Transport , "Revenue Rolls in as Vehicles Roll On," *cambodianess.com*, 2024. https://cambodianess.com/article/revenue-rolls-in-as-vehicles-roll-on

[16] Ratana, "More than 410,000 vehicles registered in 2024 - Khmer Times," *Khmer Times - Insight into Cambodia*, Jan. 14, 2025. https://www.khmertimeskh.com/501623648/more-than-410000-vehicles-registered-in-2024/ (accessed May 27, 2025).

[17] K. Clement, E. Haesen, J. Driesen, "Stochastic Analysis of the Impact of PHEVs on the Distribution Grid", CIRED, 20th International Conference on Electricity Distribution, pp. 1-4, 2009.

[18] MUHAMMADZULQARNAIN ZEB and KASHIF IMRAN, "Optimal Placement of Electric Vehicle Charging Stations in the Active Distribution Network," *IEEE Access*, Apr. 22, 2020.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Coordinated Reactive Power Control Using Genetic Algorithm for Loss Minimization in PV-Integrated Distribution Networks

Sak SREY [1*], Vannak VAI[2], Kimsrornn KHON[1,2]

[1*] *Energy Technology and Management Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*
[2] *Department of Electrical and Energy Engineering, Faculty of Electrical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[*]*Corresponding author: srey.sak13@gmail.com*

Abstract: *The increasing penetration of photovoltaic (PV) generation in distribution networks presents several operational challenges, such as voltage fluctuations, reverse power flow, and active power losses. These issues also place additional stress on conventional voltage regulation devices, including on-load tap changers (OLTCs) and capacitor banks, resulting in higher maintenance requirements and reduced system efficiency. This study proposes a Volt-Var Optimization (VVO) approach aimed at minimizing active power losses through the coordinated control of reactive power from PV inverters, OLTCs, and capacitor banks. A Genetic Algorithm (GA) is used to identify the optimal reactive power dispatch from reactive control devices while ensuring voltage constraints are satisfied throughout the network. In order to evaluate the effectiveness of this approach, the IEEE 33-bus distribution test system is used, with PV units and reactive control devices integrated at selected buses. This study proposes several reactive power control scenarios by using individual reactive control devices and their combinations are analyzed in order to assess the contribution of each control element. Simulation results confirm that the proposed coordinated reactive power control strategy enhances voltage regulation and reduces energy losses by approximately 55.5% over a 24-hour period compared to the base case, thereby improving the overall efficiency and reliability of the distribution network.*

Keywords: Genetic Algorithm (GA); Distribution networks system; Volt-Var optimization; On-load tap changer (OLTC); Photovoltaic (PV) integration

1. INTRODUCTION

The global shift toward low-carbon and renewable energy sources has led to a noticeable rise in the integration of photovoltaic (PV) systems into distribution networks. While PV generation plays an important part in improving grid sustainability and resilience, its intermittent and uncertain nature introduces practical issues. These challenges often show as node voltage fluctuations, reverse power flow, and increased active power losses. To solve these issues, strategies such as reactive power compensation, OLTC tap adjustments, and active power curtailment are crucial for preserving voltage stability and ensuring grid reliability [1]. Therefore, reactive power optimization of the distribution network becomes a key link to guarantee the economy, safety, and stability of the power system [2]. Meanwhile, it optimizes certain performances of power system operation, such as real power loss. Reactive power optimization aims to guarantee secure voltage levels while satisfying operational constraints through coordinated control of reactive devices.

Various methods have been proposed in the literature to address the VVO problem. These can generally be categorized into data-driven approaches and optimization-based approaches. Data-driven methods, including deep reinforcement learning, are highlighted as emerging techniques for VVC optimization. These techniques offer adaptability but require high-quality data and have problems connected to computational complexity and scalability [3]. In [4], the authors propose a model-augmented safe reinforcement learning approach to develop an RL framework that operates within the physical constraints of a power distribution network. Similarly, [5] shows a multi-agent deep reinforcement learning (MOMADRL) method for real-time

[1*] Corresponding author: Srey Sak
E-mail: srey.sak13@gmail.com; Tel: +855-96 417 1775

decentralized VVC using PV inverters. This approach aims to minimize network power loss, inverter degradation, and voltage deviations simultaneously. To prevent overvoltage issues caused by distributed resources, [6] presents a multi-agent DRL-based strategy to control sub-networks within a distribution system for Volt-VAR optimization. Furthermore, to obtain economic scheduling plans and fulfil the security limitations of the power systems, [7] proposes a hybrid knowledge-data- DRL approach for proper operation strategy.

In addition to data-driven methods, optimization-based methods have convex optimization methods and heuristic optimization methods. However, in the VVO problem, it is not conductive to the accuracy of the final solution when we're applying convex optimization methods. In recent years, a lot of new heuristic optimization methods have been used in power system. For instance, authors [8] proposed an improved sparrow search algorithm (SSA) to optimize reactive power, minimize power losses, and voltage fluctuations through reactive power optimization. Similarly, [9] introduced the Artificial Gorilla Troop Optimization (AGTO) algorithm to enhance voltage profiles and reduce power losses in Volt-Var Control (VVC). Furthermore, [10] proposed a parallel PSO-based VVO method using MapReduce and graph computing to enhance computational efficiency for large-scale systems. These heuristic techniques show important roles for addressing intricate VVO problems, especially when coordination among control devices is required.

Among abovementioned algorithms, genetic algorithms have been widely applied in the VVO field of distribution networks. GA really good at dealing with both discrete and continuous control variables, such as on-load tap changer (OLTC) positions, capacitor bank switching, and reactive power output of PV inverters. Authors of [1] proposes a multi-objective GA method that optimizes both real power losses and OLTC operation costs in distribution networks. In [11] a multi-objective VVO approach using NSGA-III was developed to minimize active power losses, PV curtailment, and load reduction. Similarly, [12] studied GA to optimize the placement and operation of capacitor banks, aiming to reduce real power losses and voltage deviations. Additionally, [13] presented a GA-based method to determine the optimal settings for reactive power devices, with the objective of mitigating bus voltage violations. Although many GA-based approaches have been proposed for Volt-Var Optimization, most of them address either a single control device or are formulated as multi-objective problems with limited focus on practical deployment in constrained environments. In addition, very few studies evaluate performance under time-series operation, considering realistic PV generation and load variation over 24 hours.

To address these limitations, this study proposes a Genetic Algorithm-based Volt-Var Optimization approach that coordinates OLTC, capacitor banks, and PV inverter reactive power to minimize active power losses. The algorithm is used to test on the IEEE 33-bus distribution system under eight control scenarios, and allowing comparison of each control strategy's impact. Unlike most studies, this work focuses on a single-objective formulation to isolate and evaluate the effectiveness of each device combination under realistic operating conditions.

The rest of this paper is formulated in the following form. The implementation of GA method to solve the VVO problem is indicated in Section 2. Section 3 illustrates the optimized results in the case study. A conclusions of this work is summarized in Section 4.

2. METHODOLOGY

The optimization is performed hourly over a 24-hour period to determine the best combination of control settings. The losses discussed in this paper are entirely technical, such as feeder losses, transformer load and no-load losses, and capacitor bank losses. The method ensures that voltage limitations are followed across all buses, while also increasing system efficiency through efficient reactive power dispatch.

2.1 Objective Function

The objective of this study is to minimize the total active power losses in the distribution network over a 24-hour simulation period. The power losses are calculated from the results of the power flow analysis where the line losses are calculated from the real power flows across the network branches.

$$\min F_{loss} = \sum_{h=1}^{24} \sum_{(i,j)\in\beta} P_{loss,h}^{(i,j)} \qquad \text{(Eq. 1)}$$

where F_{loss} is the total energy loss over 24 hours (kWh); $P_{loss,h}^{(i,j)}$ is the active power loss (kW) in the line connecting bus *i* and bus *j* at hour *h;* β is the set of all branches in the distribution network.

2.2 Equality constraint

This constraint ensures load balance, which means that the generation of actual and reactive power is balanced with the load.

$$P_i = V_i \sum_{j\in N_i} V_j (G_{ij}\cos\theta_{ij} + B_{ij}\sin\theta_{ij}) \quad i \in N_{B-1} \qquad \text{(Eq. 2)}$$

$$Q_i = V_i \sum_{j\in N_i} V_j (G_{ij}\sin\theta_{ij} + B_{ij}\cos\theta_{ij}) \quad i \in N_{PQ} \qquad \text{(Eq. 3)}$$

where N_i number of buses adjacent to bus *i*, including bus *i;* N_{PQ} number of Load (PQ) buses with constant P and Q

injections; N_{B-1} total number of buses, excluding slack bus; P_i, Q_i injection of active and reactive powers into the system at bus *i*; G_{ij}, B_{ij} mutual conductance, and susceptance between bus *i* and bus *j*.

2.3 Inequality constraint

- Bus voltage limits voltage levels in Cambodia [18]

$$0.95\,pu \le V_i \le 1.05\,pu \quad \text{(Eq. 4)}$$

- DG Power factor range: The power factor of the DG connected should be in the range given below equation.

$$0.8 \le PF_{DGi} \le 1 \quad \text{(Eq. 5)}$$

- Tap settings of OLTC transformer with voltage step change per tap position of *0.003125 p.u.*

$$\beta_t \in \{-16,-15,...,-1,0,1,...,15,16\}, \quad \forall t \quad \text{(Eq. 6)}$$

$$V_{sub,t} = 1.0 + \beta_t \cdot V^{Tap} \quad \text{(Eq. 7)}$$

- Reactive power injection limit of CBs

$$Q_{CB,i,t} = \sum_{j=1}^{n_i} \alpha_{i,j,t} \cdot Q_{i,j}^{step} \quad \forall i, \forall t \quad \text{(Eq. 8)}$$

$$0 \le Q_{CB,i,t} \le Q_{CB,i}^{\max}, \qquad \alpha_{i,j,t} \in \{0,1\} \quad \text{(Eq. 9)}$$

$$Q_{CB,i}^{\max} = CB_j^{\max} \times CB_j^{step} \quad \text{(Eq. 10)}$$

- Reactive power limit of smart inverter

$$Q_{i,\max}^{Inv}(t) = \sqrt{S_i^2 - [P_i^{G(PV)}(t)]^2} \quad \text{(Eq. 11)}$$

$$\left|Q_i^{Inv}(t)\right| \le Q_{i,\max}^{Inv}(t), \quad \forall i \quad \text{(Eq. 12)}$$

2.4 Simulation Workflow

To evaluate the coordinated control of reactive power devices under realistic operating conditions, a scenario-based Volt-Var Optimization (VVO) framework is developed. The proposed framework simulates system operation over a 24-hour period by incorporating time-varying load and photovoltaic (PV) generation profiles. The optimization process aiming at minimizing total active power losses by determining the optimal reactive power control settings under different scenarios is shown below:

- ***Step 1:*** Initialize the system with input data including the IEEE 33-bus distribution network, hourly load profiles, PV generation curves, control device locations (OLTCs and CBs), and their respective operational limits.
- ***Step 2:*** Formulate the optimization problem by defining:

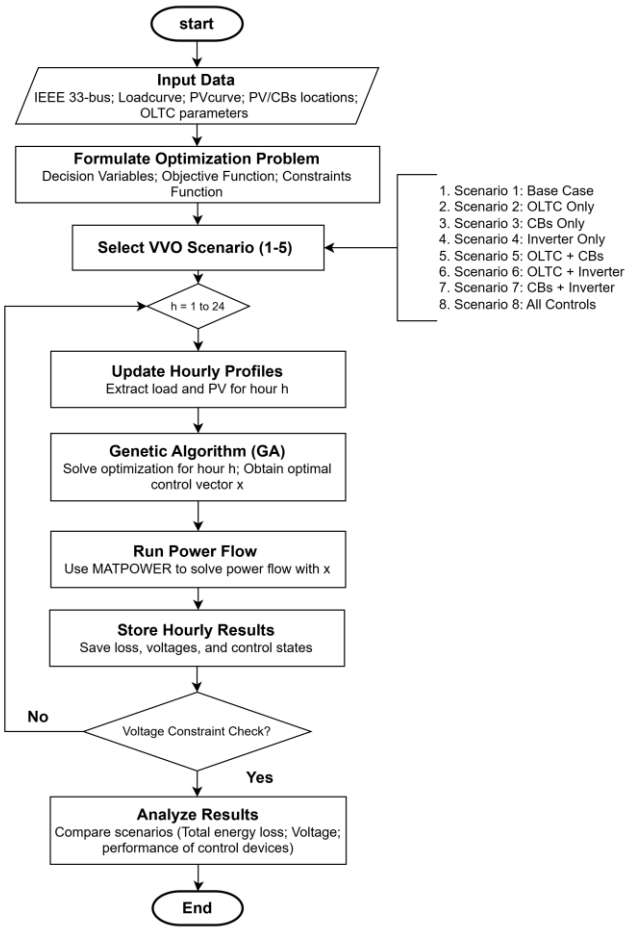


Fig.1. Flowchart of methodology

(i) decision variables (OLTC tap position, CB switching state, inverter reactive power), (ii) an objective function aiming to minimize total active power losses, and (iii) system constraints including voltage limits, power flow balance, and reactive power control device bounds.

- ***Step 3:*** Select the simulation scenario from the eight predefined control configurations.
- ***Step 4:*** For each simulation *h = 1 to h = 24*, update the load and PV generation data to reflect the time-dependent system conditions.
- ***Step 5:*** A population size of 100 and 200 generations were used to ensure convergence. GA is chosen for its ability to handle both discrete and continuous variables, making it suitable for optimizing OLTC taps, capacitor steps, and inverter reactive power.
- ***Step 6:*** Conduct a load flow analysis using the GA-optimized parameters to evaluate voltage profiles and active power losses.
- ***Step 7:*** Key results such as power loss, bus voltages, and control states are stored for hour *h*.

- ***Step 8:*** Upon completion of the daily simulation, analyze the aggregated results to compare the performance of each control scenario in terms of energy loss reduction and voltage regulation effectiveness.

3. RESULTS AND DISCUSSION

3.1 Test System and Parameter Settings

The substation voltage V_{sub} is 1.0 p.u., and the voltage limit is set to ±5% of the nominal voltage, so V_{min} and V_{max} are 0.95 p.u. and 1.05 p.u, respectively. The OLTC transformer is modeled with 32 discrete tap positions, leading to a tap step size of V^{Tap} = 0.003125 p.u., with tap positions defined from –16 to +16. Capacitor banks (CBs) are install at buses 2, 3, 6, 11, 21, and 23, each with a total capacity of 300 kVAr, following the configuration proposed by [14]. Each bank operates in 3 steps of 100 kVAr per step, as defined in this study. Also, photovoltaic (PV) units are installed at buses 13, 24, and 29 with capacities of 801.5 kW, 810.8 kW, and 877.6 kW, respectively, based on the configuration proposed by [15].

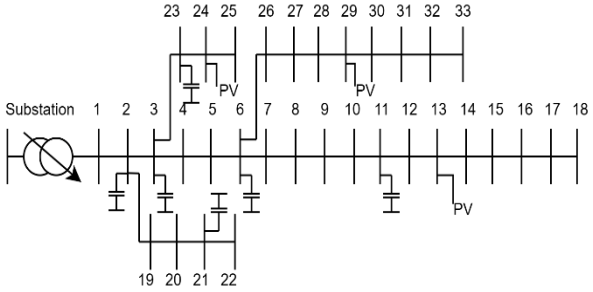


Fig.2. Topology of test system

3.2 PV and Load Profiles

To reflect realistic operating conditions, this study utilizes 24-hour profiles for both load demand and photovoltaic (PV) generation. The load profile is derived from simplified measurements collected from two households in Cambodia, as referenced in [16], while the PV generation profile is based on the data presented in [17]. These two profiles are fitted together to represent a typical day and are shown in Figure 3. The analysis is conducted over a single 24-hour period using these profiles, as longer-term measurement data is currently unavailable. It should be noted that, in practice, load demand may vary seasonally, particularly, during hotter months when the use of cooling devices such as fans increases.

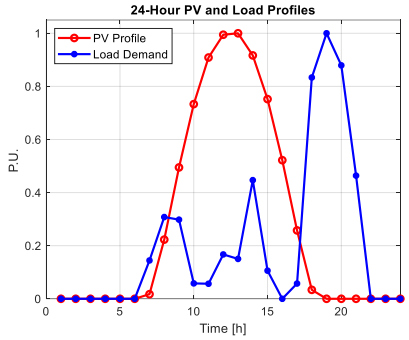


Fig.3. PV and load curves in rural area

3.3 Comparative Analysis of Control Scenarios

To evaluate the effectiveness of different control strategies for Volt-Var Optimization (VVO), eight distinct scenarios were simulated over a 24-hour period. The results of all eight scenarios are presented in Table 1, where the total 24-hour energy loss is used as the performance metric.

Table 1. A compared results of all scenarios

Scenarios	Description	Energy loss 24h (MWh)
Scenario 1	Base case	2.113
Scenario 2	OLTC only	1.958
Scenario 3	CBs only	1.104
Scenario 4	Inverter only	1.732
Scenario 5	OLTC + CBs	1.034
Scenario 6	OLTC + Inverter	1.613
Scenario 7	CBs + Inverter	1.001
Scenario 8	**All combination**	**0.940**

Based on the results, scenarios 3, 4, and 7 which involve only capacitor banks, only PV inverters, and a combination of both demonstrate limited effectiveness in maintaining voltage levels during peak demand. In particular, the coordination of PV inverters with capacitor banks in scenario 7 struggles to keep bus voltages within acceptable limits, especially during high-load hours between 18:00 and 20:00. This issue is evident at several buses, as illustrated in Figure 4.

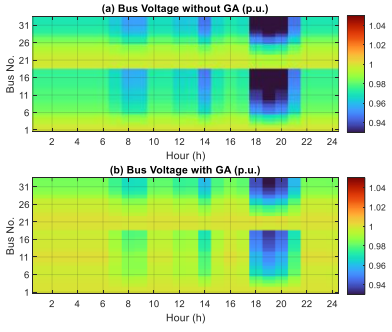


Fig.4. Bus voltage profile of IEEE 33 distribution system when (a) before and (b) after optimization of scenario 3, 4, and 7

Scenario 8, which combines OLTC, capacitor banks, and PV inverters, demonstrated the most effective performance, reducing energy losses by 55.5% to 0.94 MWh compared to 2.113 MWh in the base case. Although GA is a heuristic method and does not guarantee a global optimum, solution reliability was ensured through a robust configuration (population size of 100, 200 generations), consistent convergence behavior, and stable voltage profiles across all hours. Additionally, the results were repeatable under multiple runs, confirming that the solution is of high quality. The full operational performance of the reactive power control devices under this optimal configuration is illustrated in Figures 5 to 10.

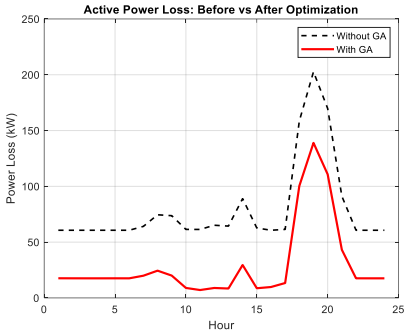


Fig.5. Active power loss over 24h

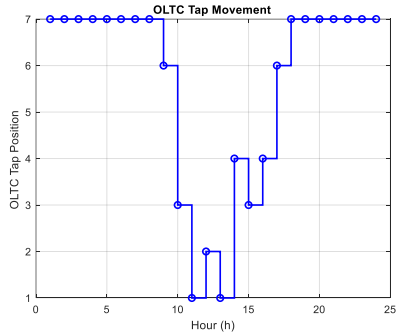


Fig.6. The tap position of OLTC over 24h

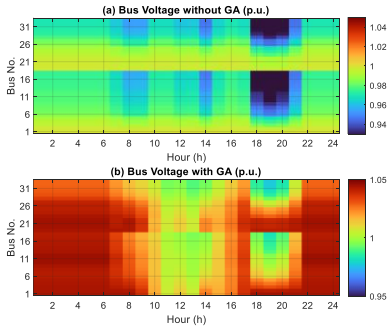


Fig.7. Bus voltage profile of IEEE 33 distribution system when (a) before and (b) after optimization

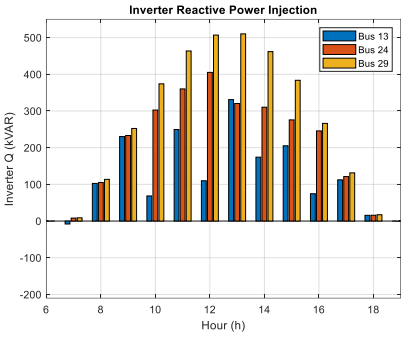


Fig.8. Inverter reactive power dispatch over 24h

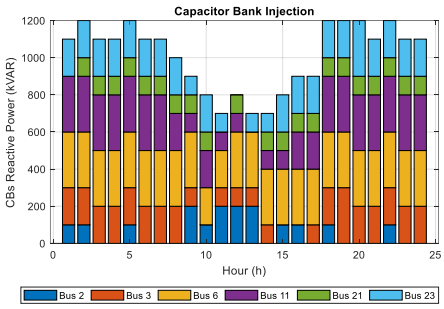


Fig.9. CBs status over 24h

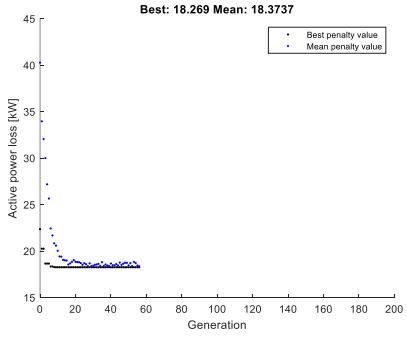


Fig.10. Evolution of the best active power loss at 24:00

4. CONCLUSIONS

This study presents a Volt-Var Optimization (VVO) approach using Genetic Algorithm (GA) to coordinate OLTCs, capacitor banks, and PV inverters for minimizing energy losses over 24 hours. The fully coordinated scenario achieved the best performance, reducing losses while maintaining voltage limits. Specifically, this scenario reduced total energy losses by approximately 55.5% compared to the base case. In contrast, the inverter-only, capacitor bank only, and combination of inverter with capacitor bank cases cannot guarantee voltage constraints at several buses. Future work will focus on minimizing OLTC operations and considering control device costs to improve overall system economy.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Electricité du Cambodge (EDC) through its Training and Research Platform, which was developed with funding from the Agence Française de Développement (AFD) and the European Union (EU), to promote advanced education and research in Cambodia's power sector.

REFERENCES

[1] Y. Xie, Z. Ding, L. Ding, and Q. Zhang, "Multi-objective Volt-Var Optimization Considering Operation Cost of OLTC in Distribution Networks," in *2024 21st International Conference on Harmonics and Quality of Power (ICHQP)*, Oct. 2024, pp. 690–694. doi: 10.1109/ICHQP61174.2024.10768816.

[2] Z. Li and J. Xiong, "Reactive Power Optimization in Distribution Networks of New Power Systems Based on Multi-Objective Particle Swarm Optimization," *Energies*, vol. 17, no. 10, Art. no. 10, Jan. 2024, doi: 10.3390/en17102316.

[3] M. Wu, L. Hong, Y. Wang, Z. Yan, and Z. Chen, "Volt-VAR control for distribution networks with high penetration of DGs: An overview," *Electr. J.*, vol. 35, no. 5, p. 107130, Jun. 2022, doi: 10.1016/j.tej.2022.107130.

[4] Y. Gao and N. Yu, "Model-augmented safe reinforcement learning for Volt-VAR control in power distribution networks," *Appl. Energy*, vol. 313, p. 118762, May 2022, doi: 10.1016/j.apenergy.2022.118762.

[5] R. Yan and Y. Xu, "Multi-Objective and Multi-Agent Deep Reinforcement Learning for Real-Time Decentralized Volt/VAR Control of Distribution Networks Considering PV Inverter Lifetime," *IEEE Trans. Power Syst.*, vol. 40, no. 2, pp. 1558–1569, Mar. 2025, doi: 10.1109/TPWRS.2024.3452154.

[6] D. Cao, J. Zhao, W. Hu, F. Ding, Q. Huang, and Z. Chen, "Attention Enabled Multi-Agent DRL for Decentralized Volt-VAR Control of Active Distribution System Using PV Inverters and SVCs," *IEEE Trans. Sustain. Energy*, vol. 12, no. 3, pp. 1582–1592, Jul. 2021, doi: 10.1109/TSTE.2021.3057090.

[7] Z. Yi, X. Wang, C. Yang, C. Yang, M. Niu, and W. Yin, "Real-Time Sequential Security-Constrained Optimal Power Flow: A Hybrid Knowledge-Data-Driven Reinforcement Learning Approach," *IEEE Trans. Power Syst.*, vol. 39, no. 1, pp. 1664–1680, Jan. 2024, doi: 10.1109/TPWRS.2023.3262843.

[8] B. He, J. Xiao, P. Yi, and S. Gu, "Reactive power optimization of distribution network with photovoltaic power based on sparrow search algorithm," *J. Phys. Conf. Ser.*, vol. 2758, no. 1, p. 012001, Apr. 2024, doi: 10.1088/1742-6596/2758/1/012001.

[9] "(PDF) Chapter 6 - Volt/var control and optimization," in *ResearchGate*. doi: 10.1016/B978-0-443-15578-9.00017-0.

[10] "Volt/Var Optimization for Active Power Distribution Systems on a Graph Computing Platform: An Paralleled PSO Approach." Accessed: Apr. 23, 2025. [Online]. Available: https://ieeexplore.ieee.org/document/9281603

[11] Y. Ai, M. Du, Z. Pan, and G. Li, "The optimization of reactive power for distribution network with PV generation based on NSGA-III," *CPSS Trans. Power Electron. Appl.*, vol. 6, no. 3, pp. 193–200, Sep. 2021, doi: 10.24295/CPSSTPEA.2021.00017.

[12] W. F. Najar, M. D. Mufti, R. Owais, and N. Rehman, "Optimised Volt/Var Control in Distribution System with Capacitor Banks Using Genetic Algorithm," in *2022 1st International Conference on Sustainable Technology for Power and Energy Systems (STPES)*, Jul. 2022, pp. 1–6. doi: 10.1109/STPES54845.2022.10006461.

[13] H. Li, G. Joshi, M. Mazhari, and T. Kopp, "A Genetic Algorithm-Based Power System Volt/Var Optimization Applied to Transmission Studies," in *2023 IEEE Texas Power and Energy Conference (TPEC)*, Feb. 2023, pp. 1–4. doi: 10.1109/TPEC56611.2023.10078585.

[14] Y. Xu, Z. Y. Dong, R. Zhang, and D. J. Hill, "Multi-Timescale Coordinated Voltage/Var Control of High Renewable-Penetrated Distribution Systems," *IEEE Trans. Power Syst.*, vol. 32, no. 6, pp. 4398–4408, Nov. 2017, doi: 10.1109/TPWRS.2017.2669343.

[15] N. R. Dagade, R. V. Dagade, and N. R. Godha, "Optimal allocation of PV based DG and capacitor in radial distribution network," in *7th IET Smart Cities Symposium (SCS 2023)*, Dec. 2023, pp. 632–636. doi: 10.1049/icp.2024.1030.

[16] "(PDF) Planning of low voltage distribution system with integration of PV sources and storage means: Case of power system of Cambodia." Accessed: May 25, 2025. [Online]. Available: https://www.researchgate.net/publication/322487198_Planning_of_low_voltage_distribution_system_with_integration_of_PV_sources_and_storage_means_Case_of_power_system_of_Cambodia

[17] "POWER | DAV." Accessed: May 25, 2025. [Online]. Available: https://power.larc.nasa.gov/data-access-viewer/

[18] N. R. King, "General requirements of electric power technical standards of the kingdom of cambodia," 2004. [Online].Available:http://eac.gov.kh/wpcontent/uploads/2014/08/General-Requirements-of-Electric-Power-Technical Standards.en_.pdf.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

A Frequency Stability Analysis for BESS placement under high penetration of IBRs (Solar PV) considering Load and Solar PV location

Pagnasatyavorn Sreynattha [1*], Vannak Vai[2], Kimsrornn Khon[1,2]

[1*] *Energy Technology and Management Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[2] *Department of Electrical and Energy Engineering, Faculty of Electrical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[*]*Corresponding author: satyavorn2013@gmail.com*

Abstract: *With the high penetration of inverter-based resources (IBRs), particularly solar photovoltaic (PV) systems, traditional power systems are facing new challenges in maintaining frequency stability due to the reduction in system inertia caused by the displacement of conventional synchronous generators. This can lead to higher rates of change of frequency (RoCoF) and lower frequency nadirs during disturbances, which results in serious consequences such as damage to equipment, load shedding, and even large-scale blackouts. To address with this issue, battery energy storage systems (BESS) offer a promising solution due to their fast response capabilities by rapidly injecting or absorbing power to stabilize the frequency during disturbance. In this study, we focus on how BESS can mitigate frequency stability issues. The objective of this study aims to bridge the gaps in previous BESS placement studies by consider overall system dynamic and using two frequency induces i.e. frequency nadir and the Rate of Change of Frequency to define the BESS placement and provide a deterministic analysis by considering the locations of the system loads and Solar PV generation. Furthermore, the proposed methodology is validated in DIgSILENT/PowerFactory on IEEE 39-bus New England test system under various scenarios, including worst contingency events under high IBRs penetration (Solar PV). The results illustrate that the BESS placement at the Solar PV side provide better frequency stability than when the BESS was placed near the Load's side in case of high penetration of IBRS (Solar PV) to grid.*

Keywords: Inverter Base Resources; Battery Energy Storage System; Frequency Nadir; Rate of Change of Frequency; DIgSILENT/PowerFactory

1. INTRODUCTION

Stability evaluation is a well-known issue [1] in the operation and planning of traditional power systems, where synchronous generators predominate. In recent years, renewable energy sources are gradually becoming one of the most promising subjects of energy systems and policies in most countries. However, with the high penetration of renewable resources, the proportion of power generation from inverter-based resources is increasing that led to introduced new challenges for power system stability, particularly frequency stability and it is expected that 36% of global energy demand will be met with renewable energy sources (RES) by 2030 [2]. Recently, photovoltaic generation plants have seen rapid growth among different renewable sources in the last decades, with capacities both of small residential application as well as of grid-connected large-scale commercial projects [3].

Moreover, with the high penetration of inverter-based resources, particularly solar photovoltaic (PV) system, has led to a reduction in system inertia due to the replacement of traditional power plants. Unlike conventional generators that naturally provide rotational inertia, which help to stabilize the frequency change, and many IBRs are not designed to provide primary and secondary frequency support [4, 21], which resulting in serious consequences such as damage to equipment's, load shedding, and even large-scale blackouts [5-6].

According with researchers [7], was introduce the impact of high-level penetration of IBRs. Their research demonstrated that the system will collapse if the level of solar PV increase to more than 40% of the total system generation

[1*] Corresponding author: Pagnasatyavorn Sreynattha
E-mail: satyavorn2013@gmail.com; Tel: +855-98 282 646

capacity. By [8], illustrated that the impact of high penetration of IBRs on frequency stability is caused by Grid-following IBRs, which results in a higher RoCoF after a contingency.

To address this problem, energy storage systems (ESS), particularly battery energy storage systems (BESS), are select as a solution to mitigate these frequency-related issues [9]. Moreover, according to researchers [10], demonstrated that a case study of the Australian power system, which retiring synchronous generators (SGs) and replaced by 80% renewable energy (RE). As a result, a BESS capacity of 15% of total generation can improved the frequency stability. An approach for BESS placement and sizing for frequency support is presented in [11]. The BESS location and size are determined according to the most severe contingencies i.e. a large generation unit outage and operating the system under different penetration levels of inverter-based renewable generation. In this study, the transmission system bus with the larger frequency decline has been identified for the BESS placement. A placement and sizing methodology of frequency responsive BESS to simultaneously maintain frequency and voltage stabilities has been proposed in [12], where the BESS is placed at the most voltage-sensitive bus. In [13], a sensitivity analysis has been conducted to identify the BESS placement where the BESS has been placed on the network based on the location of the system bus that has the minimum line loading.

However, in the previous research, as well as in most of the BESS placement studies in the literature, the BESS placement is identified based on the weakest bus in the system. Moreover, only a few studies have considered the distinction between the load side and the generation side. By author [14], presented about the BESS placement near the load side was improves the frequency stability than the Wind farm location. Therefore, there is a research gap exists in the analysis of BESS placement under high penetration of inverter-based resources, particularly Solar PV, when considering both the load side and the solar PV location. In this research, we aim to bridge these gaps in the existing BESS placement studies by considering the overall system dynamics. Specifically, we propose to use two frequency indicators, namely the frequency nadir and the Rate of Change of Frequency, to define the BESS placement and provide a deterministic analysis by considering the locations of the system loads and Solar PV generation.

This paper is composed of three major parts. The preceding section focuses on the methodology and fundamental terminologies of stability analysis, followed by Section 2. demonstrates the results and discusses. The last section provides the conclusion of the paper.

2. METHODOLOGY

In this research, an analysis has been conducted to investigate the impact of the BESS placement under high penetration of IBRs (Solar PV) on its participation in frequency stability applications. The BESS placement in this research has been chosen to be either on loads side or Solar PV sides to investigate which location the BESS would have more influence on the frequency stability. In this analysis, an initial deterministic investigation of the BESS performance has been conducted under different scenarios, where the BESS performance is assessed by comparing the obtained frequency nadir and RoCoF values from each BESS placement scenario. Fig.1. presents a detailed flow chart that shows the whole process in the analysis. For the simulation implementation of the work in this paper, the DIgSILENT/PowerFactory that is an industry recognized power system analysis tool has been used for simulation.

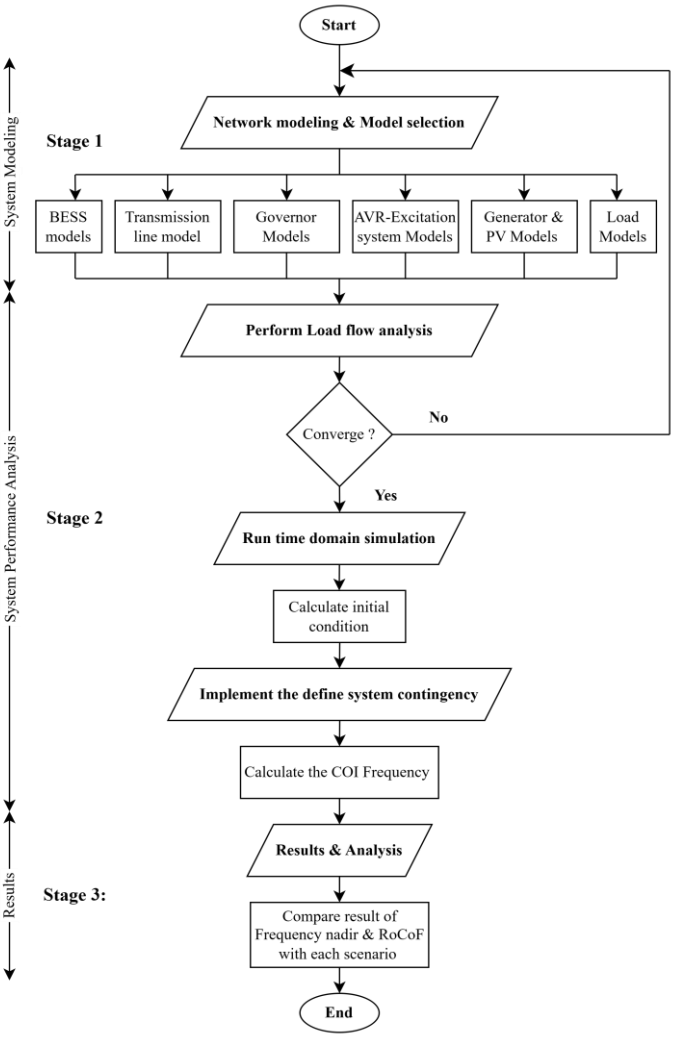


Fig.1. Flow-Chart of the proposed approach.

2.1 Frequency Stability Assessment

Fig.2. shows a typical system frequency response for generator contingency. The transient behavior of power

system during an under frequency event can be examined using two key indices: Frequency Nadir and Rate of Change of Frequency. The frequency nadir (indicated by point B) is the minimum value of frequency reached during the transient period of disturbance. The RoCoF is the declining rate of change of electrical frequency (from point A to B) represented as the gradient of frequency with time. The relationship between the frequency nadir and the active power disturbance is directly proportional, with the inverse proportional relationship to the systems's inertia. A decrease in either of these factors will lead to a decrease in the frequency nadir value and an increase in RoCoF. For the measurement of these indices, the Centre of Inertial frequency is a widely used term that represent the overall system frequency. This Centre of Inertial frequency (f_{COI} in Hz) is calculated by (1) in order to reduce small signal oscillations [12].

$$f_{COI} = \frac{\sum_{n=1}^{n=m} f_n \times H_n \times S_n}{\sum_{n=1}^{n=m} H_n \times S_n} \qquad \text{(Eq.1)}$$

Where, S_n represents the rating (in MVA), f_n is the frequency (in Hz), and H_n indicates the inertia constant (in s) of the nth synchronous generator. The total number of engaged synchronous generators is represented by m. Furthermore, RoCoF indicates how quickly the frequency of the system changes after a power disturbance. A low RoCoF value is an essential characteristic of a frequency-stable grid. This parameter can be determined using (2), according to [12].

$$RoCoF = \frac{P_d \times f_o}{2 \times IR} \qquad \text{(Eq.2)}$$

Where, IR denotes the system inertia (in MWs), P_d indicates the disturbance size (in MW), and f_o defines the nominal frequency of the system (in Hz). The limits bound for RoCoF Should lie in ± 1.5 Hz/s i.e, 0.025 p.u/second for a 60 Hz system [20].

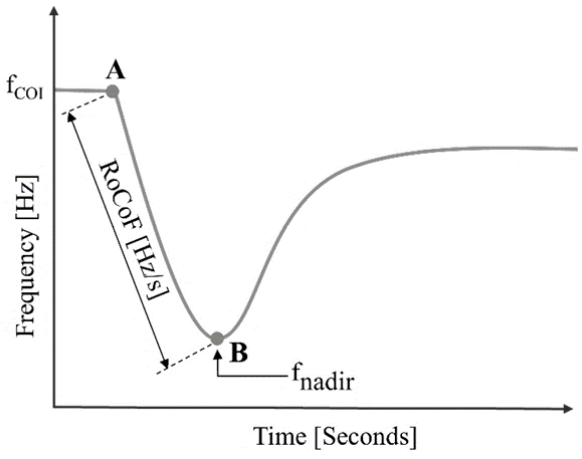


Fig.2. Frequency response of system following a generator outage

2.2 Battery Energy Storage System modelling

The BESS is capable of storing as well as providing reserve active power with a fast response time [15]. It is used to strengthen the frequency stability in power grids with limited spinning reserves due to high renewable integration. A typical BESS consists of two main components: a storage unit that regulates energy through electrochemical processes, and a bidirectional AC-DC inverter for grid integration. While BESS can provide both active and reactive power, active power is prioritized for addressing frequency stability concerns. According to [16], the model of BESS can be understood using the relationship between the DC output voltage of a battery cell, state of charge, and cell current as following (3).

$$V_{DC} = V_{\max} \times SoC + V_{\min}(1 - SoC) - I.Z \qquad \text{(Eq.3)}$$

Where, V_{DC} is the DC output voltage of a battery cell (in V), SoC is the state of charge, V_{max} and V_{min} are the cell voltages during discharged and fully charged conditions respectively. I denotes the cell current (in A), and Z is the internal resistance per cell (in ohm). The BESS model used in this study is implemented in PowerFactory software [17]. There are three controllers in total: the frequency controller, the PQ controller, and the charge controller. These three controllers operate together to ensure that the BESS functions properly when required [15]. Fig.3. shows a schematic diagram of the BESS model.

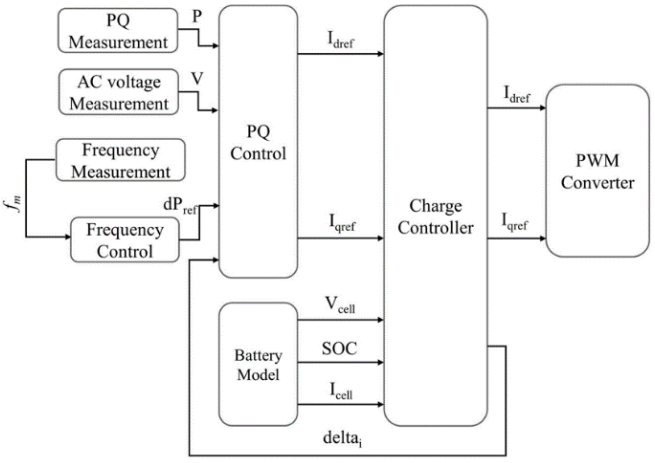


Fig.3. Schematic diagram of BESS model.

3. RESULTS AND DISCUSSION

3.1 Test System

In this study, the IEEE 39-bus New England system is used for simulation purposes with a modified model. The system consists of 10 generators, 39 buses, 12 transformer and

19 loads with the 60 Hz nominal frequency. In order to observe the dynamics of generator, each generator is equipped with proper governor setting and automatic voltage regulator (AVR) to emulate a traditional grid [18].

A single-line diagram of the modified IEEE 39-bus New England system is shown in Fig.4. A series of modifications in the model were made to create a grid with weak connections and high penetration of IBRs. Six solar PV plants with a total installed capacity of 3090 MW were included in this study for the purpose of representing high penetration of IBRs, along with 16.5 kV low-voltage buses and power transformers connecting the PV plants to the grid. Additionally, the capacity of the BESS included in this study is approximately 15% of total generation base on [10]. The Solar PV and BESS models included in the system were build using template from DIgSILENT/PowerFactory.

Moreover, the system was tested under two different load conditions: peak and off-peak. Table 1. shows generator dispatch under peak and off-peak load conditions. The total load demand during peak was 6097.7 MW and during off-peak, it was 4267.97 MW. The distribution of these loads is presented in [19].

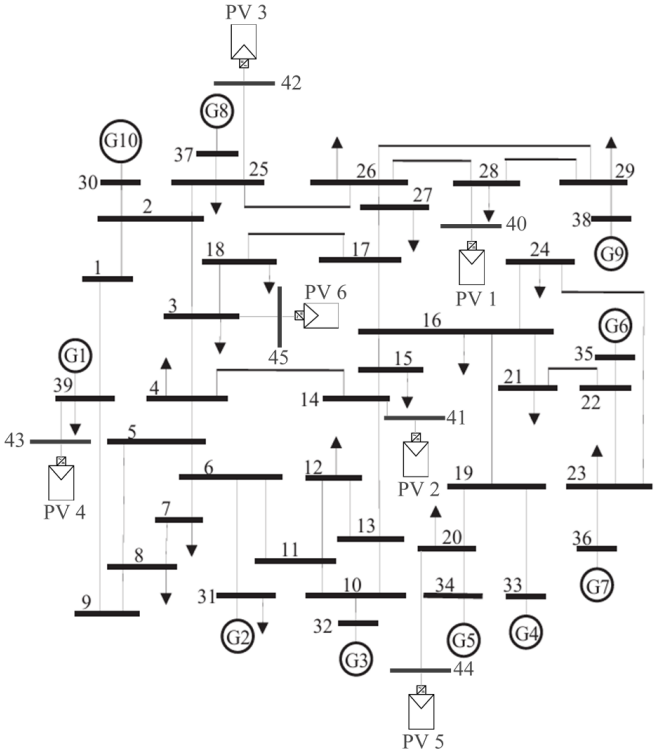


Fig.4. Single line diagram of the modified IEEE-39 buses.

3.2 Overview of Simulation Scenarios

In this study, a deterministic analysis is conducted to investigate the impact of BESS placement under different operating conditions. Two major load scenarios are considered: peak load and off-peak load conditions. Within each load scenario, two types of dynamic events are simulated: (i) the tripping of a major synchronous generator (Generator G09) and (ii) a system load variation event, where the total load is either increased by 5% along with (Generator G09) outage. To evaluate the system's dynamic response, BESS units are placed under two configurations: (1) at all load buses in the system, and (2) at the locations of six integrated solar (PV) plants.

Table 1. Generator dispatch for peak and off-peak load condition

Generator	Peak Load (MW)	Off-Peak Load (MW)
G 01** (Bus 39)	-	-
G 02* (Bus 31)	305.57	390.98
G 03 (Bus 32)	250	100
G 04 (Bus 33)	432	-
G 05 (Bus 34)	-	-
G 06 (Bus 35)	450	100
G 07 (Bus 36)	560	-
G 08 (Bus 37)	-	-
G 09 (Bus 39)	830	530
G 10 (Bus 30)	250	100
Total Synchronous Generation	3077.37	1220.95
IBRS Generation		
PV 1 (Bus 40)	500	500
PV 2 (Bus 41)	500	500
PV 3 (Bus 42)	540	540
PV 4 (Bus 43)	550	550
PV 5 (Bus 44)	500	500
PV 6 (Bus 45)	500	500
Total IBRs Generation	3090	3090
Total Generation	6167.57	4310.98
IBRs Penetration	50.1%	71.68%

* Reference Machine
** Interconnection

3.3 Analysis under Contingency condition

In this case study, the Generator G9 outage event was introduced at 5 seconds to observe the frequency response of the system for each scenario. Fig.5. and Fig.6. present the obtained results. It can be observed that without any BESS integration, the frequency nadir reaches 57.7320 Hz under peak load conditions and 58.0673 Hz under off-peak conditions. Similarly, the Rate of Change of Frequency (RoCoF) is -0.1310 Hz/s under peak load case and -0.1033Hz/s under off-peak conditions. This situation is improved with the BESS when placed at the load side, and it performs even better when placed at generation side.

Table 2. Frequency Stability indices for contingency conditions

Scenario	Peak Load		Off-Peak load	
	f_{nadir} (Hz)	*RoCof* (Hz/s)	f_{nadir} (Hz)	*RoCoF* (Hz/s)
PV Without BESS	57.7320	-0.1310	58.0673	-0.1033
PV+BESS at PV Side	59.7679	-0.1275	58.8388	-0.0941
PV+BESS at Load Side	59.7612	-0.1276	59.8447	-0.0850

Table 3. Frequency Stability indices for load variation conditions

Scenario	Peak Load		Off-Peak load	
	f_{nadir} (Hz)	*RoCof* (Hz/s)	f_{nadir} (Hz)	*RoCoF* (Hz/s)
PV Without BESS	56.9342	-0.1751	57.3674	-0.1414
PV+BESS at PV Side	59.4869	-0.0404	59.7812	-0.1311
PV+BESS at Load Side	59.4201	-0.0409	58.7890	-0.1272

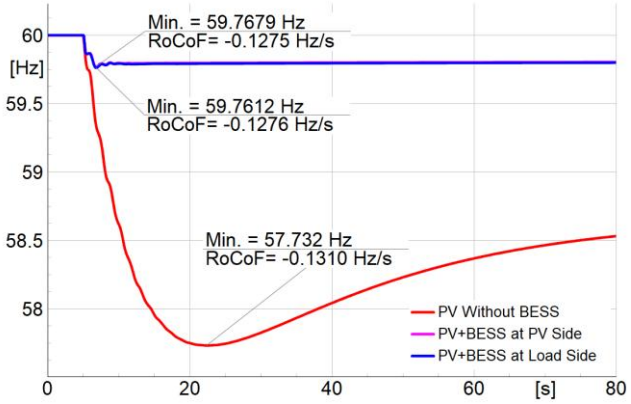


Fig.5. Frequency Response under G9 Outage at Peak Load

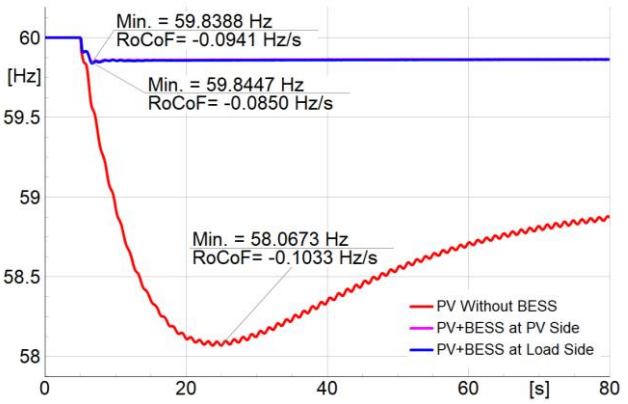


Fig.6. Frequency Response under G9 Outage at Off-Peak Load

3.4 Analysis under load variation condition

This scenario analyzes the application of the proposed method under increased load conditions. The loads are increased by 5%, and the outage of Generator G9 is introduced in 5 second to analyze the frequency response in each scenario. Fig.7. and Fig.8. illustrate the resulting frequency responses. In the absence of BESS integration, the system experiences a frequency nadir of 56.9342 Hz under peak load conditions and 57.3674 Hz under during off-peak conditions. Correspondingly, the Rate of Change of Frequency is measured at -0.1751 Hz/s for the peak load scenario and -0.1414 Hz/s for the off-peak scenario. Moreover, it is clear that placing the BESS at the (load side) results in a less effective frequency response compared to placing the BESS at the (solar PV) side.

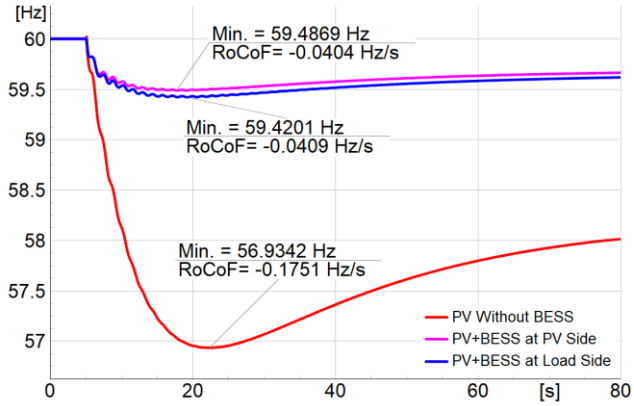


Fig.7. Frequency Response under Load Variation at Peak Load

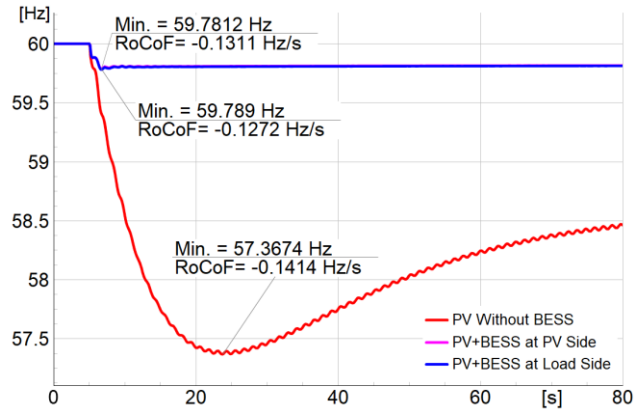


Fig.8. Frequency Response under Load Variation at Off-Peak Load

4. CONCLUSIONS

This study has presented a frequency stability analysis under high IBRs (solar PV) penetration by analyzing BESS placement at both the load and generation sides using the IEEE 39-bus New England system. Through simulation results under generator outage and load variation events, the results showed that placing the BESS at the generation (solar PV) side provides better frequency nadir and RoCoF values compared to placing the BESS at the (load side), under high penetration of IBRs (solar PV). Our future work will focus on minimizing the size of BESS and expand the work to a more large and real power system network.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of Electricité du Cambodge (EDC) through its Training and Research Platform, which was developed with funding from the Agence Française de Développement (AFD) and the European Union (EU), to promote advanced education and research in Cambodia's power sector.

REFERENCES

[1] P. Kundur et al., "Definition and classification of power system stability," IEEE Transactions on Power Systems, vol. 19, no. 3, pp. 1387–1401, Aug. 2004, doi: 10.1109/TPWRS.2004.825981.

[2] M. Shafiul Alam, F. S. Al-Ismail, A. Salem, and M. A. Abido, "High-level penetration of renewable energy sources into grid utility: Challenges and solutions," IEEE Access, vol. 8, pp. 190277–190299, 2020, doi: 10.1109/ACCESS.2020.3031481.

[3] H. M. Sultan, A. A. Zaki Diab, O. N. Kuznetsov, Z. M. Ali, and O. Abdalla, "Evaluation of the impact of high penetration levels of PV power plants on the capacity, frequency and voltage stability of Egypt's unified grid," Energies (Basel), vol. 12, no. 3, Feb. 2019, doi: 10.3390/en12030552.

[4] J. Song, X. Zhou, Z. Zhou, Y. Wang, Y. Wang, and X. Wang, "Review of Low Inertia in Power Systems Caused by High Proportion of Renewable Energy Grid Integration," Aug. 01, 2023, Multidisciplinary Digital Publishing Institute (MDPI). doi: 10.3390/en16166042.

[5] "5.Power System Stability and Control - Prabha Kundur (PowerEn.ir)".

[6] N. Ma and D. Wang, "Extracting Spatial-Temporal Characteristics of Frequency Dynamic in Large-Scale Power Grids," IEEE Transactions on Power Systems, vol. 34, no. 4, pp. 2654–2662, Jul. 2019, doi: 10.1109/TPWRS.2019.2896327.

[7] K. A. Qaid, C. K. Gan, N. Bin Salim, and M. Shamshiri, "Impacts of large-scale solar photovoltaic generation on power system frequency response," in IET Conference Publications, Institution of Engineering and Technology, 2018. doi: 10.1049/cp.2018.1309.

[8] J. Matevosyan et al., "A Future with Inverter-Based Resources: Finding Strength from Traditional Weakness," IEEE Power and Energy Magazine, vol. 19, no. 6, pp. 18–28, 2021, doi: 10.1109/MPE.2021.3104075.

[9] F. Calero et al., "A Review of Modeling and Applications of Energy Storage Systems in Power Grids," Proceedings of the IEEE, vol. 111, no. 7, pp. 806–831, Jul. 2023, doi: 10.1109/JPROC.2022.3158607.

[10] M. R. Amin, M. Negnevitsky, E. Franklin, K. S. Alam, and S. B. Naderi, "Application of battery energy storage systems for primary frequency control in power systems with high renewable energy penetration," Energies (Basel), vol. 14, no. 5, Mar. 2021, doi: 10.3390/en14051379.

[11] M. Ramírez, R. Castellanos, G. Calderón, and O. Malik, "Placement and sizing of battery energy storage for primary frequency control in an isolated section of the Mexican power system," Electric Power Systems Research, vol. 160, pp. 142–150, Jul. 2018, doi: 10.1016/j.epsr.2018.02.013.

[12] N. Al Masood, M. Nahid Haque Shazon, H. M. Ahmed, and S. R. Deeba, "Mitigation of over↓frequency through optimal allocation of bess in a low↓inertia power system," Energies (Basel), vol. 13, no. 17, Sep. 2020, doi: 10.3390/en13174555.

[13] C. K. Das et al., "Optimal sizing of a utility-scale energy storage system in transmission networks to improve frequency response," J Energy Storage, vol. 29, Jun. 2020, doi: 10.1016/j.est.2020.101315.

[14] H. Alsharif, M. Jalili, and K. N. Hasan, "A Frequency Stability Analysis for BESS Placement Considering the Loads and Wind farms Locations," in Asia-Pacific Power and Energy Engineering Conference, APPEEC, IEEE Computer Society, 2022. doi: 10.1109/APPEEC53445.2022.10072092.

[15] A. Jawad, S. A. Naim, C. Saha, and N. Al Masood, "Frequency stability enhancement of a large-scale PV integrated grid," in Proceedings of 2020 11th International Conference on Electrical and Computer Engineering, ICECE 2020, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 290–293. doi: 10.1109/ICECE51571.2020.9393073.

[16] "DIgSILENT PowerFactory Application Guide Battery Energy Storing System Template." [Online]. Available: http://www.digsilent.de

[17] O. Edition, "DIgSILENT PowerFactory User Manual," 2021. [Online]. Available: https://www.digsilent.de

[18] D. Powerfactory, "39 Bus New England System." [Online]. Available: www.digsilent.de

[19] A. F. Ramos, I. Ahmad, D. Habibi, and T. S. Mahmoud, "Placement and sizing of utility-size battery energy storage systems to improve the stability of weak grids," International Journal of Electrical Power and Energy Systems, vol. 144, Jan. 2023, doi: 10.1016/j.ijepes.2022.108427.

[20] M. Javadi, Y. Gong, and C. Y. Chung, "Frequency Stability Constrained BESS Sizing Model for Microgrids," IEEE Transactions on Power Systems, vol. 39, no. 2, pp. 2866–2878, Mar. 2024, doi: 10.1109/TPWRS.2023.3284854.

[21] "Implications of Reduced Inertia Levels on the Electricity System Technical Report on the Challenges and Solutions for System Operations with Very High Penetrations of Non-Synchronous Resources." [Online]. Available: www.epri.com

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Inverter Control Strategies for Active and Reactive Power in Photovoltaic Production for Grid Integration

Tangly Ty*, Menghorng Bun, Vannak Vai, Sievlong Suk

Department of Electrical and Energy Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

*Corresponding author: tangly4050@gmail.com, e20200599@itc.edu.kh

Abstract: *The effort to address climate change and reduce greenhouse gas emissions has led to a transition from conventional energy sources such as coal, gas, and oil to renewable energy resources like low-cost solar photovoltaic, which is increasingly integrated into the grid for more sustainable energy production. This presents significant challenges due to the variable nature of photovoltaic production, which is heavily influenced by solar irradiation and temperature fluctuation. Depending on the operational mode, the power delivered by the photovoltaic plant must align with the requirements set by the grid operator. Otherwise, a penalty cost is priced for any surplus power that exceeds the allowed limit injected into the grid. Hence, this paper proposes optimizing control of active power and reactive power for a three-phase grid-tied inverter. The methodology is conducted at the first layer, a maximum power extraction strategy to ensure obtained power is always at peak value at the photovoltaic plant using Maximum Power Point Tracking based on Perturb and Observe (P&O). At the second layer, a proportional-integral controller is implemented to control the active and reactive power on the inverter side using a three-phase system in the ABC reference frame, which is transformed into direct-quadrature-zero (dq0) components within a rotating reference frame. In order to achieve the connection of the inverter into the grid, three important parameters, such as system frequency, voltage magnitude, and phase sequence, have to be perfectly measured and fed into inverter control. The proposed method is modeled and simulated using the MATLAB/Simulink simulation tool. The simulation results demonstrate the capability of the controller to track varying active and reactive power setpoint reach within a short duration and ensure grid integration under varying solar radiation and temperature levels as the voltage and frequency are within acceptable limits.*

Keywords: Grid-Tied Inverter; Maximum Power Point Tracking (MPPT); Active Power Control; Reactive Power Control; Photovoltaic

1. INTRODUCTION

The global push for cleaner energy has significantly accelerated the integration of photovoltaic (PV) systems into modern power grids. As a result, PV installations have expanded rapidly in both developed and developing countries, including Cambodia, where solar energy plays a growing role in addressing energy access and sustainability challenges[1]. However, this widespread adoption brings technical challenges, especially at the grid interface, due to the intermittent nature of solar power generation, which is inherently dependent on solar irradiance and temperature variations[2].

In grid-connected PV systems, the mismatch between power generation and grid demand can lead to operational issues. In particular, when the active power produced by a PV system exceeds the allowable injection into the grid, it can result in curtailment or incur penalty costs as specified by the grid operator [3]. Moreover, conventional PV inverters typically lack the capability to supply or absorb reactive power, which limits their ability to support grid voltage regulation and meet modern grid code requirements [4], [5].

To overcome these limitations, advanced inverter control strategies have been developed to enable both active (P) and reactive (Q) power control. Among these, the implementation of a Proportional-Integral (PI) controller in the synchronous reference (dq) frame is widely used due to its simplicity, robustness, and effectiveness in achieving decoupled PQ control. By transforming three-phase AC quantities into a rotating reference frame, the dq control structure enables precise regulation of active and reactive power injected into the grid while maintaining grid synchronization and stability [6],[7].

This paper proposes a PI-based control strategy for a three-phase voltage source inverter (VSI) to regulate active and reactive power in a grid-connected PV system. The

[1]* Corresponding author: Ty Tangly
E-mail: e20200599@itc.edu.kh; Tel: +855-15 834 050

control framework is designed and implemented in MATLAB/Simulink, emphasizing compliance with grid requirements under varying solar irradiance and temperature conditions. The simulation results demonstrate the controller's effectiveness in tracking dynamic PQ setpoints while maintaining output voltage and frequency within acceptable limits. These findings highlight the importance of incorporating reactive power control in smart inverters to enhance the flexibility and grid-friendliness of PV systems.

2. METHODOLOGY

2.1 System Configuration

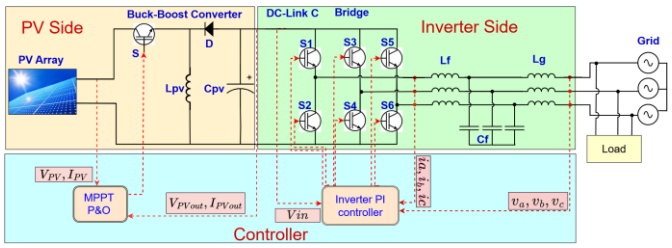


Fig. 1. System configuration of Solar Inverter for Grid Integration

The system configuration, as indicated in Fig. 1, integrates a photovoltaic (PV) array with a buck-boost converter controlled by an MPPT algorithm based on the Perturb and Observe (P&O) method, as studied in [8], to extract maximum power under varying solar conditions. The regulated DC output charges a DC-link capacitor, which supplies a three-phase voltage source inverter (VSI). The inverter, equipped with an LCL filter, converts the DC power into grid-synchronized AC. This paper focuses on the inverter side, where a PI controller in the synchronous dq-reference frame is implemented to regulate active and reactive power injection into the grid, ensuring compliance with grid requirements.

2.2 Control System

The control structure of the three-phase grid-tied inverter is based on a dual-loop strategy comprising an outer power loop and an inner current loop, all operating in the synchronous dq-reference frame. As shown in Fig. 3, the outer loop receives active (P^*) and reactive (Q^*) power references and generates the corresponding current references $i_d{}^*$ and $i_q{}^*$. These are fed into the inner current loop as detailed in Fig. 4, which regulates the inverter output currents i_d and i_q track the references accurately. The transformation between stationary abc and rotating dq frames is handled using Park and inverse Park transforms. A Phase-Locked Loop (PLL) as illustrate in Fig. 5 is used to synchronize the inverter with the grid voltage by extracting the grid angle θ and angular frequency ω from the measured three-phase voltage V_{abc}. The inverter outputs three-phase current I_{abc}, which enable controlled injection of active and reactive power into the grid.

In the following section, the design of the closed-loop control scheme using PI controllers is presented.

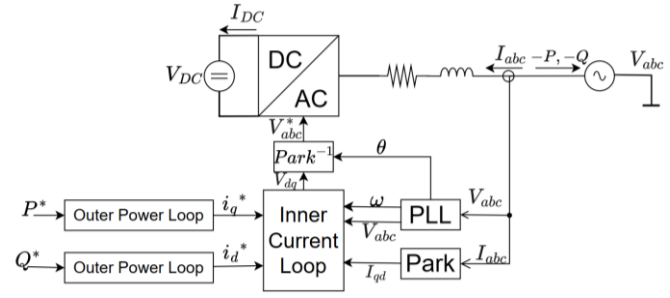


Fig. 2. Inverter Control PQ scheme

2.2.1 Outer Power Loop

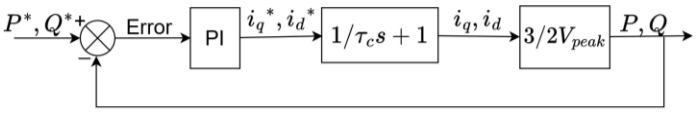


Fig. 3. Power loop design Structure

Assuming the PLL aligns the q-axis with the grid voltage $v_{dg} = 0$, the current references are derived as follows [9]:

$$i_q{}^* = \frac{2P^*}{3v_{qg}} \qquad \text{(Eq. 1)}$$

$$i_d{}^* = \frac{2Q^*}{3v_{qg}} \qquad \text{(Eq. 2)}$$

where: v_{qg} is q-axis component of the grid voltage

2.2.2 Inner Current Loop

The current controller allows for the regulation of current flowing through the inverter to the grid by regulating the magnitude. As shown in Fig. 4, the reference and actual magnitudes are subtracted from each other, producing an error input to the PI controller, and the result is applied to the AC electrical plant. Both are identical and thus:

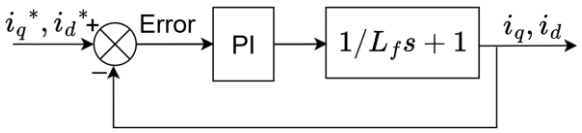


Fig. 4. Current Loop design structure

Closed loop transfer function:

$$\frac{i_q(s)}{i_q^*(s)} = \frac{G_c(s)}{L_f s + R_f + G_c(s)} \quad \text{(Eq. 3)}$$

PI controller:

$$G_c(s) = \frac{k_p s + k_i}{s} \quad \text{(Eq. 4)}$$

PI controller tuning applying IMC tuning approach:

$$P(s) = \frac{K}{\tau s + 1} \quad \text{(Eq. 5)}$$

$$K = \frac{1}{R_f}, \tau = \frac{L_f}{R_f} \quad \text{(Eq. 6)}$$

$$k_p = \frac{L_f}{\tau_c}, k_i = \frac{R_f}{\tau_c} \quad \text{(Eq. 7)}$$

$$\frac{i_q(s)}{i_q^*(s)} = \frac{i_d(s)}{i_d^*(s)} = \frac{1}{\tau_c s + 1} \quad \text{(Eq. 8)}$$

2.2.3 Phase-Locked Loop

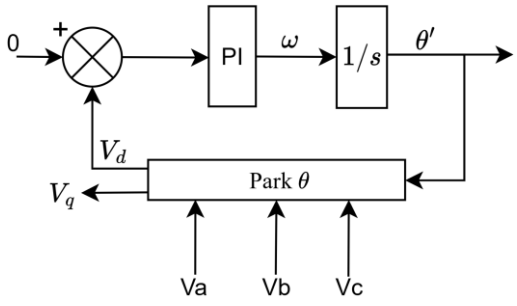
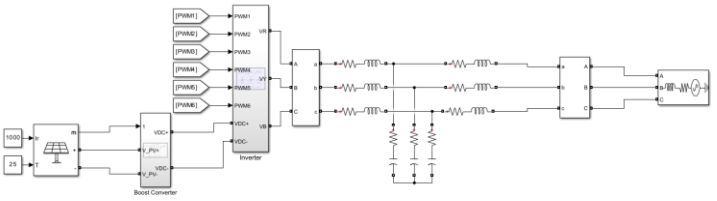


Fig. 5. PLL design structure

The Phase-Locked Loop (PLL), illustrated in Fig. 5, is designed to synchronize the inverter's control system with the grid voltage by driving the d-axis voltage v_d to zero aligning the q-axis with the grid voltage vector, as describe in [10]. It takes the three-phase voltage V_{abc} as input, transforms it into the dq-frame using the Park transformation and uses the d-axis component v_d as the input to a PI controller. The controller drives v_d to zero, ensuring alignment of the rotating reference frame with the grid voltage vector. The output of the PI controller is the estimated frequency ω, which is integrated by a Voltage-Controlled Oscillator (VCO) to produce the estimated grid angle θ'.

2.3 MATLAB Model Implementation



3. RESULTS AND DISCUSSION

Table 1. Operation parameter

Parameter	Name	Design Value	Unit
Apparent power	S_n	10	KVA
Nominal Frequency	f_n	50	Hz
Nominal AC Voltage	$V_{abc,n}$	400	V_{RMS}
Nominal DC Voltage	$V_{DC,n}$	800	V
Switching frequency	f_{sw}	10	KHz
Inductor filter Inverter size	L_{f1}	500	μf
Capacitor filter	C_f	100.28	μf
Inductor filter Grid size	L_{f2}	500	μH
Resistance	R_f	500	mΩ
Proportional gain	$K_{P,PLL}$	1332	
Integral gain	$K_{I,PLL}$	303,103	
Proportional gain	$K_{P,CL}$	5.2	
Integral gain	$K_{I,CL}$	510	
Proportional gain	$K_{P,PQ}$	0.135	
Integral gain	$K_{I,PQ}$	135	

Table 1 outlines the operational parameters for the inverter system, including apparent power S_n, nominal frequency (f_n), switching frequency (sw), filters components, and controller gain. These parameters were derived through the calculation base on the methodology in [11].

Table 2. PQ simulation reference

Time instant (s)	Active Power (KW)	Reactive Power (Kvar)
0	0	0
0.3	-3	-2
0.5	-6	-2
0.8	-4	0
0.9	-2	1

Table 2 presents the time-varying reference values for active power (in kW) and reactive power (in kVAr) applied to the inverter during a 1-second simulation period. A negative sign indicates power injection into the grid, whereas a positive sign denotes power absorption from the grid. For instance, at t = 0.3 s, the inverter is commanded to inject both active and reactive power into the grid, with reference values of -3 kW and -2 kvar, respectively. Conversely, at t = 0.9 s, the inverter absorbs 1kvar of reactive power from the grid while continuing to inject 2kW of active power. All reference values in Table 2 were pre-defined as the power set-points, as described in [9].

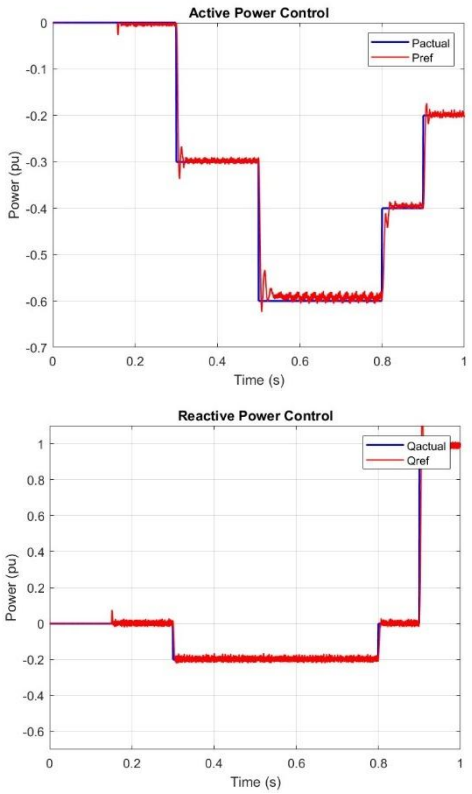


Fig. 6. Active and reactive power simulation

Fig. 6 illustrates the performance of the voltage source converter operating in P-Q control mode, tracking the reference profiles described in Table 2 over a 1-second simulation period. The active (P^*) and reactive (Q^*) power setpoints are applied as step functions at specific time instants: 0 s, 0.3 s, 0.5 s, 0.8 s, and 0.9 s. These reference transitions are followed by the inverter's actual power outputs P and Q, which respond with smooth first-order dynamics.

The response shows that the controller achieves accurate power tracking across all transitions, with minimal overshoot and acceptable settling times. Notably, the system exhibits decoupled behavior, as changes in one power component do not significantly impact the other. For instance, at $t = 0.8s$, while active power remains constant at $-4kW$, the reactive power jumps from -2Kvar to 0 and is then inverted to $1kVAr\ at\ t = 0.9\ s$. These transitions are clearly reflected in the curves actual Q and P), which closely follow their respective setpoints.

The simulation confirms that the cascaded PI control structure, with faster current control and slower outer power loops, ensures stable and accurate power delivery. This validates the effectiveness of the inverter's control scheme under dynamically changing grid conditions.

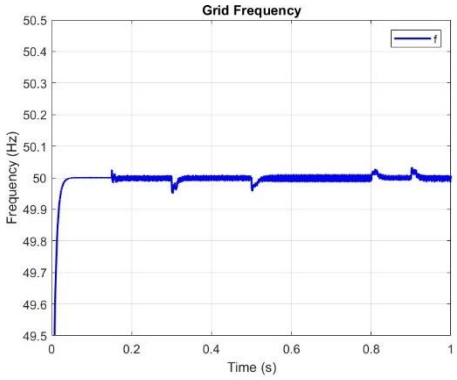


Fig. 7. Grid Frequency

The grid frequency which shown in Fig. 7 stabilizes at 50 Hz after an initial transient from 49.6 Hz, with minor fluctuations maintained throughout the 1-second period due to effective synchronization by the phase-lock loop (PLL).

Table 3 Active Power Tracking Error

Time instant (s)	Reference P (KW)	Actual P (KW)	Steady-State Error	Error (%)
0	0	-7.3	0.0007	0
0.3	-3	-2907.3	0.0092	3.0
0.5	-6	-5867.6	0.013	2.2
0.8	-4	-4055.6	0.0055	1.39
0.9	-2	-2044.5	0.0044	2.22

Table 4 Reacctive Power Tracking Error

Time instant (s)	Reference Q (KW)	Actual Q (kvar)	Steady-State Error	Error %
0	0	0	0	0
0.3	-2	- 1975.2	0.0024	1.23
0.5	-2	- 1997.6	0.00023	0.11
0.8	0	- 6.248	0.00627	0
0.9	1	968.44	0.03156	3.15

4. CONCLUSIONS

This study presented a PI-based control strategy for a three-phase voltage source inverter aimed at regulating active and reactive power in grid-connected photovoltaic systems. The proposed method, implemented in the dq-reference frame, enables precise and decoupled control of active and reactive power under varying solar irradiance and temperature conditions. Simulation results demonstrated the controller's ability to track dynamic power setpoints with minimal steady-state error and acceptable transient performance. The response curves confirm that the inverter's dual-loop control structure, consisting of fast inner current control and slower outer power control, maintains system stability while ensuring compliance with grid requirements. Additionally, the results highlight the importance of reactive power control to enhance the grid compatibility of PV inverters. Future work may explore adaptive or nonlinear control techniques for improved robustness under grid disturbances or high penetration of distributed energy resources.

ACKNOWLEDGMENTS

I would like to express my gratitude toward my teachers for their guidance and encouragement throughout the course of this project. I am truly grateful for their time, patience, and dedication in mentoring me through both the technical and academic aspects of my research.

REFERENCES

[1] R. A. Othman and O. S. A.-D. Al-Yozbaky, "Effect of reactive power capability of the PV inverter on the power system quality," *Indonesian Journal of Electrical Engineering and Informatics (IJEEI)*, vol. 10, no. 4, pp. 780–795, 2022.

[2] A. Sargolzaei, M. Jamei, K. Yen, A. I. Sarwat, and M. Abdelghani, "Active/reactive power control of three phase grid connected current source boost inverter using particle swarm optimization," in *Progress in Systems Engineering: Proceedings of the Twenty-Third International Conference on Systems Engineering*, Springer, 2015, pp. 141–146.

[3] "S. Adhikari and F. Li, 'Coordinated V-f and P-Q Control of Solar Photovoltaic Generators With MPPT and Battery Storage in Microgrids,' IEEE Transactions on Smart Grid, vol. 5, no. 3, pp. 1270–1280, May 2014. DOI: 10.1109/TSG.2014.2301157," *ResearchGate*, doi: 10.11648/j.ajece.20200402.12.

[4] M. J. M. A. Rasul and M. Kolhe, "M. J. M. A. Rasul and M. Kolhe, 'Reactive Power Control of a 3-phase Grid-Connected Building Integrated Photovoltaic System,' IOP Conf. Ser.: Mater. Sci. Eng., vol. 605, p. 012012, 2019.," *IOP Conf. Ser.: Mater. Sci. Eng.*, vol. 605, no. 1, p. 012012, Aug. 2019, doi: 10.1088/1757-899x/605/1/012012.

[5] S. Lu and K. A. Corzine, "Advanced Control and Analysis of Cascaded Multilevel Converters Based on P-Q Compensation," *IEEE Trans. Power Electron.*, vol. 22, no. 4, pp. 1242–1252, Jul. 2007, doi: 10.1109/tpel.2007.900471.

[6] P. Diaz Font, "Modelling, control and analysis of a voltage source converter," Master thesis, Universitat Politècnica de Catalunya, 2023. Accessed: May 02, 2025. [Online]. Available: https://upcommons.upc.edu/handle/2117/388346

[7] "A. Yazdani and R. Iravani, Voltage-Sourced Converters in Power Systems: Modeling, Control, and Applications, Hoboken, NJ, USA: Wiley-IEEE Press, 2010." Accessed: May 02, 2025. [Online]. Available: https://onlinelibrary.wiley.com/doi/book/10.1002/9780470551578

[8] D. Peftitsis, G. Adamidis, and A. Balouktsis, "An investigation of new control method for MPPT in PV array using DC–DC buck–boost converter," in *Proceedings of 2nd WSEAS/IASME international conference on renewable, energy sources*, 2008.

[9] A. Egea-Alvarez, A. Junyent-Ferré, and O. Gomis-Bellmunt, "Active and reactive power control of grid connected distributed generation systems," in *Modeling and Control of Sustainable Power Systems: Towards Smarter and Greener Electric Grids*, Springer, 2012, pp. 47–81.

[10] A. Egea-Alvarez, A. Junyent-Ferré, and O. Gomis-Bellmunt, "Active and reactive power control of grid connected distributed generation systems," in *Modeling and Control of Sustainable Power Systems: Towards Smarter and Greener Electric Grids*, Springer, 2012, pp. 47–81.

[11] P. Diaz Font, "Modelling, control and analysis of a voltage source converter," Universitat Politècnica de Catalunya, 2023.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Electric Field Analysis on Distribution Suspension Insulator with Different Outer Diameters

Vuth Sireyroth [1], Vuthy Chansideth[1], Va Sreylen [1], Vann Ryda [1], You Lyhour [1*]

[1] *Department of Electrical and Energy Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: lyhour.you@itc.edu.kh*

Abstract: *A suspension insulator is a type of electrical insulator used at poles or towers of distribution and power transmission lines, insulators composed of porcelain, glass, and composite to sustain above conductors. In addition, suspension insulators are used to isolate live-phase from poles or tower to avoid electric shock. The performance and reliability of suspension insulators in high-voltage transmission systems are significantly influenced by their electric field distribution. The common mode of failures is included punctures, hardware corroding, flashover voltage and cracking from cement expansion leading to punctures. It is essential to verify and test the suspension insulator beforehand to understand its electrical characteristics and mitigate potential failure. An insulator's electrical performance is greatly influenced by its physical dimension. It is critical to distinguish surface properties from bulk or volume properties. This paper represents a comprehensive study of the electric field on the suspension insulator by deeply investigating the varying impact of the outer diameter of the suspension insulator on the electric field that could lead to the flashover voltage and also system failure. To examine the electrical performance of the suspension insulator, the finite element method (FEM) was used as the main method for the electric field analysis of the suspension insulator. More than that, the software simulation ANSYS is the tool for the simulation purpose of the electric field distribution. The study focuses on changing the size of the outer diameter of the insulator and it has shown that the performance of the suspension insulator is notably affected by its geometric design; mostly, the outer diameter which is considered as one of the main effects among the other factor on the electric field.*

Keywords: FEM; Electric Field; Suspension Insulator; Outer Diameter

1. INTRODUCTION

An electrical insulator is a material that resists the flow of electric current, helping to protect electrical equipment from short circuits by preventing electrical conductors from accidentally touching each other. Some examples of insulators include materials like polymer, wood, and plastic [1]. Suspension insulators support overhead power lines, where they are mounted on poles or towers to stop any electric current from leaking. They play a key role in the safe supply of electricity. Based on the transmission voltage level and the surrounding environment, suspension insulators are used in configurations of two or more [2]. There are different types of distribution line insulators, such as pin type, suspension type, post type, strain type, spool type, ceramic type, and non-ceramic type [1].

For the insulator electric fields may be considered as the primary cause of insulation or stated more cryptically: "no field, no problem". Thus, it is clear that a detailed knowledge of the electric field associated [3]. And according to [4], Insulators must withstand testing voltages defined by standard insulation levels, but each material has a critical electric-field strength beyond which it fails through ionization, arcing, or corona discharge. Due to non-uniform electric fields, localized ionization can occur at lower voltages, leading to early flashover. The shape and characteristics of the insulator's fittings (electrodes) significantly affect electric-field strength and flashover voltage. Factors influencing flashover include the distance between electrodes, voltage waveform, electrode shape and material, as well as environmental conditions like temperature, pressure, humidity, and pollution. There are many people and theory that has been study and develop on the meaning of the electric field and according to [5], two scientists have given the meaning of the electric field as:

- Michael Faraday (1791–1867): electric fields as a ‘constant necessary condition to action in space’, which exists even when a test body is ‘not in place’.

[1*] Corresponding author: You Lyhour
E-mail: lyhour.you@itc.edu.kh ; Tel: +855-10 292 169

- Richard Feynman (1918–1988) defines the electric field similarly: ‘The positive charge, in some sense, distorts or creates a ‘condition’ in space, so that when we put the negative charge it feels a force’.

In [4], [6], [7] it was mentioned that the shape and size of the insulator has the main effect on the electric field and voltage distributions of an insulator. and the electric field on the flow though the suspension insulation could affect the insulator. So, the aim of this research is to study and understand the relationship between the electric flied performance and outer diameter dimensions of porcelain suspension insulators. This investigation will analyze how changes in the outer diameter size and shape of insulators can influence their electric filed. Therefore, any efforts made to enhance the overall performance of outdoor insulators will primarily involve modifying and optimizing on their outer dimeter.

2. METHODOLOGY

2.1 Suspension insulator

This study used ANSI 52-1 distribution suspension insulator as shown in Fig. 1 as a main equipment for simulating its electric field. The selected insulator was modeled in AutoCAD, and its outer diameter was increased incrementally from the original size. And for the length of the outer diameter will increase by 10 mm from its original size. The term 'outer diameter' is illustrated in Fig. 2. The reason that it is only increase up to 220.92 mm because of the size limitation of the insulator that can handle its own weight.



Fig. 1. Porcelain suspension insulator.

2.2 Electric Field

As mentioned earlier, the main effect of the insulator is its electric filed. Thus, the electric filed of the insulator is measure to get the information of the insulator and also to make sure the safety operation of the network. The measurement of the electric filed on the insulator is started from the bottom end of the outer diameter on the insulator and ended on the top end of the insulator. As illustrated in Fig. 2, the red line out side show how the electric field is being measured.

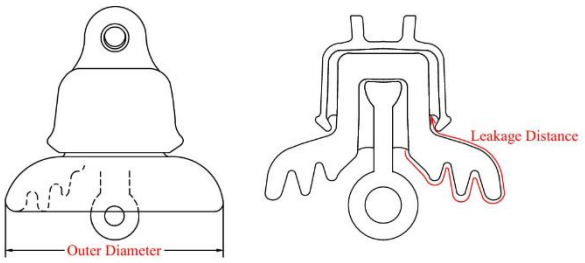


Fig. 2. The outer diameter and leakage distance of the insulator.

2.2.1 Equation of Electric Field Calculation

From [7], the equation that used for calculation of electric field distribution is determined by electric potential distribution. And it can be written as:

$$E = -\nabla \mathrm{V} \qquad \text{(Eq. 1)}$$

From Maxwell’s equation:

$$\nabla E = \frac{\rho}{\varepsilon} \qquad \text{(Eq. 2)}$$

Where,

ρ	Resistivity	$[\Omega/m]$
E	Electrical stress	$[V/m]$
ε	Dielectric constant of dielectric material $= \varepsilon_o \times \varepsilon_r$	$[F/m]$
ε_o	Air or space dielectric constant = $(8.85419 \times 10^{-12})$	$[F/m]$

Placing (Eq. 1) in (Eq .2);

$$\nabla \cdot (-\nabla V) = \frac{\rho}{\varepsilon}$$

$$\nabla^2 V = \frac{\rho}{\varepsilon_0 \times \varepsilon_r} \qquad \text{(Eq. 3)}$$

Without space charge $(\rho = 0)$, the equation become Laplace’s equation.

$$\varepsilon_0 \varepsilon_r (\nabla \cdot \nabla V) = 0 \qquad \text{(Eq. 4)}$$

And expansion of (Eq. 4) for three-dimension coordinate systems. The equation that describes the voltage potential distribution in high voltage situation were the medium is non-homogenous then equation is:

$$\varepsilon_0 \varepsilon_r \left(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} \right) = 0 \qquad \text{(Eq. 6)}$$

The above equation is to allow for the different permittivity in the different regions.

2.2.2 Properties of Materials

There are four major components to suspension insulators. They consist of an inner steel stem with grooves that holds a cap made of malleable iron and an insulating block with a conical head sealed with porcelain. The bottom end of this stem is rounded and designed to be kept in place with a pin by piercing the cap of the subsequent piece. Iron is utilized to make the HV cap and pin. To strengthen the insulation's mechanical resistance to the dielectric, these joints are finely grained. The relative permittivity of the various materials determines how the electric potential and electric field distributions are calculated in electrostatics. The electrical properties of insulation materials used to simulate on the insulator model in shown in Table 1.

Table 1. Technical parameter of the simulated insulator.

Material	**Relative permittivity ε_r**	**Electrical Conductivity, σ $[S/m]$**
Iron (Cap and Pin)	1	10,300,000
Porcelain	5.7	0
Cement [8]	5	1e-12
Air	1.0006	0

2.3 Software Simulation Selection

Numerous articles address the calculation of the electric field and potential distribution for high-voltage insulators. The distribution of the potential and electric field along insulators at different voltage ratings can be found using two primary methods. The first approach involves extracting empirical results using tools and equipment from a laboratory. This approach is costly and time-consuming. A numerical approach with software is preferable [9]. Since the numerical approach is more application there is one numerical technique that is most widely uses, it is finite element method (FEM). And the FEM is one of the most practical methods and implemented by researcher to calculate the distribution of the electric potential and electric field by determining approximate solutions of partial differential equations [10]. The main reason to use the FEM is its ability to solve quite simply problems that are distinct on complex geometries [11]. There are many FEM package like COMSOL Multiphysics and ANSYS. Moreover, in [12], the researcher used COMSOL to make of 3D simulation of cap and pin glass insulators, revealing uneven distribution of potential and electric field. The results demonstrated a decrease in electrical parameters from pin to cap. Therefore, according to this ANASYS software is used instead of COMSOL.

2.4 Finite Element Method Process

According to [7], the process of the finite element method is consisting of 7 steps:

1. Defining and discretizing region's solution in final number of regions or components.
2. Applying discretization to the calculation model, generating nodes and elements.
3. Determining electrical characteristics parameters for each element based on material characteristics.
4. Collecting and assembling elements in the solution using Maxwell equations and boundary conditions.
5. Forming finite element equations thought proper simplifying
6. Solving the system of equations using suitable calculation method.
7. Solving different solution matrixes through iterative methods.

2.5 Flowchart of Methodology

In Fig. 5, is the overall process of this paper. Frist it started by choosing one suitable insulator for simulating and then measuring all size of the insulator. After step 1 is completed next the insulation is draw in AutoCAD software. After that it is imported into ANSYS to start filling the properties and also some assumption on the material. Following this some step up was made. Succeeding the set up the analyze is started and after one result the insulation is change until the outer diameter doesn't exceed 220.92 mm.

3. RESULTS AND DISCUSSION

In order to get the result some assumption of the setting condition is used. Fig. 3 demonstrate the drawing of all the 9 insulators that will used for simulating in ANSYS software. In this context it only showing the half of the insulator but when it rotates 180° it will become a 3D insulator.

It is clearly illustrated in Fig. 3, the length of the cap diameter is started from 160.92 mm and the increase is by 5 mm on each size it means that the total amount that will increase on the cap diameter is 10 mm. The pin of insulator is applied with 10 kV and the cap is set at 0 V (ground). The simulations of each case are kept with the same air boundary surrounding it.

From Fig. 4, it can be seen that the highest electric field is mostly happened on the junctions where two objected nearly met or attached to another object.

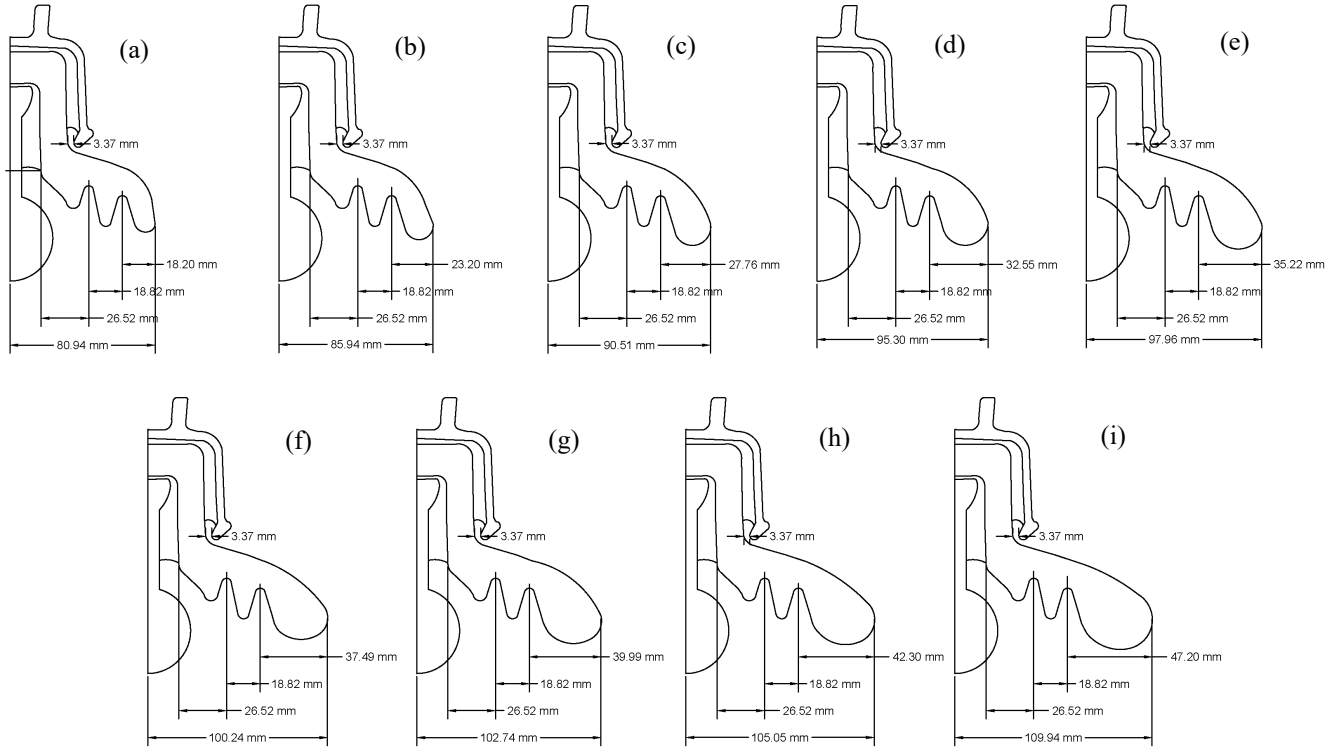


Fig. 3. All the 9 insulator used for simulation.

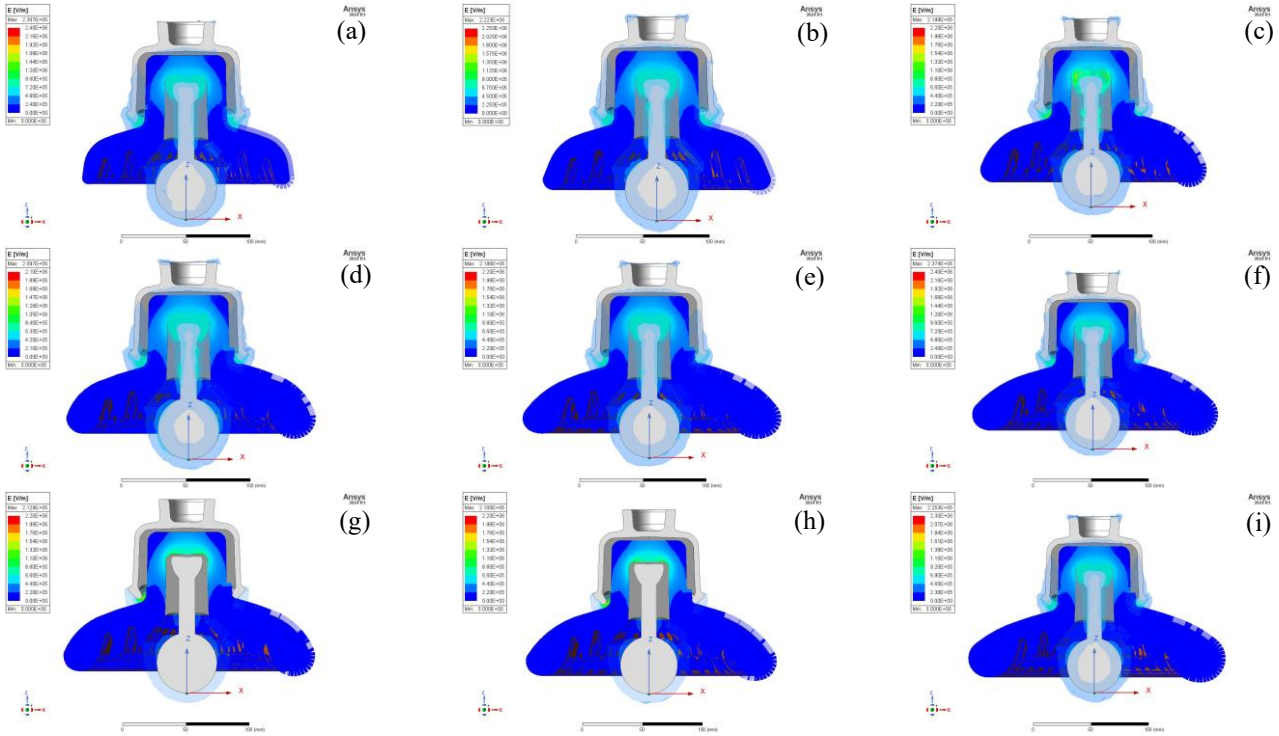


Fig. 4. 3D view of each insulator with its electric field.

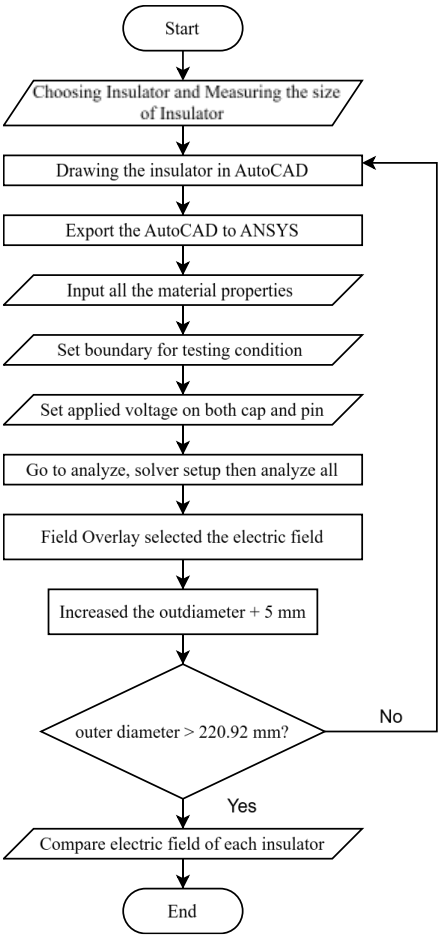


Fig. 5. Flowchart of methodology.

Fig. 6 illustrated in the electric filed distribution on the leakage distance of all the insulator that used associated with the changed of the outer diameter for simulation. The electric field simulation was carried out on nine variations of a suspension insulator model, each differing by incremental changes in the outer diameter. The results obtained from FEM analysis revealed a clear correlation between the insulator's outer diameter and its electric field distribution. However, the correlation between these two variables were not in linear relationship.

The simulations demonstrated that the electric field was most concentrated near the regions where the insulator's components were in close proximity, such as the junctions between the insulator's body and its hardware. This aligns with theoretical expectations, as sharp edges and narrow gaps typically exhibit higher electric field intensities. In Table 2, it is shown that as the outer diameter increased from 160.92 mm to 220.92 mm, the maximum electric field distribution seems to be fluctuated from 0.3661 kV/mm to 0.5815 kV/mm. Among these, the outer diameter of 200.92 mm provided lowest maximum electric field of 0.3661 kV/mm among other outer diameters.

So, two more outer diameters are formed and simulated around 200.92 mm such as 195.92 mm and 205.92, respectively. Although, the outer diameter of 195.92 mm is smaller than 200.92 mm, its maximum electric field still the highest of all. Meanwhile, the outer diameter of 205.92 mm is the second highest after that. This implies that build up the outer diameter with a specific size could disperses the electric field more uniformly across the surface of the insulator. The reduction in field concentration at sharp interface points indicates a lowered risk of flashover, which is a primary cause of insulation failure. Table 2 summarized the comparison of maximum electric field and voltage distribution of each insulator with its outer diameter.

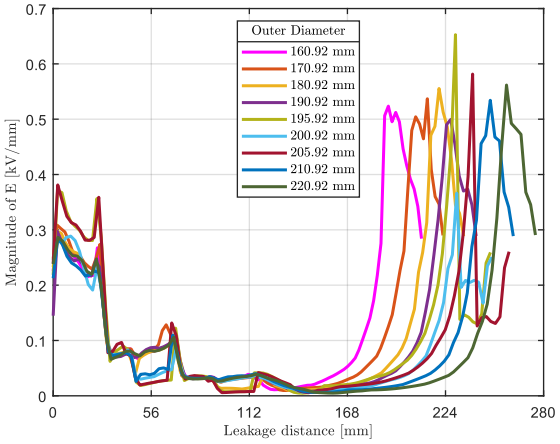


Fig. 6. Electric field distribution for individual's leakage distance of insulator.

Moreover, the voltage distribution of each insulator is shown in Fig. 7. The voltage distributions remain nearly the same in every condition. Thus, it means that the electric field distributions that calculated from ANSYS simulations are also corrected.

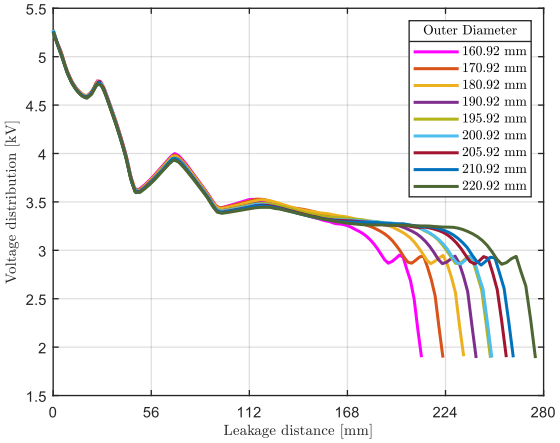


Fig. 7. Voltage distribution of each insulator.

Table 2. Maximum electric field and voltage distributions of all insulator.

Outer diameter [mm]	Max. electric field [kV/mm]	Max. voltage [kV]
160.92	0.5236	5.2723
170.92	0.5369	5.2744
180.92	0.5555	5.2754
190.92	0.4995	5.2562
195.92	0.6527	5.2739
200.92	0.3661	5.2754
205.92	0.5815	5.2574
210.92	0.5342	5.2637
220.92	0.5617	5.2500

4. CONCLUSIONS

This paper investigated the impact of the outer diameter on the electric field distribution of ANSI 52-1 porcelain suspension insulators using the FEM via ANSYS simulation. The findings shown that geometric design particularly the outer diameter plays a significant role in electric field behavior. The results demonstrated that the increased in the outer diameter from 160.92 mm to 220.92 mm, the maximum electric field varied from 0.3661 kV/mm to 0.5815 kV/mm. And it also shown that with specific size of the outer diameter could help the electric field of the insulator to drop significantly. And potentially reducing the likelihood of flashover and improving the overall safety and reliability of the insulation system. This insight is important for the design and engineering of high-voltage transmission components, where minor physical modifications can yield substantial performance improvements.

REFERENCES

[1] "Suspension Insulator: Construction, Classification and Its Applications." Accessed: Apr. 01, 2025. [Online]. Available: https://www.elprocus.com/suspension-insulator-in-process/

[2] "Suspension Insulators | Products | NGK INSULATORS, LTD." Accessed: Mar. 26, 2025. [Online]. Available: https://www.ngk-insulators.com/en/product/suspension-insulator.html

[3] I. W. McAllister, "Electric fields and electrical insulation," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 9, no. 5, pp. 672–696, 2002, doi: 10.1109/TDEI.2002.1038656.

[4] J. Hrastnik and J. Pihler, "Designing a New Post Insulator Using 3-D Electric-Field Analysis," *IEEE Transactions on Power Delivery*, vol. 24, no. 3, pp. 1377–1381, 2009, doi: 10.1109/TPWRD.2009.2022662.

[5] J. Roche, "Introducing electric fields," *Phys Educ*, vol. 51, no. 5, p. 055005, 2016, doi: 10.1088/0031-9120/51/5/055005.

[6] B. He, H. Zhao, Y. Xiu, C. Chen, Y. Meng, and Z. Peng, "Influence of flow-electric field on withstand characteristics of HV suspension insulators in sandstorm circumstance," *IEEE Transactions on Dielectrics and Electrical Insulation*, vol. 23, no. 5, pp. 2892–2898, 2016, doi: 10.1109/TDEI.2016.7736850.

[7] Lyhour You and Channarong Banmonkol, "A Comparative Study on Electrical Characteristics of Suspension Insulators," Master, Chulalongkorn University, Bangkok, 2024.

[8] B. Hani, N. M'ziou, and A. Boubakeur, "Simulation of the potential and electric field distribution on high voltage insulator using the finite element method," *Diagnostyka*, vol. 19, pp. 41–52, Mar. 2018, doi: 10.29354/diag/86414.

[9] M. S. Sanjari Nia, M. Altimania, P. Shamsi, and M. Ferdowsi, "Comprehensive Analysis for Electric Field and Potential for Polymeric and Ceramic Insulators," vol. 8, Jan. 2020, doi: 10.12691/ajeee-8-1-4.

[10] E. Akbari, M. Mirzaie, A. Rahimnejad, and M. Asadpoor, "Finite Element Analysis of Disc Insulator Type and Corona Ring Effect on Electric Field Distribution over 230-kV Insulator Strings," *International Journal of Engineering & Technology*, vol. 1, pp. 407–419, Aug. 2012, doi: 10.14419/ijet.v1i4.330.

[11] N. K. Challagondla, "Behaviour of water drop on the insulator surface and study of electric field distribution on parallel insulators under DC," BTU Cottbus-Senftenberg, 2015.

[12] R. Salhi, A. Mekhaldi, M. Teguar, O. Kherif, and M. E. A. Slama, "Cap-Pin Glass Insulator Simulation and Leakage Current waveform Extraction," in *2022 2nd International Conference on Advanced Electrical Engineering (ICAEE)*, 2022, pp. 1–4. doi: 10.1109/ICAEE53772.2022.9961979.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Economic analysis of solar rooftop at Toyota Cambodia Office in Phnom Penh by using PVsyst software: EAC OLD vs New Regulation

Prom Piseth [1*], Vannak Vai [1], Kimsrornn Khon [1,2]

[1] *Department of Electrical and Energy Engineering, Faculty of Electrical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[2] *Energy Technology and Management Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[*] *Corresponding author: piseth_prom14@yahoo.com*

Abstract: *Rooftop photovoltaic (PV) systems have become increasingly popular due to their technological, economic, societal, and environmental benefits. Business, residential, and industrial owners are eager to invest in solar energy as a way to reduce their electricity bill. The main objective of this paper is to propose 6 scenarios case studies (a grid-connected solar system with a limited peak load demand and available roof space) for Toyota Cambodia's commercial office. The study will compare six different PV system scenarios, with and without storage systems, using PVsyst software to evaluate technical and economical KPI. The objective of study propose 6 scenario as described in methodology below. For grid-connected PV systems, the generated electricity can be sent directly to the load. In contrast, storage PV systems often require energy storage devices BESS, like batteries, to store the electricity generated during the day when there is a surplus energy. To calculate the technical sizing of a grid-connected PV system with/without BESS with different scenarios under the solar PV regulation we have the method apply mathematical formular and PVsyst software, compute the energy production and compare the results with the proposed scenario (S0, S1, S2, S3, S4, and S5). For Economic analysis on the above 6 scenarios, we have method to calculate the NPV and LCOE, compare all scenarios to find which one has a good LCOE, saving, ROI, and payback period. Regarding to the Environmental analysis on the above 6 scenarios we have method to calculate the CO2 emission in all scenarios above by multiplying CO2 emission factor/kWh and compare the CO2 emission in all scenarios above to find which one is the lowest. This stored energy can then be used at night or during cloudy periods. After selecting the optimal solar system, the system will be designed and its maximum output power and efficiency will be evaluated. The first step in the study process is gathering significant data, including site location, energy consumption, meteorological information and space available for installation. The following step, PVsyst will be utilized to choose the best-case study among these two cases study which is mentioned above. Finally, in the economic part, make a comparison based on the COE and surely followed by EAC's regulation.*

Keywords: Photo Voltaic system (PVsyst); Battery Energy Storage System (BESS), Levelized Cost of Electricity (LCOE); Total Net Present Cost (NPC); Return on Investment (ROI); Total Revenues (TR); Lifecycle cost of PV project (LCC PV project); Project Lifetime (PL); Initial costs (Init costs).

1. INTRODUCTION

The energy market will become ever more important for Cambodia's economy as it continues to develop, with rapidly increasing demand for electricity. The Cambodia Energy Outlook estimates an increase by 8.63 times the last 15years [1], [2]. According to Cambodian authorities, electricity demand in the country is expected to grow from its current 1.5 GW to 3.9 GW by 2020, and 5.04 GW megawatts by 2024. So far, power imports from Laos, Vietnam and Thailand are helping the country to meet its growing demand. In Cambodia, electricity generation is dominated by hydro power and coal power plants. Power outages are especially common in the dry season, making Cambodian businesses dependent on diesel back-up systems to ensure their electricity supply. Both, high electricity costs and power

[1*] Corresponding author: Piseth Prom
E-mail: piseth_prom14@yahoo.com; Tel: +855-16 710 959

outages lead to an increasing attractiveness of investments in photovoltaic systems. Looking for ways to make energy more sustainable, solar PV systems for on-site electricity generation can be a solution to provide an alternative to conventional fossil fuels, lower electricity costs, and to increase electricity supply security. Solar PV systems for on-site electricity generation can be a solution to lower electricity costs and to increase electricity supply security.[1], [2]

2. METHODOLOGY

2.1 Flow chart of Methodology for Solar Roof Top Exiting project

The flowchart in Figure 1. shows how each stage of the process needs to be completed. Firstly, we start with the site assessment then moving to the analysis of the site information. Last but not least Design PV system for all scenarios. Finally, the best scenario is considered by the cost of efficient (LCOE) which is less than.

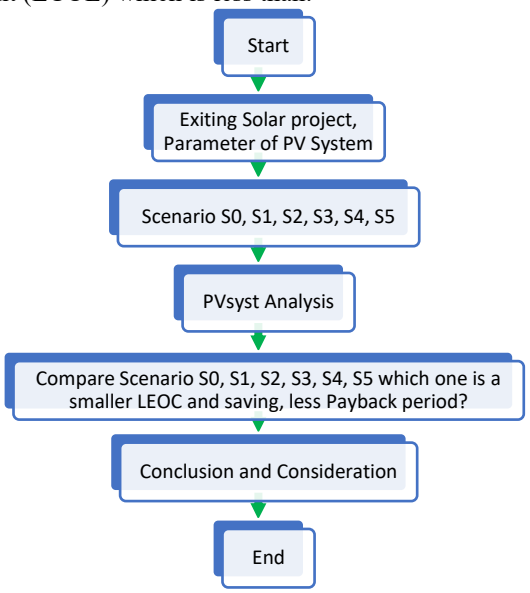


Fig. 1. Flow chart of Methodology for Solar Roof Top Exiting project

The flowchart in figure 1 shows how each stage of the process needs to be completed. First, we start the site assessment then moving to the analysis of the site information. Last but not least Design PV system and simulation with PVsyst for Scenario 0, 1, 2, 3, 4, 5. Finally the best scenario is considered by the cost of efficient (LCOE) which is less than. The step-by-step process for optimal sizing for PV systems + BESS is explained through the flowchart, as shown in Figure 1.

- Scenario 0: 100% purchase from EDC
- Scenario 1: EDC + PV (exiting) + Old Regulation
- Scenario 2: EDC+PV (exiting) + New Regulation
- Scenario 3: EDC+PV (exiting)+ BESS + New Regulation
- Scenario 4: EDC + PV (exiting)+ BESS=50% +New Regulation
- Scenario 5: EDC + PV (exiting)+ BESS=70% +New Regulation

2.2 Exiting Regulation Solar roof Top in Cambodia

Table 1. Old Regulation on Energy Tariff for consumers with PV system:

Type of Consumer and Conditions	Unit	Tariff
5. Commercial, Administration and Other Consumers Connecting to MV Distribution Line		
• For Consumers with PV System		
- Maximum Power Capacity Rate	$/kW/ month	5.80
- Electricity Rate (24H)	$/kWh	0.1500
➢ The Maximum Capacity of the Solar Power Project, as mentioned on AC side at the output of inverter based on rated inverter capacity, shall not be more than 50% of the Contract Demand in kW (if Contracted Demand is stated in kVA, it will be converted to kW at normative Power Factor of 0.90), as stated in the Power supply Agreement with the consumer.		

2.3 New Regulation on Energy Tariff for consumers with PV system

Condition 1:
Electricity users who have received permission from the Electricity Authority of Cambodia to install rooftop solar power systems and then use electricity from the national grid Combined with their roof top solar power to meet their full fill consumption needs must pay a compensation tariff to the national grid based on the number of kilowatt-hours of energy from the rooftop solar system, according to the capacity requirements as specified table below.

Table 2. New Regulation on Energy Tariff for consumers with PV system

Solar Rooftop size Capacity		Compensate Tariff
Small size	Not Over 10kWAC	No need to pay compensation tariff
Medium size	Over 10kWAC to 50kWAC	0.037 $/kWh
	Over 50kWAC to 100kWAC	0.047 $/kWh

In case the electric user has installed the BESS system combined with Solar Energy Rooftop on their own roof to fulfill the need of electric used and the Energy quantity (kWh) that discharge from the BESS system in which month have amount from 50% of Energy quantity which generate from that Solar rooftop. This is assumed that month the electric user has used their own Energy from BESS system to fulfill the power need for their electric used for at that month the electric user is not required to pay the compensate tariff as above. In contrast in which month the Energy quantity (kWh) which generate from the BESS system is less than 50% of the Energy quantity that generated from the Solar Rooftop system so, the electric user must pay the compensate tariff to the Energy quantity which is not fulfill, follow the formular below:

Energy Payment= {Energy quantity generate from Solar Rooftop system (kWh) – [2 x Energy generated from BESS system (kWh)]} x Compensate Tariff as table above.

2.4 Load Profile and PV system sizing

The Load Profile estimates the total power usage for a project over a particular time period. The greater the accuracy of the load profile, the more effective the PV system. In Cambodia, the on-grid PV system is required three important factors: Available space, as we have seen in the line graph of Toyota office the load operation almost flat, the peaking hour is around **190kW**. So, this solar system is exiting solar project designed PV system and installed is (100KW) based on which is chosen in the load calculation 50% of peak load demand capacity 200kW.

Battery capacity required is known as Ampere-hour **(**Ah).

The total ampere-hour can be obtained using the following formula:

$$C_{BAH} = \frac{Day\ excess\ Energy(kWh)}{DOD * \eta_{BAH} * V_B} \quad (1)$$

Where is the daily energy required from the battery, DOD is the permissible depth of discharge? The values of the factors considered while calculating for battery size are battery loss 0.85, depths of discharge, DOD – 0.8, and battery's nominal voltage. From our battery specification, the nominal voltage is 26V.

2.5 Economic of PV system

2.5.1 Project costing

- Total Net Present Cost (NPC)

The NPC is the present value of all the system's costs over its lifetime, minus the present value of all the revenue it earns. Costs include capital costs, replacement costs, O&M costs. Revenues include salvage value and grid sales revenue [3]

$$C_{NPC,total} = \frac{C_{ann,total}}{CRF(i,n)} \quad (2)$$

Where:

$CNPC,total$ is the total net present cost ($)

$Cann,total$ is the total annualized cost $/year

$CRF(i,n)$ is the capital recovery factor that represent a series of equal annual income and can be defined by:

$$CRF(i,n) = \frac{i(1+i)^n}{1+i^n-1} \quad (3)$$

n is the project life time(year)

i is the real discount rate % and can be defined by:

$$i = \frac{i'-f}{1+f} \quad (4)$$

i' is the nominal interest rate %

f is the annual inflation

- Discount Factor

$$Discount\ Factor = \frac{1}{1+i^N} \quad (5)$$

N is duration of a project lifetime

2.6 Project Profit

- Levelized Cost of Electricity (LCOE)

The levelized cost of electricity (LOCE), also referred as levelized cost of energy (LCOE), is used for electrical systems to calculate the cost of producing one unit of electricity (USD/kWh). The operational lifespan cost includes the initial cost of capital, operations, and maintenance and the cost of replacement (i.e., batteries and electronic equipment). With regard to the second assumption, time value of money is considered and LCOE is calculated using a discount rate. The LCOE at discount rate is calculated as follows [4]

where, N is the project lifetime and DR is the discount rate.

Considering economic aspects, the per unit cost of electricity for PV systems significantly depends on several factors including the geographical locations and discount rates [5]. For PV projects in Asia countries, this value is seen to range between 4 and 8%[6]. As an average, discount rate of 12% is used for Cambodia.

$$LCOE = \frac{Ini\cos ts + \sum_{n=1}^{N} \frac{O\&M(\$)}{(1+DR)^n} + \sum_{n=1}^{N} \frac{\mathrm{Re}\,p\cos t(\$)}{(1+DR)^n}}{\sum_{n=1}^{N} \frac{Annual\ PV\ System(kWh)}{(1+DR)^n}} \quad (6)$$

Where:

Init costs: Initial costs ($)

Rep: Replacement cost ($)

Annual PV system: Annual electricity supply to the load by PV system (kWh)

2.7 Payback Period

The Payback Period is the period required to recoup the funds expended in an investment, starting from the investment year. The payback period for Solar roof top PV systems can be calculated by following [7]:

$$Payback\,Period = \frac{LCC(\$)}{TR(\$)/\mathrm{P}\,L(year)} \quad (7)$$

Where:
PP: Payback Period
LCC: Lifecycle cost of PV project ($)
TR: Total Revenues ($)
PL: Project Lifetime (year)

2.8 Return on Investment (ROI)

Return on Investment (ROI) is a popular profitability metric used to evaluate how well an investment has performed. ROI for PV system is expressed as a percentage and is calculated by dividing project's net profit by operational lifespan cost of PV project. Mathematically, it can be expressed as follows[4], [5]:

$$ROI = \frac{T\,R(\$) - LCC(\$)}{LCC(\$)} * 100 \quad (8)$$

Where:
ROI: Return of investment (%)
TR: Total Revenues ($)
LCC: Lifecycle cost of PV project ($)
Similar to the LCOE, payback period and return on investment (ROI) for the 4 cases shall be calculated separately with and without the use of the discount rate.

3. RESULTS AND DISCUSSION

3.1. Sizing Methodology PV Systems in PVsyst

From the main result data PVsyst, solar produced energy per year =184448kWh/year, Used Energy per year =636578kWh/year. Performance Ratio=70.81%, Solar fraction =28.97%.

3.2. BESS (kWh) Calculation

By using (1):

- Calculation BESS (kWh) for scenario 3:

From PVsyst total daily Excess Energy (kWh) = 62.4kWh
By using equation (1):

$$C_{BAH} = \frac{62.4(kWh)}{0.85*0.8} = 91.76kWh \approx 100kWh$$

- Calculation BESS (kWh) for scenario 4:

50% Total daily Energy (kWh) selected highest month in June. So 50% daily (kWh) =[(18294/30)*50%]= 304.9 kWh
By using equation (2):

$$C_{BAH} = \frac{304.9kWh}{0.85*0.8} = 448.38kWh \approx 450kWh$$

- Calculation BESS (kWh) for scenario 5:

70% Total daily Energy (kWh) selected highest month in June. So 70% daily (kWh) =[(18294/30)*70%]= 426.86 kWh By using equation (3):

$$C_{BAH} = \frac{426.86kWh}{0.85*0.8} = 627.73kWh \approx 630kWh$$

3.3. Project costing

The project costing of case study 1 is approximately = 138040 x 0.785 = **108361.40$**.

Table 3. Price of Component for PV system

	Equipment	Cost
	PV modules + BOS	0.755 USD/W
	Inverter	0.03 USD/W
	Storage	0.250 USD/W

There for the Total Project costing:

- Scenario S1&S2: The project costing of case study 1 is approximately = 138040 x 0.785 = **108361.40$**.
- Scenario S3: Battery sizing with Total daily Excess Energy (kWh) = 100*1000*0.250 = **2,5000.00 $**
- Scenario S4: Battery sizing with 50% of Total daily PV Energy (kWh) = 450*1000*0.250 = **112,500.00$**
- Scenario S5: Battery sizing with 70% of Total daily PV Energy (kWh) = 630*1000*0.250 = **157,500.00$**

Table 4. Total Project investment cost (USD) for all Scenarios

Scenario	Total Investment Cost (USD)
S0	0.00 USD
S1	108361.40 USD
S2	108361.40 USD
S3	133,361.40 USD
S4	220,861.40 USD
S5	265,861.40 USD

3.4. Comparison of electricity for the 1st year operations expense saving All scenarios

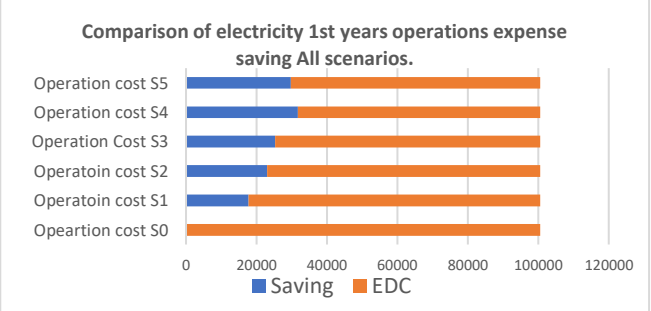


Fig. 2. Electricity 1st year operation expenses and saving all scenario

Base on electrical Tariff above Exiting regulation and new regulation the price for the grid is 0.158$ without solar and grid with solar is 0.15$.
As we have seen in the Figure 2 and Table 5 below, Scenario 4 is saving 31,793.82$ in a 1st year which is accounted for 32% while the other saving in scenario S1=18%, S2=23%, S3=25%, S5=30%.

Table 5. Operation cost for 1st year and saving ($,%)

Operation Cost	Total Expenses ($)	Saving ($)	Saving (%)
S0	100,579.32 $	0.00 $	0.00 %
S1	82,816.16 $	17,763.15 $	18 %
S2	77,533.46 $	23,045.86 $	23 %
S3	75,313.61 $	25,265.71 $	25 %
S4	68,785.5 $	31,793.82 $	32 %
S5	70,846.05 $	29,733.27 $	30 %

3.5. Payback period

The operating and maintenance cost or it was called O &M cost is the total budget that use for operating the whole system.

Every single component in the project, especially, solar PV for 25 years and inverter for 15 years is their limit lifetime during the operating of the system.
Replacement of Inver is doing every 15 years of warranty and it costs $ 4,206.88 also the BESS for Scenario 3 is 25,000.00$ and the BESS scenario 4 is 112,500.00 $ and the BESS scenario 5 is 157,500.00 $. By Using (7):

Table 6. Payback period for all scenarios

Payback Period		
Scenario 1	16.23years	= 16 Years 3 Months
Scenario 2	8.66 years	= **8 Years 8 Months**
Scenario 3	13.49 years	= 13 Years 6 Months
Scenario 4	20+ years	= Over 20 Years
Scenario 5	20+ years	= Over 20 Years

As we have seen the payback period, the case study **Scenario S2** is spending 8 years and 8 months shorter than the case study S0, S1, S3, S4, S5 to paid back the system cost. The case study **Scenario S2** yields is more annual profits than the case study after 8 years of repayment.

3.6. Net Present Cost

This part is about the economic calculation for the lifetime of the project that we counted on 20 Years of the project lifetime. Discounted Rate 12.00% base on investment economic analysis from ADB economic analysis paper [3]. Since the inflation rate is 2% will into account for economic analysis [3].

In summary, we got all parameters that can be done in (3) and (4):

$$i = \frac{0.12 - 0.02}{1 + 0.02} = 9.8\%$$

$$\Rightarrow CRF(0.098, 20yrs) = \frac{0.098(1+0.098)^{20}}{(1+0.098)^{20} - 1} = 0.116$$

We use this discount rate and for calculation the hold project for 20 years of a project lifetime to find NPC, Cann and LCOE.

As we have seen the from Fig. 3 the LCOE for scenario S2=0.086$/kWh is the lowest electricity cost than the other Scenario S0 =0.158$/kWh, S1=0.115$/kWh, S3=0.093$/kWh, S4=0.111$/kWh, S5=0.134$/kWh.

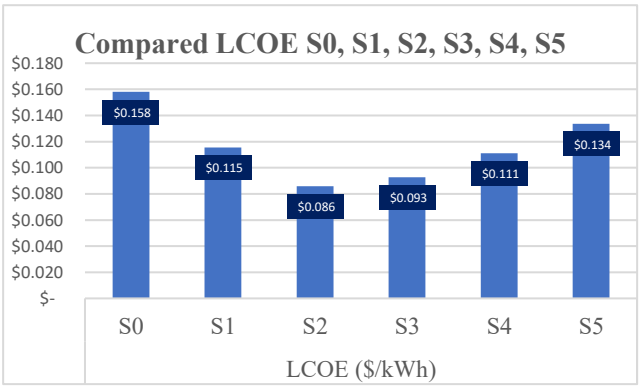


Fig. 3. LCOE comparison for all Scenarios

3.7. Project Evaluation and Discussion

Table 7. Overall system comparison initial cost ($) and O&M cost ($) between scenario S0, S1, S2, S3, S4, and S5:

	PV Capacity & BESS	Initial Cost ($)	O&M Cost ($)
S0	0.00kWp	0.00 $	100,579.32 $
S1	138.04kWp /100kWAC	108361.40 $	82,816.16 $
S2	138.04kWp /100kWAC	108361.40 $	77,533.46 $
S3	138.04kWp /100kWAC +BESS 100kWh	133,361.40 $	75,313.61 $
S4	138.04kWp /100kWAC +BESS 450kWh	220,861.40 $	68,785.5 $
S5	138.04kWp /100kWAC +BESS 450kWh	265,861.40 $	70,846.05 $

Table 8. Overall system comparison NPC cost ($) and LCOE cost ($) and 1st year saving ($) between scenario S0, S1, S2, S3, S4, and S5:

	NPC ($)	LCOE ($/kWh)	1st Year Saving ($)
S0	100,579.32 $	0.158 $/kWh	0.00$
S1	82,816.16 $	0.115 $/kWh	817,763.15$
S2	77,533.46 $	**0.086$/kWh**	23,045.86 $
S3	75,313.61 $	0.093 $/kWh	25,265.71 $
S4	68,785.5 $	0.111 $/kWh	31,793.82 $
S5	70,846.05 $	0.134$/kWh	29,733.27 $

By seeing the summary comparison from Table above, the LCOE of scenario **S2 is 0.086$/kWh** lower than the LCOE of scenario S0, S1, S3, S4, and S5. Besides that, in 1st years of saving cost, the scenario S4 could saving 31, 793.82 $ higher than the rest scenarios. Therefore, the grid-connected system is design with power peak load demand (200kWAC), (Case Study **S2** New regulation Compensate Tariff (0.047$/kWh)) which is 138kWp with an annual production of 184448kWh with 8.66 years payback period including the degradation factor. In contrast, the **S1** Old regulation Capacity Charge (5$/kW) which design the same solar capacity is 138kWp with annual production of 184448kWh with 16.23 years payback more than case S2 above also S3, S4, and S5 are applying with the new regulation are takes so long payback period. Therefore, after the recalculate the exiting solar project KPI for Toyota project the best consideration for client is to consider Case Study **Scenario S2** New regulation compensate tariff (0.047$/kWh) system which could financially profitable than the other scenarios.

4. CONCLUSIONS

This research report study of the Grid-connected photovoltaic system for the Toyota Cambodia head Quatre has been carried out and applied successfully to the clients after the recalculate the exiting solar project KPI for Toyota project the best consideration for client is to consider Case Study **S2** the LCOE of scenario S2 is 0.086$/kWh lower than the LCOE S1 which is 0.115$/kWh and LCOE scenario S3 is 0.0935$/kWh and scenario S4 is 0.111$/kWh and S5 is 0.134$/kWh and compare to EDC S0 LCOE is 0.158$/kWh. Furthermore, we can see the system of scenario S2 is much more saving cost on the capacity charge by replace with the compensate tariff. Besides that, in 20 years of project life time, the scenario S2 could saving **$ 314128.87** compare to the Electricity bill from the grid**.**

References

[1] Sevea Consulting, "giz2021_en_Sector Brief PV_GBN-Cambodia_WEB," 2020, Accessed: Jan. 21, 2025. [Online]. Available: Responsibility for the content of external websites linked in this publication always lies with their respective publishers. GIZ expressly dissociates itself from such content.

[2] "Important developments in the electricity sector of the Kingdom of Cambodia up to December2024"

[3] Homer Pro, "HOMERHelpManual," *HOMER® Pro Version 3.7*, 2016.

[4] A. Ali, M. Volatier, and M. Darnon, "Optimal Sizing and Assessment of Standalone Photovoltaic Systems for Community Health Centers in Mali," *Solar*, vol. 3, no. 3, pp. 522–543, Sep. 2023, doi: 10.3390/solar3030029.

[5] P. M. Khuong, R. McKenna, and W. Fichtner, "A cost-effective and transferable methodology for rooftop PV potential assessment in developing countries," *Energies (Basel)*, vol. 13, no. 10, May 2020, doi: 10.3390/en13102501.

[6] Chun Sing Lai and Malcolm D. McCulloch, "16 ManuscriptAppliedEnergyaccept_2".

[7] F.-E. Riakhi and A. Khaldoun, "Development of a Stand Alone Solar Carport Model Linked to Charging Stations and its Economic Analysis (Case Study)."

Food Technology and Nutrition (FTN)

3 Sub-parallel Sessions

"FTN-2: The Future of Food Production: Novel Technologies and Alternative Food Sources"

Chair: Assist. Prof. Dr. MITH Hasika, Co-Chair: Dr. PHUONG Hengsim

No.	Topic
1	Isolation and Characterization of Salmonella Bacteriophages from Chicken Breast Samples in Phnom Penh Markets as a Biocontrol Agent *Authored by: Sireyvathanak VUTH, Kongkea PHAN, Laingshun HUOY*
2	Assessment of Antimicrobial Resistance Patterns in Salmonella Isolates from Different Food Matrices in Cambodia *Authored by: Chakriya SOVANN, Navin SRENG, Hasika MITH*
3	Consumer Preferences and Development of the Sourdough Bread-making Process with Added Flavors *Authored by: Sokly CHEA, Marinich NET, Sokunthea SOM, Reasmey TAN*
4	Effects of Monoglyceride and Guar Gum on the Quality and Sensory Properties of Mango Frozen Yogurt *Authored by: Sodaneth HUOR, Pichmony EK, Sengly SROY*
5	Nutritional Quality of Farmed Striped Snakehead (C. striata): Fatty Acid Profile and Proximate Composition Analysis *Authored by: Sovannmony Lay, Caroline Douny, Marie-Louise Scippo, Vattana Mom, Hasika Mith, Chanvorleak Phat*
6	Optimisation of Method Liquid Extraction of Aroma Compound (2-acetyl-1-pyrroline) from Phka Rumdoul Rice Variety *Authored by: Sokharoath Sary, Chanvorleak Phat, Stephane Bellafiore, Sereyvath Yoeun*
7	Comparing the Physicochemical Properties of Soybean Oil Blended with Different Oils *Authored by: Monika MICH, Chanchira PEN, Manit SAY, Chin Ping TAN, Reasmey TAN*
8	Comparing the Quality of Fermented Soy Sauces Made by Using Whole Soybeans and Defatted Soybeans *Authored by: Kimmuoy SONG, Reasmey TAN, Luka LY, Parakulsuksatid PRAMUK*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Isolation and Characterization of *Salmonella* Bacteriophages from Chicken Breast Samples in Phnom Penh Markets as a Biocontrol Agent

Sireyvathanak VUTH [1], Kongkea PHAN [2], Laingshun HUOY [1,2,3*]

[1] *Department of Bioengineering, Faculty of Engineering, Royal University of Phnom Penh, Phnom Penh, Cambodia*
[2] *Faculty of Science and Technology, International University, Phnom Penh, Cambodia*
[3] *Department of Biomedical Science and Veterinary Public Health, Swedish University of Agricultural Sciences, 75007 Uppsala, Sweden*
[*] *Corresponding author: huoy.laingshun@rupp.edu.kh*

Abstract: *Bacteriophages (phages) are viruses that selectively infect bacterial cells, making them promising biocontrol agents that do not disrupt microbiota or mammalian cells. Salmonella spp., particularly Salmonella Typhimurium, is a major foodborne pathogen with antimicrobial resistance, posing significant public health risks. This study aimed to isolate and characterize lytic bacteriophages targeting S. Typhimurium from chicken breast samples. Eight phages were isolated from ten chicken samples using the double-layer agar method, producing three distinct plaque collected from three markets in Phnom Penh and assess their potential as biocontrol agents. S. Typhimurium ATCC 14028 was used as the host strain for isolation. Plaque morphologies were determined based on size and turbidity. The concentration of isolated phage stocks ranged from 0.1 × 10⁹ to 7 × 10⁹ PFU/ml. Host range analysis revealed that phages ORS02 and CAP03 had the broadest host range, infecting about 20% (6/32) of tested Salmonella strains, while other phages exhibited more specific lytic activity. In broth culture assays, four phages—SMK01, ORS02, CAP02 and CAP03—completely inhibited bacterial growth after 22 hours, as confirmed by the absence of colonies on SS agar. Notably, phages SMK01, ORS02, and CAP03 remained stable across a broad pH range (pH 3–11). When applied to artificially contaminated chicken breast samples, phage CAP03 reduced S. Typhimurium counts to below the detection limit within 24 to 48 hours at 4 °C. These findings highlight the potential of lytic phages as effective biocontrol agents against Salmonella, demonstrating stability under diverse physiological conditions and offering a sustainable alternative to antibiotics for food safety improvement.*

Keywords: lytic bacteriophage; biocontrol; *Salmonella Typhimurium*; antimicrobial resistance; host specificity

1. INTRODUCTION

Non-typhoidal *Salmonella enterica* is a major foodborne pathogen responsible for substantial global morbidity and mortality. According to the World Health Organization (WHO), non-typhoidal *Salmonella* infections caused approximately 93.8 million cases and 155,000 deaths worldwide in 2016. Notably, around 85% of these cases were attributed to exposure through contaminated food, highlighting the critical importance of food safety in preventing disease outbreaks [1].

Traditional approaches to controlling foodborne pathogens—such as the use of chemical preservatives and physical treatments—can induce undesirable alterations in food [2,3]. These include chemical and physical changes that negatively affect sensory qualities and may reduce the content or bioavailability of important nutrients. Furthermore, many preservatives have been associated with adverse effects on human health and environmental pollution. Documented risks include skin sensitivity, disruption of endocrine function, menstrual irregularities, tissue damage in the liver, kidneys, and reproductive organs, and impacts on oxidative stress, hormone balance, and body weight regulation[2].

Antibiotic use is another common intervention for controlling pathogenic microbes in the food chain. However, the widespread misuse and overuse of antibiotics—particularly in low- and middle-income countries—has become a pressing public health issue, accelerating the emergence of antibiotic-resistant bacteria. This resistance poses risks not only to

[1*] Corresponding author: Laingshun HUOY
E-mail: huoy.laingshun@rupp.edu.kh; Tel: +855-12 437 621

individual health but also to national economies, food security, and overall public health systems [4]

As concerns grow over the drawbacks of conventional methods, biocontrol strategies have emerged as promising alternatives. These include the use of natural antagonists such as probiotics, bacteriocins, competitive exclusion cultures, and bacteriophages to manage microbial contamination in food [5]. Among these agents, bacteriophage therapy is regarded as one of the most effective and sustainable options. Bacteriophages are viruses that specifically infect and lyse bacteria without harming mammalian cells or beneficial microbiota. Their advantages include high specificity, rapid bactericidal activity, environmental compatibility, and the ability to self-replicate at the infection site [6].

This study focuses on isolating lytic bacteriophages from chicken breast samples collected from markets in Phnom Penh and evaluating their in-vitro efficacy against *Salmonella enterica* serovar Typhimurium ATCC 14028 (hereafter referred to as ATCC 14028). It aims to provide baseline data and insights into the potential of phage therapy for food safety applications in Cambodia. The findings are expected to support future research, regulation, and the development of locally relevant biocontrol strategies, contributing to food safety innovation and efforts to combat antimicrobial resistance.

2. METHODOLOGY

2.1 Sampling sites

Ten boneless chicken breast samples were randomly collected from three Phnom Penh wet markets—Orussey (n=4), Chbar Ampov (n=4), and Samaki (n=2)—to represent diverse city areas. Samples were placed in sterile bags, stored on ice, and transported to the Microbiology Laboratory at the Faculty of Science and Technology, International University for analysis. Sampling locations are shown in Fig. 1.

2.2 Bacteriophage isolation

Bacteriophages were isolated by enriching 15 g of each sample with 135 mL of SM buffer (2 g $MgSO_4 \cdot 7H_2O$, 5.8 g NaCl, 50 mL of 1 M Tris-HCl) and 15 mL of log-phase ATCC 14028 culture. The mixture was incubated at 37 ± 1 °C for 24 ± 2 h, then centrifuged at 5,000 rpm for 20 min at 25 °C. The supernatant was filtered (0.22 µm syringe filter, Membrane Solutions, USA), and phage presence was assessed using a modified agar overlay method [7]. Briefly, 100 µL of filtrate and 100 µL of log-phase host were mixed with 0.8% TSA top agar and poured over 1.5% TSA base agar. Plates were swirled, solidified, and incubated at 37 ± 1 °C for 24 ± 2 h. All assays were done in triplicate.

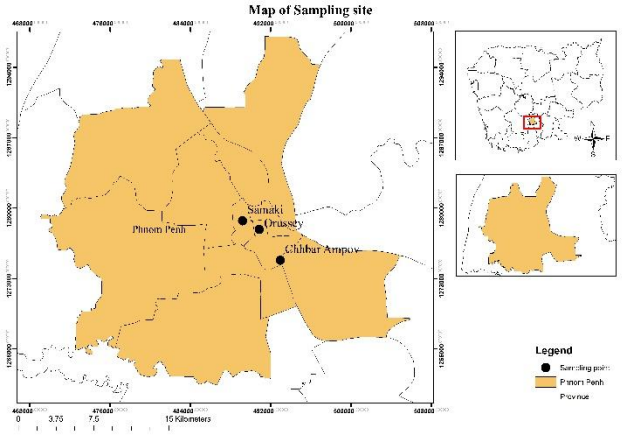


Fig. 1. Map of sampling sites

2.3 Titer d termination and stock preparation

Plaque-positive samples were selected for specificity testing, titer determination, and lysate stock preparation. Non-plaque samples were excluded. Three uniform plaques were picked using sterile pipette tips, suspended in 1 mL SM buffer, and purified via triple rounds of the double-layer agar method. Purified lysates were serially diluted and PFU/mL was determined using the overlay technique. For lysate stock, 10 mL of SM buffer was added to phage plates, incubated at room temperature for 60 min with gentle swirling, then centrifuged (5,000 rpm, 20 min), filtered (0.22 µm), and stored at 4 °C [8].

2.4 Host specificity test

The host range of the isolated bacteriophages was determined using a spot assay, following the protocol described by Petsong et al. [9]. Bacterial lawns were prepared in advance using the test strains. Subsequently, 10 µL of phage lysate (10^8 PFU/mL) was carefully spotted onto each bacterial lawn. The plates were then incubated at 37 ± 1 °C for 24 ± 2 hours. Zones of inhibition were observed and recorded, indicating successful lytic activity and host susceptibility.

2.5 Minimum inhibitory concentration test

This assay determined the minimum effective concentration of bacteriophage against ATCC 14028. Log-phase ATCC 14028 (10^5 CFU/mL) was mixed with 200–1400 µL of phage lysate (10^8 PFU/mL) and incubated at 37 ± 1 °C for 24 ± 2 h. Post-incubation, turbidity was visually assessed to indicate bacterial inhibition. Cultures were centrifuged (5,000 rpm, 20 min), pellets resuspended in 0.85% NaCl, and 10 µL plated on SS agar. This ensured recovery of bacteria only, minimizing phage interference [10].

2.6 Bacteriophage pH stability test

The pH stability of the bacteriophage was evaluated by mixing 100 µL of phage lysate (10^8 PFU/mL) with 900 µL of Buffered Peptone Water (BPW), previously adjusted to pH values of 3, 5, 7, 9, and 11 using 1N NaOH or 1N HCl, following the method described by Shende et al. [7]. The mixtures were incubated at 37 ± 1 °C for 2 hours. After incubation, the surviving phages were enumerated using the double-layer agar method. The plates were further incubated at 37 ± 1 °C for 24 ± 2 hours to assess phage viability across different pH conditions.

2.7 Virulence ability of bacteriophage on chicken breast

This experiment assessed the efficacy of bacteriophages in inhibiting ATCC 14028 on chicken stored at 4 °C. A 300 g market-sourced chicken sample was cut into 5×5 cm pieces and surface-sterilized with UV light (250 nm, 20 min). Pieces were treated with either sterile saline (0.85%) or inoculated with 1 mL of log-phase *S. Typhimurium* (10^6 CFU/mL). After 1 h of bacterial adherence at room temperature, 200 µL of phage lysate (10^8 PFU/mL, MOI 200) was applied, and samples were stored at 4 °C. At 0, 24 and 48 h, one piece was processed: suspended in saline, centrifuged, resuspended, and plated on SS agar for *Salmonella* enumeration.

2.8 Data analysis

All experimental data were recorded and organized using Microsoft Excel (Windows version 2019). Parameters including bacteriophage prevalence, phage concentration, bacterial concentration, bacterial reduction values, and percentage reduction were calculated within Excel. For statistical analysis, JASP software (version 0.18.1.0) was employed. A paired sample t-test was conducted to compare bacterial counts before and after phage treatment, with a significance level set at $p < 0.05$.

3. RESULTS AND DISCUSSION

3.1 Isolation, prevalence, and preliminary characterization of Salmonella-specific bacteriophages

A total of eight lytic bacteriophages targeting ATCC 14028 were successfully isolated, purified, and stocked from chicken breast samples collected from three wet markets in Phnom Penh (Table 1), yielding a phage recovery rate of 80%. By location, the phage-positive rate in samples from Orussey, Chba Ampov, and Samaky markets was 100%, 75%, and 50%, respectively. Following a minimum of three rounds of specific selection, phages were stored in SM buffer as stock, with titers ranging from 0.1 × 10^9 to 7 × 10^9 PFU/mL.

Table 1. Prevalence of Bacteriophage infecting *Salmonella* Typhimurium ATCC 14028 from three market in Phnom Penh

Market	Sample Code	Plaque Forming	Concentration	Prevelence
Chba Ampov	CAP01	(-)	N/A	75%
	CAP02	(+)	3 x 10^8	
	CAP03	(+)	1 x 10^8	
	CAP04	(+)	7 x 10^9	
Orreusey	ORS01	(+)	1 x 10^8	100%
	ORS02	(+)	1 x 10^9	
	ORS03	(+)	1 x 10^{10}	
	ORS04	(+)	3 x 10^9	
Samaky	SMK01	(+)	1 x 10^8	50%
	SMK02	(-)	N/A	

To our knowledge, this is the first study in Cambodia to isolate *Salmonella*-specific lytic bacteriophages from chicken meat and evaluate their potential as biocontrol agents. The 80% recovery rate observed here is notably higher than reports from Malaysia (18.2%) [11] and Japan (43%) [12]. Interestingly, chicken meat samples showed richer phage presence than commonly reported sources such as sewage, animal waste, or soil [13]. However, the observed prevalence must be interpreted cautiously due to the limited sample size (n = 10).

On the lawn of ATCC 14028, as shown in Fig. 2, three distinct bacteriophage plaque morphotypes were observed. The first type consisted of clear plaques with diameters ranging from 0.7 to 1 mm, exhibited by phages ORS01, ORS02, ORS03, ORS04, SMK01, and CAP02. The second type displayed plaques with clear centers and turbid edges, also measuring 0.7 to 1 mm in diameter, as seen with ORS03 and CAP03. The third morphotype consisted of larger clear plaques, measuring 1.5 to 2 mm in diameter, exemplified by CAP04.

Several factors influence plaque formation, including adsorption rate, latent period, and virion morphology [14]. Additionally, external conditions such as gel concentration and the presence of divalent cations (e.g., Ca^{2+}, Mg^{2+}) may affect plaque appearance [15]. However, some studies suggest that there is no consistent relationship between phage family and plaque morphology [16]. It is also important to note that plaque size does not necessarily correlate with phage lytic efficiency, as no definitive experimental evidence currently supports this association [15].

For more precise phage characterization, transmission electron microscopy (TEM) should be employed to visualize virion structure and classify phage families based on morphological features. Additionally, whole genome

sequencing is essential to identify genetic traits, confirm taxonomy, and evaluate therapeutic or biocontrol potential.

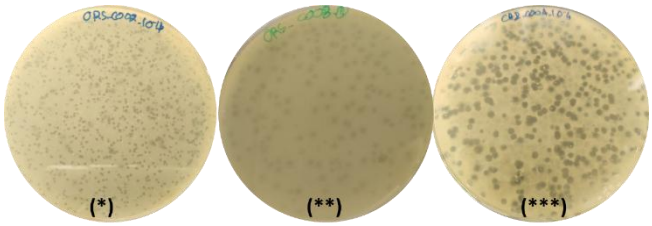


Fig. 2. Morphological characteristics of bacteriophage plaques observed on *S. Typhimurium* ATCC 14028.
(*) Clear plaques with diameters of 0.7–1 mm.
(**) Plaques with clear centers and turbid edges, 0.7–1 mm in diameter.
(***) Clear plaques with larger diameters of 1.5–2 mm.

3.2 Host range assessment reveals limited lytic activity of isolated phages

Eight phages were tested against 28 *Salmonella* strains, including two reference *S. Typhimurium* strains, one *E. coli*, and 25 food isolates. ORS02 and CAP03 showed the widest host range, lysing 6 strains each (21.4%), followed by SMK01 (5 strains, 17.9%). All phages lysed both reference strains (Fig. 3).

These results suggest that the isolated phages have a limited host range, with the broadest (ORS02 and CAP03). This is consistent with earlier findings on monovalent *Salmonella*-specific phages [15], although both broad-host-range monovalent and polyvalent phages have been reported [7]. It's important to note that 25 of the tested strains remain unidentified beyond genus level, which may account for variations in host range due to differences in phage receptors such as pili, LPS, or outer membrane structures [17].

3.3 High efficacy of selected bacteriophages against Salmonella at minimal MOI

The minimum effective concentration—phages ORS02, CAP02, CAP03, CAP04, and SMK01 successfully inhibited the growth of *Salmonella* to below the detection limit. In contrast, phage ORS04 produced a limpid broth only at the MOI of 600, while phages ORS01 and ORS03 resulted in slightly less turbid broth at the highest tested MOI of 1400 (Table 2). The effective MOI values identified from this assay were subsequently applied to treat chicken breast samples experimentally contaminated with *Salmonella* under in vivo conditions.

These findings are consistent with the phage lytic cycle theory and corroborate previous studies, which have reported that bacteriophages can significantly reduce bacterial concentrations following enrichment in broth media. Reported reductions in the study range around a 6-log cycle or reduction to below detection limits [18]. While the current study did not investigate phage burst size, infection mode, or growth kinetics, the results nonetheless support the potential of bacteriophages as an effective intervention for the reduction and control of pathogenic *Salmonella* in food products.

Table 2.MIC test of phages on host *Salmonella* Typhimurium ATCC 14028.

Phage	Volume (µl) of 10^8 PFU/mL						
	200	400	600	800	1000	1200	1400
ORS01	-	-	-	-	-	-	±
ORS02	-	+	+	+	+	+	+
ORS03	-	-	-	-	-	-	±
ORS04	-	-	-	-	-	+	+
CAP02	+	+	+	+	+	+	+
CAP03	+	+	+	+	+	+	+
CAP04	+	+	+	+	+	+	+
SMK01	+	+	+	+	+	+	+

Lytic activity: – no lysis (turbid), ± low lysis (slightly turbid), + strong lysis (clear broth culture)

Fig. 3. A heatmap illustrating the host range of each phage. The X-and Y-axes represent phages and host strains, respectively.



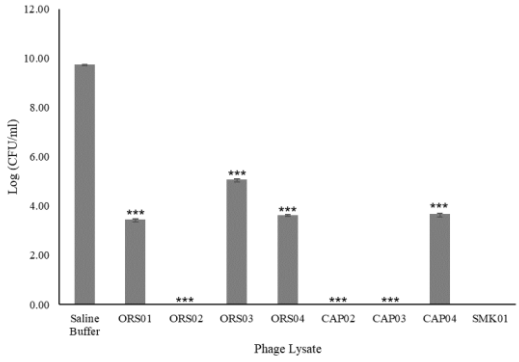


Fig. 4. Phage Inhibition Efficiency Against *S. Typhimurium* ATCC 14028 at MOI 200. (***, $p < 0.001$) indicate a statistically high significant difference.

3.4 pH stability assessment of selected bacteriophages

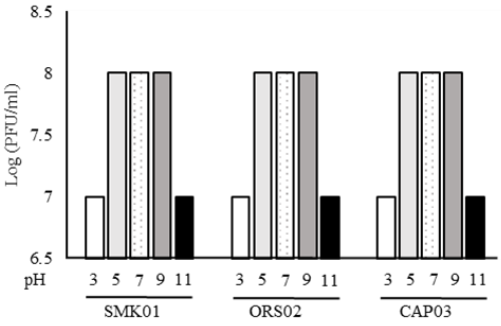


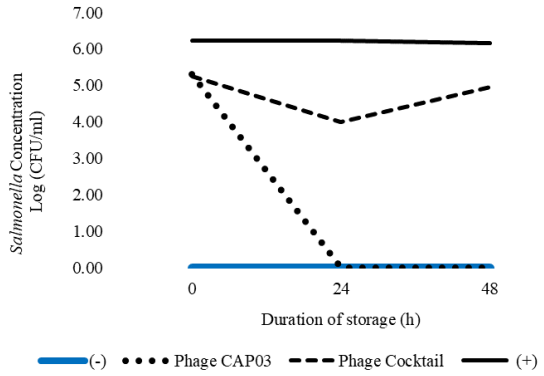
Fig. 5. Effect of pH on phage stability. Each sample and pH condition were tested once. Y-axis represents phage titer, and X-axis represents pH value and phage strain.

Three bacteriophages—SMK01, ORS02, and CAP03—were selected for the pH tolerance test and subsequent food application assays based on their broad host range and demonstrated in vitro efficacy in reducing *Salmonella* populations. These phages exhibited stability across a pH range of 5 to 9, maintaining a concentration of approximately 10^8 PFU/mL. However, under more extreme conditions (pH 3 and pH 11), the phage concentration decreased slightly to 10^7 PFU/mL. The results, presented in Fig. 4, indicate that the selected phages are tolerant to both mildly acidic and basic environments. These findings are consistent with previous studies that have reported similar pH resilience in bacteriophages [8,13].

3.5 Comparative analysis of single phage and cocktail treatments for Salmonella biocontrol under artificial contamination conditions

A phage cocktail was formulated by mixing SMK01, ORS02, and CAP03 at a 1:1:1 ratio. At 2 hours post-treatment, both CAP03 and the cocktail showed a ~15% or 1-log CFU/g reduction in bacterial concentration (Fig. 5). However, CAP03 alone demonstrated the highest inhibitory effect, achieving 100% reduction or a 6.30-log CFU/g drop on days 1 and 2 post-treatment—reducing *S. Typhimurium* to below detectable levels. This aligns with previous studies and supports the lytic phage mechanism on food sample [9]. For the cocktail, a 35% reduction was observed on day 1, but effectiveness declined over the following days, with log reductions of 1.20 at day 2, indicating lower efficacy than CAP03 alone. The limited performance may be due to incomplete characterization of the cocktail phages. Although their lysis profiles and inhibition capacities were evaluated, further molecular and microscopic analyses are needed to design a more effective, polyvalent phage cocktail targeting multiple *Salmonella* strains and minimizing resistance.

Fig. 6. *Inhibition of S. Typhimurium ATCC 14028 on chicken breast by phage CAP03 and a phage cocktail after 48 hours at 4 °C.*



3. CONCLUSION

This study presents key insights into the potential of bacteriophages as natural biocontrol agents for *Salmonella* contamination in food. The findings reveal that: (i) bacteriophages are highly prevalent in chicken meat samples, with an 80% occurrence rate—likely due to poor handling and storage practices in wet markets; (ii) the eight isolated phages display three distinct plaque morphologies and varying host range profiles, indicating their diversity and potential for further classification; (iii) selected phages—SMK01, ORS02, and CAP03—exhibit strong host specificity and pH tolerance, with minimal titer reduction observed only under extreme pH conditions (pH 3 and 11); and (iv) CAP-CP003 demonstrates potent biocontrol efficacy, reducing *S. Typhimurium* levels to below the detection limit within 24 to 48 hours post-treatment at 4 °C, at the MOI of 200.

Collectively, these results underscore the promising role of bacteriophages in enhancing food safety by naturally reducing *Salmonella* concentration during storage. Although the scope of this study is limited, it establishes a foundation for phage research in Cambodia and highlights critical factors for

consideration in future work. Continued research, particularly at the molecular level, is essential to confirm the strictly lytic nature of these phages and to rule out the potential for horizontal gene transfer—ensuring their safe application in food biocontrol strategies.

4. ACKNOWLEDGMENT

The authors gratefully acknowledge the Department of Bioengineering, Faculty of Engineering, Royal University of Phnom Penh, for providing research materials, and the Faculty of Science and Technology, International University, for access to laboratory facilities.

REFERENCES

[1] WHO, "Salmonella (non-typhoidal)," 2018. https://www.who.int/news-room/fact-sheets/detail/salmonella-(non-typhoidal).

[2] T. Mohammed Kamel, M. Bakr, and Q. Ahmed, "Side Effects of Preservatives on Human Life," *Sci. Res. J.*, vol. 2, no. October, pp. 1–14, 2022.

[3] L. Petruzzi, D. Campaniello, B. Speranza, M. R. Corbo, M. Sinigaglia, and A. Bevilacqua, "Thermal Treatments for Fruit and Vegetable Juices and Beverages: A Literature Overview," *Compr. Rev. Food Sci. Food Saf.*, vol. 16, no. 4, pp. 668–691, 2017, doi: 10.1111/1541-4337.12270.

[4] M. Farrukh, A. Munawar, Z. Nawaz, N. Hussain, A. Bin Hafeez, and P. Szweda, "Antibiotic resistance and preventive strategies in foodborne pathogenic bacteria: a comprehensive review," *Food Sci. Biotechnol.*, no. 0123456789, 2025, doi: 10.1007/s10068-024-01767-x.

[5] J. Delgado, M. Álvarez, E. Cebrián, I. Martín, E. Roncero, and M. Rodríguez, "Biocontrol of Pathogen Microorganisms in Ripened Foods of Animal Origin," *Microorganisms*, vol. 11, no. 6, 2023, doi: 10.3390/microorganisms11061578.

[6] S. A. Ranveer *et al.*, "Positive and negative aspects of bacteriophages and their immense role in the food chain," *npj Sci. Food*, vol. 8, no. 1, 2024, doi: 10.1038/s41538-023-00245-8.

[7] R. K. Shende, S. D. Hirpurkar, C. Sannat, N. Rawat, and V. Pandey, "Isolation and characterization of bacteriophages with lytic activity against common bacterial pathogens," *Vet. World*, vol. 10, no. 8, pp. 973–978, 2017, doi: 10.14202/vetworld.2017.973-978.

[8] T. Y. Thung *et al.*, "Evaluation of a lytic bacteriophage for bio-control of Salmonella Typhimurium in different food matrices," *Lwt*, vol. 105, no. November 2018, pp. 211–214, 2019, doi: 10.1016/j.lwt.2019.02.033.

[9] K. Petsong, S. Benjakul, S. Chaturongakul, A. I. M. Switt, and K. Vongkamjan, "Lysis profiles of salmonella phages on salmonella isolates from various sources and efficiency of a phage cocktail against s. Enteritidis and s. typhimurium," *Microorganisms*, vol. 7, no. 4, 2019, doi: 10.3390/microorganisms7040100.

[10] C. Lukman, C. Yonathan, S. Magdalena, and D. E. Waturangi, "Isolation and characterization of pathogenic Escherichia coli bacteriophages from chicken and beef offal," *BMC Res. Notes*, vol. 13, no. 1, pp. 1–7, 2020, doi: 10.1186/s13104-019-4859-y.

[11] T. Y. Thung *et al.*, "Use of a lytic bacteriophage to control Salmonella Enteritidis in retail food," *LWT - Food Sci. Technol.*, vol. 78, pp. 222–225, 2017, doi: 10.1016/j.lwt.2016.12.044.

[12] H. M. Duc, H. M. Son, K. ichi Honjoh, and T. Miyamoto, "Isolation and application of bacteriophages to reduce Salmonella contamination in raw chicken meat," *LWT - Food Sci. Technol.*, vol. 91, no. January, pp. 353–360, 2018, doi: 10.1016/j.lwt.2018.01.072.

[13] C. Huang *et al.*, "Isolation, characterization, and application of Bacteriophage LPSE1 against Salmonella enterica in Ready to Eat (RTE) Foods," *Front. Microbiol.*, vol. 9, no. MAY, pp. 1–11, 2018, doi: 10.3389/fmicb.2018.01046.

[14] R. Gallet, S. Kannoly, and I. N. Wang, "Effects of bacteriophage traits on plaque formation," *BMC Microbiol.*, vol. 11, 2011, doi: 10.1186/1471-2180-11-181.

[15] A. Ghasemian, M. Bavand, and Z. Moradpour, "A broad-host range coliphage against a clinically isolated E. coli O157: Isolation and characterization," *J. Appl. Pharm. Sci.*, vol. 7, no. 3, pp. 123–128, 2017, doi: 10.7324/JAPS.2017.70320.

[16] and M. M. Lorena Rodríguez-Rubio, Pedro Blanco-Picazo, "Chapter 7: Are Phages Parasites or Symbionts of Bacteria?," in *Biocommunication of Phages*, Springer International Publishing Switzerland, 2020, pp. 143–162.

[17] C. Kocharunchitt, T. Ross, and D. L. McNeil, "Use of bacteriophages as biocontrol agents to control Salmonella associated with seed sprouts," *Int. J. Food Microbiol.*, vol. 128, no. 3, pp. 453–459, 2009, doi: 10.1016/j.ijfoodmicro.2008.10.014.

[18] N. K. El-Dougdoug *et al.*, "Control of Salmonella Newport on cherry tomato using a cocktail of lytic bacteriophages," *Int. J. Food Microbiol.*, vol. 293, no. September 2018, pp. 60–71, 2019, doi: 10.1016/j.ijfoodmicro.2019.01.003

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessment of Antimicrobial Resistance Patterns in *Salmonella* Isolates from Different Food Matrices in Cambodia

Chakriya SOVANN [1,2], Navin SRENG [2], Hasika MITH [1,3*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Laboratory of Environment and Food Safety, Institut Pasteur du Cambodge, 5 Preah Monivong Blvd., P.O. Box 983, Phnom Penh, Cambodia*

[3] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: hasika@itc.edu.kh*

Abstract: *Salmonellosis is one of the most common foodborne diseases and it is increasingly occurring worldwide including Cambodia. Since antibiotics are essential for treating illnesses, the emergence of antimicrobial resistance in Salmonella-contaminated food poses a major public health risk. The current study was therefore conducted to examine the occurrence of Salmonella in various food matrices and to determine the antimicrobial resistance profiles of the Salmonella isolates. From January to December 2023, a total of food samples (n=579) were submitted to the Laboratory of Environment and Food Safety for Salmonella testing using the standard method NF EN ISO 6579-1:2017. All of the samples were categorized into four sample types, including raw egg (n=2), fish and fish products (n=80), ready-to-eat foods (n=125), and cooked foods (n=372). The isolated Salmonella were subjected to antimicrobial susceptibility testing by using the disc diffusion method against 22 antibiotics from eight different antibiotic families following the European Committee on Antimicrobial Susceptibility Testing guidelines. Microbiological testing revealed that 11/579 food samples (1.9%) were positive for Salmonella. All positive Salmonella are originated from fish and fishery products (100%). Antimicrobial susceptibility assay revealed that 7/11 (63.6%) of the Salmonella isolates were fully susceptible to all 22 antibiotics tested. 27.3% of the isolates were resistant to cotrimoxazole, piperacillin, amoxicillin-clavulanic acid, ticarcillin, and amoxicillin. Meanwhile, 18.2% were resistant to levofloxacin, ciprofloxacin, and nalidixic acid. Multidrug resistance (MDR) was seen in 2/11 (18.2%) and 1/11 (9.1%) were ESBL-producing Salmonella of the isolates. Although this study found Salmonella at a low frequency, the presence of this bacterium and antibiotic resistance in fish and fish products is a warning sign. As a recommendation, the use of antibiotics in fish farming should be closely monitored to reduce antibiotic resistance to foodborne pathogens and to protect consumers.*

Keywords: *Salmonella*; food; fish and fish products; antibiotics resistance; Cambodia

1. INTRODUCTION[1]

Salmonellosis is a foodborne illness caused by *Salmonella* bacteria that is transmitted through the consumption of contaminated food. It frequently manifests as symptoms such as diarrhea, nausea, vomiting, stomach pain, and fever. According to the Centers for Disease Control and Prevention (CDC), foodborne *Salmonella* infections cause approximately 1.35 million cases that occur in the US annually, resulting in 26,500 hospitalizations and 420 deaths [1]. *Salmonella* is a Gram-negative, rod-shaped bacterium belonging to the Enterobacteriaceae family, with two main species *Salmonella bongori* and *Salmonella enterica* [2]. Among these two species, *Salmonella enterica* contributes significantly to the burden of disease worldwide [3].

Antimicrobial resistance (AMR) is a major threat to global health concerns causing an estimated 1.27 million deaths and contributing to 4.95 million deaths worldwide [4]. The misuse and overuse of antibiotics in human medicine, veterinary practices, and agriculture have accelerated the emergence and spread of drug-resistant bacteria, including *Salmonella*. Particularly, the use of antibiotic growth promotants (AGPs) in food animal production has been linked

[1]* Corresponding author: Hasika Mith
E-mail: hasika@itc.edu.kh; Tel: +855-17 887 472

to the rise of multidrug-resistant (MDR) *Salmonella* strains [5]. A major contributing factor to the emergence of antibiotic-resistant *Salmonella* is the subtherapeutic or preventative use of antibiotics in food animals. This practice encourages the development of antimicrobial-resistant strains on farms, significantly increasing the risk to human health from consuming contaminated animal products [6]. Additionally, MDR *Salmonella* has been found in vegetables, likely due to contamination from irrigation water [7].

Over the past two decades, Extended Spectrum Beta-lactamase (ESBL)-producing *S. enterica* strains have emerged among animals and animal food products. ESBL are enzymes that enable bacteria to resist the effects of penicillins, first-, second-, and third-generation cephalosporins (3GC), and aztreonam (excluding cephamycins and carbapenems) by hydrolysis these antibiotics. However, their activity can be inhibited by beta-lactamase inhibitors like clavulanic acid [8]. These ESBL-producing strains are concerning due to their resistance to most beta-lactam antibiotics, including cephalosporins, which are commonly used to treat severe salmonellosis [9]. The use of antibiotics in the production of food animals is spreading and this practice is known to encourage antibiotic-resistant *S. enterica*, which can infect humans through contaminated foods [10]. However, this practice is not currently monitored for antibiotic use in livestock farming [11]. Previous work in Cambodia has documented a high prevalence of *S. enterica* contamination (80%) among retail meats sold in Phnom Penh markets [12]. Another study in Cambodia found ESBL-producing *Salmonella* in fish (17%), pork (25%), and chicken (3%) [13]. However, the prevalence of antimicrobial-resistant *Salmonella* has not been investigated on different food matrices specifically. Thus, this study aimed to determine the occurrence of *Salmonella* and its antibiotic resistance profile from various food matrices in Cambodia.

2. METHODOLOGY

2.1 Food sample collection

During the period from January to December 2023, a total of 579 different food samples were obtained from different clients such as hotels, food catering, food factories, and public and private companies in Cambodia. All of the samples were categorized into four sample types, including raw eggs (n=2), fish and fish products (n=80), ready-to-eat foods (n=125), and cooked foods (n=372). Approximately 50 g of each sample was collected in a sterile plastic pouch and transported to the Laboratory of Environment and Food Safety (LEFS), Institut Pasteur du Cambodge (IPC). Perishable products were refrigerated at 2-8 °C during the transport and the analyses were performed within 24 hours after sampling.

2.2 Salmonella detection

The samples were analyzed using the standard method, ISO 6579-1:2017. The amount of 25 g of sample was first pre-enriched in 225 mL of buffered peptone water (BPW), then incubated at 37 °C for 18 hours. After the incubation, 0.1ml and 1ml were inoculated into 10ml of Rappaport-Vassiliadis medium (RVS) and 10 ml of Muller-Kauffmann tetrathionate-novobiocin broth (MKTTn), respectively. Then RVS broth was incubated at 41.5 °C for 24 h and the MKTTn broth at 37 °C for 24 h. Each culture from selective enrichment broth was plated out on two selective solid media of Xylose Lysine Deoxycholate agar (XLD) and RAPID' *Salmonella* chromogenic agar (R'SAL). Then incubated at 37 °C for 24 h. Lactose-negative colonies on XLD and magenta colonies on R'SAL were selected for confirmation by using MALDI Biotyper®. All positive *Salmonella* strains were stored in skim milk medium with glycerol at -20 °C for further studies.

2.3 Antimicrobial susceptibility testing

The isolated *Salmonella* were subjected to antimicrobial susceptibility testing by using the disc diffusion method following the European Committee on Antimicrobial Susceptibility Testing guidelines (EUCAST) [14]. The twenty-two commercial antibiotic discs from eight different classes were tested such as (1) Penicillins: amoxicillin (AMO; 20µg), amoxicillin-clavulanic acid (AMC; 20/10µg), piperacillin (PIL; 30µg), piperacillin-tazobactam (PTZ; 30/6µg), ticarcillin (TIC; 75µg), ticarcillin-clavulanic acid (TCC; 75/10µg), (2) Cephalosporins: cefepime (FEP; 30µg), cefotaxime (COX; 5µg), cefoxitin (FOX; 30µg), ceftazidime (CZD; 10µg), (3) Carbapenems: ertapenem (ETP; 10µg), imipenem (IPM; 10µg), (4) Monobactams: aztreonam (ATM; 30µg), (5) Fluoroquinolones: ciprofloxacin (CIP; 5µg), levofloxacin (LVX; 5µg), nalidixic acid (NAL; 30µg), (6) aminoglycosides: amikacin (AKN; 30µg), gentamicin (GMN; 10µg), tobramycin (TMN; 10µg), (7) Miscellaneous: fosfomycin (FOS; 200µg), cotrimoxazol (SXT; 1.25/23.75µg), (8) Tetracyclines: tigecycline (TGC; 15µg). *E. coli* ATCC 25922 was used as the quality control. Diameter interpretations were based on 2024 EUCAST recommendations. The strains were considered multi-drug resistant (MDR) if they demonstrated resistance to three or more classes of antibiotics (at least one antibiotic of each class) [15].

3. RESULTS AND DISCUSSION

3.1 Prevalence of Salmonella in food matrices

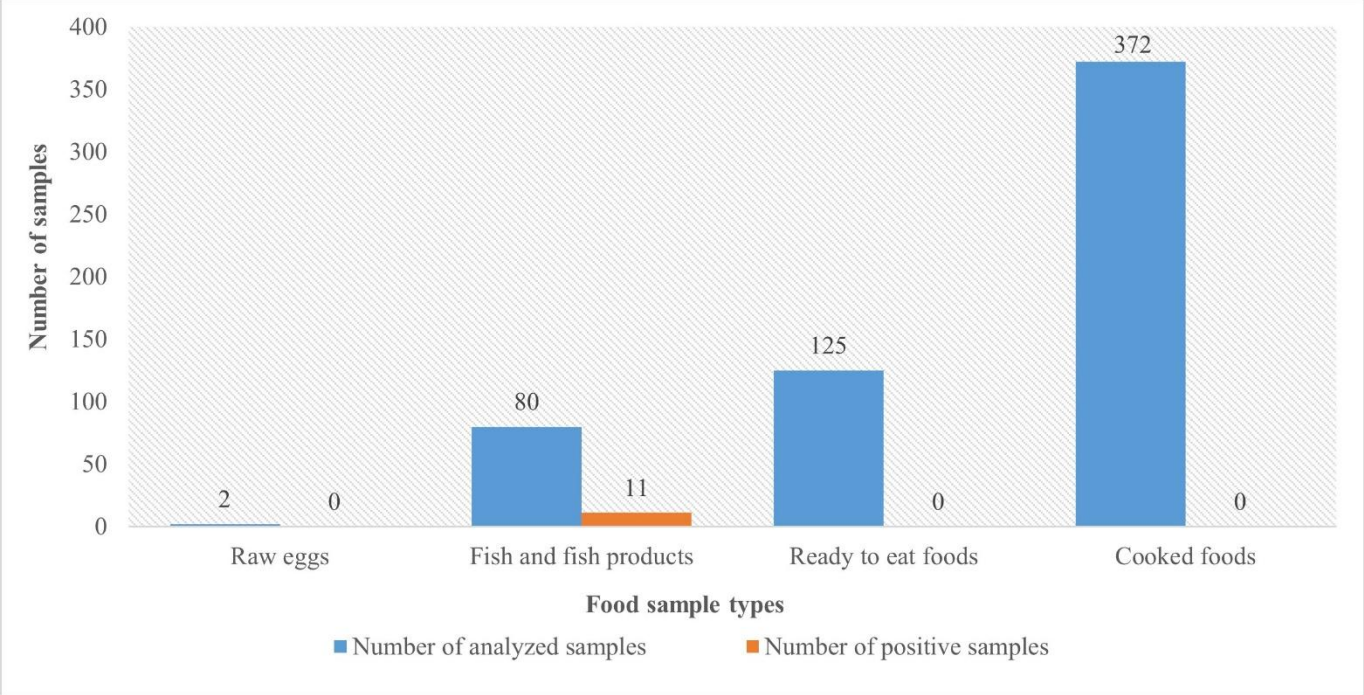


Fig. 1. Prevalence of *Salmonella* in the analyzed food samples from January to December 2023

Out of five hundred and seventy-nine examined at LEFS from January to December 2023, 11 (1.9%) samples were found positive for *Salmonella* as shown in Figure 1. Among the four food sample types, only fish and fish products were found positive for *Salmonella* at 13.8% (11/80). No *Salmonella* was detected in raw eggs, cooked foods, and ready to eat food. This overall prevalence is considerably lower than those reported in previous Cambodian studies. For instance, a 2024 study analyzing 285 food samples from Phnom Penh markets found *Salmonella* in 48.4% of samples, including 64% of seafood samples [16]. Similarly, a systematic review covering data from 2000 to 2022 reported a 3.5% prevalence of *Salmonella* in fermented fish products like prahok [17]. The lower prevalence observed in our study may reflect improved hygiene practices, differences in sampling methods, or seasonal variations. On the other hand, a previous study mentioned that *Salmonella* spp. is a pathogen responsible for severe foodborne infections that can be introduced into the fish production chain through inadequate handling or hygiene or contact with contaminated water [18]. In comparison to another study conducted in Cambodia, our findings are also lower, as that study reported 17% of fish samples (10 out of 60) were contaminated with ESBL-producing *S. enterica* [13]. The absence of *Salmonella* in other food sample types, such as raw eggs, cooked foods, and ready to eat foods, is a positive sign. Nonetheless, these results should be interpreted with caution, as the occurrence of *Salmonella* can be unpredictable and influenced by factors like seasonal variation and the food's geographic source. Continuous monitoring and stringent food safety practices are key to reducing the risk of contamination in these food categories.

3.2 Antimicrobial resistance patterns of Salmonella isolates

Antimicrobial susceptibility assay revealed that 63.6% (7/11) of the *Salmonella* isolates investigated were fully susceptible to all tested twenty-two antibiotics while 36.4% (4/11) were resistant to various antibiotics. This indicates that the overall resistance levels within these samples are relatively low. However, a significant number of isolates showed resistance to multiple antibiotics that are critical for clinical use. In this study, the most common resistance observed was to cotrimoxazole, piperacillin, amoxicillin-clavulanic acid, ticarcillin, and amoxicillin at 27.3% of the isolates, followed by 18.2% resistant to levofloxacin, ciprofloxacin, and nalidixic acid. Although the information regarding to antimicrobial resistance of *Salmonella* isolated from fish and fish products was limited in Cambodia, these findings raise public health concerns, especially given the critical role of fluoroquinolones and β-lactams in the treatment of invasive salmonellosis. Previous study in Cambodia on retail meats sold in Phnom Penh markets showed high levels of multiple antimicrobial resistance of isolated *Salmonella*, ranging from 8.3% to 90.0%, to various antibiotics including amoxicillin, nalidixic acid, sulfonamide, and tetracycline [12]. Moreover, the study revealed that in Malaysia, Thailand, and Vietnam, *Salmonella* isolates from food, animals, and humans show high resistance to commonly used antibiotics, with resistance rates ranging from 23-100% for ampicillin, tetracycline, sulfamethoxazole, and streptomycin [19]. Additionally, it was also evident from the results that 18.2% (2/11) of the isolates showed multidrug

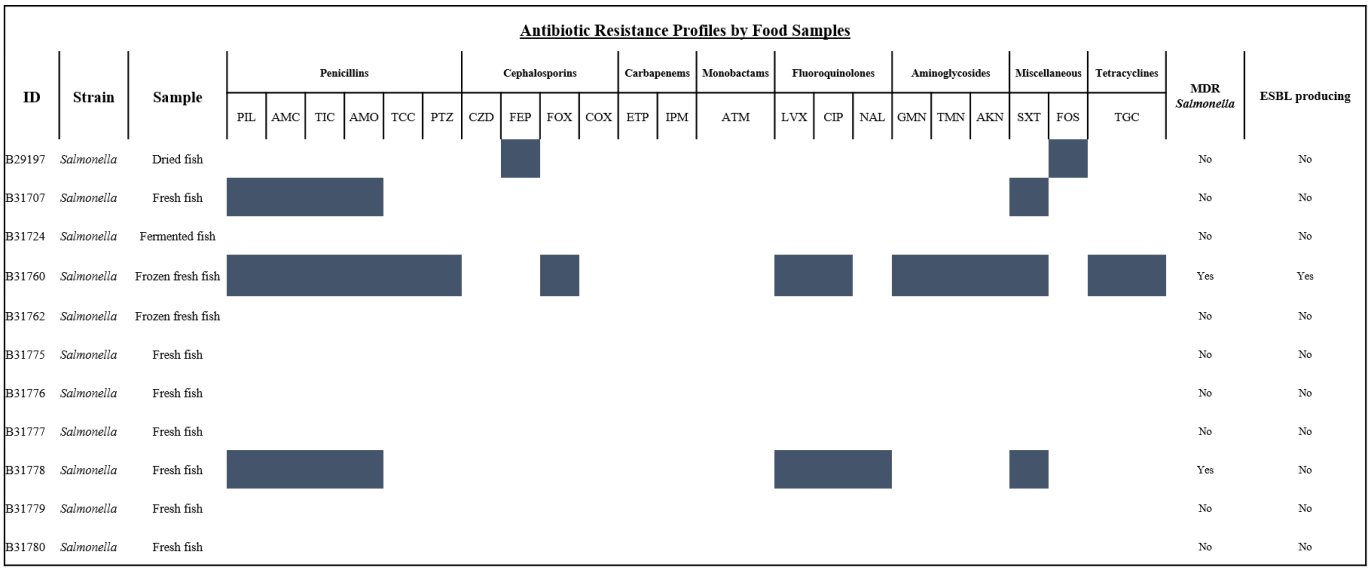


Fig. 2. Antibiotic-resistant *Salmonella* recovered from fish and fishery products. Black squares denote the presence of antibiotic resistance.

resistance, which was resistant to at least three classes of antibiotics. The emergence of these MDR strains represents a critical public health challenge because they can limit effective treatments and cause treatment failures in both human and veterinary medicine. One strain was suspected to be ESBL-producing *Salmonella* at 9.1% (1/11) when the inhibition zone around any cephalosporin increased toward the central disk with amoxicillin-clavulanic acid as shown in Figure 2. However, the genotypic confirmation of this ESBL-producing *Salmonella* is required. These findings suggest a comparatively lower prevalence of ESBL-producing *Salmonella* than a report from Cambodia, which identified a high level of ESBL-producing *Salmonella* in various food sources with prevalence rates of 17% in fish, 25% in pork, and 3% in chicken [13]. This lower prevalence may be attributed to the differences in sample sources. In our study, samples were obtained from individuals with hygienic handling and well-known food safety procedures, while the samples previously reported in Cambodia were randomly collected from the local markets, potentially the areas with higher AMR prevalence. Notably, the horizontal gene transfer mechanism is currently a global concern, potentially the area with high AMR load [20]. In our study, even though only one suspected ESBL-producing *Salmonella* was isolated from fish, it remains a concern due to the possibility of a horizontal gene transfer mechanism happening. Therefore, further study to better understand this sort of mechanism should be encouraged.

4. CONCLUSIONS

A low prevalence of *Salmonella* (1.9%) originating from fish and fish products was found, and the majority of isolates were responsive to antibiotics. Although this study found *Salmonella* at a low frequency, the presence of this bacterium and antibiotic resistance in fish is a warning sign, including multidrug resistance and possible ESBL development. As a recommendation, the use of antibiotics in fish farming should be closely monitored to reduce antibiotic resistance to foodborne pathogens and to protect consumers. In addition, molecular analysis such as whole genome sequencing or PCR should be used for further study to detect resistance genes and monitor the spread of resistant strains.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Laboratory of Environment and Food Safety (LEFS) of the Institut Pasteur du Cambodge (IPC) for their support in providing the equipment, reagents, consumables, and financial resources that were essential to the successful completion of this study.

REFERENCES

[1] Get the Facts about Salmonella. Available online: https://www.fda.gov/animal-veterinary/animal-health-literacy/get-facts-about-salmonella (accessed on 13 May 2024).

[2] WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007–2015. Available online: https://iris.who.int/handle/10665/199350 (accessed on 09 May 2024).

[3] O. Gal-Mor, E. C. Boyle, and G. A. Grassl. “Same species, different diseases: how and why typhoidal and non-typhoidal Salmonella enterica serovars differ.” *Front. Microbiol.*, 2014, vol. 5. Available online: https://doi.org/10.3389/fmicb.2014.00391

[4] Antimicrobial resistance. Available online: https://www.who.int/news-room/fact-sheets/detail/antimicrobial-resistance (accessed on 13 May 2024).

[5] B. M. Marshall and S. B. Levy. "Food Animals and Antimicrobials: Impacts on Human Health." *Clin. Microbiol. Rev.*, 2011, vol. 24, no. 4, pp. 718–733. Available online: https://doi.org/10.1128/cmr.00002-11

[6] Y. Muktar, G. Abera, and J. Kemal. "Antibiotic Resistance in Salmonella Species, a Serious Public Health Problem: A Review." *Global Veterinaria*, 2015, vol. 15, no. 5, pp. 469–479. Available online: https://doi.org/10.5829/idosi.gv.2015.15.05.96225

[7] D. T. A. Nguyen et al. "Prevalence, Serovar, and Antimicrobial Resistance of Nontyphoidal Salmonella in Vegetable, Fruit, and Water Samples in Ho Chi Minh City, Vietnam." *Foodborne Pathog. Dis.*, 2021, vol. 18, no. 5, pp. 354–363. Available online: https://doi.org/10.1089/fpd.2020.2891

[8] Y. Caron et al. "Beta-lactam resistance among Enterobacteriaceae in Cambodia: The four-year itch." *Int. J. Infect. Dis.*, 2018, vol. 66, pp. 74–79. Available online: https://doi.org/10.1016/j.ijid.2017.10.025

[9] F. Medalla et al. "Estimated Incidence of Antimicrobial Drug–Resistant Nontyphoidal Salmonella Infections, United States, 2004–2012." *Emerg. Infect. Dis.*, 2017, vol. 23, no. 1, pp. 29–37. Available online: https://doi.org/10.3201/eid2301.160771

[10] D. G. White et al. "The Isolation of Antibiotic-Resistant Salmonella from Retail Ground Meats." *N. Engl. J. Med.*, 2001, vol. 345, no. 16, pp. 1147–1154. Available online: https://doi.org/10.1056/NEJMoa010315

[11] C. Om and M.-L. McLaws. "Antibiotics: practice and opinions of Cambodian commercial farmers, animal feed retailers and veterinarians." *Antimicrob. Resist. Infect. Control*, 2016, vol. 5, no. 1, p. 42. Available online: https://doi.org/10.1186/s13756-016-0147-y

[12] K. S. Lay, Y. Vuthy, P. Song, K. Phol, and J. L. Sarthou. "Prevalence, Numbers and Antimicrobial Susceptibilities of Salmonella Serovars and Campylobacter spp. in Retail Poultry in Phnom Penh, Cambodia." *J. Vet. Med. Sci.*, 2011, vol. 73, no. 3, pp. 325–329. Available online: https://doi.org/10.1292/jvms.10-0373

[13] M. Nadimpalli et al. "CTX-M-55-type ESBL-producing Salmonella enterica are emerging among retail meats in Phnom Penh, Cambodia." *J. Antimicrob. Chemother.*, 2019, vol. 74, no. 2, pp. 342–348. Available online: https://doi.org/10.1093/jac/dky451

[14] AST of bacteria. Available online: https://www.eucast.org/ast_of_bacteria (accessed on 06 May 2024).

[15] A.-P. Magiorakos et al. "Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance." *Clin. Microbiol. Infect.*, 2012, vol. 18, no. 3, pp. 268–281. Available online: https://doi.org/10.1111/j.1469-0691.2011.03570.x

[16] L. Huoy et al. "Prevalence of Salmonella spp. in meat, seafood, and leafy green vegetables from local markets and vegetable farms in Phnom Penh, Cambodia." *Food Microbiol.*, 2024, vol. 124, p. 104614. Available online: https://doi.org/10.1016/j.fm.2024.104614

[17] S. P. San et al. "Biological Hazards and Indicators Found in Products of Animal Origin in Cambodia from 2000 to 2022: A Systematic Review." *Int. J. Environ. Res. Public Health*, 2024, vol. 21, no. 12, p. 1621. Available online: https://doi.org/10.3390/ijerph21121621

[18] D. V. G. S. Fernandes, V. S. Castro, A. da Cunha Neto, and E. E. de S. Figueiredo. "Salmonella spp. in the fish production chain: a review." *Ciênc. Rural*, 2018, vol. 48, p. e20180141. Available online: https://doi.org/10.1590/0103-8478cr20180141

[19] T. T. H. Van, H. N. K. Nguyen, P. M. Smooker, and P. J. Coloe. "The antibiotic resistance characteristics of non-typhoidal Salmonella enterica isolated from food-producing animals, retail meat and humans in South East Asia." *Int. J. Food Microbiol.*, 2012, vol. 154, no. 3, pp. 98–106. Available online: https://doi.org/10.1016/j.ijfoodmicro.2011.12.032

[20] J. Yan, B. Doublet, and A. Wiedemann. "Trends in horizontal gene transfer research in Salmonella antimicrobial resistance: a bibliometric analysis." *Front. Microbiol.*, 2024, vol. 15, Sep. 2024. Available online: https://doi.org/10.3389/fmicb.2024.1439664

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Consumer Preferences and Development of the Sourdough Bread with Different Flavors

Sokly CHEA[1], Marinich NET[1,2*], Sokunthea SOM[1], Reasmey TAN[1,2*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: rtan@itc.edu.kh*

Abstract: *Sourdough bread is the product of fermentation with a mixed microbial culture, called sourdough starter. A sourdough starter is combined with flour, water, and other ingredients to make dough. Sourdough bread is known due to its rich nutritional and organoleptic qualities. Therefore, the objective of this study was to analyze consumer preferences by conducting an online survey, and develop sourdough bread with different flavors. The questionnaire survey was shared via social media platforms. For the first section, the profiles of the respondents' information was collected, while the second section focused on sourdough consumption and preferences. For making sourdough bread, the control sample without adding sourdough starter and flavor was used to select the temperature and baking time. After selecting a good temperature and baking time (200℃, 25min), sourdough bread (SDB0) was developed with added starter culture and other ingredients by selecting the best formulation, and then only one of the best formulas, three formulations of sourdough bread were developed with different flavors including chocolate, matcha and banana. The control sample and formulated sourdough bread were analyzed for their proximate composition and physical characteristics. According to online survey, the results showed that 66.5% of respondents were female, 84.2% were aged 18 to 25, and 4.2% were in the highest age group 25. Additionally, in the survey of 1000 people, were covered for 67.5% of the total participants had known about sourdough, with mostly known from friends/family members (30.1%), and bakery (12.8%). Consumption frequency varied, with 18.5% eating sourdough once a week, 5.3% consuming it 2-3 times per week, and 2.8% having it once a month. Bakeries (34.4%), supermarkets (24.7%), and coffee shops (11.4%) were the most common purchase locations. The key factors influencing the purchase included shape, size, flavors, and price, with 798 (79.6%) preferring medium-sized loaves and 537 (53.7%) favoring prices between $1 to $1.5. For sourdough bread, the results showed that the moisture and ash contents of the control sample and three formulated sourdough breads had no significant difference ($p>0.05$), ranging from 29.68±0.02 to 38.2±0.13 and 0.31±0.13 to 0.97±0.01, respectively. The pH value and total titratable acidity varied from 4.21±0.014 to 5.94±0.04 and 0.03 to 0.063% ($P>0.05$), respectively. And the result of color, the crumb L* decreased, a* and b* increased with positive values, when different flavors were added. For the result of the crust L* and b* increased with positive value, but a* decreased with negative value ($P<0.05$). The findings of this study demonstrate that incorporating flavors into sourdough bread production can increase consumer preferences while improving the bread's nutritional composition. It is suggested that future research on the microbiological quality and shelf-life of sourdough bread should be investigated.*

Keywords: Sourdough; Preferences; Starter culture; Consumption; Organoleptic characteristics;

1. INTRODUCTION

Nowadays, consumers are very interested in choosing unique or traditional foods that contain less or no chemical preservatives [1,2]. Similarly, this new trend has been recently developed in the bread industry, particularly with sourdough applications [3]. Sourdough's usage in bread manufacturing dates back to antiquity, and there is currently a renewed interest in sourdough applications [4,5]. Sourdough is one of the oldest bioprocessing techniques, characterized

[1,2*] Corresponding author: Reasmey TAN
E-mail: rtan@itc.edu.kh; Tel: +855-12 602 202

by a mutual relationship between a complex microbiota dominated by lactic acid bacteria (LAB) and yeast population. It was widely used in baking applications as an acidifying and leavening bread dough agent until the late nineteenth century [6]. Sourdough bread is produced when a sourdough starter is combined with flour and water to make dough. In recent years, increased consumer awareness of health-related hazards of food chemical preservatives has resulted in a trend toward increased preference for natural preservative-free processed foods [7]. According to Catzeddu.[8] sourdough bread also has superior nutritional properties, rich in nutrients and vitamins was then kneaded by hand for an additional 2 minutes until smooth and non-sticky, allowing gluten development. Indeed, an increase in the appreciation for more flavorful and healthy breads over the last decade has raised consumer demand for sourdough breads. The current renaissance of sourdough as a bioprocessing is encouraged by its beneficial effect on flavor, texture, shelf life, and health-promoting properties [9]. Therefore, the objective of this study was to analyze consumer preferences by conducting an online survey, and develop sourdough bread with different flavors.

2. METHODOLOGY

2.1 Online survey

The online survey was conducted using Google Forms and shared on social media to collect data on consumer preferences and awareness of sourdough bread. The questionnaire consisted of two types of questions: open-ended and closed-ended, and was divided into two sections: the profiles of the respondents and sourdough bread consumption and preferences.

2.2 Materials

The sourdough starter used in this study was purchased online. The instant dry yeast, wheat flour, and other ingredients were sourced from local markets, supermarkets, and bakery shops in Phnom Penh.

2.3 Sourdough bread preparation

Sourdough bread was prepared by mixing a starter with other ingredients and selected flavors, including chocolate, matcha, or banana. One formulation was developed using without sourdough starter and flavors (control), and four other formulations used different combinations with and without added flavors of sourdough bread, such as SDB0 (original), SDB1 (chocolate), SDB2 (Matcha), and SDB3 (banana). After mixing all ingredients, the dough was kneaded, fermented for 4 hours, proofed for 2 hours, and baked at 200°C for 25 minutes. After cooling to room temperature, the bread was ready for further analysis.

2.4 pH value

The pH value of each sample was measured using a digital pH meter (LAQUA F-72-HORIBA, China), following the AOAC (2004) method. Each test used a 5g sample that was put into beaker and 45 mL of distilled water and homogenized. Before testing sample, the pH meter was calibrated with standard buffer of known pH values 7 and 4.

2.5 Total titratable acidity

Total titratable acidity was measured using the AOAC (942.15) volumetric method. A 5g sample was diluted with 45 mL of distilled water and titrated with 0.1 N NaOH using phenolphthalein as an indicator. The volume of NaOH used was recorded to calculate:

$$\% \text{ TTA} = \frac{\text{N} \times \text{V} \times 0.09}{\text{V}_{\text{Sample}}} \times 100 \qquad \text{(Eq. 1)}$$

where:
N_{NaOH} = concentration normality of NaOH (N)
V_{NaOH} = volume of NaOH used (ml)
V_{sample} = volume of sample used (ml)
0.09 = conversion factor of citric acid

2.6 Moisture content

Moisture content was determined following AOAC (2000) methods. A 5g sample was cut into small pieces and placed in an aluminum cap. It was dried in an oven at 105°C for about 3 hours, then cooled in a desiccator for 15 minutes. The final weight was recorded and calculated by using the following equation:

$$\% \text{ Moisture} = \frac{(M1-M2)}{Msample} \times 100 \qquad \text{(Eq. 2)}$$

where:
M_1 = mass in grams of plate and sample after drying (g)
M_2 = mass in grams of plate and sample before drying (g)

2.7 Water activity

Water activity was measured with an AQUALAB PAWKIT meter. A 2g bread sample was placed in the meter, and results were obtained after 5 minutes at ambient temperature.

2.8 Ash content

The ash content was determined using the AOAC (2004) method. A 5g bread sample was placed in a pre-dried porcelain cup and incinerated at 550°C for 3 hours. After cooling, the final mass was recorded, and the ash content was calculated:

$$\% \text{ Ash content} = \frac{M1-M2}{Msample} \times 100 \qquad \text{(Eq. 3)}$$

where:
M_1 = mass in grams of plate and sample after drying
M_2 = mass in grams of plate and sample before drying

2.9 Color

Color was evaluated using a Chroma CR400 portable colorimeter (Konica-Minolta). The instrument was calibrated with a white tile before measuring. L*, a*, and b* values were recorded for the crust (center) and crumb (middle slice) of the sourdough bread, with triplicate measurements taken. The final results were the average of all measurements [9].

2.10 Sensory evaluation

Sensory evaluation using a 9-point hedonic scale was conducted with 115 panelists to assess acceptability based on appearance, color, pore size, odor, aroma, texture, saltiness, sourness, flavor, and overall preference [10].

2.11 Statistical analysis

All tests were performed in duplicate, with results presented as mean ± standard deviation. Statistical analysis was done using SPSS (version 25). One-way ANOVA and Tukey's post-hoc test were used to evaluate differences, with significance set at $p<0.05$.

3. RESULTS AND DISCUSSION

3.1 Profile of the respondents

In survey of 1000 people, 33.5% of the total participants were male and 66.5% were female. Most respondents (84.2%) were between 18 and 25 years old, while 11.6% were under 18, and 4.2% were over 25 years old, as shown in **Table 1.**

Table 1. Profile of the respondents

Age	Participants	Percentages (%)
Under 18	116	11.6
18 to 25	842	84.2
Over 25	42	4.2
Total	1000	100

3.2. Sourdough bread consumption by respondents

According to the survey data, as shown in **Table 2**, 29.6% of participants consumed sourdough bread very rarely, while 18.5% ate it once per week. Only 5.3% consumed it 2- 3 times per week, 2.8% once per month, and 2.7% consumed it daily.

Table 2. Frequency of sourdough bread by respondents

Frequency	Participants	Percentages (%)
Daily	27	2.7
2-3 times per week	53	5.3
1 time per week	185	18.5
1 time per month	28	2.8
Very rarely	296	29.6

3.3 Major drivers for the preferences of respondents

According to the survey data, there are different ways to choose sourdough bread products by respondents. The 4 factors that affected the consumer's decision mentioned by the respondents are shown in **Fig. 1**. The most important factors for consumers were the shape and flavors of sourdough bread. The loaf slices were 527 (52.5%) among respondents.

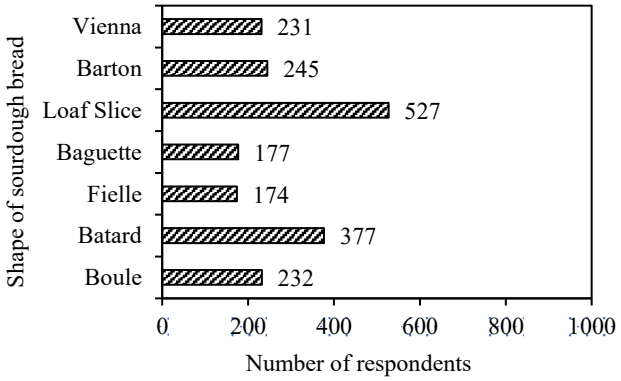


Fig 1. Consumer preferences for the shape of sourdough bread.

Fig. 2. illustrates the consumer preferences for the flavors of sourdough bread.

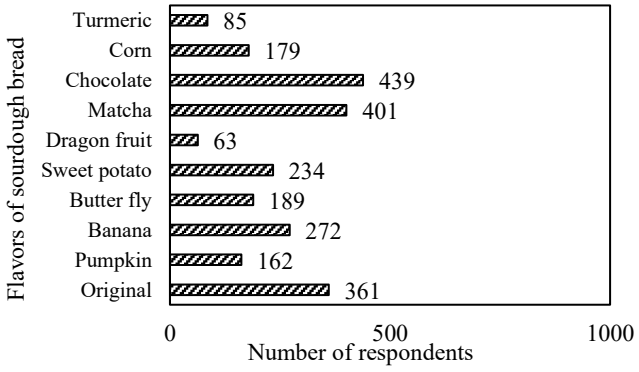


Fig. 2. Consumer preferences for the flavors of sourdough bread.

According to the survey data about the consumer preferences of the respondents, as shown in **Table 3,** followed by other factors, including price, flavor, and size, which accounted for 537 (53.7%), 872 (87.2%), and 798 (79.8%) for size respondents, respectively.

Table 3. Preference of the respondents for sourdough bread.

Preferences of the respondents for sourdough bread		
Price ($)	Participants	Percentages (%)
0.5 to 1	537	53.7
1 to 1.5	378	37.8
Higher than 1.5	41	4.1
Others	44	4.4
Size		
Small	141	14.1
Medium	798	79.8
Large	37	3.7
Others	24	2.4
Flavor		
Sourness	29	2.9
Sour-sweet taste	872	87.2
Others	99	9.9
Total	1000	100

3.4 Physicochemical analysis

The results of moisture, a_w, ash, pH, acidity, and color of sourdough bread with different flavors are shown in **Table 4.** There was no significant difference (p>0.05) in different flavors. The pH value of sourdough bread ranged from 4.215 to 5.95 for sample controls, SDB0, SDB1, SDB2, and SDB3. The sourdough bread with added flavor had a lower pH value than the control. The results were in agreement with the report [11,12]. Casado et al. [13] report that the pH of the fermentation process in sourdough bread plays a crucial role in determining the quality and characteristics of the final product.

The value of titratable acidity in all different conditions ranged from 0.034 to 0.081 for sample control, SBD0, SDB1, SDB2, and SBD3. The sourdough bread had a higher total titratable acidity than the control. In addition, in all sourdough bread, a pH drop creates an increase in total titratable acidity. The results were in agreement with the report of [13,14]. According to Azni et al. [16] The total titratable Acidity plays a crucial role in determining the quality of sourdough bread. Moreover, the pH and total titratable acidity (TTA) of sourdough significantly impact the fermentation process and final product characteristics [17].

Table 4. Mean values of the physicochemical quality of the sourdough bread with different flavors. Different letters within the same column had no significant difference from each other *(p>0.05)*.

Code of samples	Parameters (mean value ± standard deviation (SD))				
	Moisture (%)	a_w	Ash (%)	pH	Acidity (%)
Control	29.68 ± 0.02^{d}	0.8382 ± 0.001^{a}	0.315 ± 0.07^{cd}	5.94 ± 0.014^{a}	0.036 ± 0.006^{d}
SDB0	30.33 ± 0.04^{c}	0.8327 ± 0.001^{a}	0.425 ± 0.09^{c}	4.865 ± 0.007^{a}	0.081 ± 0.049^{a}
SDB1	37.45 ± 0.01^{a}	0.8494 ± 0.001^{a}	0.35 ± 0.09^{d}	4.825 ± 0.007^{a}	0.063 ± 0.007^{a}
SDB2	38.2 ± 0.13^{a}	0.8501^{a}	0.52 ± 0.13^{b}	4.605 ± 0.007^{a}	0.054 ± 0.006^{c}
SBD3	36.67 ± 0.09^{b}	0.8387^{a}	0.835 ± 0.17^{a}	4.215 ± 0.049^{b}	0.054 ± 0.049^{c}

The results of moisture, water activity, and ash content of sourdough bread are presented in **Table 4**. The moisture content of sourdough bread ranged from 29.68% to 36.67% for samples Control, SDB0, SDB1, SDB2, and SDB3 (p>0.05). The result was similar to the study of Regulation (EU) No 432/2012 and Regulation (EC) No 1924/2006. Ogunsakin et al. [18] The moisture all samples were within the standard of freshly sourdough bread, ranging from 34.68 to 40.03%. Based on the result of this study, for all conditions, such as Control, SDB0, SDB1, SDB2, and SDB3, was values ranged from 0.315% to 0.835%. The results were in agreement with the report [19] according to Tokatlidis et al. [20]. Ash content refers to the mineral or inorganic material present in the flour. The ash content of the flour used to make sourdough bread is a critical factor that impacts the development of sour flavor and overall bread quality. The water activity of sourdough bread with an average of water activity of 0.8382, 0.8327, 0.8496, 0.8501, and 0.8387 for sample controls, SDB0, SDB1, SDB2, and SDB3.This result is in agreement with [21]. According to Tapía et al. [22]. water

activity in bread is indeed a crucial parameter that significantly influences the product's shelf-life Water activity limits microbial growth, stabilizes food matrices, and extends shelf life. Its measurement is key to analyzing and assessing product quality. Additionally, it provides information regarding the possibility of microbiological growth on the surface during the storage period [23].

3.5 Color

There was a significant difference (p<0.05) in crumb and crust in terms of all five sourdough as shown in **Table 5.** Based on the result of the crust L*, a*, and b* values were higher in control, SDB0, and SDB3 compared to SDB1 and SDB2, with SDB2 showing a negative a* value. Furthermore, crumb color had higher lightness (L*) in the control, SDB0, and SDB3, while SDB1 and SDB3 had higher a* values. The negative a* in SDB0 reflects the green color of consistent with the green flour used. The result was in agreement with the study of [24].

Table 5. Crust and crumb color for sourdough bread, the different letters within the same column differed and had significant differences from each other (p<0.05).

Sample	Bread crust color		
	L*	a*	b*
Control	43.29 ± 0.07^{b}	13.93± 0.03^{a}	30.81 ± 0.12^{b}
SDB0	45.77 ± 0.02^{a}	11.93±0.03^{b}	32.58 ± 0.05^{a}
SDB1	9.89 ± 0.07^{c}	15.04± 0.21^{a}	33.60 ± 0.38^{a}
SDB2	12.64 ± 0.36^{b}	1.65±0.03^{c}	34.05 ±0.03^{a}
SDB3	47.39± 0.27^{a}	11.48±0.1^{b}	30.58 ± 0.48^{b}
Sample	Bread crumb color		
	L*	a*	b*
Control	71.22 ± 0.03^{a}	1.49± 0.02^{c}	12.86 ± 0.13^{d}
SDB0	66.13 ± 0.01^{b}	1.33±0.01^{c}	11.33 ± 0.1^{d}
SDB1	12.07 ± 0.26^{c}	13.73± 0.21^{a}	37.84 ± 0.25^{a}
SDB2	10.11 ± 0.03^{d}	-1.09±0.03^{d}	27.18 ±0.02^{c}
SDB3	66.71± 0.23^{b}	12.48±0.02^{b}	32.58 ± 0.07^{b}

3.6 Sensory evaluation

Sensory evaluation of the three best conditions SDB1, SDB3, and the control, showed that SDB3 was the most acceptable. There was no significant difference (p>0.05) between the control and SDB1 (**Fig. 3**). Samples SDB0 and SDB2 had a lower result in some attributes. SDB3 had the highest pore size, and increasing sourdough levels positively impacted the pore size of the final bread [25]. Whereas, the other two conditions had the lowest score, approximately point 6. Overall acceptability is very important for sensory evaluation. All the bread samples had higher overall acceptability, with acceptable scores receiving more than seven. Additionally, condition SDB3 had higher overall acceptability compared with other conditions of sourdough bread. The final product's texture and sensory attributes are significantly influenced by the sourdough and instant dry yeast, with the fermented dough process [26].

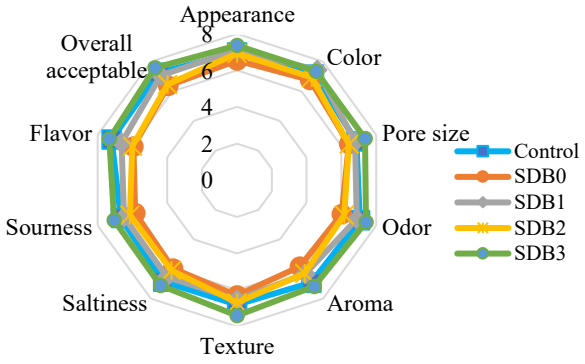


Fig. 3. Sensory evaluation of sourdough bread

4. CONCLUSIONS

In conclusion, the quality testing of sourdough bread with different flavors showed no significant difference (p>0.05) in pH value, total titratable acidity, moisture, water activity, and ash content of sourdough bread (p>0.05). As a result of color, the crumb L* decreased, and a* and b* increased with positive values when different flavors were added. For the sensory evaluation, the incorporation of banana flavors resulted in greater consumable acceptance. Future research on microbiological quality and shelf life is recommended.

REFERENCES

[1] I. Gutiérrez-del-Río, J. Fernández, and F. Lombó, "Plant nutraceuticals as antimicrobial agents in food preservation: terpenoids, polyphenols and thiols," *Int. J. Antimicrob. Agents*, vol. 52, no. 3, pp. 309–315, 2018, doi: 10.1016/j.ijantimicag.2018.04.024.

[2] S. Bordeleau, H. Asselin, M. J. Mazerolle, and L. Imbeau, "'Is it still safe to eat traditional food?' Addressing traditional food safety concerns in aboriginal communities," *Sci. Total Environ.*, vol. 565, pp. 529–538, 2016, doi: 10.1016/j.scitotenv.2016.04.189.

[3] C. G. Rizzello, R. Coda, and M. Gobbetti, *Use of Sourdough Fermentation and Nonwheat Flours for Enhancing Nutritional and Healthy Properties of Wheat-Based Foods*. Elsevier Inc., 2017. doi: 10.1016/B978-0-12-802309-9.00018-2.

[4] C. A. Batt and M. L. Tortorello, *Encyclopedia of Food Microbiology: Second Edition*. 2014.

[5] A. Pasqualone, “Traditional flat breads spread from the Fertile Crescent: Production process and history of baking systems,” *J. Ethn. Foods*, vol. 5, no. 1, pp. 10–19, 2018, doi: 10.1016/j.jef.2018.02.002.

[6] C. Graça *et al.*, “Yoghurt as a starter in sourdough fermentation to improve the technological and functional properties of sourdough-wheat bread,” *J. Funct. Foods*, vol. 88, p. 104877, 2022, doi: 10.1016/j.jff.2021.104877.

[7] M. Gobbetti and M. Gänzle, *Handbook on sourdough biotechnology*. 2013. doi: 10.1007/978-1-4614-5425-0.

[8] P. Catzeddu, “Chapter 14 - Sourdough Breads,” in *Science direct*, 2019, pp. 177–188.

[9] R. Peñalver, W. Díaz-Vásquez, M. Maulén, and G. Nieto, “Sustainable Processes and Physico-Chemical Characterization of Artisanal Spontaneous Gluten Free Sourdough (Quinoa, Amaranth and Brown Rice) Compared to Wheat Sourdough,” *Sustain.* , vol. 16, no. 8, 2024, doi: 10.3390/su16083297.

[10] M. T. Rahman, M. A. Zubair, K. Shima, and M. P. Chakma, “Development of Quality Parameters for Yogurt with Strawberry Juice,” *Food Nutr. Sci.*, vol. 11, no. 12, pp. 1070–1077, 2020, doi: 10.4236/fns.2020.1112075.

[11] L. De Vuyst and P. Neysens, “The sourdough microflora: Biodiversity and metabolic interactions,” *Trends Food Sci. Technol.*, vol. 16, no. 1–3, pp. 43–56, 2005, doi: 10.1016/j.tifs.2004.02.012.

[12] C. Thiele, S. Grassl, and M. Gänzle, “Gluten Hydrolysis and Depolymerization during Sourdough Fermentation,” *J. Agric. Food Chem.*, vol. 52, no. 5, pp. 1307–1314, 2004, doi: 10.1021/jf034470z.

[13] A. Casado, A. Álvarez, L. González, D. Fernández, J. L. Marcos, and M. E. Tornadijo, “Effect of fermentation on microbiological, physicochemical and physical characteristics of sourdough and impact of its use on bread quality,” *Czech J. Food Sci.*, vol. 35, no. 6, pp. 496–506, 2017, doi: 10.17221/68/2017-CJFS.

[14] S. A. Vogelmann and C. Hertel, “Impact of ecological factors on the stability of microbial associations in sourdough fermentation,” *Food Microbiol.*, vol. 28, no. 3, pp. 583–589, 2011, doi: 10.1016/j.fm.2010.11.010.

[15] I. Banu, I. Vasilean, and I. Aprodu, “Effect of Select Parameters of the Sourdough Rye Fermentation on the Activity of Some Mixed Starter Cultures,” *Food Biotechnol.*, vol. 25, no. 4, pp. 275–291, 2011, doi: 10.1080/08905436.2011.617251.

[16] I. N. Azni, A. T. A. Prabawati, I. Basriman, and J. R. Amelia, “Quality Characteristics of Sourdough Bread with The Addition of Water Yeast and Wheat-Mocaf Flours Combination,” *J. Tek. Pertan. Lampung (Journal Agric. Eng.*, vol. 12, no. 2, p. 350, 2023, doi: 10.23960/jtep-l.v12i2.350-362.

[17] T. Lebedenko, V. Kozhevnikova, O. Kotuzaki, and T. Novichkova, “Determining the efficiency of spontaneous sourdough for stabilizing the quality of bread products in bakeries and catering enterprises,” *Eastern-European J. Enterp. Technol.*, vol. 4, no. 11–100, pp. 22–35, 2019, doi: 10.15587/1729-4061.2019.174289.

[18] A. O. Ogunsakin *et al.*, “Evaluation of functionally important lactic acid bacteria and yeasts from Nigerian sorghum as starter cultures for gluten-free sourdough preparation,” *Lwt*, vol. 82, pp. 326–334, 2017, doi: 10.1016/j.lwt.2017.04.048.

[19] M. Carocho *et al.*, “Comparison of different bread types: Chemical and physical parameters,” *Food Chem.*, vol. 310, no. November 2019, 2020, doi: 10.1016/j.foodchem.2019.125954.

[20] M. A. Islam and S. Islam, “Sourdough Bread Quality: Facts and Factors,” *Foods*, vol. 13, no. 13, 2024, doi: 10.3390/foods13132132.

[21] R. Supasil *et al.*, “Improvement of Sourdough and Bread Qualities by Fermented Water of Asian Pears and Assam Tea Leaves with Co-Cultures of Lactiplantibacillus plantarum and Saccharomyces cerevisiae,” *Foods*, vol. 11, no. 14, 2022, doi: 10.3390/foods11142071.

[22] M. Tapia, S. Alzamora, and J. Chirife, “Effects of Water Activity (a w) on Microbial Stability as a Hurdle in Food Preservation,” 2020, pp. 323–355. doi: 10.1002/9781118765982.ch14.

[23] M. S. Rahman, “Food stability determination by macro-micro region concept in the state diagram and by defining a critical temperature,” *J. Food Eng.*, vol. 99, no. 4, pp. 402–416, 2010, doi: 10.1016/j.jfoodeng.2009.07.011.

[24] J. Pico, J. Bernal, and M. Gómez, “Wheat bread aroma compounds in crumb and crust: A review,” *Food Res. Int.*, vol. 75, pp. 200–215, 2015, doi: 10.1016/j.foodres.2015.05.051.

[25] Z. Zaidiyah, B. M. Nur, Y. M. Lubis, I. S. Nasution, F. Nadia, and U. Khaira, “Loaf volume and porosity for sourdough bread based modified cassava flour by linier regression and sensory evaluation approach,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1116, no. 1, 2022, doi: 10.1088/1755-1315/1116/1/012068.

[26] A. Martín-Garcia, M. Riu-Aumatell, and E. López-Tamames, “Influence of Process Parameters on Sourdough Microbiota, Physical Properties and Sensory Profile,” *Food Rev. Int.*, vol. 39, no. 1, pp. 334–348, 2023, doi: 10.1080/87559129.2021.1906698.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Effects of monoglyceride and guar gum on the quality and sensory properties of mango frozen yogurt

Sodaneth HUOR[1], Pichmony EK [1*], Sengly SROY[1]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: pichmony@itc.edu.kh*

Abstract: *Frozen yogurt is a fermented dairy dessert that contains probiotics with smooth, and creamy texture. However, the selection of yogurt starter culture and the addition of stabilizers and emulsifiers might affect its quality and sensory properties. This study aimed to evaluate the effects of commercial local versus imported yogurt starters and the influence of monoglyceride and guar gum on the physicochemical and sensory characteristics of frozen yogurt with added Keo Romeat mango as natural flavor. Initially, two trials without additives were developed using commercial local and imported plain yogurt starters. The frozen yogurt samples were subjected to physicochemical analysis including pH, acidity, overrun, and melting rate. A sensory evaluation (hedonic test) was also conducted to determine consumer preference. The results showed that yogurt made with the imported plain yogurt starter had no significant difference in terms of pH and overrun ($p > 0.05$), yet showed significant difference in titratable acidity and melting rate, compared with the local plain yogurt starter. Sensory evaluation results were not significantly different in all attributes ($p > 0.05$). However, 62.7% of panelists preferred the frozen yogurt made with commercial imported yogurt starter. Thus, it was further studied by adding monoglyceride and guar gum at different levels: T0 (control, no additives), T0.2 (0.2% combined monoglyceride and guar gum, ratio 1:1), and T0.4 (0.4% combined monoglyceride and guar gum, ratio 1:1). The results revealed that there were no significant different of the three trials in term of pH and titratable acidity. Nonetheless, the additives significantly affect overrun, consistency, and melting rate of the product. The present of food additives improved overrun and consistency except melting rate. In term of sensory properties, trial T0.2 received the highest score in overall acceptable attribute. The results suggest that the use of a combined emulsifier and stabilizer can enhance the texture and consumer preference of mango frozen yogurt, making it a promising approach for commercial product development.*

Keywords: Frozen yogurt; emulsifier; stabilizer; sensory evaluation; probiotics

1. INTRODUCTION

Frozen yogurt is a fermented dairy frozen dessert known for its sour flavor and good health benefits due to the presence of probiotics [1]. Currently, the increasing demand for healthier and functional foods has driven innovation in frozen dessert formulations [2]. Moreover, adding natural flavor into the frozen yogurt also increase a positive profile to the product along with its flavor. In Cambodia, a mango type called Keo Romeat, a tropical fruit rich nutrition and attractive color and flavor [3] is a promising ingredient option in the development of frozen yogurt products.

Generally, stabilizers and emulsifiers are widely used in frozen dessert production in order to enhance product quality by improving texture, mouthfeel, and storage stability [4]. However, the selection and optimization of these additives are critical, as they can influence both physicochemical and sensory properties of the final product. A stabilizer, guar gum, which is a natural polysaccharide, and monoglycerides, a commonly used emulsifiers, are often added to stabilize the mixture, enhance viscosity, and prevent phase separation [5]. The synergistic use of stabilizers and emulsifiers has been reported to improve the overall quality of frozen dairy products[6].

Previous studies have explored the effect of stabilizers and emulsifiers on the properties of frozen desserts. For example, Rezaei *et al.*[7] demonstrated that combinations of arabic gum and guar gum improved viscosity, overrun, and

[1*] Corresponding author: Pichmony Ek
E-mail: pichmony@itc.edu.kh; Tel: +855-92 882 055

better overall sensory and physicochemical characteristics in frozen yogurt. Similarly, Muse and Hartel learned that emulsifiers reduced ice crystal size and improved meltdown properties in ice cream [8]. Despite these valuables' findings, the specific effects of combining monoglyceride and guar gum in mango flavored frozen yogurt remain underexplored.

This study aims to investigating the influence of monoglyceride and guar gum on the quality and sensory properties of mango frozen yogurt. The work was conducted in two phases. In the initial phase, two formulations were prepared using different commercial plain yogurt starter cultures (local and imported), without adding monoglyceride or guar gum. Based on the superior physicochemical and sensory attributes, the selected formulation was subjected to optimization with guar gum and monoglyceride. In the second phase, three trials were performed: control (T0) without additive, and two experimental treatments for 0.2% (T0.2) and 0.4% (T0.4) concentrations of a mixture of guar gum and monoglyceride (1:1). The samples were subjected to quality analysis including pH, acidity, overrun, melting rate, consistency, and sensory characteristics.

This research contributes not only to the further understanding of stabilizer and emulsifier functionality in frozen yogurt science but also assists in local fruit utilization and dairy innovation. By optimizing additive levels, the research aims at improving the sensory acceptability, textural stability, and general acceptability of mango frozen yogurt, with a potential chance of developing new functional frozen dessert products.

2. METHODOLOGY

2.1 Yogurt preparation process

Two brand of commercial plain yogurt starter were bought from the market. One is a local and other is an imported product. The specific bacteria strains were not identified. To prepare yogurt, commercial UHT cow milk containing 3.5% fat was used. The milk was heated to 80-90 ºC to denature protein for 5-10min, and then cool to 45 º C. At this point, 2% (2g in 100g milk) of plain yogurt starter was added. The inoculation was done at 45 °C, which is ideal for the growth of beneficial bacteria. The bacteria ferment lactose into lactic acid, increasing acidity and lowering pH, causing casein proteins in the milk to coagulate and form the thick texture of yogurt[9]. Next, the mixture was then incubated at 45 º C until the pH reached 4.6. When the desired pH was achieved, the yogurt was taken out to mix with ice cream mix. These 2 trials were performed in three batches.

2.2 Mango puree preparation

Fresh ripe mangoes were washed thoroughly, peeled, and the seeds were removed. The pulp was collected and blended using a kitchen blender until a smooth and uniform purée was obtained. The purée was passed through a fine mesh sieve to remove fibrous particles. It was then pasteurized at 85 °C for 5 minutes to ensure microbial safety and improve shelf stability. After pasteurization, the purée was cooled and stored in clean containers at 4 °C until use.

2.3 Frozen yogurt preparation

According a study by Goof & Hartel [10], The process of making ice cream mix involve mixing all the ingredients including milk, milk powder, sugar, and combined guar gum and monoglyceride. Generally, emulsifier and stabilizer blends levels in frozen desserts typically range from 0.2% to 0.5%[10]. Thus, the concentration of combined guar gum and monoglyceride of 0, 0.2, 0.4% were studied. However, for the first phase, which is plain yogurt starter selection, monoglyceride and guar gum were excluded. Next, the mix was homogenized using electric food mixer with 500W power (Sokany brand, China) for 2 min to ensure even distribution of solid ingredients. After mixing, the mix was subjected to pasteurization at 80-90 ºC for 5-10 min, then cooled to room temperature. After that, the ice cream mix was thoroughly mixed well with the yogurt and was aged in the refrigerator overnight. Once the mixture was done aging, the mango puree was added to the mixture. The final mix was overrun in an ice cream maker machine (Whynter, model ICM-200Ls, China) for 20 min. Finally, the frozen yogurt was packed in a fine plastic container and frozen at -18 ºC for 24h. Each trial was performed in three batches.

2.4 pH

Frozen yogurt samples were melted to ambient temperature before measuring the potential hydrogen. The pH value was measured by a pH meter (LAQUA PH210 Handheld pH, Japan) through AOAC official (AOAC, 1990). Firstly, the pH meter was calibrated with three pH levels of buffer solution (pH = 7, pH = 4). After that, the prop was immersed in the sample and the data were recorded.

2.5 Total acidity

Titratable acidity was determined by titrating the sample with 0.1N of sodium hydroxide. Phenolphthalein was used as indicator to indicate the end point of titration [7]. The total acidity is expressed as percentage of lactic acid (%l.a).

$$\%Total\ acidity = \frac{N_{NaOH} \times V_{NaOH} \times M_{eq}}{m_{sample} \times 1000} \times 100 \text{ (Eq. 1)}$$

N_{NaOH}: normality of NaOH (N)

V_{NaOH}: volume of NaOH (ml)
M_{eq}: mass equivalent of lactic acid (90mg/mEq)
m_{sample}: initial mass of sample (g)

2.6 Consistency

The consistency of yogurt was measured using Bostwick consistometer (Endecotts, England). A sample of yogurt mix at room temperature (before frozen) was placed in the Bostwick consistometer, and the gate was opened to allow the sample to flow. The distance traveled in 30 seconds was recorded as a measure of flow consistency (cm/s).

2.7 Overrun

The overrun test is for the determination of the air volume in ice cream samples. Firstly, the yogurt mix was measured to a volume of 50 mL, and its weight was recorded. Secondly, the yogurt mix was processed in the ice cream machine to incorporate air until it was frozen. After freezing, the frozen yogurt was measured to a volume of 50 mL, and its weight was recorded. The tests were duplicated in one sample, and an average value was calculated [6].

$$Overrun\ (\%) = \frac{W_1 - W_2}{W_2} \times 100 \quad \text{(Eq. 2)}$$

W1: weight of mix (g)

W2: weight of same volume of finished product (g)

2.8 Melting rate

Melting rate was performed following Isik *et al.* with some modification [11]. A frozen yogurt sample of 50 g was weighed and placed on a fine wire mesh with 0.7cm × 0.7cm over a collection container. The experiment was carried out at 26.4 ± 0.28°C with 56 ± 5.65% humidity in Food hall Technology of ITC to ensure controlled temperature. Next, the sample was allowed to melt naturally. The weight of the collected sample in the container was recorded at 15, 30, and 45 minutes of melting. The ratio of these values to the initial weight of frozen yogurt was calculated.

Calculation

$$\text{Melting Rate}(\%) = \frac{\text{Weight collected at time t (g)}}{\text{Initial sample weight (g)}} \times 100 \text{ (Eq. 3)}$$

2.9 Sensory evaluation

Mango frozen yogurt samples were taken out from the freezer, then scooped and place in a small plastic cup for panelists to evaluate their color, smoothness, creaminess, flavor, and overall acceptability. The evaluation was conducted in two phases. In the first phase, two samples made with different commercial plain yogurt starters were tested. Next, in the second phase three sample with different concentration of guar gum and monoglyceride were evaluated. Moreover, panelists were instructed to drink water before tasting each new sample to cleanse their mouth and avoid any influence from the previous sample. The sensory evaluation was conducted in a sensory evaluation room at ambient temperature of Institute of Technology of Cambodia using online plate form (Google Form). Hedonic test was performed with 9 scale points from 1= extremely dislike to 9= extremely like. There were 75 panelists in the first phase and 80 in the second phase participated.

2.10 Statistical analysis

Results are reported as the mean ± standard deviation. Statistical analysis was performed using IBM SPSS statistics version 27 software (SPSS Inc., Chicago, IL, USA). Differences between groups were analyzed using one-way analysis of variance (ANOVA) followed by Tukey's post hoc test to determine significant differences between means. A p-value less than 0.05 was considered statistically significant. Means sharing the same superscripts do not differ significantly from one another.

3. RESULTS AND DISCUSSION

3.1 Selection of plain yogurt starter culture

Table 3.1. Physicochemical of mango frozen yogurt made with two different commercial plain yogurt starters

Trial	pH	Titratable acidity(%l.a)	Overrun (%)	Melting rate at 30min (%)
TI	4.69 ± 0.04^a	0.67 ± 0.02^a	25.76 ± 4.77^a	65.74 ± 3.62^a
TL	4.70 ± 0.03^a	0.61 ± 0.02^b	23.46 ± 2.33^b	48.47 ± 4.27^b

TI: Trial with Imported plain yogurt starter; TL: Trial with Local plain yogurt starter; l.a: Lactic acid. Means with different superscripts (a, b) within the same column differ significantly at $p < 0.05$

Table 3.1 revealed that the pH and overrun values between the two yogurt starters (TI and TL) were not significantly different ($p > 0.05$). In contrast, titratable acidity and melting rate showed highly significant differences ($p < 0.001$). The sample made with the imported yogurt starter (TI) had a higher acidity (0.67 ± 0.02% l.a) and a notably higher melting rate (65.74 ± 3.62%) compared to the local starter (0.61 ± 0.02% l.a and 48.47 ± 4.27%, respectively). The yogurt made with the imported starter (TI) exhibited a higher titratable acidity (0.67 ± 0.02% LA) compared to the local starter (TL), which produced 0.61 ± 0.02% l.a. This

difference may be due to the presence of more active acid-producing strains, such as *Lactobacillus delbrueckii* subsp. *bulgaricus* and *Streptococcus thermophilus*, in the imported culture. These bacteria are known to produce lactic acid more efficiently during fermentation, resulting in a more pronounced tangy flavor and lower pH [12]. The higher acidity may also be the reason why the melting rate of TI was higher. When acidity increases, it can affect the structure of proteins in the yogurt, making it weaker and causing the product to melt more quickly[8].This means that while the imported yogurt starter gave better acidity, it may make the frozen yogurt less stable when left at room temperature.

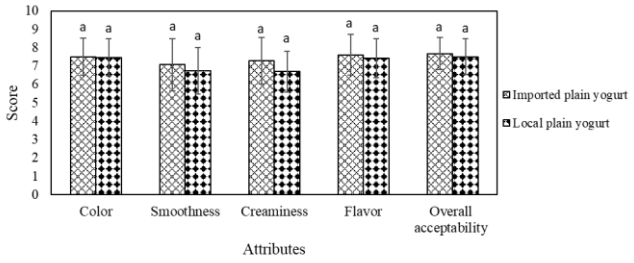


Fig. 3.1. Sensory properties of mango frozen yogurt made with different commercial plain yogurt starters. n=75

Figure 3.1 presents the evaluated results of sensory attributes of mango frozen yogurt made with different yogurt starter cultures imported and local. The attributes assessed included color, smoothness, creaminess, flavor, and overall acceptability. The results indicate that there were no statistically significant differences ($p > 0.05$) in any of these sensory attributes between the two trials.

Although the differences were not statistically significant, some variations were observed. The bar shows that frozen yogurt made with imported yogurt culture received slightly higher mean scores in all attributes compared to the product made with the local starter. Especially, the most notable difference was in smoothness, and creaminess, where the commercial imported starter scored 7.07, 7.28 versus 6.73, 6.69, respectively for the local starter. This result suggests that panelists slightly preferred the texture of the product made with the imported culture. Moreover, texture score received the lowest compare to other three attribute, indicating the need to improve texture.

The result of the survey revealed a clear tendency that 63% of the panelists preferred the product made with the imported starter. This preference may be attributed to its slightly higher average scores in texture, flavor, and overall acceptability, even though these differences were not statistically significant.

In accordance to consumer preference result in **Fig. 3.2**, the imported starter was selected for further research and formulation development. In addition, some panelists suggested that the texture of both samples could be improved for a creamier and smoother mouthfeel. Therefore, future work will focus on adding and adjusting stabilizer and emulsifier levels to enhance the texture and overall sensory quality of the frozen yogurt.

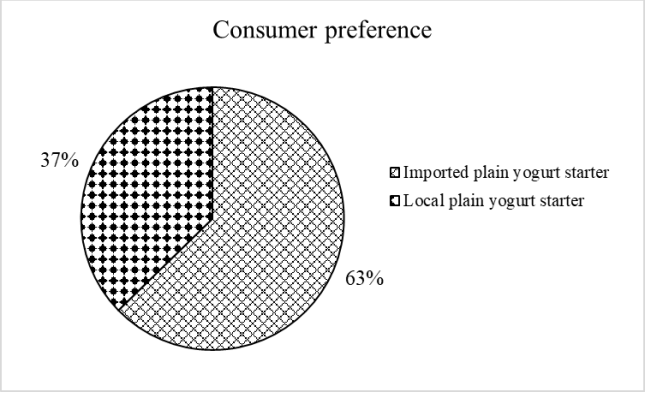


Fig. 3.2. Consumer preference of mango frozen yogurt made with different commercial plain yogurt starters. n=75

3.2 Effects of combined monoglyceride and guar gum

Table 3.2. Physicochemical of mango frozen yogurt made with different concentration of food additives

Trial	pH	Titratable acidity(%l.a)	Overrun (%)	Consistency (cm/s)
T0	4.71 ± 0.04^{a}	0.63 ± 0.09^{a}	25.1 ± 1.61^{b}	0.67 ± 0.01^{a}
T0.2	4.71 ± 0.01^{a}	0.62 ± 0.02^{a}	53.5 ± 2.38^{a}	0.63 ± 0.02^{b}
T0.4	4.70 ± 0.04^{a}	0.65 ± 0.01^{a}	55.2 ± 2.88^{a}	0.52 ± 0.00^{c}

T0: Trial control without additive; T0.2: Trial 0.2% of combined monoglyceride and guar gum, 1:1; T0.4: Trial 0.4% of combine monoglyceride and guar gum,1:1; Means with different superscripts (a, b, c) within the same column differ significantly at $p < 0.05$ (Tukey's HSD).

The effect of combined monoglyceride and guar gum concentrations on the physicochemical properties of mango frozen yogurt was evaluated using one-way ANOVA. The results are presented in **Table 3.2**.

There were no statistically significant differences ($p > 0.05$) observed in pH and titratable acidity among the samples. All treatments showed similar pH values, ranging from 4.70 ± 0.04 to 4.71 ± 0.04, and titratable acidity ranged between $0.62 \pm 0.02\%$ l.a and $0.65 \pm 0.01\%$ l.a. A previous study stated that the pH value of frozen yogurt samples can be in a range of 4.3 to 5.3 to meet the consumer satisfaction [11]. Moreover, similar to our results, E. Milani and A. Koocheki's result showed that the addition of guar gum did not have much effect on the final acidity of the frozen yoghurt

[5]. This indicates that the addition of monoglyceride and guar gum at different concentrations did not notably affect the acidity or fermentation profile of the frozen yogurt.

In contrast, significant differences were observed ($p < 0.05$) in overrun and consistency. Overrun, which is referred to the amount of air in ice cream, is important because it influences product quality and profits. Results showed that overrun increased substantially with the addition of stabilizers and emulsifiers, from 25.07 ± 1.69% in the control (T0) to 55.20 ± 2.88% in T0.4. However, overrun between T0.2 and T0.4 is not significantly different, meaning that the increase concentration of additives does not affect overrun. According to Moeenfard & Tehrani, the author revealed that overrun increased significantly with the addition of stabilizer which is similar to our study [6]. This suggests that monoglyceride and guar gum enhanced air incorporation, likely improving the smoothness and lightness of the product.

Consistency which was measured in centimeter per seconds, significantly decreased as the concentrations increased, from 0.67 ± 0.01 cm in the control to 0.52± 0.001 cm in T0.4. A lower flow distance indicates higher viscosity and improved structural stability, which is often desirable in frozen yogurt products. The same author also stated that stabilizers also have the function to improve the viscosity of mixture[6]. This thickening effect is attributed to the water-binding properties of guar gum and the emulsifying action of monoglycerides, which together stabilize the yogurt matrix[13].

Overall, the addition of monoglyceride and guar gum having no impact on pH or acidity. In contrast, it significantly improved the air incorporation and viscosity of mango frozen yogurt. These functional ingredients can be considered beneficial for improving product texture and stability.

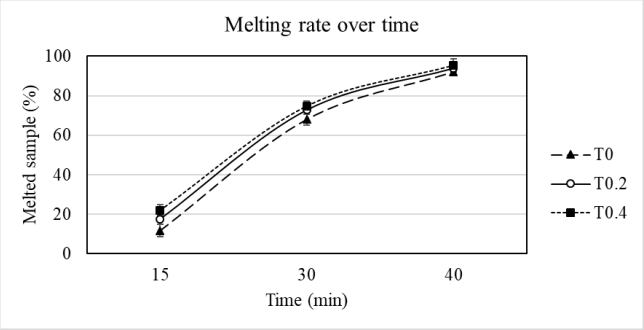


Fig. 3.3. Melting rate of mango frozen yogurt at different time intervals. T0: Trial control without additive; T0.2: Trial 0.2% of combined monoglyceride and guar gum, 1:1; T0.4: Trial 0.4% of combined monoglyceride and guar gum,1:1

The melting characteristics of frozen desserts play a crucial role in determining product quality because they affect the sensory attributes in shaping consumer perception and satisfaction. **Fig. 3.2.** illustrates the melting rate of mango frozen yogurt increased significantly ($p < 0.05$) with the addition of additives, especially at 15 and 30 minutes. Samples T0.2 and T0.4 had higher melting rates than the control (T0). One possible reason is that these samples had higher overrun, which means they contained more air. When more air is whipped into the product, it can make the structure lighter and less dense, so it melts faster [8]. Another reason might be the present of monoglyceride affecting the melting behavior. This could be due to the emulsifying properties of monoglycerides, which can improve fat dispersion but may reduce the overall solid structure, making the product melt faster [10]. However, at 40 minutes, the difference in melting rate was not significant, suggesting that monoglyceride and guar gum has a stronger effect in the early stage of melting. In addition, some of study revealed that the addition of some food additive such as stabilizers or sweeteners including sodium alginate, carrageenan, locust bean gum, stevia, and syrup in frozen desert improve the melting stability of ice cream[5], [6], [14]. Therefore, the combination of monoglyceride and guar gum at these concentrations might not be able to improve melting stability of frozen yogurt.

Overall, this result is important for understanding how stabilizer and emulsifier levels and type affect frozen yogurt stability during serving. It is suggested to use another type of stabilizers and/or syrup in order to improve melting stability.

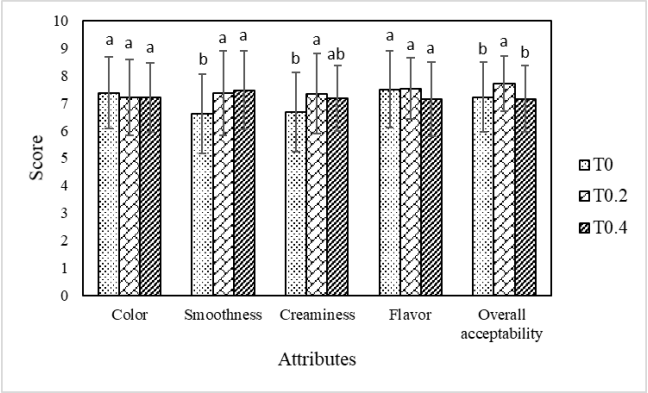


Fig. 3.4. Sensory properties of mango frozen yogurt made with different concentration of food additives. T0: Trial control without additive; T0.2: Trial 0.2% of combined monoglyceride and guar gum, 1:1; T0.4: Trial 0.4% of combined monoglyceride and guar gum,1:1; Means with different superscripts (a, b) within the same attribute differ significantly at $p < 0.05$ (Tukey's HSD). n=80

The sensory evaluation result in **Fig.3.4.** shows similarities and differences of each attribute based on the level of emulsifier and stabilizer added. In terms of color, and flavor, there were no significant different in all trials. This

suggests that the addition of food additive does not affect consumer preference of color and flavor. However, for smoothness, and creaminess T0.2 scored significantly the highest, indicating that the emulsifier and stabilizer effectively improved texture by enhancing mouthfeel and reducing ice crystal formation. In terms of overall acceptability, T0.2 was rated the highest and showed a significant preference over both T0 and T0.4. This suggests that a moderate level (0.2%) of combined monoglyceride and guar gum offers the best balance of improved texture while maintaining desirable sensory characteristics.

4. CONCLUSIONS

Mango frozen yogurt with different commercial yogurt starters slightly affect the quality of the product. There was no significant different in sensory evaluation in all attributes. Moreover, this study demonstrated that the incorporation of monoglyceride and guar gum with both concentrations (T0.2, T0.4) did not affect the pH and total acidity but it had significant effect on overrun, consistency, and melting rate at 15 and 30minuits. The results suggest that the use of a combined monoglyceride and guar gum can enhance the texture of mango frozen yogurt. However, the melting stability still requires improvement. Sensory evaluation's results revealed that trial 0.2% of food additives received the most preference from consumer. Nonetheless, this approach remains promising for commercial product development. Further research could analyze the microbiological quality as well as the nutritional value of the final product.

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to the Institute of Technology of Cambodia for providing the facilities that supported the "Techno Ice" ice cream project. We also greatly appreciate the valuable assistance of the laboratory staff from the Faculty of Chemical and Food Engineering.

REFERENCES

[1] R. Cedillos *et al.*, "Influence of Hesperidin on the Physico-Chemical, Microbiological and Sensory Characteristics of Frozen Yogurt †," *Foods*, vol. 13, no. 5, 2024, doi: 10.3390/foods13050808.

[2] M. Akbari, M. H. Eskandari, and Z. Davoudi, "Application and functions of fat replacers in low-fat ice cream: A review," *Trends Food Sci. Technol.*, vol. 86, no. January 2017, pp. 34–40, 2019, doi: 10.1016/j.tifs.2019.02.036.

[3] EuroCham Cambodia, "Sourcing from Cambodia: MANGO Product & Supplier Brochure." 2021. [Online]. Available: https://www.eurocham-cambodia.org/uploads/9054a-sourcing-from-cambodia_mango-latest.pdf

[4] F. Javidi, S. M. A. Razavi, F. Behrouzian, and A. Alghooneh, "The influence of basil seed gum, guar gum and their blend on the rheological, physical and sensory properties of low fat ice cream," *Food Hydrocoll.*, vol. 52, pp. 625–633, 2016, doi: 10.1016/j.foodhyd.2015.08.006.

[5] E. Milani and A. Koocheki, "The effects of date syrup and guar gum on physical, rheological and sensory properties of low fat frozen yoghurt dessert," *Int. J. Dairy Technol.*, vol. 64, no. 1, pp. 121–129, 2011, doi: 10.1111/j.1471-0307.2010.00631.x.

[6] M. Moeenfard and M. M. Tehrani, "Effect of Some Stabilizers on the Physicochemical and Sensory Properties of Ice Cream Type Frozen Yogurt," *J. Agric. Environ. Sci*, vol. 4, no. 5, pp. 584–589, 2008.

[7] R. Rezaei, M. Khomeiri, M. Aalami, and M. Kashaninejad, "Effect of inulin on the physicochemical properties, flow behavior and probiotic survival of frozen yogurt," *J. Food Sci. Technol.*, vol. 51, no. 10, pp. 2809–2814, 2014, doi: 10.1007/s13197-012-0751-7.

[8] M. R. Muse and R. W. Hartel, "Ice cream structural elements that affect melting rate and hardness," *J. Dairy Sci.*, vol. 87, no. 1, pp. 1–10, 2004, doi: 10.3168/jds.S0022-0302(04)73135-5.

[9] S. Nagaoka, "Yogurt production," *Methods Mol. Biol.*, vol. 1887, pp. 45–54, 2019, doi: 10.1007/978-1-4939-8907-2_5.

[10] H. D. Goff and R. W. Hartel, "Ice cream, seventh edition," *Ice Cream, Seventh Ed.*, pp. 1–462, 2013, doi: 10.1007/978-1-4614-6096-1.

[11] U. Isik, D. Boyacioglu, E. Capanoglu, and D. Nilufer Erdil, "Frozen yogurt with added inulin and isomalt," *J. Dairy Sci.*, vol. 94, no. 4, pp. 1647–1656, 2011, doi: 10.3168/jds.2010-3280.

[12] T. Dan, H. Hu, J. Tian, B. He, J. Tai, and Y. He, "Influence of Different Ratios of Lactobacillus delbrueckii subsp. bulgaricus and Streptococcus thermophilus on Fermentation Characteristics of Yogurt," *Molecules*, vol. 28, no. 5, 2023, doi: 10.3390/molecules28052123.

[13] R. Rezaei, M. Khomeiri, M. Kashaninejad, and M. Aalami, "Effects of guar gum and arabic gum on the physicochemical, sensory and flow behaviour characteristics of frozen yoghurt," *Int. J. Dairy Technol.*, vol. 64, no. 4, pp. 563–568, 2011, doi: 10.1111/j.1471-0307.2011.00705.x.

[14] A. Arslaner, M. A. Salik, S. Özdemir, and A. Akköse, "Yogurt ice cream sweetened with sucrose, stevia and honey: Some quality and thermal properties," *Czech J. Food Sci.*, vol. 37, no. 6, pp. 446–455, 2019, doi: 10.17221/311/2018-CJFS.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Nutritional quality of Farmed striped snakehead (*C. striata*): Fatty acid profile and proximate composition analysis

Sovannmony Lay[1,2,3], Caroline Douny[3], Marie-Louise Scippo[3], Vattana Mom[1,2,3], Hasika Mith[1,2], Chanvorleak Phat[1,2*]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *University of Liège, Laboratory of Food Analysis, FARAH -Veterinary Public Health, Quartier Vallée 2-Avenue de Cureghem 10, 4000, Liège, Belgium*

[*] *Corresponding author: phatchanvorleak@itc.edu.kh*

Abstract: *Fish is a vital source of high-quality protein and essential fatty acids, particularly long-chain omega-3 polyunsaturated fatty acids (LC-PUFA), which play crucial roles in cardiovascular health, brain function, and inflammation regulation. The farmed striped snakehead (Channa striata) is a commercially significant freshwater fish species widely cultured in Cambodia due to its high nutritional value and economic importance. While it is traditionally recognized for its high protein content, its lipid composition, particularly its fatty acid profile, is of increasing interest due to its implications for human health. This study aims to assess the proximate composition and fatty acid profile of Farmed striped snakehead from two Farms in Kampong Thom province, Cambodia, to evaluate its nutritional quality. Proximate composition, including moisture, protein, lipid, and ash content, was determined using standard analytical methods, while fatty acid composition was analyzed by gas chromatography-mass spectrometry (GC-MS). Proximate analysis revealed that Farmed C. striata contains high protein content (ranging from 17.24% to 18.71% W.W.), making it a significant protein source in the diet. Its lipid content ranged from 3.14% to 4.11% W.W., and fatty acid profiling indicated that C. striata is rich in essential fatty acids, particularly omega-3 (n-3) and omega-6 (n-6) polyunsaturated fatty acids (PUFAs). Notably, high levels of eicosapentaenoic acid (EPA) (up to 2.6 g/100 g of total fatty acids) and docosahexaenoic acid (DHA) (up to 6.7 g/100 g of total fatty acids) were detected. The semi-essential fatty acids (EPA+DHA) ranged from 2.2 to 9.3 g/100 g of total fatty acids. The daily consumption requirements of EPA + DHA recommended by the European Food Safety Authority (EFSA) for the prevention of primary cardiovascular illnesses were satisfied by fish species in the current investigation. These findings highlight the nutritional significance of Farmed C. striata, emphasizing its potential as a high-quality dietary protein and essential fatty acid source. Dietary diversification through increased farmed fish consumption can help address potential nutrient deficiencies and improve Cambodian people's public health.*

Keywords: Striped snakehead; *Channa striata*; fatty acid profile; omega-3 fatty acids; proximate composition

1. INTRODUCTION

Food security is a severe issue that has been faced in developing countries. Malnutrition and starvation are major challenges for people in rural areas. Moreover, protein energy malnutrition specifically presents a significant barrier in achieving nutritional security for the majority of South East Asian countries [1]. Fish is consumed worldwide and contributes significantly to our daily diet because of its high protein, amino acids, polyunsaturated fatty acid (PUFA), minerals, and vitamin content [2], [3].

In Cambodia, fish is the primary source of macro and micronutrients, accounting for up to 37% of total protein intake. In 2014, Cambodians consumed an average of 63 kilograms of freshwater fish per capita, but this decreased to 33 kg/capita/year in 2019 [4], [5]. In general, inland fisheries account for most of fish consumption and supply to local and international markets, with marine fisheries accounting for a small percentage; however, the collapse in inland fisheries

[1*] Corresponding author: Chanvorleak Phat
E-mail: phatchanvorleak@itc.edu.kh; Tel: +855-93 418 000

fish stocks prompted the aquaculture industry to contribute to rural livelihoods [6].

Based on our knowledge, currently, there is a lack of information on the nutritional and physicochemical composition, especially EPA and DHA of farmed freshwater fish in Cambodia. Therefore, the objective of this study aimed to assess the proximate composition and fatty acid profile of Farmed striped snakehead from two farms in Kampong Thom province, Cambodia, to evaluate its nutritional quality. Proximate composition, including moisture, protein, lipid, and ash content.

2. METHODOLOGY

2.1 Sample preparation

The fish samples were collected from Farmers during their harvested periods. *Channa Striata* (Trey Ros) was collected from two different farms in Kampong Thom (Farm 1: E12°55′19″N 104°33′23.48″, Farm 2: 12°42′32.43″N 104°53′19.42″E). 3 kg of *C. striata* from each farm were packed in a plastic bag on which the location of sampling, date of collection and the local name of the fish species were recorded. Then, the samples were immediately stored under ice in a polystyrene box. After that, they were transferred to the Institute of Technology of Cambodia (ITC) laboratory. Fish samples were prepared by removing the intestine, scale, fins and gills. Moreover, fish preparation was based on the consumable part of eating behaviors for Cambodian people. Then, samples were cleaned with distilled water before grinding them by using a grinding machine. For proximate composition analysis, fish samples were frozen at -20 ℃ for the lyophilization process. Frozen samples were dried by using a freeze dryer machine (Christ, Alpha 1-4 LDplus) at the temperature -50 ℃ and pressure 0.04 mbar for 44 ± 2 h. After drying, the lyophilized samples were stored in the freezer (-20 ℃) for further analysis.

2.2 Moisture analysis

The moisture content was analyzed by the AOAC method [7]. 5 g of fish sample were put in aluminium caps and dried in an oven at 105 ± 2°C for 3 hours or until a stable mass. After drying, a sample was taken from the oven, cooled in a desiccator for 15 min and weighed until a constant mass. Each sample was analyzed in triplicate. The moisture content was determined by the formula below:

$$\% \text{ Moisture} = [(m2-m3)/(m2-m1)]x\ 100 \qquad \text{(Eq. 1)}$$

where:
m1 = mass of empty cup (g)
m2 = mass of cup with test portion before drying (g)
m3 = mass of cup with test portion after drying (g)

2.3 Ash content analysis

The ash content was analyzed by AOAC method [7]. The dried samples from the moisture determination were placed into a furnace. The temperature was increased and maintained at 550°C for 3 hours. The samples were cooled down in the desiccator. The final weight of each sample was recorded. Each sample was analyzed in triplicate. The formula for calculating ash content was:

$$\% \text{ Ash} = [(m2-m1)/(m3-m1)]x\ 100 \qquad \text{(Eq. 2)}$$

where:
m1 = mass of porcelain crucible (g)
m2 = mass of test sample after drying (g)
m3 = mass of porcelain with test portion after drying (g)

2.4 Protein analysis

Protein content was determined by using the validated Kjeldahl method (AOAC, 2012) with the help of distillation units. Briefly, total nitrogen was detected in the following procedure. First, the mineralization step, where organic nitrogen is transformed into mineral nitrogen in the form of $(NH_4)_2SO_4$. The organic matter was digested by a mixture of (H_2SO_4 + H_3PO_4 and H_2O_2) having an oxidative action, in the presence of a catalyst. This catalyst is composed of $CuSO_4$, TiO_2 and K_2SO_4 which increases the mineralization temperature. There was a reduction of organic nitrogen into ammonia NH_3, retained in the acid digested in the form of ammonium sulfate $(NH4)_2SO_4$. Then, it was moved to the distillation and titration steps. The ammonia in the digested was displaced by adding an excess of a strong NaOH base. During distillation, it was entrained by steam and was titrated in the presence of a colored indicator, directly by a recovery in boric acid. Each sample was analyzed one time since the method was validated. Total protein content was calculated as the following formula and expressed as a percentage (%) in wet weight basis.

$$\text{Protein (g/100g w.b.)} = N \times 6.25 \qquad \text{(Eq. 3)}$$

2.5 Fat extraction and fatty acid analysis by GC-MS

Total fat was extracted followed by the method of Folch et al., [8]. Briefly, approximately 1 g of lyophilized sample was weighed, 1 ml of water was added, it was then vortexed and let stand for 10 min. After that, 40 ml of chloroform/methanol (2/1; v/v) was added into the test tubes and the samples were placed on the rotary shaker for an overnight agitation. Next, the samples were filtered through a funnel and paper filter containing Na_2SO_4. Furthermore, 8 ml of 0.88% KCl were added, and the samples were centrifuged at 4000 rpm for 10 minutes. Subsequently, the upper phase (methanol/water) was removed with a water pump and a

precise volume of the chloroform phase extract was poured into a test tube and stored in -80℃ for fatty acid analysis.

The fatty acid profile was determined by analysis of the fatty acid methyl esters (FAME) by gas chromatography-mass spectrometry (GC-MS) [4]. The method consists of the methylation of the fat extracted from samples in the presence of the internal standard nonadecanoic acid (C19:0). FAME were separated on Focus GC gas chromatograph (Thermo Fisher Scientific) using a CP-Sil88 column for FAME (100 m × 0.25 mm, 0.2 μm, Varian; Agilent Technologies, Santa Clara, CA). The GC conditions were: inlet temperature set at 250°C, spitless injection, helium used as the carrier gas at 1.5 mL min−1, temperature program initiated at 55°C and hold for 1 min, increased at 5°C min^{-1} to 180°C, then 10°C min-1 to 200°C and hold for 15 min, then an increase of 10°C min^{-1} to 225°C and hold for 14 min; total run time was 59.50 min. The injection volume was 1 μL. The MS conditions were: transfer line set at 250°C and ion source at 220°C. The FAME were detected using selected ion monitoring (SIM) mode. The peaks were identified by comparing their mass spectrum and retention times with those of the corresponding standards (C10:0 to C24:0). In each chromatographic run, two ions were monitored for each fatty acid analyzed, which facilitated detection and quantitative analysis: m/z 74 and 143 for saturated fatty acids (SFA), 79 and 91 for monounsaturated fatty acids (MUFA) and polyunsaturated fatty acids (PUFA). The sum of SFA, MUFA, and PUFA were individually expressed as the percentage of the total fatty acids. For quantification, an 8-points calibration curve containing standard solutions of each fatty acid to be analyzed and an internal standard (C19:0) was methylated, extracted, and injected with each series of analysis. The response (ratio between fatty acid methyl esters and the internal standard peak areas) was plotted against standard concentrations.
A quality control sample consisting of a CRM of a blend of beef and pork fat was also analyzed with each series of analysis.

2.6 Nutritional quality indices

The results from the fatty acid profile were used to compute the lipid quality indices, using three indicators such as the thrombogenicity index (TI), atherogenicity index (AI), and the hypocholesterolemic/hypercholesterolemic (h/H) ratio according to the methods [9], [10]. The nutritional quality indices (NQI) were calculated as follows:

$$AI = (4 \times C14{:}0 + C16{:}0 + C18{:}0)/(\Sigma MUFA + \Sigma n\text{-}3PUFA + \Sigma n\text{-}6PUFA) \quad \text{(Eq.4)}$$

$$TI = (C14{:}0 + C16{:}0 + C18{:}0)/((0.5 \times \Sigma MUFA) + (0.5 \times \Sigma n\text{-}6PUFA) + (3 \times \Sigma n\text{-}3PUFA) + (\Sigma n\text{-}3PUFA/\Sigma n\text{-}6PUFA)) \quad \text{(Eq.5)}$$

$$h/H = (C18{:}1n\text{-}9 + C18{:}2n\text{-}6 + C20{:}4n\text{-}6 + C18{:}3n\text{-}3 + C20{:}5n3 + C22{:}5n3 + C22{:}6n3)/(C14{:}0 + C16{:}0) \quad \text{(Eq.6)}$$

3. RESULTS AND DISCUSSION

3.1 Proximate composition of Channa striata

Proximate composition including moisture, ash, protein, and fat are presented in Table 1.

Table 1. Proximate composition of Farmed *Channa Striata*

Parameters	Farm 1	Farm 2
Moisture (%)	76.3 ± 0.6	77.7 ± 1.1
Ash (%)	1.33 ± 0.2	0.69 ± 0.2
Protein (%)	18.7 ± 0.0	17.2 ± 0.0
Fat (%)	4.1 ± 0.0	3.2 ± 0.0

Data are presented as mean ± standard deviation.

Moisture content of these two samples ranged from 76.3 ± 0.6 % (Farm 1) to 77.7 ± 1.1 % (Farm 2), ash content ranged from 1.33 ± 0.2 % (Farm 1) and 0.69 ± 0.2% (Farm 2), protein content ranged from 18.7 ± 0.0 % (Farm 1) to 17.2 ± 0.0 % (Farm 2), and lipid content ranged from 4.1 ± 0.0 % (Farm 1) and 3.2 ± 0.0 % (Farm 2). According to this result, Farm 1 showed higher content in ash, protein, and lipid content, except for moisture content, compared to Farm 2. The variations of proximate composition in these two farms might be due to life cycle variations, types of cultured system, and types of diet of these two farms [11].

In general, fish were categorized into distinct groups as lean fish (<2%), low fat (2 to 4 %), medium fat (4 to 8 %), and high fat (>8 %) based on their fat content [12]. Based on the result, fat content of *C. Striata* in Farm 1 was higher than *C. Striata* in Farm 2. The difference in fat content in these two farms may be attributed to the usage of different type of feed.

Moreover, the protein content in these two samples was considerable and would be suitable for a healthy diet, based on the European Regulation No. 1924/2006, which defines a food product as 'high in protein' if at least 20% of its energy value is provided by protein [13]. In addition, in terms of availability and affordability, fish are advantageous over other plant and animal protein sources and can be important for fulfilling the protein requirements of people in developing countries [12].

3.2 Fatty acid profile

Twenty-nine fatty acids were identified in farmed *C. striata* (Table 2). These two samples showed a similar fatty acid pattern (Saturated fatty acids: SFA > Polyunsaturated fatty acids: PUFA > Monounsaturated fatty acids: MUFA). SFA were dominant in these two samples: 49.3 g/100 g fatty acids in Farm 1 and 41.7 g/100 g fatty acids in Farm 2, followed by PUFA with levels between 36.1 g/100 g fatty acids (Farm 1) and 33.5 g/100 g fatty acids (Farm 2), while MUFA were the minor fatty acid category, varied from 14.6

g/100 g fatty acids (Farm 1) to 24.9 g/100 g fatty acids (Farm 2).

Table 2. Fatty acid profile of farmed *Channa Striata* (expressed as g/100 g of total fatty acids).

Fatty acid	Farm 1	Farm 2
C10:0	<LOQ	<LOQ
C12:0	<LOQ	<LOQ
C13:0	<LOQ	<LOQ
C14:0	2.7	0.9
C16:0	33.7	29.5
C17:0	1.9	<LOQ
C18:0	11.0	11.2
C20:0	<LOQ	<LOQ
C22:0	<LOQ	<LOQ
C24:0	<LOQ	<LOQ
C16:1	5.0	1.8
C17:1	<LOQ	<LOQ
C18:1,9t	<LOQ	<LOQ
C18:1,9c	9.6	20.5
C18:1,11c	<LOQ	2.3
C22:1	<LOQ	0.4
C18:2,9c,12c	17.0	24.8
C18:2,9t,12t	<LOQ	<LOQ
C18:2,9c,11t	<LOQ	<LOQ
C18:2,9t,11t	<LOQ	<LOQ
C18:3 (n-6)	0.9	2.5
C20:2	<LOQ	0.7
C20:4	3.3	1.7
C18:3 (ALA)	3.3	1.6
C18:4	<LOQ	<LOQ
C20:3	<LOQ	<LOQ
C20:5 (EPA)	2.6	<LOQ
C22:5 (DPA)	2.3	<LOQ
C22:6 (DHA)	6.7	2.2
SFA	49.3	41.7
MUFA	14.6	24.9
PUFA	36.1	33.5
Total n-6	21.2	29.7
Total n-3	14.9	3.8
n-6/n-3	1.4	7.8
PUFA/SFA	0.7	0.81
EPA+DHA	9.3	2.2
AI	0.9	0.7
TI	0.7	1.04
h/H	1.0	1.0

SFA: Saturated fatty acid; MUFA: Monounsaturated fatty acid; PUFA: Polyunsaturated fatty acid; AI: Atherogenicity index; IT: Thrombogenicity index; (h/H); hypocholesterolemic/hypercholesterolemic ratio; <LOQ: below the limit of quantification.

Within SFA, palmitic acid (C16:0) was the predominant in both samples, varying from 29.5 g/100 g fatty acids (Farm 2) to 33.7 g/100 g fatty acids (Farm 1). While oleic acid (C18:1 n-9) was the predominant of MUFA for both fish. This result is similar to what was reported in previous studies [14], [15], [16].

For omega-3 fatty acids, ALA (C18:n-3), EPA (C20:5 n-3) and DHA (C22:6 n-3) exhibited the highest values for the two samples with a higher DHA content (2.2 to 6.7 g/100 g fatty acids), followed by ALA (1.6 to 3.3 g/100 g fatty acids), and EPA was 2.6 g/ 100 g fatty acids in Farm 1, while Farm 2 was below the LOQ value of the validated method.

The result of fatty acid compositions showed that there were considerable differences among the farmed fish species in this study. These results may explain that the fatty acid composition was affected by several variables, including the species, fish size and age, seasonal conditions, geographic location, and type of feed [16], [17].

PUFA/SFA, n-6/n-3, AI, TI, and h/H were evaluated to assess the lipid quality of farmed *C. Striata* in this study. PUFA/SFA ratio varied from 0.7 to 0.81, and ratio of n-6/n-3 varied from 1.4 to 7.8. In addition, the AI and TI index values ranged from 0.7 to 0.9, and 0.7 to 1.04, respectively. The value of the h/H index in these two samples was 1.0. A minimum PUFA/SFA ratio of 0.45 and a maximum n-6/n-3 ratio of 4.0 are recommended to achieve a balanced fatty acid intake [18]. Furthermore, the indices of AI and TI are related to pro- and anti-atherogenic and pro- and antithrombogenic fatty acids. High AI and TI values might stimulate platelet aggregation and subsequent thrombus and atheroma formation in the cardiovascular system. Thus, lower AI and TI values are desirable to prevent cardiovascular disorders. The h/H index refers to the ratio of hypocholesterolemic over hypercholesterolemic fatty acids and is related with the specific effects of fatty acids on cholesterol metabolism [19].

Being essential nutrients, EPA and DHA need to be obtained through diet, and therefore, searching for affordable sources of these PUFAs is important for consumer guidance [12]. Likewise, *C. Striata* from Farm 1 was one of the major sources of EPA and DHA (EPA+DHA: 9300 mg/100 g). This value is suitable and corresponds to the European Food Safety and Authority (EFSA) recommendation for adults to have a daily maximum dietary intake of 500 mg of EPA and DHA [20].

4. CONCLUSIONS

This study provides the first characterization of the protein source, lipid nutritional values, and fatty acid profiles of the economically significant farmed striped snakehead (*C. striata*) in Cambodia.

The results revealed that Cambodia striped snakehead used in this study provided a considerable source of proteins up to 18%, and a high level of essential fatty acids, especially EPA and DHA (9300 mg/100 g). Therefore, greater encouragement should be given to the consumption of farmed striped snakehead, which would benefit the aquaculture

system in Cambodia to improve the processing and marketing of all fish species as well as improving the nutritional profile of the Cambodian population. In future studies, it would be useful to determine the nutritional profile of the striped snakehead feed commonly used in Cambodia and investigate the impact of feed composition on nutritional profile of striped snakehead.

ACKNOWLEDGMENTS

The authors would like to express thanks to the CAMBOFISH project, funded by ARES-CCD, for financially supporting this research and the laboratory of the Faculty of Chemical and Food Engineering, the Institute of Technology of Cambodia, and the Laboratory of Food Analysis, FARAH-Veterinary Public Health, the University of Liège, Belgium for facilitating laboratory equipment for this experiment.

REFERENCES

[1] S. R. Jarapala, "Fish nutrition, potential health benefits and heavy metal research in India: Steps towards new horizon," vol. 6, 2017.

[2] B. Abraha, H. Admassu, A. Mahmud, N. Tsighe, X. W. Shui, and Y. Fang, "Effect of processing methods on nutritional and physico-chemical composition of fish: a review," *MOJFPT*, vol. 6, no. 4, Aug. 2018, doi: 10.15406/mojfpt.2018.06.00191.

[3] S. NurSyahirah and A. Rozzamri, "Effects of frying on fish, fish products and frying oil – a review," *Food Res.*, vol. 6, no. 5, pp. 14–32, Sep. 2022, doi: 10.26656/fr.2017.6(5).608.

[4] Douny C., El Khoury R., Delmelle J., Brose F., Degand G., Moula N., Farnir F., Clinquart A., Maghuin-Rogister G., Scippo M.L. Effect of storage and cooking on the fatty acids profile of omega-3 enriched eggs and pork meat marketed in Belgium. *Food Science & Nutrition* (2015), 3(2), 140-152 (https://doi.org/10.1002/fsn3.197)

[5] C. Vilain, E. Baran, G. Gallego, and S. Samadee, "Fish and the nutrition of rural cambodians," *Asian Journal of Agriculture and Food Sciences*, vol. 04, no. 01, 2016.

[6] R. Kunthy, "Current status of sustainable aquaculture and resource enhancement in cambodia," 2021.

[7] AOAC, *Official Methods of Analysis of AOAC International*, 19th ed. Gaithersburg, Md: AOAC International, 2012.

[8] J. Folch, M. Lees, and G. H. S. Stanley, "A simple method for the isolation and purification of total lipides from animal tissues," *Journal of Biological Chemistry*, vol. 226, no. 1, pp. 497–509, May 1957, doi: 10.1016/S0021-9258(18)64849-5.

[9] S. Chakma, M. A. Rahman, M. A. B. Siddik, M. S. Hoque, S. M. Islam, and I. N. Vatsos, "Nutritional profiling of wild (*Pangasius pangasius*) and farmed (*Pangasius hypophthalmus)* pangasius catfish with implications to human health," *Fishes*, vol. 7, no. 6, Art. no. 6, Dec. 2022, doi: 10.3390/fishes7060309.

[10] T. L. V. Ulbricht and D. A. T. Southgate, "Coronary heart disease: seven dietary factors," *The Lancet*, vol. 338, no. 8773, pp. 985–992, Oct. 1991, doi: 10.1016/0140-6736(91)91846-M.

[11] S. Islam, S. Bhowmik, P. R. Majumdar, G. Srzednicki, M. Rahman, and Md. A. Hossain, "Nutritional profile of wild, pond-, gher- and cage-cultured tilapia in Bangladesh," *Heliyon*, vol. 7, no. 5, p. e06968, May 2021, doi: 10.1016/j.heliyon.2021.e06968.

[12] B. P. Mohanty *et al.*, "DHA and EPA Content and Fatty Acid Profile of 39 Food Fishes from India," *BioMed Research International*, vol. 2016, pp. 1–14, 2016, doi: 10.1155/2016/4027437.

[13] European, Regulation (EC) No 1924/2006 of the European Parliament and of the Council of 20 December 2006 on nutrition and health claims made on foods, vol. 404. 2006. Accessed: Apr. 19, 2025. [Online]. Available: http://data.europa.eu/eli/reg/2006/1924/oj/eng

[14] M. Akter, H. Faruque, M. Hasan, and M. S. Rahman, "Proximate composition, amino acids and fatty acids profiles of wild and cultured Climbing Perch, Anabas testudineus (Bloch, 1795)," *Bangladesh Journal of Zoology*, vol. 48, pp. 365–378, Apr. 2021, doi: 10.3329/bjz.v48i2.52376.

[15] S. Sroy, E. Arnaud, A. Servent, S. In, and S. Avallone, "Nutritional benefits and heavy metal contents of freshwater fish species from Tonle Sap Lake with SAIN and LIM nutritional score," *Journal of Food Composition and Analysis*, vol. 96, p. 103731, Mar. 2021, doi: 10.1016/j.jfca.2020.103731.

[16] N. N. Memon, F. N. Talpur, M. I. Bhanger, and A. Balouch, "Changes in fatty acid composition in muscle of three farmed carp fish species (Labeo rohita, Cirrhinus mrigala, Catla catla) raised under the same conditions," *Food Chemistry*, vol. 126, no. 2, pp. 405–410, May 2011, doi: 10.1016/j.foodchem.2010.10.107.

[17] S. González, G. J. Flick, S. F. O'Keefe, S. E. Duncan, E. McLean, and S. R. Craig, "Composition of farmed and wild yellow perch (Perca flavescens)," *Journal of Food Composition and Analysis*, vol. 19, no. 6–7, pp. 720–726, Sep. 2006, doi: 10.1016/j.jfca.2006.01.007.

[18] B. L. Rodrigues, M. L. G. Monteiro, A. C. Vilhena Da Cruz Silva Canto, M. P. D. Costa, and C. A. Conte-Junior, "Proximate composition, fatty acids and nutritional indices of promising freshwater fish species from Serrasalmidae family," *CyTA - Journal of Food*, vol. 18, no. 1, pp. 591–598, Jan. 2020, doi: 10.1080/19476337.2020.1804463.

[19] X. Zhang *et al.*, "Fatty acid composition analyses of commercially important fish species from the Pearl

River Estuary, China," *PLoS ONE*, vol. 15, no. 1, p. e0228276, Jan. 2020, doi: 10.1371/journal.pone.0228276.
[20] EFSA, "Scientific Opinion on Dietary Reference Values for fats, including saturated fatty acids, polyunsaturated fatty acids, monounsaturated fatty acids, trans fatty acids, and cholesterol," *EFS2*, vol. 8, no. 3, Mar. 2010, doi: 10.2903/j.efsa.2010.1461.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Optimisation of Method Liquid Extraction of Aroma Compound (2-acetyl-1-pyrroline) from Phka Rumdoul Rice Variety

Sokharoath Sary [1], Chanvorleak Phat [1,2], Stephane Bellafiore [1,3], Sereyvath Yoeun [1,2*]

[1] *Food Technology and Nutrition Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Plant Health Institute of Montpellier (PHIM), IRD, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France*

[*] *Corresponding author: svath@itc.edu.kh*

Abstract: *The fragrance of rice significantly influences consumer preference, with the jasmine-like aroma of aromatic rice. 2-Acetyl-1-pyrroline (2-AP) is the key aroma compound responsible for the characteristic perfume of fragrant rice. Optimisation of its extraction method is crucial to improve analytical accuracy and maximise yield. This study aims to optimise the liquid extraction method of 2-AP from samples of Phka Rumdoul rice variety collected in Preah Vihear province by investigating the effects of three key factors: type of solvent, temperature, and sonication time. Four solvents - methanol, ethanol, propanol, and acetonitrile were evaluated for their extraction efficiency.. The effect of temperature was studied at 40°C, 60°C, and 80°C. Sonication time was varied at 30 min, 45 min, and 60 min, and the amount of 2-AP in all samples was analyzed by gas chromatography-mass spectrometry (GC-MS). The results showed that extraction by ethanol at a temperature of 80°C and a sonication time of 30 min provides the optimal condition for 2-AP extraction, thus improving the extraction efficiency for quality control and research in aromatic rice production.*

Keywords: Phka Rumdoul Rice variety; Jasmine-like fragrance; 2-acetyl-1-pyrroline; Liquid Extraction; Gas chromatography-mass spectrometry.

1. INTRODUCTION

Rice (*Oryza sativa*) is one of the world's most important staple foods, nourishing over half of the global population. Fragrant rice is highly valued for its superior grain quality, distinct aroma, and rich flavor [1]. As consumers become increasingly concerned about the quality of the rice they consume, ensuring high-quality rice plays a crucial role in connecting farmers with consumers [2].

Among various rice varieties, Phka Rumduol is a Cambodian fragrant rice that has gained popularity for its jasmine-like scent, soft grain, and cultural significance. In addition to its role in food security, aromatic rice also plays a major part in the economic development of rice-producing countries by enhancing export value and farmer income [3]. The market demand for high-quality aromatic rice continues to rise globally, particularly in Asia, Europe, and North America, where consumers prefer rice not only for its taste and texture but also for its aromatic properties.

The fragrance of aromatic rice is primarily attributed to the presence of 2-acetyl-1-pyrroline (2-AP), a volatile heterocyclic compound that imparts a characteristic popcorn- or pandan-like scent. Even in trace concentrations, 2-AP significantly contributes to the overall aroma profile and

Consumer perception of rice quality [4, 5]The amount of 2-AP in aromatic rice ranges from 0.1 to 1.9 µg/mL [6]. The formation of 2-AP in rice is influenced by genetic, environmental, and post-harvest factors. Research has shown that a mutation in the Betaine Aldehyde Dehydrogenase 2 (BADH2) gene is associated with 2-AP accumulation, particularly through the disruption of the betaine aldehyde dehydrogenase pathway, which normally suppresses 2-AP accumulation in non-aromatic rice [7].

However, 2-AP is highly volatile and chemically unstable, making it susceptible to degradation under various conditions such as heat, pH, light, and oxidative environments. These characteristics present significant analytical challenges in accurately quantifying the compound in rice. Furthermore, the natural concentration of 2-AP in rice is extremely low, typically in the range of microorganisms per

[1]* Corresponding author: Sereyvath Yoeun
E-mail: svath@itc.edu.kh ; Tel: +855-10 348 588

kilogram, necessitating the use of highly sensitive extraction and analytical techniques [7,8].

2-Acetyl-1-pyrroline has an IUPAC name of 1-(3,4-dihydro-2H-pyrrol-5-yl) ethenone [10]. The structural homolog of this compound is 6-acetyl-2,3,4,5-tetrahydropyridine. 2-AP is a substituted pyrroline, a cyclic imine, and a ketone. Its molecular formula is CH_9NO, with a molecular weight of 111.14 g/mol. The compound has a relatively low melting point and a high boiling point, with values of 19°C at 760 mm Hg and between 182°C and 183°C at 760 mm Hg. Its vapor pressure is 0.793 mm/Hg at 25°C, and its flash point is 151°F. 2-AP appears as a colorless to yellow solid with an assay ranging from 95% to 100%. It is classified as a flavor and fragrance chemical compound [11].

Gas chromatography-mass spectrometry (GC-MS) has become the gold standard for detecting and quantifying 2-AP due to its superior sensitivity, specificity, and reliability. However, the accuracy of GC-MS analysis greatly depends on the efficiency of the sample preparation and extraction method used. Among the various approaches for extracting volatile compounds from plant matrices, liquid extraction coupled with sonication is widely recognized for its simplicity, cost-effectiveness, and ability to disrupt cellular structures, thus enhancing compound release [12].

Several studies on the extraction method of 2AP have been conducted. The extraction method utilized is the solvent extraction method. One of the favourable methods is the solvent extraction method. Therefore, the study of the solvent effects is very important for the screening and solvent selection of the extraction, fractionation, and purification steps in the plant processing. By understanding the solvent properties, component (solute) properties, and solvent-solute interaction, rapid fractionation and isolation of desired components can be achieved. The quantity of extracted component was dependent on its stability in the solvent used. [11]. In addition, the selection of an appropriate solvent is critical in liquid extraction, as it influences the solubility and recovery of target compounds [13]. Ethanol is considered an effective solvent for extracting 2-acetyl-1-pyrroline from rice due to its optimal polarity, which enhances the solubilization of this key aroma compound. Studies have shown that ethanol performs better than other solvents like methanol and propanol, likely because its polarity closely matches that of 2-AP, resulting in more efficient extraction [11]. Moreover, ethanol is less toxic and more environmentally friendly than other organic solvents, making it suitable for food-grade extractions and analytical purposes [14]. In addition, sonication parameters such as temperature and time significantly affect extraction efficiency. Sonication facilitates the release of intracellular contents by generating acoustic cavitation, which breaks down cell walls and enhances solvent penetration. However, improper control of sonication parameters may lead to the loss or degradation of heat-sensitive compounds like 2-AP [15].

Despite various efforts to standardize the extraction of 2-AP, there is still a need for optimized, reproducible methods that can be applied in both research and industrial quality control settings, particularly in developing countries where analytical resources are limited. This study aims to optimize the liquid extraction method for 2-Acetyl-1-pyrroline from Phka Rumdoul rice samples by evaluating the effects of solvent, temperature, and sonication time to improve analytical accuracy and maximize extraction yield.

2. METHODOLOGY

2.1 Materials and chemicals

Ultrasonic Bath (Elmasonic Select 30, Elma Schmidbauer, Germany) with a frequency wave sound 50/60 Hz was used. Rice Polishing machine model K-1 and Rice Husking machine model P-1 were purchased from Lek Seng Huat, Thailand. The analysis was conducted using HPLC-grade solvents of ethanol (Fisher ChemAlart Guide, Korea), methanol anhydrous for analysis (Merck KGaA, Germany), acetonitrile gradient grade for liquid chromatography (Merck KGaA, Germany), and 2-propanol gradient grade for liquid chromatography (Merck KGaA, Germany).

2.2 Experimental design

A total of 500 grams of Phka Rumdoul rice was collected from Tmat Paey, Choam Khsan district, Preah Vihear province, Cambodia (13° 58′ 25.83″ N, 104° 53′ 1.27″ E). The experiment was designed in two phases with 3 factors. The first phase focused on the effect of solvent (methanol, ethanol, propanol, and acetonitrile). The second phase focused on sonication time of 30, 45, and 60 minutes (The times were selected to capture the typical effective range of ultrasonic extraction, to allow solvent penetration and acoustic cavitation to break down the rice cell structure and release 2-AP). Shorter than 30 minutes might result in incomplete extraction, while longer than 60 minutes may cause degradation of 2-AP and reduce efficiency, and temperatures 40°C, 60°C, and 80°C (The temperatures were selected to cover a scientifically meaningful range that allows efficient extraction without degrading the heat-sensitive 2-AP, Lower temperatures would reduce extraction efficiency, while higher temperatures would risk aroma loss due to volatility and solvent evaporation). All samples were analyzed in triplicate to ensure accuracy and reliability.

2.3 Sample preparation

Approximately 100 grams of rice seeds were dehusked using a rice husking machine and then polished into white rice using a rice polishing machine. The polished rice was subsequently ground into a fine powder with an average particle size of 0.5 mm. The samples were packed in ziplock bags and kept at -20°C for further analysis.

2.4 Solvent extraction

Two grams of the powdered sample were weighed and transferred into a Falcon tube. To this, 2 ml of solvent was added, and the tube was sealed with parafilm. The mixture was then shaken and vortexed thoroughly to ensure uniform distribution. The tube was subsequently placed in an ultrasonic bath at a set temperature, where it underwent sonication at maximum speed for different sonication times. During sonication, the sample was vortexed every 2 minutes to maintain even dispersion. After the sonication process, the sample was stored at -20°C for 15 minutes to allow the aroma compounds to stabilize. Following this, the mixture was gently inverted to ensure complete homogenization. The sample was then centrifuged at 40000g for 5 minutes, and the liquid phase was carefully separated using liquid extraction. The liquid was passed through a syringe filter with a pore size of 0.45 µm and filtered into a vial tube, which was sealed with parafilm. Finally, the vial was stored at -20°C until it was ready for GC-MS analysis.

2.5 Gas chromatography-mass spectrometry analysis

The analysis procedure was adapted from the protocol presented at the *Workshop on the Analysis of 2-Acetyl-1-Pyrroline in Rice Samples by Headspace GC-MS*, organized by SA Interfood Rice Products Co., Ltd. in collaboration with Ubon Ratchathani Rice Research Center. The GC-MS analysis was conducted on a Shimadzu GC-MS-TQ8040 coupled to a GC-2010 Plus equipped with an AOC-20S autoinjector and -20i autosampler (Shimadzu, Japan). The analysis was performed using an HP-5MS column (25 m × 250 µm × 0.25 µm) with helium as the carrier gas at a constant flow rate of 1.2 mL/min. The injection was carried out in split mode with a split ratio of 10:1 and a 1 µL injection volume. The inlet temperature was set at 255°C, while the MSD transfer line was maintained at 290°C. The oven temperature program started at 50°C with a 0.5 min hold, increased at 10°C/min to 150°C, which was held for 0.5 min, resulting in a total run time of 11 minutes. The mass spectrometer operated in SIM (selected ion monitoring) mode, with a scan time of 0.1 s and a scan range from 40 to 400 m/z. The SIM mode started at 83.00 minutes and ended at 115.00 minutes. The electron ionization energy was set to 70 eV, with the ion source temperature at 250°C and the quadrupole temperature at 150°C. The data treatment system and the computer identified aroma, which calculated the monoisotopic mass, predicted the structural formula of compounds, and compared them using the MS database.

2.6 Statistical Analysis

All statistical analyses were conducted using R Studio version 4.4.2. Statistical analysis using one-way ANOVA was used to assess the differences statistically significant among solvents. A two-way ANOVA was used to assess the effects of sonication temperature and time on 2-AP concentration, followed by Tukey's HSD post hoc tests for pairwise comparisons. Descriptive statistics (means ± SD) were calculated from triplicate experiments. Statistical packages such as *agricolae*, *ggplot2*, and *dplyr* were used for analysis and visualization, with significance set at $p < 0.05$.

3. RESULTS AND DISCUSSION

This study investigated the effect of three key parameters of ultrasound-assisted extraction on the yield of 2-acetyl-1-pyrroline (2-AP) extracted from rice samples: solvent type, temperature, and the duration of sonication. The results offer valuable insights for optimizing extraction conditions while striking a balance between efficiency and resource utilization.

3.1 Effect of solvent on the extraction yield of 2-AP

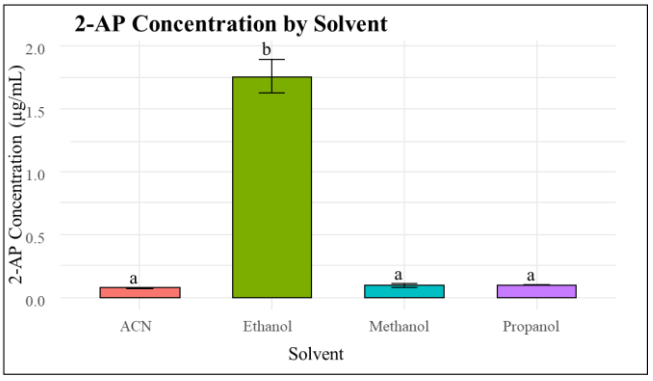


Fig 1. Effect of solvent on the extraction yield of 2-AP.

Based on **Fig. 1**, ethanol exhibited the highest extraction efficiency, yielding an average concentration of 1.75908 µg/mL of 2-AP, although it showed a relatively higher standard deviation (±0.13506 µg/mL) compared to other solvents. Despite the slight variation among replicates, ethanol's extraction yield was significantly higher than that of methanol, propanol, and acetonitrile. This suggests that ethanol not only effectively solubilizes 2-AP but also

penetrates the rice matrix efficiently during ultrasonic-assisted extraction. Methanol, although structurally similar to ethanol and more polar [16], extracted a much lower amount of 2-AP (0.09251 µg/mL) and showed a moderate standard deviation (±0.01774 µg/mL). The lower extraction performance of methanol compared to ethanol may be attributed to differences in solvent-rice matrix interactions or possible losses of volatile compounds during the high-temperature sonication process [17]. Propanol resulted in a slightly higher extraction yield (0.09667 µg/mL) compared to methanol, with a smaller standard deviation (±0.00581 µg/mL), indicating good repeatability. However, due to its lower polarity compared to ethanol and methanol, propanol might be less effective in dissolving the polar 2-AP molecules, limiting its extraction efficiency [11]. Acetonitrile (ACN) demonstrated the lowest extraction efficiency, with an average concentration of 0.07405 µg/mL and the smallest standard deviation (±0.00226 µg/mL), suggesting high repeatability but poor extraction capability. Despite its high polarity, ACN's inability to form hydrogen bonds may hinder its interaction with 2-AP, making it less suitable for extracting this aroma compound from rice [18].

Statistical analysis using one-way ANOVA confirmed that the differences among solvents were statistically significant ($p < 0.001$). Furthermore, the result showed that ethanol was significantly different from the other solvents ($p < 0.05$), confirming that protic solvents, especially ethanol, are superior for 2-AP extraction. The results also indicated that there was no significant difference between methanol and propanol, but acetonitrile was significantly different from propanol.

3.2 Effect of temperature and sonication time on the extraction yield of 2-AP

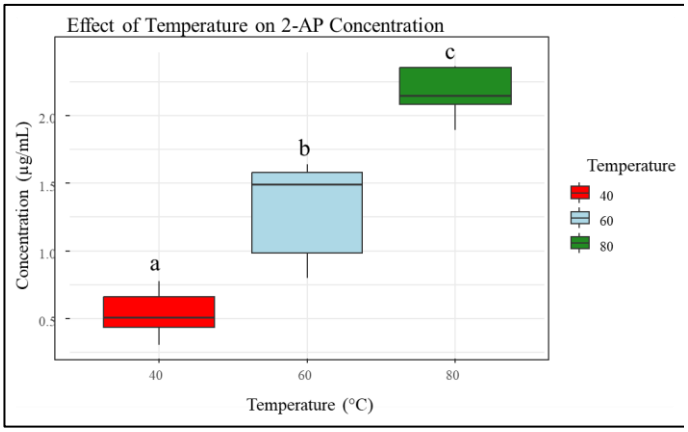


Fig 2. Effect of temperature on the extraction yield of 2-AP.

Based on **Fig. 2**, the data demonstrated that sonication temperature had a highly significant influence on the concentration of 2-AP extracted ($p < 0.001$). The mean 2-AP concentrations at different temperatures, at 40°C, 60°C, and 80°C, were 0.542 ± 0.157 µg/mL, 1.31 ± 0.324 µg/mL, and 2.17 ± 0.188 µg/mL, respectively. This steady increase in 2-AP concentration with rising temperature suggests that elevated temperatures enhance the breakdown of rice cell structures, thereby facilitating a more efficient release of volatile compounds such as 2-AP [19]. Furthermore, the differences in concentration between all pairs of temperature treatments were statistically significant. The difference between 60°C and 40°C was 0.77 µg/mL, with a *p*-value of 1e-06, indicating a significant increase in 2-AP yield at the mid-range temperature. Comparing 80°C and 40°C, the concentration difference was even more remarkable, at 1.63 µg/mL, with a *p*-value less than 0.001, clearly establishing 80°C as the most effective temperature in this study. Additionally, a significant increase of 0.86 µg/mL was observed between 80°C and 60°C ($p = 2e-07$). These results highlight 80°C not only as the most effective temperature but also as a consistently performing condition, as demonstrated by its low standard deviation of 0.188 µg/mL.

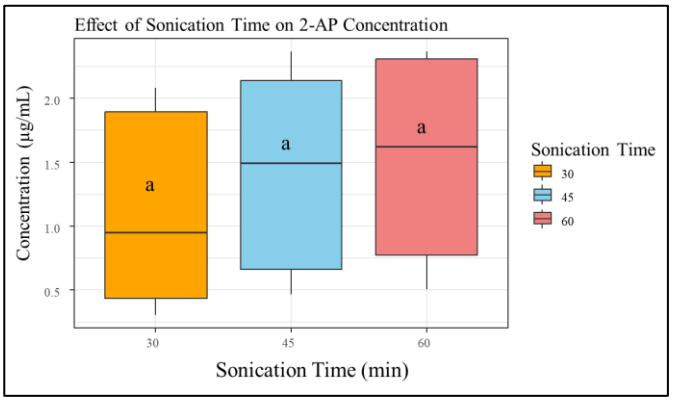


Fig 3. Effect of sonication time on the extraction yield of 2-AP.

Based on **Fig. 3**, the impact of sonication time on 2-AP concentration was not statistically significant, despite showing a slight upward trend. The mean concentrations obtained at 30, 45, and 60 minutes were 1.08 ± 0.697 µg/mL, 1.41 ± 0.722 µg/mL, and 1.53 ± 0.726 µg/mL, respectively. Although the data suggested that longer sonication might lead to a modest increase in 2-AP yield, the large standard deviations indicate considerable variability, and the differences were not statistically meaningful. In addition, there was no significant difference between 45 and 30 minutes ($p = 0.607$), between 60 and 30 minutes ($p = 0.393$), or between 60 and 45 minutes ($p = 0.929$). These results imply that increasing the sonication duration beyond 30 minutes does not significantly enhance 2-AP yield and may not be a cost-effective approach, especially considering potential

compound degradation or unnecessary energy expenditure during prolonged treatments.

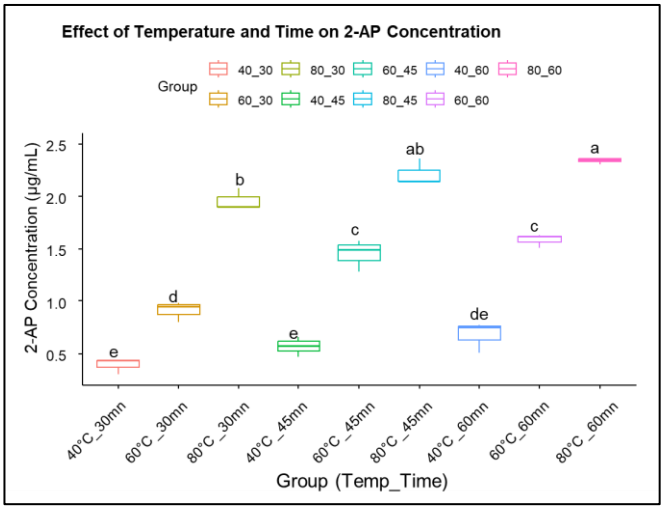


Fig 4. Effect of temperature and sonication time on the extraction yield of 2-AP.

Based on **Fig. 4**, the combination of statistical outcomes and practical considerations, the optimal condition for extracting 2-AP using ultrasonic-assisted extraction appears to be 80°C for 30 minutes. This treatment yielded the highest average concentration of 2.17 µg/mL and required the shortest processing time among the top-performing groups. Although slightly higher concentrations were observed at longer durations. In comparison, previous studies reported 2-AP concentrations ranging from 0.0194 to 0.1240 µg/mL using ultrasound-assisted ethanol extraction [14], 0.01 to 1.104 µg/mL with methylene chloride extraction and GC analysis [20], and 0.005 to 5.0 µg/mL via stable isotope dilution [21], indicating that the 2.17 µg/mL obtained in this study is relatively high and suggests the effectiveness of the current extraction method. Therefore, 80°C for 30 minutes represents the most efficient and effective extraction condition within the scope of this study.

4. CONCLUSIONS

This study tests the different extraction processes to identify the concentration yield of 2-acetyl-1-pyrroline (2-AP) from the Phka Rumdoul rice variety. We analyze the extraction with different types of solvents, temperatures, and times. Among the tested conditions, extraction with ethanol at 80°C for 30 minutes provides the optimal condition for 2-AP extraction, demonstrating both effectiveness and efficiency. While increasing sonication time beyond 30 minutes showed a slight increase in yield, the improvement was not statistically significant. We suggest that, if we have a large sample, we can do the extraction with ethanol at 80°C for 30 minutes. These findings provide a reliable extraction protocol for future studies and quality control applications related to aromatic rice, supporting improved standardization in the evaluation of rice aroma quality.

ACKNOWLEDGMENTS

The authors thank the French Research Institute for Sustainable Development (IRD) through the project "Promoting Integrated Disease Management and Sustainable Quality of Fragrant Rice in Cambodia by Enhancing the Value of native microbiota- FEF Healthyrice" for a master grant fellowship, the Institute of Technology of Cambodia, hosting the Plant health laboratory, the Ministry of Europe and Foreign Affairs via the project "LMI-LEAD", and CIRAD and GDA-DALRM managing the biplots systems as part of the ASSET project (AFD-EU).

REFERENCES

[1] A. Kongpun *et al.*, "Abiotic and Biotic Factors Controlling Grain Aroma along Value Chain of Fragrant Rice: A Review," Mar. 01, 2024, *Elsevier B.V.* doi: 10.1016/j.rsci.2023.11.004.

[2] R. Chakraborty, T. S. Roy, and J.-I. Sakagami, "Thermal and packaging materials for enhancing the longevity and aroma integrity of fragrant rice during storage", doi: 10.4081/jbr.2025.13140.

[3] FAO, "Moving forward on food loss and waste reduction," *State Food Agric. 2019*, pp. 1–182, 2019, [Online]. Available: https://www.fao.org/3/ca6030en/ca6030en.pdf

[4] R. G. Buttery, L. C. Ling, J. G. Turnbaugh, and B. O. Juliano, "Cooked Rice Aroma and 2-Acetyl- 1-pyrroline," *J. Agric. Food Chem.*, vol. 31, no. 4, pp. 823–826, 1983, doi: 10.1021/jf00118a036.

[5] R. Widjaja, J. D. Craske, and M. Wootton, "Comparative Studies on Volatile Components of Scented and Non-Scented Rice," *J. Food Sci. Technol.*, vol. 20, no. 2, pp. 43–47, 1996.

[6] L. J. Romanczyk, C. A. McClelland, L. S. Post, and W. M. Aitken, "Formation of 2-Acetyl-1-pyrroline by Several Bacillus cereus Strains Isolated from Cocoa Fermentation Boxes," *J. Agric. Food Chem.*, vol. 43, no. 2, pp. 469–475, Feb. 1995, doi: 10.1021/jf00050a040.

[7] S. X. Sun *et al.*, "Genetic analysis and gene fine mapping of aroma in rice (Oryza sativa L. cyperales, poaceae)," *Genet. Mol. Biol.*, vol. 31, no. 2, pp. 532–538, 2008, doi: 10.1590/S1415-47572008000300021.

[8] T. Yoshihashi, T. T. H. Nguyen, and N. Kabaki, "Area dependency of 2-acetyl-1-pyrroline content in an aromatic rice variety, Khao Dawk Mali 105," *Japan Agric. Res. Q.*, vol. 38, no. 2, pp. 105–109,

2004, doi: 10.6090/jarq.38.105.
[9] R. J. Bryant and A. M. McClung, “Volatile profiles of aromatic and non-aromatic rice cultivars using SPME/GC-MS,” *Food Chem.*, vol. 124, no. 2, pp. 501–513, 2011, doi: 10.1016/j.foodchem.2010.06.061.
[10] X. Wei, D. D. Handoko, L. Pather, L. Methven, and J. S. Elmore, “Evaluation of 2-acetyl-1-pyrroline in foods, with an emphasis on rice flavour,” *Food Chem.*, vol. 232, pp. 531–544, 2017, doi: 10.1016/j.foodchem.2017.04.005.
[11] N. Ngadi and Y. Yahya, “Jurnal Teknologi Extraction of 2-Acetyl-1-Pyrroline (2AP) in Pandan Leaves (Pandanus Amaryllifolius Roxb.) Via Solvent Extraction Method: Effect of Solvent,” 2014. [Online]. Available: www.jurnalteknologi.utm.my
[12] T. Itani, M. Tamaki, Y. Hayata, T. Fushimi, and K. Hashizume, “Variation of 2-acetyl-1-pyrroline concentration in aromatic rice grains collected in the same region in Japan and factors affecting its concentration,” *Plant Prod. Sci.*, vol. 7, no. 2, pp. 178–183, 2004, doi: 10.1626/pps.7.178.
[13] L. F. Ballesteros, J. A. Teixeira, and S. I. Mussatto, “Selection of the Solvent and Extraction Conditions for Maximum Recovery of Antioxidant Phenolic Compounds from Coffee Silverskin,” *Food Bioprocess Technol.*, vol. 7, no. 5, pp. 1322–1332, 2014, doi: 10.1007/s11947-013-1115-7.
[14] J. Peng *et al.*, “Direct Determination of 2-Acetyl-1-Pyrroline in Rice by Ultrasound-Assisted Solvent Extraction Coupled with Ultra-performance Liquid Chromatography-Tandem Mass Spectrometry,” *Food Anal. Methods*, vol. 16, no. 5, pp. 900–908, Chromatographic Technique for Quantifying 2-Acetyl-1-Pyrroline and Hexanal in Rice (Oryza sativa, L.).”
[21] T. Yoshihashi, “Quantitative analysis on 2-acetyl-1-pyrroline of an aromatic rice by stable isotope dilution method,” *J. Food Sci.*, vol. 67, no. 2, pp. 619–622, 2002.
2023, doi: 10.1007/s12161-023-02478-5.
[15] F. Chemat, N. Rombaut, A. G. Sicaire, A. Meullemiestre, A. S. Fabiano-Tixier, and M. Abert-Vian, “Ultrasound assisted extraction of food and natural products. Mechanisms, techniques, combinations, protocols and applications. A review,” *Ultrason. Sonochem.*, vol. 34, pp. 540–560, 2017, doi: 10.1016/j.ultsonch.2016.06.035.
[16] pp 122-128. APLIKASI METODE GEOMAGNET DALAM EKSPLORASI PANASBUMI Sudaryo. Vol. 29, No.3, “No TitleВорона,” *Птицы*, vol. 1, no. 2, pp. 12–17, 1988.
[17] A. Ghasemzadeh, H. Z. E. Jaafar, A. S. Juraimi, and A. Tayebi-Meigooni, “Comparative evaluation of different extraction techniques and solvents for the assay of phytochemicals and antioxidant activity of hashemi rice bran,” *Molecules*, vol. 20, no. 6, pp. 10822–10838, 2015, doi: 10.3390/molecules200610822.
[18] N. A. Abd Aziz, R. Hasham, M. R. Sarmidi, S. H. Suhaimi, and M. K. H. Idris, “A review on extraction techniques and therapeutic value of polar bioactives from Asian medicinal herbs: Case study on Orthosiphon aristatus, Eurycoma longifolia and Andrographis paniculata,” *Saudi Pharm. J.*, vol. 29, no. 2, pp. 143–165, 2021, doi: 10.1016/j.jsps.2020.12.016.
[19] Z. H. Prodhan, G. Faruq, K. A. Rashid, and R. M. Taha, “Effects of temperature on volatile profile and aroma quality in rice,” *Int. J. Agric. Biol.*, vol. 19, no. 5, pp. 1065–1072, 2017, doi: 10.17957/IJAB/15.0385.
[20] C. J. Bergman, J. T. Delgado, R. Bryant, C. Grimm, K. R. Cadwallader, and B. D. Webb, “Rapid Gas

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Comparing the Physicochemical Properties of Soybean Oil Blended with Different Oils

Monika MICH[2], Chanchira PEN[2], Manit SAY[1,2], Chin Ping TAN[3], Reasmey TAN[1,2*]

[1] *Food Technology and Nutrition Research Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Department of Food Technology, Faculty of Food Science and Technology, University Putra Malaysia, 43400 UPM Serdang Selangor, Malaysia*

[*] *Corresponding author: rtan@itc.edu.kh*

Abstract: *Cooking oil is a foundational ingredient in kitchens worldwide, utilized for frying, baking, and as a component in salad dressings and sauces. However, repeated high-temperature cooking can degrade oil quality and pose health risks, prompting a need for safer formulations. Various oil modification techniques, including microencapsulation, interesterification, hydrogenation, fractionation, and blending, are employed to address these concerns, with blending emerging as an economically efficient and widely adopted solution in the food industry. This study aimed to compare the thermal oxidative stability of soybean oil and its blends during deep-frying. Blended oils were prepared by mixing soybean oil with sunflower oil (T1), rice bran oil (T2), coconut oil (T3), and palm olein oil (T4) in a fixed ratio (% v/v). Banaba slices (~1 cm thick, 25 g per batch) were deep-fried at 180 °C for 2 minutes over 10 frying cycles, with a 5-minute resting interval between cycles. Oil samples were analyzed for acid value, peroxide value, iodine value, p-anisidine value, and color to assess oxidative stability. The study revealed that blending soybean oil with sunflower oil, rice bran oil, coconut oil, or palm olein oil in a fixed ratio enhanced the thermal stability during deep-frying. After crossing the tenth cycle (around one hour), blended oils exhibited greater resistance to thermal oxidation compared to non-blended soybean oil, as shown by improved chemical parameters such as acid value, peroxide value, p-anisidine value, iodine value, and total oxidation TOTOX. Blended oils also contributed to improved oil color and physicochemical properties. However, further research is recommended to study the nutritional quality, fatty acid composition, oxidative stability, shelf life, and sensory evaluation of blended oils to enhance understanding and application.*

Keywords: Blended oils; Physicochemical properties; High-temperature frying; Frying Stability.

1. INTRODUCTION

Vegetable oils are integral to human diets, providing essential fatty acids, including saturated, monounsaturated, and polyunsaturated fats, which are crucial for cellular metabolism, energy reserves, and energy production [1]. While most cooking oils remain liquid at room temperature, saturated-fat-rich oils, such as coconut oil and palm oil, solidify at room temperature. However, the health benefits of oils can diminish when exposed to high temperatures, as oxidation of fatty acids occurs, particularly in highly unsaturated oils. This poses health risks, including the production of potentially harmful compounds [2].

To enhance the stability, nutritional value, and utility of vegetable oils, several modification techniques are employed, including hydrogenation, interesterification (IE), fractionation, microencapsulation, and blending. While methods like hydrogenation and IE offer improved stability, they involve drawbacks like the formation of trans fats and high costs [3, 4]. Among these approaches, blending stands out as a cost-effective, efficient method to improve oil functionality and profile without producing harmful residues. This practice enables the combination of oils with complementary properties, creating oils with balanced fatty acid profiles, enhanced oxidative stability, enriched bioactive compounds, and improved health benefits [5].

Generally, blending refers to the mixing of different types of substances having a similar composition to make a product of desirable quality. Blended oils are increasingly recognized as a transformative innovation in the food industry. The process allows for the tailoring of oil attributes to meet both nutritional demands and functional requirements, making it a promising solution for optimizing edible oil formulations [6].

The objective of this study was to compare the thermal oxidative stability and quality of soybean oil with its blended oils (coconut oil, rice bran oil, sunflower oil, and palm olein oil) during the deep-frying process in order to evaluate the

[1*] Corresponding author: Reasmey TAN; E-mail: *rtan@itc.edu.kh; Tel: +855-12 602 202*

physicochemical characteristics of the oils and analyze how blending influences their performance and stability under high-temperature conditions.

2. METHODOLOGY

2.1 Sample collection and preparation

Five types of cooking oils, such as soybean oil, sunflower oil, rice bran oil, coconut oil, and palm olein oil, were purchased from Makro Supermarkets (Sen Sok branch) located in Phnom Penh. All samples had identical expiration dates and were labeled with nutritional information.
A fresh, unripe banana was purchased from the local market. To remove surface dust, the banana was thoroughly washed with water. After cleaning, it was peeled and soaked in water to prevent discoloration until further use. Using a sharp knife, the banana was carefully sliced into thin pieces, approximately 1 mm in thickness.

2.2 Blending of oils

Refined soybean oil (SB) was blended with different edible oils such as refined sunflower oil (SO), rice bran oil (RO), coconut oil (CO), and palm olein oil (PO) in a fixed ratio (%v/v). The mixtures were stirred at 60℃ for 5 minutes using a magnetic stirrer.

2.3 Frying procedure

Physicochemical properties of blended oil were analyzed before and after the 1st, 5th, and 10th frying cycles. Freshly sliced bananas (25g) were deep-fried in blended oil at 170℃ for 2 minutes, following the method described by [7]. The experiments were conducted over 10 batches, with a 5-minute interval between each batch. Fried bananas and used oils from the 1st, 5th, and 10th batches were collected, packed, and stored in the refrigerator for further analysis [8].

2.4 Physicochemical properties of oil

All the parameters of physicochemical analysis of cooking oils were performed using the standard method of oil analysis by the Association of Official Analytical Chemists (AOAC) method with minor modifications. The experiment was conducted in duplicate.

2.4.1 Acid value (AV)

Five grams of oil sample was weighed and pre-heated before being mixed with 100 mL of 95% freshly neutralized hot ethyl alcohol. Then, the mixture of oil and solvent was titrated with 0.1N NaOH solution by using phenolphthalein as an indicator. The acid value was calculated as follows:

$$\text{Acid value (AV)} = \frac{V \times N \times M}{W} \qquad \text{(Eq. 1)}$$

Where:
AV= acid value (mg KOH/g oil)
V = volume of standard NaOH solution (mL)
N = normality of standard NaOH solution
M = molecular weight of NaOH (56.1 g/mol)
W = weight of oil sample (g)

2.4.2 Peroxide value (PV)

Five grams of the oil sample was weighed and dissolved in 30 mL of acetic acid/chloroform (3:2 ratio) solvents. Then, the solution was further reacted with 0.5 mL of 15% potassium iodide (KI) and stand with occasional shaking for 1 min. The mixture was stoppered and stored in the dark for 5 min before adding 30 mL of distilled water. The liberated iodine was titrated with 0.1 N sodium-thiosulphate $Na_2S_2O_3$ using 0.5 mL of 1% starch as an indicator. Blank titration was performed. The peroxide value was determined as follows:

$$\text{Peroxide value (PV)} = \frac{(S\text{-}B) \times M \times 1000}{W} \qquad \text{(Eq. 2)}$$

Where:
PV= peroxide value (meq O_2/kg oil)
S = volume of $Na_2S_2O_3$ consumed by the oil sample (mL)
B = volume of $Na_2S_2O_3$ used for blank sample (mL)
M = the molarity of $Na_2S_2O_3$ solution
W = weight of the oil sample (g)

2.4.3 Iodine value (IV)

The oil sample was weighted to 0.5 g and put in a 500 mL flask. Fifteen mL of cyclohexane and 15 mL of Wijs solution were added. The blank was determined without the oil sample. The solution was shaken and put in a dark place for 30 minutes for a complete reaction between iodine and the unsaturated bonds of oils. Twenty mL of 10 % KI was added, followed by 100 mL of distilled water. The final content was titrated with 0.1 mol/L sodium-thiosulphate solutions using starch as an indicator. Iodine value was calculated as follows:

$$\text{Iodine value (IV)} = \frac{12.69 \times (B - S) \times M}{W} \qquad \text{(Eq. 3)}$$

Where:
B = volume in mL of standard sodium thiosulphate solution required for the blank
S = volume in mL of standard sodium thiosulphate solution required for the sample
M = molarity of the standard sodium thiosulphate
W = weight in g of the sample

2.4.4 p-anisidine value (AnV)

One gram of oil sample was dissolved and adjusted to 25 mL by using 2,2,4-trimethylpentance. The unreacted test solution was then prepared by adding 1 mL of glacial acetic

acid to 5 mL of the preceding solution in the test tube. Then the mixture was kept in a dark place at 23 °C for 10 minutes. After the prescribed time, the solution was transferred to a clear and dry spectrophotometer cell. At 350nm, the absorbance of the unreacted test solution (A0) was measured. Following the unreacted test solution, the absorbance of the reacted test solution (A1) was measured. However, an anisidine reagent was used in place of glacial acetic acid. The absorbance of the blank (A2) was observed according to the reacted test solution method by substituting the sample with 5 mL of iso-octane. The anisidine value of the vegetable oil sample was calculated by the following equation:

$$\text{Anisidine (AnV)} = \frac{100 \times Q \times V \times 1.2\,(A_1 - A_2 - A_0)}{m} \qquad \text{(Eq. 4)}$$

Where:
A_0 = absorbance of the unreacted test solution
A_1 = absorbance of reacted test solution
A_2 = absorbance of the blank
m = mass (g) of the test portion
Q = sample content (g/mL, Q = 0.1 g/mL)
V = 25 mL, the volume in which the test sample is dissolved.

The correction factor for the dilution of the 5 mL of the test solution with 1 mL of the reagent = 1.2

2.4.5 TOTOX value

The TOTOX value is considered advantageous because it combines information about the oil's past condition, indicated by the AnV, with its current state (shown by the PV). TOTOX is giving a value of the total oxidation status in oil and is equal to:

$$\text{TOTOX Value} = 2PV + AnV \qquad \text{(Eq. 5)}$$

2.4.6 Color

The coloration of the cooking oil samples was measured by using a Chroma Meter (CR-400, KONICA MINOLTA, Japan). The color parameters were presented as L^* (Lightness extends from 0 to 100 for black to white, respectively), a^* (Represent redness to greenness), and b^* (Represent yellowness to blueness). The instrument was standardized using a white calibration plate.

2.5 Statistical analysis

All physicochemical parameters of oil samples obtained from the deep-frying process were analyzed by 24 one-way ANOVA followed by SPSS software (25, US). By using Duncan's new multiple range test, significant variations between the means were identified (p<0.05).

3. RESULTS AND DISCUSSION

3.1 Acid value

The acid value is commonly used as a general indicator of the oil's quality and edibility. According to Codex Alimentarius, 2019 recommendations, the maximum acid value for all edible oils should be less than 0.6 mg KOH/g oil [2]. According to the results presented in **Table 1**, the acid value of all types of blended oils after the 10th frying cycle remained below 0.6 mg KOH/g oil. Among the blends at different frying cycles, the highest acid value was 0.55 mg KOH/g oil, which were observed in the mixture of soybean oil and rice bran oil (T2) at 10th batch, while the lowest acid value was found in the blend of soybean oil and coconut oil (T4) at 1st batch which were 0.07 mg KOH/g oil. Moreover, at the 1st and 5th cycles, the acid value of T3 and T4 is not significantly different (p<0.05). This result indicates that both blends of soybean with coconut and soybean with palm oil help to reduce the acid value during frying than the use of T0, T1, and T2. It indicates that the soybean blend with coconut oil is more stable in acid value with repeated frying. Higher values indicate that triglycerides of oil are converted into fatty acids and glycerol, which cause rancidity of the oil [3]. The acid value of oils indicates the amount of free fatty acid present in the oil. It determines the purity of oils. The higher the values, the lower the possibility of the oils being used for cooking purposes.

3.2 Peroxide value (PV)

Peroxide value is used as a measure of the extent to which rancidity reactions have occurred during storage. The quality and stability of fats and oils can be indicated by using the peroxide value [9]. According to FAO/WHO recommendations, the peroxide value for refined oils should be ≤10 meq O_2/kg, which serves as an indicator of good oil quality. In the present study, as shown in **Table 2**, the analyzed peroxide value in oils before frying ranged from 0.69 to 1.69 meq O_2/kg. However, after conducting the 10th frying cycle, only the peroxide values of blends T2 and T3 remained below the recommended limit of FAO/WHO, while T1, T4, and T5 exceeded this limit after just the 5th frying cycle. Notably, there were no significant differences among the peroxide values of T2, T3, T4, and T5 after frying, suggesting that blending soybean oil with sunflower oil, rice bran oil, coconut oil, and palm oil effectively enhanced the oxidative stability of the oil. This improved stability makes the blended oils more suitable for repeated frying compared to using soybean oil alone. Moreover, the increase in the peroxide value after every cycle of frying is primarily caused by the chemical degradation of fats and oils during the frying process.

3.3 p-anisidine value (AnV)

The p-anisidine value is a reliable indicator of secondary lipid oxidation products, including aldehydes and ketones, formed during the oxidative degradation of oils at elevated frying temperatures. These compounds contribute to rancid odors and flavors in fried foods. The results presented in **Table 1** indicate that the p-anisidine value of oils before frying ranged from 2.96 to 3.06, with gradual increases observed over successive frying cycles, reflecting the accumulation of lipid oxidation products. Moreover, after the 10th cycle of frying, the p-anisidine value of used oils ranged from 32.53 to 65.14, while the lowest value was found in the blends of soybean oil with palm oil (T4) and soybean oil with coconut oil (T3), and T3 and T4 are not significantly different. And the highest p-anisidine value was found in blends of soybean oil with rice bran oil (T2). These results show that the effect of frying temperatures and repeated use accelerates hydroperoxide decomposition, leading to secondary oxidation. The interaction of oxidation products with fried foods can slow the rate of anisidine value increments [10].

3.4 TOTOX

TOTOX is a third way to express oxidation and is a calculation combining PV and AnV. This parameter was conceived as a way to give a more complete picture of oxidation by including primary and secondary oxidation measurements [11]. Based on the results shown in **Table 1**, the TOTOX values of the oil before frying, 1st, 5th, and 10th cycles were ranged from 4.34 – 18.95, 17.10 – 36.16, 30.49 – 47.93, and 34.48 – 68.51, respectively. Regarding to the result above shows that the blended oil (Soybean oil with coconut oil) and (Soybean oil with palm olein oil) received a lower oxidation if it was compared to the pure soybean oil. The rate of oxidation increases with the degree of fatty acid unsaturation.

Table 1. Chemical properties of different oil blends with different frying cycles

Parameters	Code	Blend Oils	Frying cycle			
			Before frying	1st	5th	10th
AV (mg KOH/g oil)	T0	SB	0.10 ± 0.02^{b}	0.12 ± 0.01^{a}	0.16 ± 0.01^{b}	0.18 ± 0.02^{b}
	T1	SB : SO	0.11 ± 0.00^{b}	0.13 ± 0.01^{a}	0.13 ± 0.00^{b}	0.20 ± 0.01^{b}
	T2	SB : RO	0.43 ± 0.00^{c}	0.46 ± 0.11^{b}	0.47 ± 0.00^{c}	0.55 ± 0.00^{c}
	T3	SB : CO	0.13 ± 0.02^{a}	0.07 ± 0.01^{a}	0.08 ± 0.01^{a}	0.13 ± 0.01^{a}
	T4	SB : PO	0.13 ± 0.01^{b}	0.10 ± 0.03^{a}	0.08 ± 0.02^{a}	0.14 ± 0.00^{a}
PV (meq O_2/kg)	T0	SB	0.99 ± 0.00^{b}	4.95 ± 0.28^{ab}	13.58 ± 1.06^{cd}	15.58 ± 0.95^{c}
	T1	SB : SO	1.175 ± 0.01^{b}	6.42 ± 0.95^{bc}	12.91 ± 1.44^{bc}	9.51 ± 0.28^{a}
	T2	SB : RO	1.69 ± 0.14^{c}	8.23 ± 1.32^{c}	9.135 ± 0.56^{a}	9.43 ± 1.20^{a}
	T3	SB : CO	0.69 ± 0.13^{a}	4.36 ± 0.25^{a}	9.8 ± 1.05^{ab}	12.61 ± 0.66^{a}
	T4	SB : PO	0.97 ± 0.02^{b}	4.31 ± 0.30^{a}	16.75 ± 0.23^{d}	10.68 ± 0.23^{ab}
AnV	T0	SB	3.06 ± 0.12^{a}	8.19 ± 0.08^{b}	14.52 ± 0.86^{b}	40.75 ± 1.15^{c}
	T1	SB : SO	5.12 ± 0.20^{c}	12.04 ± 0.29^{c}	16.51 ± 2.12^{b}	38.26 ± 0.26^{bc}
	T2	SB : RO	15.57 ± 1.03^{d}	19.71 ± 0.25^{d}	29.66 ± 3.55c	65.14 ± 8.74^{d}
	T3	SB : CO	2.96 ± 1.36^{a}	8.38 ± 0.21^{b}	10.89 ± 1.28^{b}	34.79 ± 2.53^{ab}
	T4	SB : PO	4.17 ± 0.22^{b}	7.156 ± 0.07^{a}	10.01 ± 3.44^{b}	32.53 ± 12.57^{a}
TOTOX	T0	SB	5.04 ± 0.15^{a}	18.097 ± 0.45^{b}	41.68 ± 3.23^{b}	42.73 ± 1.46^{a}
	T1	SB : SO	7.477 ± 0.24^{b}	24.89 ± 2.26^{bc}	42.66 ± 0.13^{bc}	40.62 ± 0.11^{a}
	T2	SB : RO	18.95 ± 1.62^{c}	36.16 ± 2.83^{c}	47.93 ± 3.46^{c}	68.51 ± 3.07^{b}
	T3	SB : CO	4.34 ± 0.09^{a}	17.1 ± 0.77^{a}	30.49 ± 0.68^{a}	36.17 ± 1.70^{a}
	T4	SB : PO	6.12 ± 0.02^{ab}	15.78 ± 0.53^{bc}	43.5 ± 0.02^{bc}	34.48 ± 3.20^{a}
IV (g I_2/100g oil)	T0	SB	87.18 ± 24.88^{a}	67.89 ± 12.86^{a}	81.00 ± 10.35^{ab}	72.8 ± 0.89^{a}
	T1	SB : SO	87.78 ± 18.42^{a}	98.36 ± 1.77^{b}	70.77 ± 5.69^{a}	85.1 ± 14.25^{a}
	T2	SB : RO	89.82 ± 20.76^{a}	91.79 ± 4.78^{b}	89.46 ± 14.83^{ab}	77.31 ± 21.97^{a}
	T3	SB : CO	77.51 ± 7.24^{a}	92.97 ± 6.57^{b}	81.12 ± 9.31^{ab}	89.32 ± 3.37^{a}
	T4	SB : PO	77.42 ± 4.00^{a}	84.6 ± 1.38^{ab}	99.16 ± 0.29^{b}	76.55 ± 3.57^{a}

Values are means of duplicated ± standard deviation. Within a column with the same lowercase letter, the values are not significantly different at p<0.05.

Table 2. Color of different oil blends with different frying cycles

Parameters	Code	Oils	Frying cycle			
			Before frying	1st	5th	10th
L*	T0	SB	24.617 ± 0.01[b]	24.237 ± 0.35[a]	31.150 ± 0.35[b]	23.973 ± 0.25[a]
	T1	SB : SO	24.583 ± 0.23[b]	28.840 ± 0.04[b]	31.467 ± 0.13[b]	23.767 ± 0.02[a]
	T2	SB : RO	23.683 ± 0.04[a]	31.480 ± 0.56[d]	24.220 ± 0.45[a]	29.820 ± 0.13[b]
	T3	SB : CO	24.470 ± 0.30[b]	29.867 ± 0.28[c]	31.003 ± 0.06[b]	24.050 ± 0.06[a]
	T4	SB : PO	23.533 ± 0.32[a]	24.187 ± 0.07[a]	32.207 ± 0.45[c]	31.330 ± 0.32[c]
a*	T0	SB	-0.320 ± 0.02[d]	-0.510 ± 0.01[c]	-0.683 ± 0.04[b]	-0.640 ± 0.04[b]
	T1	SB : SO	-0.447 ± 0.02[c]	-0.430 ± 0.07[c]	-0.683 ± 0.03[b]	-0.700 ± 0.05[b]
	T2	SB : RO	-0.747 ± 0.01[b]	-0.893 ± 0.05[a]	-0.837 ± 0.04[a]	-0.843 ± 0.10[b]
	T3	SB : CO	-0.313 ± 0.07[d]	-0.467 ± 0.07[c]	-0.803 ± 0.06[a]	-0.497 ± 0.03[c]
	T4	SB : PO	-0.943 ± 0.04[a]	-0.810 ± 0.06[b]	-0.723 ± 0.01[b]	-1.047 ± 0.02[a]
b*	T0	SB	3.507 ± 0.23[a]	4.803 ± 0.71[a]	5.910 ± 0.15[a]	5.007 ± 0.09[b]
	T1	SB : SO	3.780 ± 0.11[a]	4.547 ± 0.01[ab]	5.933 ± 0.01[a]	6.363 ± 0.02[c]
	T2	SB : RO	5.210 ± 0.09[b]	7.463 ± 0.26[d]	6.073 ± 0.30[ab]	6.363 ± 0.02[c]
	T3	SB : CO	3.470 ± 0.01[a]	5.040 ± 0.19[b]	6.487 ± 0.16[b]	4.370 ± 0.11[a]
	T4	SB : PO	5.387 ± 0.07[b]	5.683 ± 0.06[c]	6.320 ± 0.01[ab]	7.237 ± 0.02[d]

Values are means of duplicated ± standard deviation. Within a column with the same lowercase letter, the values are not significantly different at $p<0.05$.

Coconut oil has the highest level of saturation, followed by palm and sunflower oils. This means that palm and sunflower oils are more likely to oxidize when used for frying at high temperatures compared to coconut oil [12].

3.5 Iodine value (IV)

Iodine value (IV) measures the degree of unsaturation in a vegetable oil. It determines the stability of oils to oxidation and allows the overall unsaturation of the fat to be determined qualitatively [9]. According to the results indicated in **Table 2**, the iodine values (IV) of various blended oils across multiple frying cycles, Before to frying, there were no significant differences (p<0.05) among the treatments, T1 (pure soybean oil), T1 (soybean-sunflower oil blend), T2 (soybean-rice bran oil blend), T3 (soybean-coconut oil blend), and T4 (soybean-palm olein oil blend) with iodine values of 87.18, 87.78, 89.82, 77.51, and 77.42 g I2/100g oil, respectively. At the first frying cycle, significant differences (p<0.05) were observed. T0 exhibited the lowest IV (67.89 g I2/100g oil), while blended oils showed increases such as T1 at 98.36, T2 at 91.97, T3 at 92.97, and T4 at 84.6 g I2/100g oil. By the 5th frying cycle, the IV of T4 significantly increased to 99.16 g I2/100g oil compared to T0 (81), T1 (70.77), T2 (89.46), and T3 (81.12). At the 10th frying cycle, IV differences were not significant for all oil types. Higher iodine values indicate a larger proportion of unsaturated fatty acids, making oils more susceptible to oxidation and rancidity. Abdel-Razek et al. [13] found that blending soybean or sunflower oil with olive oil reduced IV proportionally. For example, a 40% olive oil blend reduced IV by 16.2% for soybean oil and 12.0% for sunflower oil, while a ternary blend (30% soybean, 30% sunflower, 40% olive oil) reduced IV by 15.7%. This suggests that blended oils have higher unsaturation than pure soybean oil.

2.6 Color

Color is a key factor in assessing oil quality in the food industry, with CIE L*, a*, and b* coordinates used to measure lightness (L*), redness-greenness (a*), and yellowness-blueness (b*). As a primary physical characteristic, color affects consumer acceptance, varying widely depending on the production source [14]. According to the results in **Table 2**, the initial lightness (L*) values of the oils before frying were relatively consistent, reflecting similarities in color properties among blended and non-blended oils. Minor variations in L* values were likely due to the intrinsic properties of blending. After the 1st frying cycle, significant changes were observed, with blends T2 and T3 showing the highest lightness values, while T4 remained stable. By the 10th frying cycle, T2 and T4 demonstrated noticeable increases in lightness (30.09 and 31.07, respectively). Among the blends, T2 and T4 showed superior stability, highlighting the benefits of incorporating rice bran oil and palm olein oil due to their high oxidative stability and antioxidant content.

The a* values of blended oils increased as frying cycles progressed, with T4 showing the most notable rise by the 10th cycle. Similarly, b* values of blended oils also increased, particularly for T4, which significantly increased to 7.24,

indicating its high carotenoid content. These results suggest that palm olein oil inherently provides better resistance to degradation during frying due to its carotenoid and vitamin E composition. As frying progresses, L* values decrease while a* and b* values increase, attributed to pigment leaching and chemical changes such as hydrolysis, oxidation, and polymerization [15]. Although chlorophyll and carotenoid levels were not directly measured in this study, the stability observed in the soybean-palm olein oil blend (T4) highlights its potential in maintaining better color and oxidative stability. These findings emphasize the importance of selecting appropriate oil blends to enhance frying performance and durability.

4. CONCLUSIONS

This study highlights the improved thermal stability and physicochemical properties of soybean oil when blended with sunflower oil, rice bran oil, coconut oil, and palm olein oil. After undergoing deep-frying for approximately one hour, blended oils performed better than non-blended soybean oil in chemical parameters such as acid value, peroxide value, iodine value, p-anisidine value, and total oxidation (TOTOX). Treatment T3 emerged as the most effective blend, exhibiting superior performance in both physical and chemical parameters, followed by treatments T4 and T1. Furthermore, blended oils also improved the color quality after frying, enhancing overall oil safety and usability during deep-frying. Future studies should explore the nutritional quality of blended oils, focusing on fatty acid compositions, oxidative stability factors such as shelf life, and sensory evaluations to provide a comprehensive understanding.

ACKNOWLEDGMENTS

The work was funded by the Cambodia Higher Education Improvement Project (Credit No. 6221-KH).

REFERENCES

[1] A. Kurmar, A. Sharma and K. C. Upadhyaya, "Vegetable Oil: Nutritional and Industrial Perspective," *Current Genomics,* vol. 17, no. 3, pp. 230-240, 2016.

[2] S. Abrante-Pascual, B. Nieva-Echevarria and E. Goicoechea-Oses, "Vegetable Oils and Their Use for Frying: A Review of Their Compositional Differences and Degradation," *Foods,* vol. 13, no. 24, p. 27, 2024.

[3] M. Yazilikaya, F. Yemiscioglu and V. Cisse, "Chemical Interesterification of Fractionated Palm Oil, Sunflower Oil and Soybean Oil Blends and its effect," *International Journal of Agriculture Environment and Bioresearch,* vol. 4, no. 3, pp. 276-288, 2019.

[4] B. R. Boger, S. R. Georgetti and L. Kurozawa, "Microencapsulation of Grape Seed Oil by Spray Drying," *Food Science and Technology,* vol. 38, no. 2, pp. 263-270, 2018.

[5] A. Dhyan, R. Chopra and M. Garg, "A Review on Blending of Oils and Their Functional and Nutritional," *Chemical Science Review and Letters,* vol. 7, no. 27, pp. 840-847, 2018.

[6] S. Grover, P. Kumari, A. Kumar, A. Soni, S. Sehgal and V. Sharma, "Preparation and Quality Evaluation of Different Oil Blends," *Letters in Applied NanoBioScience,* vol. 10, no. 2, pp. 2126-2137, 2021.

[7] E. Choe and D. B. Min, "Chemistry of Deep-fat Frying Oils," *Journal of Food Science,* vol. 72, no. 5, pp. 77-86, 2007.

[8] L. Xu, X. Li, J. Huang, P. Gao, G. Jin and X. Wang, "Rapid Measuring Flavour Quality Changes of Frying Rapeseed Oils Using a Flash Gas Chromatography Electronic Nose," *European Journal of Lipid Science and Technology,* vol. 121, no. 1, pp. 1-6, 2019.

[9] L. Geng, W. Zhou, X. Gu, R. Sa, J. Liang, X. Wang and M. Sun, "Iodine Values, Peroxide Values and Acid Values of Bohai Algae Oil Compared With Other Oils During the Cooking," *Heliyon,* vol. 9, no. 4, pp. 1-10, 2023.

[10] F. A. Aladedunye and R. Przybylski, "Degradation and Nutritional Quality Changes of Oil During Frying," *Journal of the American Oil Chemists Society,* vol. 86, pp. 149-156, 2009.

[11] T. W. Wai, B. Saad and B. P. Lim, "Determination of TOTOX Value in Palm Oleins Using a FI-potentiometric Analyzer," *Food Chemistry,* vol. 113, no. 1, pp. 285-290, 2009.

[12] S. K. and S. S., "Comparative Study on Stability of Coconut Oil, Sunflower Oil and Palm Oil During Continuous Deep Frying," *Ceylon Journal of Science,* vol. 51, no. 4, pp. 359-368, 2022.

[13] A. G. Abdel-razek, S. M. El-shami, M. A.Hassan and M. Mahmoud, "Blending of Virgin Olive Oil with Less Stable Edible Oils to Strengthen Their Antioxidative Potencies," *Australian Journal of Basic and Applied Sciences,* vol. 5, no. 10, pp. 312-318, 2010.

[14] K. Kittipongpittaya, A. Panya, T. Prasomsri and P. Sueaphet, "Tropical Oil Blending and Their Effects on Nutritional Content and Physicochemical Properties During Deep Fat Frying," *Journal of Nutritional Science and Vitaminology,* vol. 66, pp. 206-214, 2020.

[15] P. K. Nayak, "Physico-chemical Changes During Repeated Frying of Cooked Oil: A review," *Journal of Food Biochemistry,* vol. 40, no. 3, pp. 371-390, 2015.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Comparing the Quality of Fermented Soy Sauces Made by Using Whole Soybeans and Defatted Soybeans

Kimmuoy SONG[1], Reasmey TAN[1,2*], Luka LY[1], Parakulsuksatid PRAMUK[3]

[1] *Food Technology and Nutrition Research Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Department of Biotechnology, Faculty of Agro-Industry, Kasetsart University, Bangkok, Thailand*

[*]*Corresponding author: rtan@itc.edu.kh*

Abstract: *In East and Southeast Asian cuisines, soy sauce is a traditional ingredient used in cooking and as a condiment. The soy sauce fermentation is made from whole or defatted soybeans, wheat, brine solution, Aspergillus oryzae, and Zygosaccharomyces rouxii. The main purpose of this study was to compare the fermented soy sauces made by using whole soybeans and defatted soybeans using Aspergillus oryzae during koji fermentation, with and without the addition of yeast during moromi fermentation. For soy sauce production in this study, 4 different conditions were used with 22% salt concentration and fermented for 32 weeks. To assess the quality of each sample, the physicochemical quality parameters such as moisture content, and protease activity of koji fermentation, and TA, AN, and RS content during moromi fermentation, with the evaluation of consumer preference using sensory evaluation. As the results, the initial moisture content of WS1 (Whole soybeans without adding yeast) and WS2 (whole soybeans with adding yeast) (49.01% and 46.67%) was higher than DS1 (Defatted soybeans without adding yeast) and DS2 (Defatted soybeans with adding yeast) (39.87% and 35.59%). Enzyme activity for protease was from 14.17Unit/g to 635.77Unit/g. Whole soybeans had a lower than defatted soybeans in terms of protease activity at 72 hours, respectively. Moreover, TA, AN, and RS increased throughout moromi fermentation. After 32 weeks of fermentation, DS2 showed the highest TA (2.57g/100ml) and lower AN (0.70g/100ml) among DS1, WS1, and WS2. In addition, WS1 had a higher value of RS (5.72g/100ml). The sensory evaluation for DS2 got the highest overall scores among the other 3 soy sauce samples (DS1, WS1, and WS2). Thus, defatted soybeans could be considered for use to replace whole soybeans in soy sauce production.*

Keywords: Defatted soybeans; Koji fermentation; Moromi fermentation; *Zygosaccharomyces rouxii*; *Aspergillus oryzae*

1. INTRODUCTION

Soy sauce, a salty condiment made from soybeans (and sometimes wheat), is used across East Asia with various names. Traditionally, it adds flavor and provides protein and salt. Since 1950, production has modernized, especially in Japan, but much of the production in other Asian countries still follows traditional methods [1]. Soybean cake, a byproduct of oil extraction, contains 49.3% crude protein, higher than whole soybeans. With low oil content, it serves as a sustainable and affordable alternative for soy sauce production. This study compared whole and defatted soybeans, using *A. oryzae* for the koji stage and *Z. rouxii* for the moromi stage, without adding lactic acid bacteria. The moromi fermentation lasted 7 months. Furthermore, the scope of this experiment studied enzyme activity and physicochemical properties parameters such as moisture content and protease activity in the koji stage, whereas Total acidity (TA), Amino nitrogen (AN), and Reducing sugar (RS) content in the moromi stage including sensory evaluation. While sensory evaluation was conducted with untrained panelists (Naïve) as the students of ITC. The parameters enzyme activities of

[1]* Corresponding author: Reasmey Tan
E-mail: rtan@itc.edu.kh ; Tel: +855-12 602 202

the koji quality analysis were done only for neutral pH and not the analysis of volatile compounds for moromi quality.

2. METHODOLOGY

2.1 Koji fermentation

For koji fermentation, 0.5 kg of whole soybeans were cleaned and soaked in water (1:1.2 w/v) for 4 hours, while defatted soybeans were mixed with water in the same ratio without soaking. Both were then steamed at 121°C for 30 minutes to sterilize and denature proteins, aiding enzymatic digestion. In parallel, 0.5 kg of whole wheat was cleaned, roasted to a golden color for flavor and enzyme activation, and then crushed. The soybeans (whole or defatted) were mixed with roasted wheat in a 1:1 ratio, and inoculated with *A. oryzae* spores (10^8 spores/kg). The mixture was covered with cheesecloth and fermented at room temperature for 72 hours until mold growth produced a greenish-yellow color, indicating the koji was ready for moromi fermentation.

The study was conducted on four conditions with different soy base raw materials: whole soybeans and whole wheat without adding yeast (WS1), whole soybeans and whole wheat with adding yeast (WS2), defatted soybeans and whole wheat without adding yeast (DS1), and defatted soybeans and whole wheat with adding yeast (DS2). For these 4 conditions, mold *A. oryzae* (from Japan) was added at koji fermentation, and yeast *Z. rouxii* was added at moromi fermentation. The koji was submerged in 22% brine solution with ration 1:1.5 (w/v) during moromi.

2.2 Moromi fermentation

After 72 hours of koji fermentation, moromi fermentation was performed in the glass jar (3 L), in which koji mash was mixed with 22% brine solution at a ratio of 1:1.5 (w/v) and allowed to ferment at room temperature. This is called moromi fermentation stage. The brine solution (22%) was prepared by dissolving salt into distilled water. In the moromi fermentation was stirred every day to allow more oxygen into the glass jar during aging. Moreover, *Z. rouxii* was added to moromi fermentation in the 7 weeks of moromi fermentation of all samples, with the pH of moromi mash below 5.2. *Z. rouxii* was added due to its a vigorous alcoholic fermentation, due to the availability of starch from wheat.

2.3 Koji quality analysis

2.3.1 Moisture content

The moisture content of the koji sample was determined using the gravimetric method based on AOAC (2001). First, aluminum plates were dried at 105°C for 3 hours and cooled in a desiccator for 15 minutes, after which their mass was recorded as m_1. Then, 3 g of koji was placed on the dried plate, and the combined mass was recorded as m_2. The sample was dried in an oven at 105°C for 24 hours, then cooled in a desiccator for 10 minutes and weighed again to obtain m_3. Weighing continued until a constant mass was reached. Moisture content was calculated using the following formula:

$$\text{Moisture content (\%)} = \frac{m2-(m3-m1)}{m2} \times 100 \quad \textbf{(Eq. 1)}$$

where $m1$ is the mass of the dried plate for 3 hours (g), $m2$ is the mass of the koji sample (g), and $m3$ is the mass of the sample and plate after dried for 24 hours (g)

2.3.2 Protease activity

Protease activity was analyzed according to [2]. The preparation of chemicals involved preparing a sodium phosphate buffer solution, preparing 1.5g of casein, 0.44M trichloroacetic acid (TCA), Folin-Ciocalteu's phenol reagent, and 0.55M Sodium Carbonate (Na_2CO_3) solution. Four test tubes were used for the preparation of samples. The samples were incubated at 37°C for 10 minutes, followed by enzyme samples, TCA solution, and starch solution. After 30 minutes, each tube was filtered and added to new test tubes. The absorbance of the test sample and blank sample was measured using a UV-VIS Spectrophotometer at 660 nm. The results showed that the enzyme samples were insoluble and the TCA solution stopped the reaction. To determine the protease activity in units/ml, apply the following equation (**Eq. 2)**

$$\text{Protease activity} = \frac{mg\ of\ tyrosine\ release \times V_{assay}}{V_{sample} \times time} \times \text{df} \quad \textbf{(Eq. 2)}$$

where mg of tyrosine release (m_{assay} and m_{blank}) were obtained from the standard curve of tyrosine standard 0.1g/ml using the absorbance value, V_{assay} was the volume used for colorimetry determination (11 mL), $time$ is referred to the reaction of the enzyme (min), V_{sample} was the volume of the sample (extracted enzyme)

2.4 Moromi quality analysis

2.4.1 Total acidity

Total acidity (TA) was determined by titration method according to [3]. First of all, the sample was diluted 20 times by adding 1 ml of sample with 19 ml of distilled water, then mixed using a vortex mixer. Next, pump 20 ml of the diluted sample was mixed with 60 ml of DW in a beaker. Then the mixture was titrated with (0.05 mol/L) NaOH solution to pH = 8.2. The volume consumed of NaOH was then noted down as a sample test. After that, 80 ml of distilled water was used to titrate with a (0.05 mol/L) NaOH solution to pH = 8.2, and the volume of NaOH was noted for the blank sample. So, the volume of NaOH consumed was used to calculate the TA content. The TA content of the sample was determined following the calculation below.

$$\text{TA (g/100mL)} = \frac{C_{NaOH} \times V_{NaOH} \times 0.09}{V_{sample}} \times 100 \qquad \textbf{(Eq. 3)}$$

where C_{NaOH} is the concentration of NaOH used (0.05 mole/L), V_{NaOH} is the volume of NaOH used in the sample test (mL), and V_{sample} is the initial volume of the sample used (1 mL)

2.4.2 Amino nitrogen

Amino nitrogen (AN) was determined by the titration method according to [3]. The sample was diluted 20 times by mixing 1 ml of the sample with 19 ml of distilled water, followed by vortex mixing. 20 ml of the diluted sample was combined with 60 ml of distilled water in a beaker and titrated with 0.05 mol/L NaOH solution to pH 8.2. After adding 10 ml of formaldehyde solution, titration continued with NaOH to pH 9.2. The volume of NaOH used was recorded. A similar procedure was followed for the blank sample (80 ml of distilled water), and the volume of NaOH consumed was used to calculate the amino nitrogen (AN) content (g/mL). The content of the sample was determined following the calculation below.

$$\text{AN} = \frac{C_{NaOH} \times [V_{test} - V_{blank}]_{NaOH} \times 0.014}{V_{sample}} \times 100 \qquad \textbf{(Eq. 4)}$$

where C_{NaOH} is the concentration of NaOH used (0.05 mole/L), V_{test} is the volume consumed of NaOH used in the sample test (mL), V_{blank} The volume consumed of NaOH was used in the blank (mL), and V_{sample} is the initial volume of the sample used (1 mL)

2.4.3 Reducing sugar

Reducing sugar (RS) was determined by the 3,5-dinitrosalicylic acid method (DNS) according to [4]. First of all, to prepare the DNS solution, 1 g of DNS was dissolved in 20 ml of 2N NaOH (Solution A) and 30 g of potassium sodium tartrate tetrahydrate in 50 ml of water (Solution B). The two solutions were mixed and diluted to 100 ml with distilled water. For the moromi samples, 1 ml was diluted with 49 ml of water, mixed with 2 ml of DNS solution, and heated at 100°C for 5 minutes. After cooling, 7 ml of distilled water was added, and absorbance was measured at 540 nm using a UV-VIS Spectrophotometer. The absorbance value of each tube was recorded.

2.4.4 Sensory evaluation

The sensory evaluation of soy sauce involved 80 naïve panelists from the Institute of Technology of Cambodia. Panelists rated five soy sauce samples on various attributes using a nine-point hedonic scale. The samples included a control from a local market and four made from whole and defatted soybeans, with and without yeast, in 22% salt brine.

2.4.5 Statistic analysis

Data analysis was conducted using IBM SPSS (version 25.0). Results are presented as means ± SD. ANOVA and Duncan's test were used to determine significant differences at a 95% confidence interval.

3. RESULTS AND DISCUSSION

3.1 Koji quality analysis

3.1.1 Moisture content

The moisture content is one of the important parameters during the koji fermentation stage. According to **Fig. 1,** the moisture content at 0h for DS1, DS2, WS1, and WS2 was 39.87%, 35.59%, 49.01%, and 46.67%, respectively. After 72h of fermentation, it decreased to 23.41%, 21.47%, 25.68%, and 23.51%, showing significant differences ($p \leq 0.05$). Accompanying the results of [5], faster moisture reduction leads to a richer soy sauce with longer shelf life, while whole soybeans, with slower moisture loss, create a more complex flavor. According to [6], the moisture decreases as the mold uses water to grow mycelia.

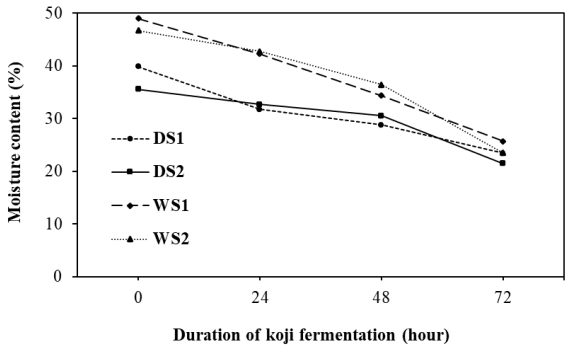


Fig. 1. Moisture content of koji during koji fermentation from 0 to 72 h.

3.1.3 Protease activity

Proteases are a fundamental group of enzymes known for their diverse physiological roles and biotechnological applications [7]. Protease activity increased in DS1 and DS2 due to higher protein content in defatted soybeans, promoting *A. oryzae* growth and enzyme production during koji fermentation [8]. Neutral protease activity increased over 40 hours, aligning with *A. oryzae* growth stages: spore germination (0–12 h), hyphal growth (12–32 h), and sporulation (32–44 h). The rise in protease activity may be due to increased pH and enzyme production by *A. oryzae*, which breaks down proteins into amino acids as fermentation progresses [2].

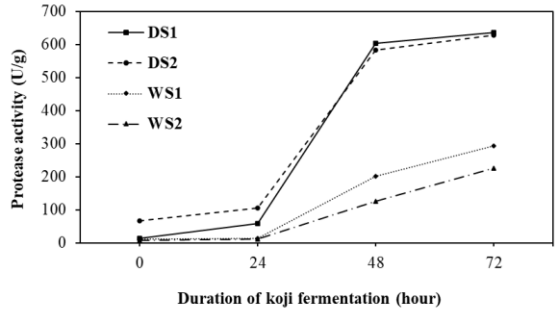


Fig. 2. Protease activity of koji during koji fermentation from 0 to 72 h.

3.2 Moromi quality analysis

3.2.1 Total acidity

Total acidity (TA) is a crucial marker of soy sauce quality, as it directly reflects the generation of organic acids during the fermentation process [9]. According to **Fig. 3,** the result illustrated that the TA value at week 0 of moromi fermentation, the TA values of DS1, DS2, WS1, and WS2 were significantly different ($p \leq 0.05$). DS1 had an initial acidity of 0.43 ± 0.03g/100ml, DS2 started at 0.29 ± 0.03g/100ml, while WS1 and WS2 had lower acidity levels of 0.20 ± 0.03g/100ml and 0.14 ± 0.13g/100ml, respectively. These differences are due to the higher lipid content in whole soybeans, which acts as a buffer, delaying acid production, while defatted soybeans, with lower lipid content, allow fermentation to begin more quickly. Based on the studies of [10]. Additionally, in the early stage (0-40 days), bacterial fermentation drove the rapid increase in total acidity. This was due to the metabolism of carbon sources producing acids, and free amino acids generated from protease hydrolysis, which led to an initial rise in total acid content [11].

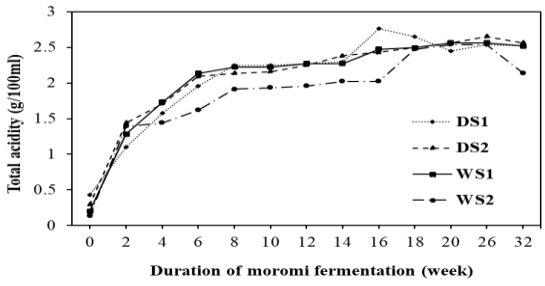


Fig. 3. Total acidity of moromi during moromi fermentation from 0 to 32 weeks.

3.2.2 Amino nitrogen

AN is a crucial metric for assessing the quality of soy sauce products, as it includes both sensory attributes and nutritional values. According to **Fig. 4,** the initial amino nitrogen of moromi fermentation for week 0 samples DS1, DS2, WS1, and WS2 was 0.16 ± 0.00g/100ml, 0.10 ± 0.00g/100ml, 0.09 ± 0.00g/100ml, and 0.05 ± 0.13g/100ml respectively with significant difference ($p \leq 0.05$). Based on [12], the amino nitrogen (AN) content increased continuously due to the microbial breakdown of proteins during fermentation. However, high brine content reduced enzyme activity, and the metabolism of salt-tolerant yeast and lactic acid bacteria further lowered TN and AN during prolonged fermentation [10].

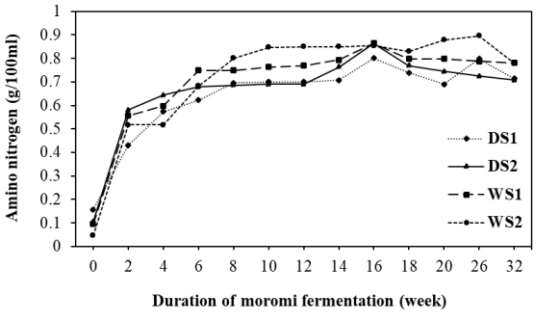


Fig. 4. Amino nitrogen of moromi during moromi fermentation from 0 to 32 weeks.

3.2.3 Reducing sugar

Reduced sugar in soy sauce moromi is a quality indicator, while Maillard reaction products contribute to its distinct flavor [13]. According to **Fig. 5,** at the start of moromi fermentation, RS content varied significantly: 0.52 ± 0.01g/100ml (DS1), 0.46 ± 0.01g/100ml (DS2), 1.07 ± 0.03g/100ml (WS1), and 1.63 ± 0.47g/100ml (WS2) ($p \leq 0.05$). Fungal amylase activity, which breaks down starch, primarily influenced RS levels, though high salt inhibited enzyme activity [10]. Additionally, Over the 32 weeks of fermentation, RS levels in DS1 and DS2 slightly decreased, particularly between weeks 10-18 for DS1 and weeks 14-20 for DS2.

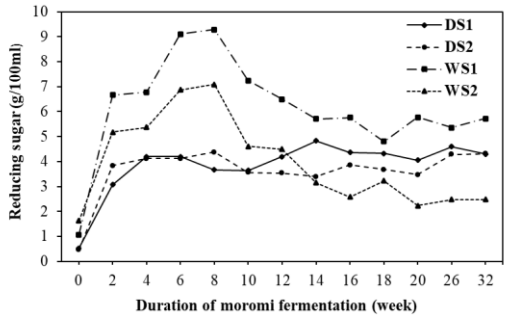


Fig. 5. Reducing sugar of moromi during moromi fermentation from 0 to 32 weeks.

3.2.4 Sensory evaluation

A sensory evaluation of soy sauce samples from four different conditions was conducted with 80 panelists, who scored color, texture, odor, aroma, sweetness, saltiness, flavor, and overall acceptability using a 9-point hedonic scale. ZH1, a commercial soy sauce, scored the highest across all attributes, with scores ranging from 6.93 to 7.18, indicating a strong preference for it over the other samples. Among the experimental samples, DS2 (defatted soybeans with yeast) performed the best, scoring 6.57 overall, followed by DS1 (6), WS1 (6.49), and WS2 (5.91). Furthermore, DS2 performed better than the other three other samples in most out of the samples tested. Overall, the study shows that, when compared to other samples, the soy sauce made from defatted soybeans with yeast (DS2) has better sensory qualities, though it still falls short of commercial control.

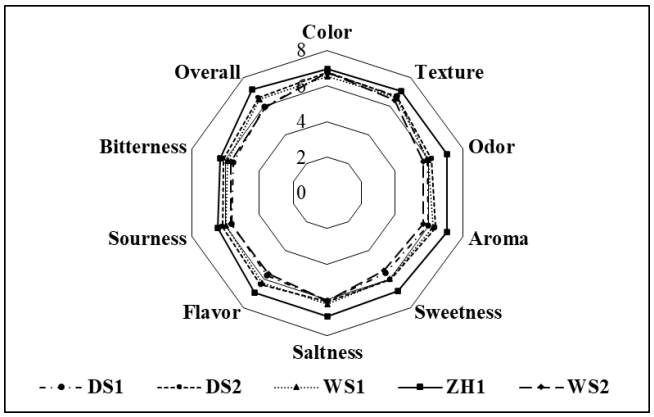


Fig. 6. Sensory evaluation of different soy sauces after 32-week moromi fermentation.

4. CONCLUSIONS

This study explored soy sauce fermentation under four conditions using *A. oryzae* mold and yeast over 32 weeks, analyzing the physicochemical properties of koji and moromi. Key differences were observed in moisture content, and protease activity during koji fermentation, as well as Total acidity, amino nitrogen, and reducing sugar during moromi fermentation. The results indicated that defatted soybeans and whole wheat without adding yeast and defatted soybeans and whole wheat with adding yeast, which used defatted soybeans, exhibited higher protease activity compared to whole soybeans and whole wheat without adding yeast and whole soybeans and whole wheat with adding yeast, suggesting better enzyme production crucial for soy sauce. Sensory evaluations favored defatted soybeans and whole wheat with adding yeast, with whole soybeans and whole wheat without adding yeast also scoring well, while defatted soybeans and whole wheat without adding yeast and whole soybeans and whole wheat with adding yeast were deemed acceptable. The results met China's soy sauce standards, suggesting that defatted soybeans could be a viable alternative to whole soybeans in soy sauce production. As a recommendation, moromi fermentation for over 6 months, identifying the specific species of the isolated mold and yeast, as

well as low-salt content fermentation, the aroma profile of soy sauce, and the nutritional value of the final product should be investigated in future research.

ACKNOWLEDGMENTS

The work was funded by Cambodia Higher Education Improvement Project (Credit No. 6221-KH, SGA#05).

REFERENCES

[1] D. K. O'toole, *The role of microorganisms in soy sauce production*, 1st ed., vol. 108. Elsevier Inc., 2019. doi: 10.1016/bs.aambs.2019.07.002.

[2] N. Xuan Hoang, C.-H. Ting, and C.-K. Hsu, "Optimization of Protease and Amylase Activities Derived from Soy Sauce Koji," *J. Food Sci. Nutr. Res.*, vol. 03, no. 03, pp. 153–170, 2020, doi: 10.26502/jfsnr.2642-11000046.

[3] X. Jiang *et al.*, "Isolation, identification and application on soy sauce fermentation flavor bacteria of CS1.03," *J. Food Sci. Technol.*, vol. 56, no. 4, pp. 2016–2026, 2019, doi: 10.1007/s13197-019-03678-w.

[4] Y. Zhang *et al.*, "Impact of steam explosion pretreatment of defatted soybean meal on the flavor of soy sauce," *Lwt*, vol. 156, p. 113034, 2022, doi: 10.1016/j.lwt.2021.113034.

[5] C. C. Chou and M. Y. Ling, "Biochemical changes in soy sauce prepared with extruded and traditional raw materials," *Food Res. Int.*, vol. 31, no. 6–7, pp. 487–492, 1998, doi: 10.1016/S0963-9969(99)00017-4.

[6] C. Chancharoonpong, P.-C. Hsieh, and S.-C. Sheu, "Enzyme Production and Growth of Aspergillus oryzae S. on Soybean Koji Fermentation," *APCBEE Procedia*, vol. 2, pp. 57–61, 2012, doi: 10.1016/j.apcbee.2012.06.011.

[7] R. J. S. De Castro and H. H. Sato, "Advantages of an acid protease from Aspergillus oryzae over commercial preparations for production of whey protein hydrolysates with antioxidant activities," *Biocatal. Agric. Biotechnol.*, vol. 3, no. 3, pp. 58–65, 2014, doi: 10.1016/j.bcab.2013.11.012.

[8] Y. Zhao, D. Sun-Waterhouse, M. Zhao, Q. Zhao, C. Qiu, and G. Su, "Effects of solid-state fermentation and proteolytic hydrolysis on defatted soybean meal," *Lwt*, vol. 97, pp. 496–502, 2018, doi: 10.1016/j.lwt.2018.06.008.

[9] W. Zhou, D. Sun-Waterhouse, J. Xiong, C. Cui, W. Wang, and K. Dong, "Desired soy sauce characteristics and autolysis of Aspergillus oryzae induced by low temperature conditions during initial moromi fermentation," *J. Food Sci. Technol.*, vol. 56, no. 6, pp. 2888–2898, 2019, doi: 10.1007/s13197-019-03742-5.

[10] N. X. Hoang, S. Ferng, C. H. Ting, W. H. Huang, R. Y. Y. Chiou, and C. K. Hsu, "Optimizing the initial moromi fermentation conditions to improve the quality of soy sauce," *Lwt*, vol. 74, pp. 242–250, 2016, doi: 10.1016/j.lwt.2016.07.049.

[11] B. Liu, Z. Cao, L. Qin, J. Li, R. Lian, and C. Wang, "Investigation of the synthesis of biogenic amines and quality during high-salt liquid-state soy sauce fermentation," *Lwt*, vol. 133, no. March, p. 109835, 2020, doi: 10.1016/j.lwt.2020.109835.

[12] N. X. Hoang *et al.*, "Effect of initial 5 days fermentation under low salt condition on the quality of soy sauce," *Lwt*, vol. 92, pp. 234–241, 2018, doi: 10.1016/j.lwt.2018.02.043.

[13] S. Ferng *et al.*, "A Controlled Fermentation Environment for Producing Quality Soya Sauce," *Adv. Technol. Innov.*, vol. 4, no. 1, pp. 11–20, 2019.

Food Technology and Nutrition (FTN)

3 Sub-parallel Sessions

"FTN-3: Building Resilient and Sustainable Food Supply Chains for the Future"

Chair: Assist. Prof. Dr. KHOEURN Kimleang, Co-Chair: Dr. NGET Sovannmony

No.	Topic
1	Identification of Cuticle-binding Bacteria from Meloidogyne graminicola as Potential Biocontrol Agent Affecting the Pest Mobility *Authored by: Sreynich YAN, Michel BARBIER, Malyna SUONG, Stéphane BELLAFIORE*
2	Evaluation of Nutrient Profiles of Single and Mixed Cover Crops Species Collected from Rovieng, Preah Vihear Province *Authored by: Menghuy HAN, Kimleang KHOEURN, Peanny HOUNG, Monika MICH, Soukheng MENG, Mathilde SESTER, Florent TIVET*
3	Evaluation of Rice Seed Quality of Cambodian Varieties Produced under Conventional and Agroecology Practices *Authored by: Singy HO, Malyna SUONG, Kakada OEUM, Vichny VICHETH1, Rada KONG, Florent TIVET, Myriam ADAM, Lionel MOULIN*
4	Influence of the Soil Physicochemical Properties on Phka Rumdoul Rice Quality and Yield in Cambodia *Authored by: Vichny VICHETH, Stephane BELLAFIORE, Sereyvath YOEUN, Sreymom SIENG, Socheat KEO, Sambo PHEA, Chanvorleak PHAT*
5	Analysis of Pesticide Residues in Rice Grains using QuEChERS Method Followed by GC-MS *Authored by: Aykim HENG, Chanvorleak PHAT, Leanghok BUN, Mathilde SESTER, Florent TIVET, Thanon NUON, Sereyvath YOEUN*
6	Assessment of Pesticide Residues in Rice and Water from Kanghot Rice Fields in Battambang *Authored by: Leanghok BUN, Chanvorleak PHAT, Mathilde SESTER, Florent TIVET, Thanon NUON, Sereyvath YOEUN*
7	Assessment of Pesticide Residues in Rice and Irrigation Canal: A Study of Agrochemical-Free Farming in Rovieng District, Preah Vihear Province *Authored by: Rethireach LY, Sereyvath YOEUN, Leanghok BUN, Mathilde SESTER, Florent TIVET, Rama SRY, Chanvorleak PHAT*
8	Seasonal Pesticide Residues in Surface Waters of Banteay Meanchey Province, Cambodia *Authored by:* Vivath TIT, Leanghok BUN, Sereyvath YOEUN, Chanvorleak PHAT
9	Development of Dishwashing Liquid Through Utilization of Various Active Ingredients *Authored by: Kosal VAPHAT, Peany HOUNG, Keakaknika LY, Samphois NUT*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Identification of cuticle-binding bacteria from *Meloidogyne graminicola* as potential biocontrol agent affecting the pest mobility

Sreynich YAN [1], Michel BARBIER [1, 2], Malyna SUONG [1, 3*], Stéphane BELLAFIORE [1, 2]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Plant Health Institute of Montpellier (PHIM), IRD, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France*

[3] *International Relations Office, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: malyna.suong@itc.edu.kh*

Abstract: *Meloidogyne graminicola is a root-knot nematode that threatens global food security. International regulations limit the use of nematicides, which are harmful to the environment and human health. The use of biocontrol agents is one of the most reliable methods of managing this pest. Suppressive soil has been described to have the ability to reduce the nematode by direct antagonism or indirectly through the microorganisms present in the soil, which are closely related to soil's physico-chemical properties. Meloidogyne spp. has been shown to interact with soil microorganisms before invading the plants and causing disease. This study aimed to survey the nematode cuticle to search for biocontrol agents as an alternative to chemical nematicides. In this study, the juvenile 2 (J2) of Meloidogyne graminicola was first disinfected before being exposed to the suppressive soil suspension in vitro. The bacteria attached to the nematodes were then extracted from the nematode cuticle and isolated individually. Once the bacteria had been isolated, their gDNA was extracted their identity was determined using PCR barcoding and Sanger sequencing. All isolated bacteria were tested for their potential antagonist activity against* M. graminicola *in an* in vitro *assay developed in the FEF Healthy Rice laboratory. As a result, sixteen genera of bacteria were identified from the nematode cuticle. Among these 16 bacteria genera, nine were found to be antagonistic to M. graminicola. This preliminary finding could be used for further development of biocontrol to minimize the nematode pest movement and reduce its impact on rice crop, contributing to sustainable rice production.*

Keywords *Meloidogyne graminicola*; Plant parasitic nematode; Nematode antagonist; Suppressive soil; Sustainable agriculture.

1. INTRODUCTION

Plant parasitic nematode is global threat to the global food security. *Meloidogyne graminicola (Mg),* a soil-borne pest, is a major cause of cereal yield loss especially rice in many country [1, 2, 3]. *M. graminicola* parasitic on the plant root and forming a root swelling call gall. The gall disrupts water and nutrient uptake resulting in poor growth and reproduction of plant which affects yield losses in the crop. In response to these challenges, various strategies have been used during the green revolution, including the development of high-yield crop varieties, the implementation of irrigation systems, and the application of fertilizers and pesticides. However, the extensive use of chemical fertilizers, pesticides, and intensive farming methods has resulted in a multitude of long-term environmental and human health concerns, including soil degradation, water contamination, emerging pesticide-resistant pathogens and biodiversity loss [4, 5, 6].

Finding environmentally friendly alternative methods to control parasites is a priority for sustainable development and ensuring food safety. Biocontrol is a promising alternative and ecologically friendly methods to control plant parasitic nematodes. Biocontrol strategies consist of using beneficial microbial antagonists or resistant plants to reduce the incidence of pathogens [7]. Several types of organisms, including fungi, bacteria, viruses, protists, nematodes, and other invertebrates, have been found to parasitize or prey upon plant parasitic nematodes [8]. Biological control agents of

[1*] Corresponding author: Malyna Suong
E-mail: malyna.suong@itc.edu.kh; Tel: +855-86 936 525

plant-parasitic nematodes are typically specific natural enemies. *Pasteuria penetrans* is an endospore-forming bacterium that persists in the soil until a suitable nematode host encounters the spore [9]. The *Pasteuria* spores adhere to the nematode cuticle (outer surface), infect the nematode and develop inside of the nematode body. In addition, several species of *Burkholderia* spp. have been documented for their nematicidal activity against *Meloidogyne* spp. by reducing root gall formation and the mobility of stage 2 juveniles [10, 11]. The objectives of this study were to isolate and identify the bacteria that bind to the nematode cuticle and to test the nematicidal potential of these bacteria on nematode mobility.

2. METHODOLOGY

2.1 Soil sampling

The soil sample was collected from Stung Chinit rice study field which located in Sontuk district, Kampong Thom province of Cambodia (12°32'55.2"N 105°08'50.0"E). The soil showed suppressive activity against *Meloidogyne graminicola* in a previous study (Barbier et al., unpublished).

2.2 Bacteria extracted from the nematode cuticle

The nematodes were first extracted from the infested root. First, the rice was removed from the column. Then, the root was carefully washed with tap water to remove the sand. Then, the roots were cut into 2-centimeter pieces and placed into a beaker. Approximately 50 ml of 0.7% sodium hypochlorite was added to the beaker, covering all roots, and left for five minutes. The solution was then passed through three sieves of different sizes in the micron range to rinse roots and nematodes from the bleach with tap water (1 liter) and 11 to collect clean nematodes on the layer from the smaller micron sieve (25 microns). The largest sieve was placed on top of the smaller sieves (250 microns, 100 microns, and 25 microns). The roots were gently pressed against the 250-micron sieves in order to release the eggs and juveniles that were present inside the root. Nematodes are typically 300–400 micrometers long and 20 micrometers wide, which they can hold above the last 25-micrometer sieve. The water that was left over after passing through the 25-micron sieve was collected in a beaker. The number of nematodes was then counted using an inverted phase contrast microscope. This microscope uses 1 ml of well-homogenized solution in a Sedgwick Rafter cell S52 counting cell. The nematodes were leave to hatch for 48h then about 10000 J2 of freshly hatch nematodes were then surface sterilized with chlorhexidine gluconate (0.004%) for 5 minutes at room temperature before rinsing twice with 1 ml sterile distilled water. Nematodes were suspended in 100 µL of an antibiotic mixture consisting of streptomycin at 2 µg/mL and kasugamycin at 2 µg/mL and left for 24 hours at room temperature. The nematodes were then washed twice in 1 ml sterile distilled water. At same time, soil suspensions were obtained by suspended ten gram of soil in 40 mL of the mineral water and incubating at 25°C with shaking at 100 rpm for 14 hours. After incubation, the soil suspensions were filtered with a coffee filter (20 µm pore size) to remove the largest soil particles and collect the solute with microbiota. After that, the suppressive soil suspension was incubated with the disinfected J2 of *M. graminicola* for 24 hours at 25°C. The soil-bacteria suspension and nematodes were then passed over a 10um diameter sieve to retain only the nematodes (animal approx. 400 µm long and 10 µm in diameter). They were then rinsed with 40 ml sterile distilled water to remove microorganisms not specifically bound to their cuticle. The nematodes and their cuticle-bound microbiota were then resuspended in 1 ml sterile distilled water and transferred to a 1.5 ml Eppendorf tube containing a sterile 5 mm ceramic bead. The bacterial suspension was then extracted from the nematode cuticle by vortex with sterile 5 mm ceramic bead at maximum speed for 2min. The suspension of shredded nematodes was diluted 10 and 100 times. Then 100 µL of the non-diluted suspension of shredded nematodes and the two dilutions were spread out with 15 sterile glass beads on Actinomycetes, Luria-Bertani (with low salt), Middlebrook, R2A and Tryptone Soya Agar (all from Sigma-Aldrich®) culture media. After two days of growth at 25°C, the isolated colonies were spread out in TSA culture media, and after two days of growth were cultivated Tryptone Soya Broth (Sigma-Aldrich®) incubated one night at 25°C with a 100 rpm agitation. Then the liquid cultures were cryo-conserved at -80°C in 20% glycerol. Following this, a few microliters of each bacteria were used to perform the colony PCR. The 16S rRNA genes of the purified bacterial isolates were amplified by PCR (directly on boiled colonies): Two microliters of bacterial liquid culture in Tryptone Soya Broth (Sigma-Aldrich®) were placed into a PCR mix containing the universal couple of forward FGPS6 (5'-GGAGAGTTAGATCTTGGCTCAG-3') and reverse FGPS1509 (5'-AAGGAGGGGATCCAGCCGCA-3') primers [12]. The PCR amplifications were performed in 25 µl final reaction mixture in sterile water containing DreamTaq Buffer (5X), 2.5 mM of each dNTP, 0.4 µM of each primer and 0.125 µl of DreamTaq DNA Polymerase (Thermo Fisher Scientific). The PCR thermocycler performed an initial denaturation at 94°C for 2 min, 35 cycles of denaturation, annealing, and elongation at 94 °C for 30 sec, 56°C for 30 sec, and 72°C for 30 sec, respectively, and a final elongation step at 72°C for 10 min. To check the amplicon size (1,480 bp), they were visualized under UV light after gel electrophoresis (1% agarose, TAE buffer, 100 V for 25 min). The PCR products were sent to Macrogen (Seoul, South Korea) for sequencing using the internal amplicon primer 16S-1080_r (5'-GGGACTTAACCCAACATCT-3').

2.3 Bacteria identification

Amplicon sequenced using primer 16S-1080_r, allowing access to a sequence of approximately 1000 base pairs. The 16s rRNA gene sequences were checked and cleaned using Chromas version 2.2.6, and the bacteria from each sequence were then identified to genus level using NCBI Blast. The Phylogenetic and molecular evolutionary analyses were conducted using MEGA version 12 [13].

2.4 Nematicide bioassay

Based on genus identification, sixteen of the nematode-binding bacterial strains representing each genus were tested for nematicidal activity against *Meloidogyne graminicola.* The physiological water was served as control. **Table 1.** Represent the bacteria strains which were used for the nematocidal bioassay against J2 of *M. graminicola*. Each of the bacterial strains was plated on Tryptic Soy Agar 50% (TSA50) and incubated at 28°C for two days. The isolated colonies from each strain were then cultured in Tryptic Soy Broth 50% (TSB50) liquid culture for 24 hours at 180 rpm. The bacterial cultures were then diluted with physiological water (NaCl 0,9%) to obtain 0.8 optical density (OD) and mixed with 1 ml of freshly hatched *M. graminicola* J2 larvae in a glass tube and incubated at room temperature for five hours. The mixture was then oxygenated with air pump before being placed in microcolumns and incubated at room temperature for 20 hours. Finally, the microcolumns were removed, the suspension was concentrated and oxygenated again before the number of mobile and non-mobile larvae was counted and the remaining volume was measured to evaluate the experiment. The experiment was performed twice independently with three replicates for each experiment.

Table 1. List of bacteria strains that used to test of nematicide effect against J2 *M. graminicola*

Bacteria codes	Genus
ctrl	Physiological water
1	*Moraxella*
14	*Sphingomonas*
21	*Rothia*
25	*Priestia*
30	*Microbacterium*
57	*Methylobacterium*
58	*Micrococcus*
60	*Acinetobacter*
68	*Janibacter*
72	*Pantoea*
77	*Sinomonas*
79	*Agrobacterium*
85	*Paenibacillus*
89	*Rhizobium*
94	*Burkholderia*
101	*Bosea*

2.5 Statistical analysis

The statistical analysis was performed using RStudio software (version 4.4.2, 2024). The comparison between the 16 different bacteria strain was made using the non-parametric Kruskal-Wallis Test, with a p-value threshold of 0.05. Below this threshold, the observed differences were considered significant.

3. RESULTS AND DISCUSSION

A total seventy-nine bacteria strains were isolate and purified from nematode binding bacteria. All the strains were preserved in 20% glycerol at -80°C for long term storage. Each strain was identified to the genus level by using the 16S rRNA gene sequence. Sixteen bacteria genera were identified and blasted with NCBI GenBank database. The phylogenetic tree **Fig. 2.** Each label on the tree represents a bacteria genus and the number on the branch show the bootstrap percentage, value above 70 are considered common ancestor. The scale bar shows genetic distance 0.05. On the top cluster, *Sphingomonas* have a huge clade of sequences in tightly grouped with short internal branches, this suggest that this species has closely related strains with the 99% values of bootstrap value. *Agrobacterium*, *Rhizobium* and *Methylobacterium* are group together in the middle cluster show moderate relatedness and evolutionary distance from *Sphingomonas.* The lower cluster, *Acinetobacter, Moraxella, Pantoea* and *Burkholderia* are group further from *Sphingomonas* which indicates that they are more distantly related. The bottom cluster, *Micrococcus, Rothia, Janibacter, sinomonas and Paenibacillus* were place in the different branch with 60% bootstrap which shows greater evolutionary distance from the middle cluster bacteria.

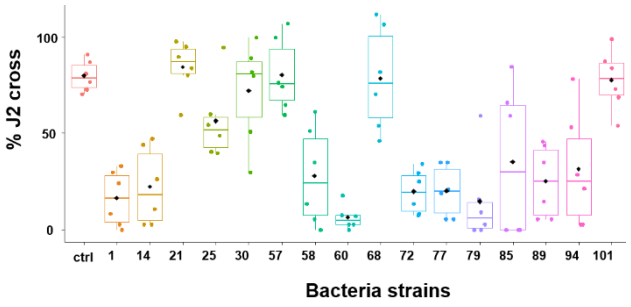


Fig. 1. Evaluation of nematicide effect of 16 bacterial strains on the mobilities of juvenile stage 2 of *M. graminicola*.

A box plot illustrating the number of cross nematodes after 20 hours incubation in bacteria strains or physiological water is shown in **Fig. 1.** This figure shows the nematicide activities of 16 different bacteria strains against the juvenile stage 2 *M. graminicola* on their mobilities compared to the control physiological water. The control physiological water has high mean value almost 90% of cross J2. Some strains such as 1, 60, 72, 77, 79, 89 and 94 which were respectively from the genus of *Moraxella, Acinetobacter, Pantoea, Sinomonas, Agrobacterium, Rhizobium* and *Burkholderia* were shown to have mean value less than 50% of cross nematode. This indicate that this bacterium has an ability to reduce the mobilities of the nematode. However, the other bacteria strains such as 21, 25, 30, 57, 58, 68 and 101 which were respectively from the genus of *Rothia*, *Priestia, Microbacterium, Methylobacterium, Micrococcus, Janibacter* and *Bosea* showed to have a mean value more than 50%. Strain 60 (*Acinetobacter)* showed to have the lowest mean value 15% and less tendency compare to a genus from *Agrobacterium, Rhizobium* and *Burkholderia.*

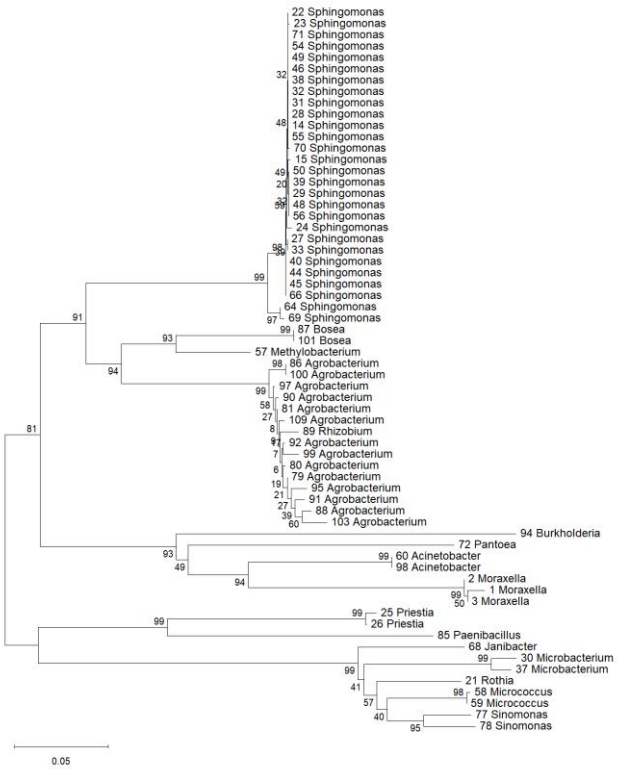


Fig. 2. Phylogenetic analysis of 64 bacteria strains identified from nematode cuticle. This phylogenetic tree was done using the Neighbor-Joining method based on 16S rRNA gene sequence relate. Numbers at each branches indicate the bootstrap percentage of 1,000 replications. The tree construct using MEGA 12 software.

Acinetobacter sp. was found attached to the cuticle of the PPN *M. hapla*, *M. incognita* and *P. penetrans* [13, 14]. It has also been reported to impose a nematicide activity against *C. elegans* [13, 15, 16]. *Acinetobacter* sp. also affects J2 mortality, J2 motility and egg hatch *M. haphla* in vitro bioassay [18]. A species from *Burkholderia sp. JB-2* efficiently controlled the nematode population in soil and egg mass formations of *M. incognita* in tomato [11].

4. CONCLUSIONS

After isolated the bacteria that binding to nematode cuticle, 16 genera of bacteria were identified by sequencing. We identified the impact of these bacteria on J2 *M. graminicola in vitro*. It is interesting to note that several of these bacteria could be used as biocontrol agents after verifying their harmlessness to other organisms. Among them, *Actinetobacter* shows a highly effective nematicidal effect on J2 mobility. Further study is therefore needed to determine whether the nematicidal effect of this bacteria is specific to the J2 stage of *M. graminicola*, or whether it acts on the different larval stages of the nematode. It will also be necessary to test whether this bacterium can indeed reduce parasitism by *M. graminicola* under real-life conditions, by carrying out experiments to infect rice with *M. graminicola* using soil previously inoculated with *Actinetobacter*.

ACKNOWLEDGMENTS

The authors would like to thank the Ministry of Europe and Foreign Affairs through the project "FEF-Healthyrice" for a master grant fellowship, the institute of Technology of Cambodia hosting the Plant health laboratory, the French Research Institute for Sustainable Development (IRD) via the project "Laboratory of Excellence in co-Engineering for Sustainable Agrosystems-LMI LEAD", and CIRAD and GDA-DALRM managing the biplots systems as part of the ASSET project (AFD-EU).

REFERENCES

[1] S. Mantelin, S. Bellafiore, and T. Kyndt, "*Meloidogyne graminicola*: A major threat to rice agriculture," *Molecular plant pathology*, vol. 18, Mar. 2016, doi: 10.1111/mpp.12394.

[2] D. Brand, C. Soccol, S. Abdulhameed, and S. Roussos, "Production of fungal biological control agents through solid state fermentation: A case study on *paecilomyces lilacinus* against root-knot nematodes, *Micología Aplicada. Internacional*, vol. 22, Jan. 2010.

[3] R. Plowright and J. Bridge, “Effect of Meloidogyne Graminicola (Nematoda) On the Establishment, Growth and Yield of Rice Cv Ir36,” Jan. 1990, doi: 10.1163/002925990X00059.
[4] J. Desaeger, C. Wram, and I. Zasada, “New reduced-risk agricultural nematicides - rationale and review,” *J Nematol*, vol. 52, pp. e2020-91, Sep. 2020, doi: 10.21307/jofnem-2020-091.
[5] A. L. Taylor, “NEMATOCIDES AND NEMATICIDES - A HISTORY,” *Nematropica*, pp. 225–232, Dec. 2003, Accessed: Feb. 03, 2025. [Online]. Available: https://journals.flvc.org/nematropica/article/view/69679
[6] J. G. Jones, N. M. Kleczewski, J. Desaeger, S. L. F. Meyer, and G. C. Johnson, “Evaluation of nematicides for southern root-knot nematode management in lima bean,” *Crop Protection*, vol. 96, pp. 151–157, Jun. 2017, doi: 10.1016/j.cropro.2017.02.015.
[7] D. Kour *et al.*, “Chapter 11 - Potassium solubilizing and mobilizing microbes: Biodiversity, mechanisms of solubilization, and biotechnological implication for alleviations of abiotic stress,” in *New and Future Developments in Microbial Biotechnology and Bioengineering*, A. A. Rastegari, A. N. Yadav, and N. Yadav, Eds., Elsevier, 2020, pp. 177–202. doi: https://doi.org/10.1016/B978-0-12-820526-6.00012-9.
[8] G. R. Stirling, *Biological control of plant-parasitic nematodes: soil ecosystem management in sustainable agriculture*, 2nd ed. Wallingford: CABI, 2014.
[9] R. Dyrdahl-Young and P. DiGennaro, “*Pasteuria penetrans* (*Bacilli: Bacillales: Pasteuriaceae*): EENY718/IN1228, 9/2018,” *EDIS*, vol. 2018, no. 6, Nov. 2018, doi: 10.32473/edis-in1228-2018.
[10] R. Zhang *et al.*, “Nematicidal Activity of *Burkholderia arboris J211* Against *Meloidogyne incognita* on Tobacco,” *Front Microbiol*, vol. 13, p. 915546, Jun. 2022, doi: 10.3389/fmicb.2022.915546.
[11] J.-H. Kim *et al.*, “Assessment of nematicidal and plant growth-promoting effects of *Burkholderia sp. JB-2* in root-knot nematode-infested soil,” *Front Plant Sci*, vol. 14, p. 1216031, Jul. 2023, doi: 10.3389/fpls.2023.1216031.
[12] A. Sy *et al.*, “Methylotrophic MethylobacteriumBacteria Nodulate and Fix Nitrogen in Symbiosis with Legumes,” *Journal of Bacteriology*, vol. 183, no. 1, pp. 214–220, Jan. 2001, doi: 10.1128/jb.183.1.214-220.2001.
[13] S. Kumar, “MEGA12: Molecular Evolutionary Genetic Analysis Version 12 for Adaptive and Green Computing,” *Molecular Biology and Evolution*, Dec. 2024, doi: https://doi.org/10.1093/molbev/msae263.
[14] M. Adam, A. Westphal, J. Hallmann, and H. Heuer, “Specific Microbial Attachment to Root Knot Nematodes in Suppressive Soil,” *Applied and Environmental Microbiology*, vol. 80, no. 9, pp. 2679–2686, May 2014, doi: 10.1128/AEM.03905-13.
[15] A. Elhady *et al.*, “Microbiomes associated with infective stages of root-knot and lesion nematodes in soil,” *PLOS ONE*, vol. 12, no. 5, p. e0177145, May 2017, doi: 10.1371/journal.pone.0177145.
[16] Y. Deng *et al.*, “Biosafety assessment of *Acinetobacter* strains isolated from the Three Gorges Reservoir region in nematode Caenorhabditis elegans,” *Sci Rep*, vol. 11, no. 1, p. 19721, Oct. 2021, doi: 10.1038/s41598-021-99274-0.
[17] M. Adam, A. Westphal, J. Hallmann, and H. Heuer, “Specific Microbial Attachment to Root Knot Nematodes in Suppressive Soil,” *Applied and environmental microbiology*, vol. 80, Feb. 2014, doi: 10.1128/AEM.03905-13.
[18] O. Topalović, A. Elhady, J. Hallmann, K. R. Richert-Pöggeler, and H. Heuer, “Bacteria isolated from the cuticle of plant-parasitic nematodes attached to and antagonized the root-knot nematode *Meloidogyne hapla*,” *Sci Rep*, vol. 9, no. 1, p. 11477, Aug. 2019, doi: 10.1038/s41598-019-47942-7.

Evaluation of Nutrient Profiles of Single and Mixed Cover Crops Species Collected from Rovieng, Preah Vihear Province

Menghuy HAN [1*], Kimleang KHOEURN [1], Peanny HOUNG[1], Monika MICH[1], Soukheng MENG[1], Mathilde SESTER[2], Florent TIVET[2]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2]*CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[*]*Corresponding author: hanmenghuy@gmail.com*

Abstract: *Cover crops are sustainable agricultural practice that involves plant specific crops to protect and enhance soil health and manage levels of nitrogen (N), phosphorus (P), and potassium (K). These crops were planned to improve soil quality, support plant growth, and provide essential nutrients to agricultural lands in Preah Vihear province in Cambodia. When the land of farm free, it usually lost some nutrient of land at this time the crops was planned to remain the quality of land. The study focused on the analysis of nutrient contents in single and mixed cover crops including moisture, ash content, nitrogen (IC), potassium (IC), and total nitrogen (kjeldalh). Samples were collected from four different locations and analyzed moisture and ash contents were determined through oven drying and furnace ashing at 550°C for 4 h, respectively. Total nitrogen content was analyzed by the Kjeldahl digestion method, while potassium content was determined using IC after deionized water extraction. As results, the moisture content in the cover crops ranged from 4.93% to 7.56%, exhibiting levels suitable for seed storage after drying. Ash content, which reflects mineral accumulation, was highest in Cowpea (up to 12.07%), suggesting its potential for enhancing soil mineral content, while Sesbania showed the lowest ash content (around 6.53%). Nitrogen analysis revealed that leguminous crops such as Cowpea3 (1.15% N) and Sunhemp3 (1.66% N) had higher nitrogen concentrations than non-leguminous crops like Stylo9 (1.46% N) and Kenaf1 (0.67% N), supporting their role in biological nitrogen fixation. Moreover, differences between the Kjeldahl method and Ion Chromatography (IC) were observed in nitrogen quantification, indicating the importance of methodological consideration in nutrient analysis. Regarding potassium content, Kenaf1 demonstrated the highest value (3.46% K), highlighting its potential to improve soil potassium levels, while other crops such as Sumhemp2 (2.78% K) and Sesbania2 (2.85% K) also contributed significantly. The results revealed that Cowpea2 exhibited the highest ash and nitrogen content, highlighting its strong nutrient accumulation capabilities. Kenaf1 showed the highest potassium content, making it beneficial for enhancing soil potassium levels. These findings provide valuable insights into selecting appropriate cover crops for improving soil health, nutrient management, and agricultural productivity.*

Keywords: Nutritional contents; Nitrogen fixation; Soil fertility; Sustainable agriculture.

1. INTRODUCTION

Agricultural has long been a key component of human development, and its sustainability depends largely on healthy soil systems that supply essential nutrients to crops [1]. A cover crop is a plant that is intentionally grown on farmland, not primarily for harvest, but to improve soil health, manage erosion, control weeds, enhance soil fertility, improve biodiversity, and retain moisture. Farmers use cover crops as part of sustainable agricultural practices, often growing them during periods when the primary cash crops are not being cultivated, such as in the off-season or between planting cycles [2]. In modern agriculture, soil degradation and declining fertility are growing concerns, especially in tropical regions like Cambodia. Cover crop plants grown not for harvest but to enhance soil health offer a sustainable solution to these challenges. They help fix atmospheric nitrogen, reduce erosion, restore nutrients from deeper soil layers, and enhancesoil structure through their root system [3]. Leguminous crops like Cowpea and Sunhemp fix atmospheric nitrogen, while others like Kenaf are efficient in potassium accumulation and improving soil quality. Although global research supports the benefits of cover crops, data under Cambodian conditions are limited, making local evaluation essential.

Therefore, this study investigates the nutrients in six cover crops, such as Stylo, Cowpea, Kenaf, Sunhemp, Sesbania, and a mixed cover crop, which is collected from the Rovieng district, Preah Vihear. The primary objective is to evaluate their moisture content, ash content, nitrogen (via Ion

[1]* Corresponding author: Mr. Menghuy Han
E-mail: *hanmenghuy@gmail.com*; Tel: +855-882229057

Chromatography), potassium content (via Ion Chromatography), and total nitrogen (using the Kjeldahl method).

2. METHODOLOGY

2.1 Sample collection and preparation

The experimental site was located in Rovieng District, Preah Vihear Province, in Northern Cambodia. Samples were collected from four field locations with six different cover crop types studied: Stylosanthes guianensis (Stylo), Cowpea (Vigna unguiculata), Sunhemp (Crotalaria juncea), Kenaf (Hibiscus cannabinus), Sesbania (Sesbania rostrata), and a mixed crop containing multiple species. All of the Samples were collected with four quadrat samples (1 m × 1 m) taken per location to ensure consistent sampling size and ensure representative biomass sampling.

Once dried, samples were ground into a fine powder using a stainless-steel grinder. The ground samples were immediately placed into airtight zip-lock bags to prevent contamination from dust, moisture, or other environmental factors. The packed sample was carefully labeled with the sample's location, cover crop type, and field number to ensure traceability. The prepared samples were then transported to the laboratory at the Institute of Technology of Cambodia (ITC) for further analysis.

2.2 Moisture content

Moisture content of cover crop was determined according to the Association of Official Analytical Chemists (AOAC, 2012). The aluminum dish was dried in an oven at a temperature of 105°C for 1h and transferred to a desiccator for 15 min. Then, 1.0g of sample was transferred to an aluminum dish and placed in an oven (Model 100-800, Garment GmbH) at 105°C until a constant weight was achieved. Moisture percentage was calculated using the following equation:

$$\% \text{ Moisture content} = \frac{M_1-(M_2+M_0)}{M_1} \times 100 \quad \text{(Eq. 1)}$$

Where:

M_0 = mass in grams of the empty aluminum dish;
M_1 = mass in grams of sample before drying;
M_2 = mass in grams of aluminum with sample after drying.

2.2 Ash content

Ash content of cover crop was determined according to the Association of Official Analytical Chemists [4]. Prior to analysis, the crucible was dried in an oven at a temperature of 105°C for 1h and transferred to a desiccator for 15 min. 1.0 g of sample was placed in a crucible and incinerated in a muffle furnace at 550°C for 4 h and transferred to a desiccator. It was then allowed to cool down in the desiccators for 15 min before weighing. (%) Ash content was calculated using the following equation:

$$\% \text{ Ash content} = \frac{M_2-M_0}{M1} \times 100 \quad \text{(Eq. 2)}$$

Where:

M_0 = weight in grams of sample;
M_1 = weight in grams of empty crucible;
M_2 = weight in grams of crucible with ash.

2.4 Total nitrogen analysis using Kjeldahl

Total nitrogen in cover crop samples was determined by Kjeldahl method, following the AOAC 2001.11 standard. One gram of sample was transferred into 250 ml kjeldahll flash, then it added 9g of potassium sulfate (K_2SO_4), 1g of copper sulfate ($CuSO_4$) as catalyst, and 20 ml of sulfuric acid (H_2SO_4 98%) and heated for 2h in digestion machine (Model 4000508). After that, allowed to cool down at room temperature. And then, connected to the kjeldahl of distillation unit (Model Pro-NitroM). After 10-15min, long neck flask was added 20 ml of 4% of boric acid solution in to Erlenmeyer flask. Lastly, collected distillated in Elenmeyer flask and add 3 to 4 drops of combined red and blue indicators. Then, using 0.25 M of sulfuric acid, it titrated to endpoint. The determine of nitrogen content following equation below:

$$\%N = \frac{(V_{sample} - V_{blank}) \times Z \times C \times F \times M_N}{m_s \times 10} \quad \text{(Eq.3)}$$

Where:
V_{sample} = volume of titrant for the sample (mL);
V_{blank} = volume of titrant for the blank (mL);
Z equal to 2 for the molar reaction factor for H_2SO_4;
C is the concentration of titrant (mol/L);
F is the factor of titrant (0.9969);
M_N is the mass molecular weight of the nitrogen (14.0 g/mol);

2.5 Ion Chromatography (IC) Analysis

Ion chromatography (IC) is a powerful analytical technique used for the separation and quantification of ionic species in aqueous samples. The method, first introduced according to Small et al., 1975 [5], IC was developed as an efficient alternative to labor-intensive wet chemistry techniques for inorganic ion analysis. It uses ion-exchange resins as the stationary phase and conductivity detection for quantification. Oxalic acid was used as the cation mobile phase for the IC instrument. This method enabled accurate

detection of ionic forms of nitrogen (e.g., NO_3^-, NH_4^+) and potassium (K^+).

2.5.1 Sample Digestion

One gram of the sample was transferred into 100 ml of Erlenmeyer flask with 50 mL of deionized water (DI H_2O) and shaken with an ultrasonic shaker for 30 minutes to facilitate ion release from the plant material. Then it was centrifuged at 3000 rpm for 10 minutes to separate solids from the liquid extract [6]. The supernatant was carefully filtered through a 0.45 µm membrane filter to remove any residual particulates. After filtration, 5 mL of the sample extract was transferred into clean vials. These vials were then sealed and stored in a refrigerator at 4 °C until further analysis ion chromatography (IC).

2.5.2 Potassium and Nitrogen using IC Analysis

The Determination of nitrogen (N) and potassium (K) in the cover crop extracts was carried out using Ion Chromatography (IC) with a Shimadzu Prominence High-Performance Liquid Chromatograph (HPLC) system. The IC system was equipped with a conductivity detector and two types of ion exchange columns: one for cation separation and the other for anion separation. Potassium (K^+) and ammonium (NH_4^+) analysis was performed using a Shim-pack IC-C3 column, which is specifically designed for separating monovalent and divalent cations. The detection was based on electrical conductivity, and calibration was done using standard solutions. For nitrogen analysis, a Shim-pack IC-A3 column was used to separate nitrate (NO_3^-), whichare common form of nitrogen found in plant extracts. This column is highly selective and suitable for non-suppressed IC, making it ideal for environmental and agricultural samples. Oxalic acid (2.5 mM) was used for the cation mobile phase to support potassium analysis. All eluent solutions were filtered through a 0.45 µm membrane and degassed before use. The instrument operated under isocratic conditions with a flow rate of 1.2 mL/min and a column temperature of 40°C. Standard solutions were prepared with deionized water and used to generate calibration curves for the quantification of cations. The data obtained from the IC analysis were used to determine the concentrations of key nutrients in the cover crop samples, helping to assess the nutrient dynamics important for soil and crop management studies.

3. RESULTS AND DISCUSSION

This study evaluated the nutrient content, specifically moisture, ash, nitrogen content (N), potassium content (K), and total nitrogen (kjeldalh) of six cover crops cultivated in Rovieng district, Preah Vihear province. The results reflect both species-specific nutrient accumulation patterns and the potential for optimizing soil fertility through informed cover crop selection.

3.1 Moisture Content

Moisture content in cover crop samples ranged from 4.93% to 7.56%, which is well below the recommendation of 12% for safe seed storage [7]. These low values confirm that the tested crops, including Cowpea, Mix, and Stylo, are suitable for post-harvest storage in dry form. Notably, Sesbania 2 (7.56%) and Stylo 9 (7.29%) had the highest moisture contents, while Stylo 2 (4.93%) and Cowpea 3 (5.08%) had the lowest. These differences may reflect species-specific water retention capacities or local environmental conditions during growth and harvest. Moisture content is critical depending on use: seeds require <12% moisture, while haylage or silage crops are typically harvested at 50–70% [3,6]. Since all tested samples were below 8%, they are more suitable for seed storage rather than fermentation-based uses. Proper post-harvest drying and monitoring are therefore essential for maintaining seed quality and minimizing spoilage.

3.2 Ash Content

Ash content in a cover crop sample refers to the total mineral content or inorganic residue remaining after the organic matter in the sample is burned off completely at high temperatures, typically in a muffle furnace at 500–600°C. It is a key indicator of the nutritional value of the cover crop, particularly for soil health, as it reflects the presence of essential minerals like calcium, potassium, phosphorus, magnesium, and trace elements. It is an important parameter for assessing the nutritional quality of a cover crop, especially for soil health, as it indicates the presence of essential minerals like calcium, potassium, phosphorus, magnesium, and trace elements. According to the result shown in **Table 1**, the ash content, indicative of total mineral accumulation, ranged from 6.10% to 12.07%, consistent with the 5–15% typical range for cover crops [8]. The highest values were recorded in Cowpea 2 (12.07%) and Cowpea 1 (11.42%), confirming Cowpea's high mineral-accumulating capacity. In contrast, Mix 8 (6.10%) and Sesbania 3 (6.53%) showed the lowest ash values, suggesting less mineral uptake or different growth rates. Higher ash content is desirable for improving soil fertility through organic residue incorporation. Crops such as Cowpea could be particularly valuable in systems where mineral restoration is a goal, either as a green manure or compost input [9].

3.4 Potassium content

Potassium (K) is the most abundant cation in plant tissue and is required by cotton in an amount similar to nitrogen (N). It plays a role in plant water relations through the regulation of stomata functioning, the rate of CO2 fixation, enzymatic activation, nutrient transport and homeostasis, stress tolerance, and eventually plant growth, fiber yield, and quality. Based on the results shown in **Table 1**, potassium

content across samples ranged from 1.52% to 3.46%, with the highest recorded in Kenaf 1 (3.46%). This result aligns with prior research suggesting that non-leguminous crops such as Kenaf are efficient scavengers of potassium due to their deep-rooting system. Sesbania and Cowpea also demonstrated high K values, making them effective contributors to potassium cycling. In potassium deficient soils, integrating crops like Kenaf, Cowpea, or Sesbania could reduce dependency on synthetic fertilizers while maintaining soil productivity. The broader range in Mix crops also suggests a composite advantage, as nutrient traits from multiple species can complement each other within the same system [10].

Table 1. Comparison range of moisture, ash, and potassium content (IC) in Cover Crops

Cover Crop	Moisture (%)	Ash Content (%)	Potassium (K%)
Stylo	4.93–7.29	6.63–10.71	1.54–2.06
Mix	5.45–7.23	6.10–8.78	1.52–2.89
Sunhemp	5.19–7.10	8.38–10.01	1.95–2.78
Sesbania	5.58–7.56	6.53–7.52	2.52–2.85
Cowpea	5.08–6.12	11.42–12.07	2.25–2.66
Kenaf	5.17–5.74	7.40–7.97	2.13–3.46

3.5 Nitrogen content (IC) and Total nitrogen (Kjeldahl)

Nitrogen is a primary plant nutrient that is an important factor in determining a crop's health and productivity. Cover crops can significantly vary in their total nitrogen content, ranging from less than 1% in some cereals after head emergence to 4-5% in young, leafy legumes. Legumes, especially, tend to have higher nitrogen content, making them excellent sources of nitrogen for subsequent crops. The nitrogen content in a cover crop can influence how quickly it decomposes and releases nitrogen to the soil.

As the result shown in **Table 2**, a representative separation of standard inorganic nitrogen species (such as NH_4^+ and NO_3^-) using ion chromatography (IC) was achieved with effective resolution, confirming the method's suitability for analyzing plant tissue extracts. The ion chromatography analysis was conducted using a Shim-pack IC-A3 analytical column**,** employing a mobile phase composed of oxalic acid (2.5 mM)**.** This configuration allowed accurate quantification of inorganic nitrogen ions present in aqueous extracts of cover crop tissues. All operations were performed under strict conditions to ensure precision and minimize contamination. In Table 2, the quantitative values of the identified cations in the sample analyzed are reported. Results show that considerable variability in nitrogen content was observed, reflecting both species-specific nutrient uptake and biochemical composition. Table 2 further summarizes the range of total nitrogen (Kjeldahl) and soluble nitrogen content (IC analysis) across six cover crop types. The Kjeldahl method, a standard procedure for measuring total nitrogen in biological samples, includes both organic nitrogen and ammonium ions after sulfuric acid digestion. In contrast, the IC method selectively detects water-soluble nitrogen ions (e.g., NH_4^+), offering a more immediate insight into the bioavailable nitrogen fraction. According to the data, Cowpea exhibited the highest total nitrogen content (2.02–2.51%), followed by Sunhemp (1.71–2.16%) and Stylo (1.50–2.04%). These results are consistent with their roles as leguminous cover crops capable of symbiotic nitrogen fixation. Kenaf, a non-leguminous species, displayed the lowest total nitrogen values (1.33–1.71%). In terms of water-soluble nitrogen determined by IC, Nitrogen content varied widely, from 0.0243 % to 0.9006%. Since legume tissues tend to have high N contents (e.g., 4–7% N in nodules, 2.5–5% N in leaves; [10]) Mix samples showed the widest variability (0.0243–0.9006%), possibly due to species diversity and differential decomposition rates. Sunhemp showed consistently high IC nitrogen values (0.748–0.8959%), implying efficient mineralization and release of nitrogen into the soluble pool. Stylo, Cowpea, and Sesbania showed moderate IC nitrogen values, whereas Kenaf again had the lowest soluble nitrogen content (0.1795–0.351%). These results underscore that total nitrogen content does not always correspond with the amount of readily available nitrogen in the cover crop. The combination of Kjeldahl and IC data highlights the distinction between total nutrient reserves and their soluble, plant-available forms. From an agronomic standpoint, this dual approach provides critical insight for evaluating the nutrient contribution of cover crops to soil fertility and sustainable cropping systems.

Table 2. Comparison range of Nitrogen Content (IC), and Total Nitrogen (Kjeldahl) in Cover Crops

Cover Crop	Number of sample	Total Nitrogen (%) (Kjeldahl)	Nitrogen Content (%) (IC)
Stylo	3	1.50 – 2.04	0.0593 – 0.7873
Mix	4	1.36 – 1.98	0.0243 – 0.9006
Sunhemp	5	1.71 – 2.16	0.748 – 0.8959
Sesbania	3	1.25 – 1.69	0.1436 – 0.6197
Cowpea	2	2.02 – 2.51	0.252 – 0.6007
Kenaf	5	1.33 – 1.71	0.1795 – 0.351

4. CONCLUSIONS

This study provides essential baseline data on the nutritional composition of six cover crops. Results that all tested species maintained moisture contents below the threshold for seed storage, making them suitable for dry preservation. Cowpea exhibited the highest ash content, indicating its strong mineral accumulation potential, while Kenaf and Sesbania showed superior potassium levels,

reinforcing their role in nutrient scavenging and soil potassium enrichment. Nitrogen content was assessed through both Kjeldahl and ion chromatography (IC) methods. Cowpea, Sunhemp, and Stylo recorded the highest total nitrogen (Kjeldahl), consistent with their nitrogen-fixing capabilities. IC analysis revealed that Mix and Sunhemp had the highest soluble nitrogen content, emphasizing differences in bioavailable nitrogen fractions among species. Notably, the mismatch between total and soluble nitrogen concentrations highlights the importance of evaluating both parameters to fully understand nutrient availability.

Overall, the findings suggest that strategic use of these cover crops, especially Cowpea for mineral accumulation, Kenaf for potassium uptake, and Sunhemp or Mix for bioavailable nitrogen, can significantly contribute to sustainable soil fertility management in Cambodian agricultural systems. These insights are valuable for guiding cover crop selection and integrated nutrient management practices in similar tropical agroecological zones.

ACKNOWLEDGMENTS

The project was funded by the European Union and the French Development Agency.

REFERENCES

[1] V. W. Ruttan, "Productivity growth in world agriculture: Sources and constraints," *J. Econ. Perspect.*, vol. 16, no. 4, pp. 161–184, 2002, doi: 10.1257/089533002320951028.

[2] W. Hsiung (Xiong), S. Yang, and Q. Tao, "Historical development of agroforestry in China," *Agrofor. Syst.*, vol. 30, no. 1–2, pp. 277–287, 1995, doi: 10.1007/BF00708926.

[3] Andy Clark, *Managing cover crops profitably*, Third. 2012. Accessed: Apr. 06, 2025. [Online]. Available: https://www.sare.org/wp-content/uploads/Managing-Cover-Crops-Profitably.pdf

[4] N. Thiex, L. Novotny, and A. Crawford, "Determination of ash in animal feed: AOAC Official Method 942.05 revisited," *J. AOAC Int.*, vol. 95, no. 5, pp. 1392–1397, 2012, doi: 10.5740/jaoacint.12-129.

[5] H. Small, T. S. Stevens, and W. C. Bauman, "Novel Ion Exchange Chromatographic Method Using Conductimetric Detection," *Anal. Chem.*, vol. 47, no. 11, pp. 1801–1809, 1975, doi: 10.1021/ac60361a017.

[6] T. R. I. Cataldi, G. Margiotta, A. Del Fiore, and S. A. Bufo, "Ionic content in plant extracts determined by ion chromatography with conductivity detection," *Phytochem. Anal.*, vol. 14, no. 3, pp. 176–183, 2003, doi: 10.1002/pca.700.

[7] F. Magdoff, *Building Soils for Better Crops*, vol. 156, no. 5. 2021. doi: 10.1097/00010694-199311000-00014.

[8] H. J. Poffenbarger *et al.*, "Biomass and nitrogen content of hairy vetch-cereal rye cover crop mixtures as influenced by species proportions," *Agron. J.*, vol. 107, no. 6, pp. 2069–2082, 2015, doi: 10.2134/agronj14.0462.

[9] H. Blanco-Canqui, M. M. Claassen, and D. R. Presley, "Summer cover crops fix nitrogen, increase crop yield, and improve soil-crop relationships," *Agron. J.*, vol. 104, no. 1, pp. 137–147, 2012, doi: 10.2134/agronj2011.0240.

[10] N. K. Fageria and A. Moreira, *The Role of Mineral Nutrition on Root Growth of Crop Plants*, 1st ed., vol. 110, no. C. Elsevier Inc., 2011. doi: 10.1016/B978-0-12-385531-2.00004-9.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Evaluation of rice seed quality of Cambodian varieties produced under conventional and agroecology practices

Singy HO[1], Malyna SUONG[1,2], Kakada OEUM[1], Vichny VICHETH[1], Rada KONG[3,4], Florent TIVET[3,4], Myriam ADAM[5,6], Lionel MOULIN[1,7*]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., Phnom Penh, Cambodia*
[2] *International Relations Office, Institute of Technology of Cambodia, Russian Federation Blvd., Phnom Penh, Cambodia*
[3] *Cambodian Conservation Agriculture Research for Development Centre, Department of Agricultural Land Resources Management, General Directorate of Agriculture, MAFF, Phnom Penh, Cambodia*
[4] *CIRAD, UPR AIDA, University of Montpellier, Montpellier, France,*
[5] *CIRAD, UMR AGAP Institut, F-34398 Montpellier, France,*
[6] *Faculty of Agriculture and Food Processing, National University of Battambang, Battambang, Cambodia*
[7] *Plant Health Institute of Montpellier (PHIM), IRD, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France*

[*] *Corresponding author: lionel.moulin@ird.fr*

Abstract: *Rice (*Oryza sativa*) is a major food crop in Southeast Asia. In Cambodia, farmers primarily use conventional practices that rely on soil tillage and high agrochemical inputs (fertilizers, pesticides), which lead to decreased in soil fertility and environmental pollution. For several years, agroecology practices have been introduced, for example, in Preah Vihear in organic-input production systems, with two production systems being tested, either conventional tillage with no chemical inputs (abbreviated as CT), or green manure by introducing cover crops between rice cycles to improve soil fertility and rice yield (abbreviated as GM). In this study, we aimed to evaluate and compare the quality of rice from Cambodian local varieties (Phka rumduol, Kraches, Prich, Neang Sar, Nean Ork, Neang Om) produced in GM and CT systems, to qualify GM systems on the rice quality criteria. We assessed seed physical quality (biomass and fragmentation following unhusking), seed nutritional level (amylose and protein content using Near Infrared Spectroscopy-NIRS), and seed microbiological quality (detection by Polymerase Chain Reaction of seed-borne phytopathogens and mycotoxin-producing fungi). We analysed the physical and nutritional quality of rice seeds produced in biplot systems (CT/GM) in 17 farmers' plots, and the presence of several seed phytopathogens and mycotoxin-producing fungi was assessed in five CT/GM biplots. Biomass analysis of 100 rice seeds showed significant differences between varieties, but not between GM and CT production systems. For seed breakage, there were no significant differences between varieties or treatments, but there were differences between farmers' plots. Significant differences in seed protein content were observed between CT and GM practices when compared altogether, but not by variety or practice. Significant differences in amylose content were detected between varieties but not between practices. We detected several phytopathogens in seeds by PCR technique, including* Xanthomonas oryzae, Burkholderia glumae*, as well as mycotoxin-producing* Alternaria species*, with no clear impact of the agricultural practice.*

Keywords: Rice crop; Agricultural practices; Seed quality; Near Infrared Spectroscopy; Phytopathogens

1. INTRODUCTION[1]

Rice (*Oryza sativa*) is an important food crop in the world, with nearly 40% of the world's population consuming rice as their main staple food. Asia accounts for about 86% of the world's rice area and contributes about 90% of global rice production [1]. In Cambodia, farmers produce the fragrant indica Phka Rumduol variety for export during the rainy season, while short-cycle indica varieties are grown for local or neighboring countries' markets at the end of the rainy season [2]. Farmers use mainly conventional practices (which rely on soil tillage and agrochemicals inputs leading to a decrease in soil fertility and environmental pollution.

Achieving sustainable rice production requires overcoming many challenges, including climate change, access to arable land, pressure from pests and diseases, and

[1]* Corresponding author: Lionel Moulin
E-mail: lionel.moulin@ird.fr; Tel: +33-69 803 4312

limited water resources. Many phytopathogens reduce rice yield [3], and several fungi reduce rice quality by producing mycotoxins. For several years, agroecology practices have been introduced, for example, in Preah Vihear in organic production systems, where two systems are tested: conventional tillage with no inputs (CT), or green manure by growing cover crops between rice cycles that are rolled into soil by tillage to improve soil fertility and rice yield (GM) [4].

Various options are available for improving the quality, safety and nutritional value of rice: biofortification, improving the yield of high-quality traditional varieties through breeding, improvements in milling technologies, the promotion of healthy rice products, as well as agronomic practices as organic or conservation agriculture that improve soil fertility and rice yields and seed quality [5].

In order to promote sustainable rice production, it is important to assess the quality of rice seeds produced in organic systems and evaluate their suitability for the market. Rice quality relies on physical [6], nutritional (amylose, protein) [7], and microbiological (presence of seedborne phytopathogens and mycotoxin-producing fungi) [8] properties.

In this study, we compared the quality of rice seeds from six local varieties produced using organic methods either with conventional tillage or green manure in Preah Vihear province in Cambodia. We analyzed the physical properties (size, weight, breakage, and moisture content), nutritional content (amylose and protein content) and the presence of seed-borne phytopathogens and mycotoxin-producing fungi.

2. METHODOLOGY

2.1 Rice fields and seed sample collection

The seed samples were harvested between October and November 2023 (depending on rice cycle length) from rice fields located at GPS coordinates 13°25'00.1"N 105°08'42.3"E, close to Boh village, Rik Reay commune, Rovieng district, Preah Vihear province, Cambodia. Seventeen rice fields were cultivated by 17 farmers under supervision by GDA-DALRM and CIRAD agronomists, as part of the ASSET project (AFD-EU). Each rice field included six local rice varieties (Table 1), which were grown in biplot systems with two organic agricultural practices: conventional tillage (CT) with the use of tillage and no chemical inputs, and green manure (GM) with cover crops grown between rice cycles that are incorporated into soil by tillage, and no chemical inputs. Around 100 g of seed samples per variety/rice field/practice were kept at room temperature until further analysis.

2.2 Seed sample preparation

One hundred seeds were counted and weighed using a precision balance to assess the average seed mass. To dehusk seeds, we used an Automatic Rice Husker TR-260 (Kett Electric Laboratory Co. Ltd), followed by additional manual separation for seeds that did not dehusk well. After dehusking, seeds were stored in ziplock bags and stored at room temperature until the next step.

Table 1. Rice varieties and their characteristics

Rice Variety	Cycle length	Characteristics	Market
Kraches	medium	Long slender grain, soft texture	Local
Neang Om	medium	Slightly sweet taste, dry texture	Local
Neang Ork	medium	Long slender grain, medium texture	Local
Neang Sar	short	Bold grain, hard texture	Local
Prich	medium	Medium grain, low shattering	Local
Phkar Rumdoul	long	Extra-long grain, soft texture, distinct aromatic flavor	Highest value; best-selling, export

Note: short: 100−120 days; medium: 120−140 days; long: 160 days plus.

2.3 Analysis of seed breakage

To evaluate seed brokenness, 5 g of dehusked seeds were separated manually on a green A4 plastic sheet, and pictures were captured and then analyzed for their number and size of full and broken seeds using ImageJ (version Java 1.8.0_345). After adjusting the scale, seeds were detected by adjusting the color threshold, and the particle analysis function was used to calculate the size and number of full and broken seeds.

2.4 Analysis of moisture content and nutritional value of seeds using Near InfraRed Spectroscopy

The moisture, protein, and amylose content from the dehusked rice seeds was evaluated using a Near Infrared Rice Composition Analyser AN-920 (Kett Electric Laboratory Co. Ltd.) calibrated for long rice composition, and following the manufacturer's instructions. Means of three replicates were used for each sample.

2.5 DNA extraction of rice seed

In order to recover DNA from seeds for detection of phytopathogens by PCR, we used a published protocol [9]. It consists of an incubation of the seeds in a lysis buffer composed of 2M KOH and 60% PEG200, then 1 µl of the lysate is used as template for PCR. For each sample, 10 seeds were used for DNA extraction.

2.6 Molecular detection of phytopathogens in rice seed

Polymerase Chain Reaction (PCR) was used to detect the presence of DNA of rice pathogens and toxin-producing fungi

in 6 varieties grown in 5 biplot systems. Each PCR reaction contained 14.875 µl of sterile pure water, 5 µl of PCR buffer 5X (Go-Taq, Promega), 2 µl of dNTP mix (10 mM each), 1 µl of each primer (forward and reverse, at 10 uM), and 0.125 µl of GoTaq DNA Polymerase (5U/µl). 1 µl of DNA was added per reaction. A negative control with no DNA was always added to evaluate contamination by environmental DNA. A positive control was also added that contains primers and a known DNA targeted by the primers, to evaluate PCR mix performance. All reactions were amplified in a PCR thermocycler (Labnet International, Inc.) with its specific program depending on primers for the target pathogen (Table 2).

2.7 Gel electrophoresis of PCR products

PCR products were visualized by gel electrophoresis of 1.5% agarose gel in Tris-Acetate-EDTA (TAE) 1X buffer (Sigma-Aldrich). 10 µl was loaded in each well and the electrophoresis was run at 80 volts for 50 mins. A ladder (5 µl of SmartLadder, Eurogentec) was included to assess the PCR band sizes. After migration, the gel was incubated in 50 ml of water with 5 ul of 10,000X Gelgreen fluorescent dye for 10 min, then rinsed in water for 15 min and viewed and pictured in a Bluebox Transilluminator (minipcr©).

Table 2. PCR primers and programs used to detect phytopathogens and mycotoxin-producing fungi

Primer name	Sequence (5'-3')	Target	Size (bp)	PCR program*	References
Xo3756F Xo3756R	CATCGTTAGGACTGCCAGAAG GTGAGAACCACCGCCATCT	*Xanthomonas oryzae* (hyp. protein)	331	35 cycles (94°C for 30 s, 55°C for 30 s, 72°C for 45 s).	[10]
toxB_F toxB_Rd	GCATTTGAAACCGAGATGGT TCGCATGCAGATAACCRAAG	*Burkholderia glumae* (*toxB*)	508	35 cycles (94°C for 30 s, 55°C for 30 s, 72°C for 45 s).	[11]
Dir1ITSAlt Inv1ITSAlt	TGTCTTTTGCGTACTTCTTGTTTCCT CGACTTGTGCTGCGCTC	*Alternaria* species (ITS)	370	35 cycles (94°C for 30 s, 55°C for 30 s, 72°C for 45 s).	[12]

Note: All PCR programs had an initial step at 94°C for 2 min, and a final extension step for 10 min at 72°C.

2.8 Statistical analysis

R version 4.4.0 was used for the statistical analyses, using the packages rstatix, ggplot2, and multcompView. As all of our data did not follow a normal distribution (p<5% with the Shapiro-Wilk test), non-parametric Kruskal-Wallis tests (α=5%) followed by multiple comparison post-hoc Dunn test (p.adjust.method = "bonferroni") were used to assess the significance of differences between means. A letter code was used to represent the different statistical groups in the boxplot figures. Box plots with identical letters are not statistically different, while box plots with no letters in common are statistically different at α=5%.

3. RESULTS AND DISCUSSION

In this study, we evaluated the rice seed quality of six Cambodian varieties grown in 17 GM and CT organic biplot systems in Preah Vihear to evaluate the suitability of GM as an alternative to CT on the rice quality criteria. We evaluated several criteria: the physical, nutritional, and microbiological properties of the seeds produced by the 17 farmers.

3.1 Assessment of physical quality of seeds in 6 rice varieties grown under CT and GM

The physical quality of the rice seeds was determined by their biomass (on 100 seeds) and brokenness. In Figure 1, we show the results of biomass of 100 seeds as boxplots organized by varieties and agricultural practice (A), by variety only (B), and by practice (C). The weight of 100 seeds is a proxy for the average biomass of seeds.

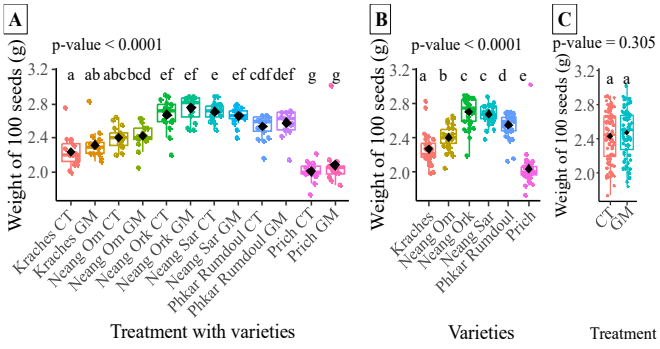


Fig. 1. Comparison of the biomass of 100 seeds of 6 rice varieties produced in 17 CT and GM biplots. With A: by variety and practice, B: by variety only, C: by practice only (CT/GM). The p-value above each plot is the result of a Kruskal-Wallis test. The letters above each box plot represent the statistical groups from pairwise Wilcoxon tests.

The average biomass of 100 seeds was 2.44 ± 0.264 g. Overall, there were significant differences in the biomass of 100 seeds between varieties (**Fig. 1.A** and **1.B**) but not between GM/CT treatments by variety (**Fig. 1.A**) or by treatment only (**Fig. 1.C**). Almost every variety was different on the 100 seed biomass criteria, with Nean Ork and Neang Sar producing the heaviest seeds. Neang Ork produced the heaviest weight (2.71 ± 0.132 g) and Prich produced the lightest weight (2.04 ± 0.199 g). These differences can be explained by their different genetic background [13] and the intrinsic

characteristics of each variety [14]. On the other hand, there was no significant difference between CT and GM treatment (p = 0.305, **Fig. 1.C**), which means that GM or CT treatments did not affect the biomass of 100 rice grains for each variety.

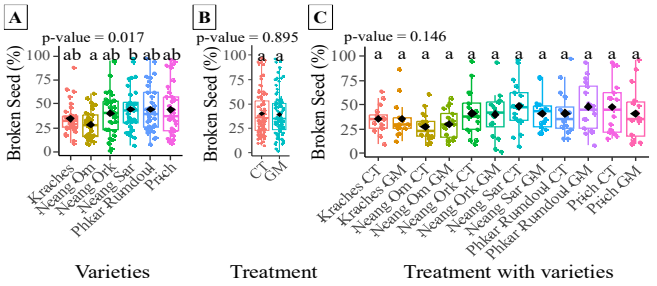


Fig. 2. Comparison of seed brokenness of rice varieties produced in CT and GM biplots, between varieties (A), agricultural practice (B), and varieties + practice (C). Same legend for statistics as in **Fig. 1**.

The second physical parameter that we investigated is the breakage of seeds following dehusking. We plotted the % of broken seeds as boxplots in **Fig. 2**. As shown in **Fig. 2.A**, several varieties differ in their % of broken seeds (p = 0.017) but there was no significant difference between the CT and GM groups (**Fig. 2.B and 2.C**). The rice variety Neang Om was the least broken out of the six varieties, with a breakage rate between 14% and 28%. The brokenness of rice was classified as lower or 5% (premium rice), 10% (grade 1), 25% (grade 2), and more than 35% (grade 3) [15]. The broken rice rate is a critical metric that exerts a considerable influence on the appearance, processing, and economic value of rice [16]. A multitude of factors, including hormones, enzymes, nutrients, and environmental influences, have been demonstrated to play a role in the incidence of seed brokenness [17]. Furthermore, additional factors that have the potential to compromise the integrity of rice include weakness of rice [14], the timing of planting, water management practices, soil nitrogen levels, the cultivar of rice, the time of harvest, the handling of paddy, and the moisture content of the grain [18]. In our results, the CT or GM practice did not affect the broken rice rate.

3.2 Assessment of seed nutritional quality in rice varieties grown under CT and GM

The nutritional quality of rice seeds was assessed by evaluating the protein and amylose content with a NIRS precalibrated for rice.

The protein and amylose contents are presented as boxplots in **Fig. 3**. The average protein content of rice seeds was 8.2 ± 0.8 %. As shown in **Fig. 3A**, the boxplots of variety in CT or GM share a common letter code corresponding to a statistical group, meaning they are not statistically different. When all varieties were mixed to analyze the global effect of the practice (**Fig. 3C**), the seed protein content was significantly higher (p=0.007) in CT (8.36%) compared to GM (8.09%), with a 3.3% increase. The decomposition and nutrient release of green manure are primarily influenced by the soil's physical, chemical and biological characteristics. In comparison to CT, green manures with minimal tillage may not fully integrate with the soil, hindering their breakdown [19]. In addition, significant differences (p<0.01) in protein content were also observed between varieties (**Fig. 3.B**). The rice variety that contained the highest protein content was Neang Om with 8.72 ± 0.6 % and the lowest was Prich with 7.47 ± 0.8 %. Usually, the protein content of rice range from 7-10% [20], which is in the range of our measurements. The differences observed between varieties is a known feature [7]. The difference in protein content between rice varieties may be due to genetic composition, agronomic management techniques or environmental variables [21]. Rice's protein levels have a significant impact on its nutritional value and processing capabilities [6].

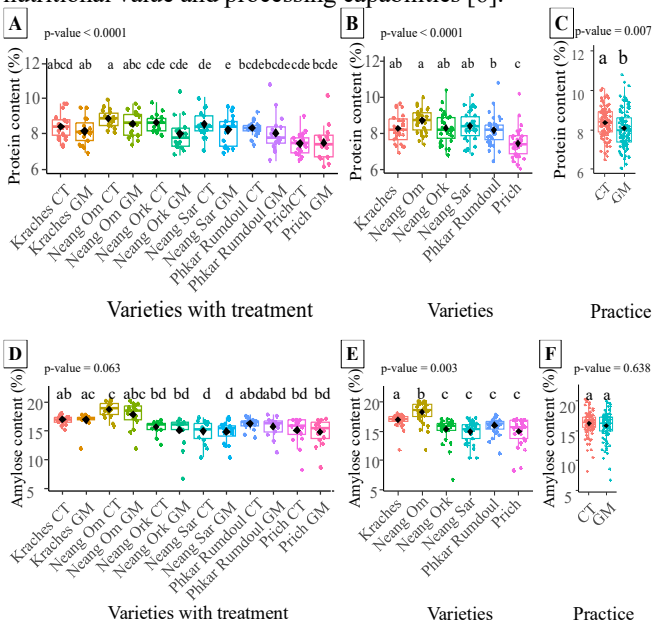


Fig. 3. Protein and amylose content in rice varieties produced in CT and GM. A,B,C: protein content (%) in samples by variety and agricultural practice (A); by variety (B) and practice (C), D, E and F: amylose content variety and agricultural practice (D) and by variety (E) and practice (F). Same legend for statistics as in **Fig. 1**.

The amylose content of the rice samples is shown as boxplots in **Fig. 3.D** (by variety and practice), **3.E** (variety) and **3.F** (practice). The average amylose content of rice seeds was 16.04 ± 2.1 %. The boxplots in **Fig. 3.D** showed no significant difference in amylose content between seed samples from the same variety mutant produced in CT or GM practice (each variety in CT or GM shares a common letter in their statistical group). A significant difference (p = 0.003) was observed in amylose content between varieties. Neang Om had the highest amylose content (18.1 ± 1.8 %) while Prich had the lowest (15.1 ± 2.3 %). The amylose content of rice varieties can be classified as waxy (0–2%), very low (5–12%), low (12–20%),

intermediate (20–25%), or high (25– 33%) [7]. Here, all our varieties can be considered as low. A low amylose content of rice offers a good cooking quality with its soft and sticky texture, slow digestibility, and reduced sugar release [22].

In this study, we have evaluated the impact of variety and agricultural practice on several rice quality properties. The statistical results are summarized in **Table 3**.

Table 3. Sum-up of statistical analyses of the impact of variety and practice on rice quality parameters.

Rice seed	Biomass	Breakage	Protein	Amylose
Variety	***	*	***	***
Practice GM/CT)	ns	ns	**	ns

Note: ns (not significant): p>0.05, *: p <0.05, ** : p<0.01, *** : p<0.001

3.3. Detection of seed-borne microorganisms

To evaluate the presence of seed-borne phytopathogens and mycotoxin-producing fungi, we focused on two frequent seed-borne bacterial rice pathogens, *Xanthomonas oryzae* and *Burkholderia glumae*, and on *Alternaria* species, which produce various toxic compounds, including mycotoxins in rice [23]. We used PCR with specific primers of each microbial target (**Table 2**), and we tested 10 seeds for the 6 varieties in 5 biplots systems (GM and CT), and showed the results as a percentage of positive PCR signals in **Fig. 4**.

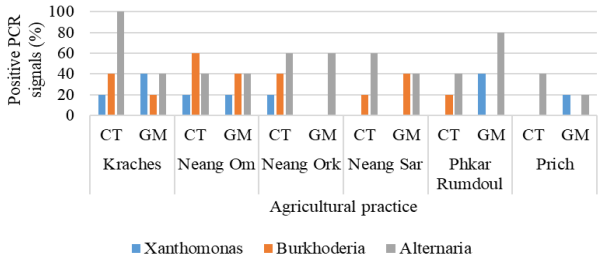


Fig. 4. Bar plot of positive PCR signal for the detection of the seed-borne phytopathogens *Xanthomonas oryzae*, *Burkholderia glumae,* and mycotoxin producer *Alternaria* species.

Xanthomonas oryzae was detected in 5 varieties (no detection in Neang Sar) and in low frequency (0-40% of samples). It was detected on Phka rumduol and Prich in GM biplots but not in their corresponding CT system, indicating GM might increase the *X. oryzae* presence on these varieties. *Burkholderia glumae* was also detected in 5 varieties (not in Prich) in 0-50% of the samples, and on Phka rumduol in CT but not GM, indicating this time CT increased *B. glumae* compared to GM. *Alternaria sp.* was detected in all varieties with some variation (30% in Prich and 70% in Kraches) and practices (100% in Kraches-CT, to 20% in Prich-GM, with no clear impact of variety or practice on its frequency.

4. CONCLUSIONS

The type of variety impacted all parameters tested, from biomass to amylose content. Concerning the agricultural practice (CT-conventional versus GM-green manure), it did not impact the biomass, breakage, or amylose content of the rice seeds, but it impacted the protein content by reducing it by 3%, which is a low modification.

We also tested by PCR the detection of two common seed-borne phytopathogens and could detect them in almost all varieties with different frequencies, but no clear impact of the practice. Alternaria species were, and further work is required to quantify if the seeds contain mycotoxins at a level that could impact their access to market.

Thus, we could conclude that green manure practices performed in the investigated biplots in Preah Vihear did not impact the quality of rice compared to conventional practices. The values of biomass, breakage and nutrient contents measured are also in the range of acceptability for the market. Green manure from cover crops in rice organic production can thus be recommended as an alternative to conventional practice due to its sustainability and eco-friendly properties [4], including the absence of chemical inputs, for better soil, environmental and human health.

ACKNOWLEDGMENTS

The authors thank the French Research Institute for Sustainable Development (IRD) through the project "Laboratory of Excellence in co-Engineering for Sustainable Agrosystems-LMI LEAD" for a master grant fellowship to SH, the Institute of Technology of Cambodia, hosting the Plant health laboratory, the Ministry of Europe and Foreign Affairs via the project "FEF-Healthyrice", and CIRAD and GDA-DALRM managing the biplots systems as part of the ASSET project (AFD-EU).

REFERENCES

[1] FAO, “A regional rice strategy for sustainable food security in Asia and the Pacific.” Food and Agriculture Organization of the United Nations, 2014. [Online]. Available: https://www.fao.org/fileadmin/templates/rap/files/aprc32/regional-rice-strategy-for-asia-pacific_final-edition.pdf

[2] Cambodia Rice Federation, “5th World Best Rice Award,” Cambodia Rice Federation. [Online]. Available: http://www.crf.org.kh

[3] S. Savary, L. Willocquet, S. J. Pethybridge, P. Esker, N. McRoberts, and A. Nelson, “The global burden of pathogens and pests on major food crops,” *Nat Ecol Evol*, vol. 3, no. 3, pp. 430–439, Mar. 2019, doi: 10.1038/s41559-018-0793-y.

[4] FAO, MAFF, and CASIC, *Bottom-up solutions to promote conservation agriculture in Cambodia*. FAO ;, 2022. Accessed: Apr. 22, 2025. [Online]. Available:

https://openknowledge.fao.org/handle/20.500.14283/cc2698en

[5] P. Nickens, D. R. Ader, M. Miller, P. Srean, T. Gill, and S. Huot, "Conservation agriculture and cover crop adoption by smallholder farmers in Cambodia: Understanding perceptions, challenges, and opportunities for soil improvement," *Advancements in Agricultural Development*, vol. 4, no. 4, pp. 39–48, 2023, doi: 10.37433/aad.v4i4.387.

[6] R. C. Bautista and P. A. Counce, "An Overview of Rice and Rice Quality," *Cereal Foods World*, vol. 65, no. 5, 2020, doi: 10.1094/cfw-65-5-0052.

[7] A. A. Abera, E. E. Tadesse, B. B. Abera, and N. Satheesh, "Effect of rice variety and location on nutritional composition, physicochemical, cooking and functional properties of newly released upland rice varieties in Ethiopia," *Cogent Food and Agriculture*, vol. 7, no. 1, pp. 1–23, 2021, doi: 10.1080/23311932.2021.1945281.

[8] S. M. Ebrahim Zainali, "Evaluation Seed-born Fungi of Rice [Oryza sativa L.] and that Effect on Seed Quality," *Journal of Plant Pathology & Microbiology*, vol. 05, no. 04, 2014, doi: 10.4172/2157-7471.1000239.

[9] F. Zhang *et al.*, "Rapid detection of Fusarium fujikuroi in rice seeds and soaking water samples based on recombinase polymerase amplification-lateral flow dipstick," vol. 73, no. 3, 2023.

[10] J. M. Lang *et al.*, "Genomics-Based Diagnostic Marker Development for Xanthomonas oryzae pv. oryzae and X. oryzae pv. oryzicola," *Plant Disease*, vol. 94, no. 3, pp. 311–319, Mar. 2010, doi: 10.1094/PDIS-94-3-0311.

[11] M. Bangratz *et al.*, "Design of a new multiplex PCR assay for rice pathogenic bacteria detection and its application to infer disease incidence and detect co-infection in rice fields in Burkina Faso," *PLOS ONE*, vol. 15, no. 4, p. e0232115, Apr. 2020, doi: 10.1371/journal.pone.0232115.

[12] M. Á. Pavón, I. González, M. Rojas, N. Pegels, R. Martín, and T. García, "PCR Detection of Alternaria spp. in Processed Foods, Based on the Internal Transcribed Spacer Genetic Marker," *Journal of Food Protection*, vol. 74, no. 2, pp. 240–247, Feb. 2011, doi: 10.4315/0362-028X.JFP-10-110.

[13] P. E. R. Prahardini, E. Fidiyawati, S. S. Antarlina, and T. Sudaryono, "Characteristic of grain and rice quality from five superior varieties support food sustainability," *IOP Conference Series: Earth and Environmental Science*, vol. 1253, no. 1, pp. 1–8, 2023, doi: 10.1088/1755-1315/1253/1/012016.

[14] R. T. Adhiguna, S. Sutrisno, and R. Thahir, "Comparison of physical properties, proximate composition and milling quality of rice grains from different branches within a panicle," *Int'l Journal of Scientific & ...*, vol. 7, no. 9, pp. 1647–1652, 2016.

[15] FAO, "NFA standard specification for milled rice (Second revision) TRED · SQAD No. 2; 1980 - Google Search," Food and Agriculture Organization. Accessed: May 30, 2025. [Online]. Available: https://www.fao.org/4/x5048e/x5048e03.htm

[16] J. Ye, Z. Hu, Y. Chen, D. Fu, and J. Zhang, "Identification of broken rice rate based on grading and morphological classification," *Lwt*, vol. 215, no. November 2024, p. 117175, 2025, doi: 10.1016/j.lwt.2024.117175.

[17] H. Wu, Q. He, and Q. Wang, "Advances in Rice Seed Shattering," *International Journal of Molecular Sciences*, vol. 24, no. 10, pp. 1–17, 2023, doi: 10.3390/ijms24108889.

[18] O. Da, "Factors affecting cracking and breakage of rice (Oryza sativa L.) during milling: A Review," *Journal of Food Technology and Preservation*, vol. 3, no. 1, pp. 1–4, 2021, doi: 10.36266/JFTP/117.

[19] X. Gao, A. Lv, P. Zhou, Y. Qian, and Y. An, "Effect of green manures on rice growth and plant nutrients under conventional and no-till systems," *Agronomy Journal*, vol. 107, no. 6, pp. 2335–2346, 2015, doi: 10.2134/agronj15.0225.

[20] Y. Yang *et al.*, "Natural variation of OsGluA2 is involved in grain protein content regulation in rice," *Nature Communications*, vol. 10, no. 1, 2019, doi: 10.1038/s41467-019-09919-y.

[21] F. M, D. John, and M. Raman, "Physicochemical properties, eating and cooking quality and genetic variability: a comparative analysis in selected rice varieties of South India," *Food Prod Process and Nutr*, vol. 5, no. 1, p. 49, Nov. 2023, doi: 10.1186/s43014-023-00164-x.

[22] P. Srinang, S. Khotasena, J. Sanitchon, S. Chankaew, S. Jogloy, and T. Monkham, "New Source of Rice with a Low Amylose Content and Slow In Vitro Digestion for Improved Health Benefits," *Agronomy*, vol. 13, no. 10, 2023, doi: 10.3390/agronomy13102622.

[23] C. Schwarz, C. Tiessen, M. Kreutzer, T. Stark, T. Hofmann, and D. Marko, "Characterization of a genotoxic impact compound in Alternaria alternata infested rice as Altertoxin II," *Arch Toxicol*, vol. 86, no. 12, pp. 1911–1925, Dec. 2012, doi: 10.1007/s00204-012-0958-4.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Influence of the Soil Physicochemical Properties on Phka Rumdoul Rice Quality and Yield in Cambodia

Vichny Vicheth [1], Stephane Bellafiore [1,3], Sereyvath Yoeun [1,2], Sreymom Sieng [4], Socheat Keo [5]
Sambo Pheap [6], Chanvorleak Phat [1,2*]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Plant Health Institute of Montpellier (PHIM), IRD, CIRAD, INRAE, Institut Agro, Univ Montpellier, Montpellier, France*
[4] *Department of Agricultural Land Resources Management, 395-656 Street, Toeuk Laak3, Tuol Kork, Phnom Penh, Cambodia*
[5] *Executive Director of Sansom Mlup Prey Organization, 456 Street 44c, Sangkat Tuol Tumpong Ti Pir Khan Chamkarmon, Phnom Penh, Cambodia*
[6] *Faculty of Agricultural Science, Royal University of Agriculture, P.O. Box 2696, Kan Dangkor, Phnom Penh, Cambodia*

[*] *Corresponding author: phatchanvorleak@itc.edu.kh*

Abstract: *For at least 2,000 years, Cambodian farmers have cultivated rice over large areas. Phka Rumdoul is one of the most widely grown aromatic rice varieties in the country. It is known for its high resilience and great aroma test value, which makes it an outstanding world variety, winning up to six times the world's best rice award. This rice is therefore a "gold mine" for Cambodian agriculture. Resilient, it limits the use of chemical fertilizers and offers great export opportunities. Soil serves as the primary medium for nutrient supply, root development, and microbial interactions that directly affect rice growth, grain composition, and yield. This present study aims to see the potential effect of the physico-chemical quality of soil and its influence on the nutrients and physical properties of rice grains. The soil samples from nine provinces were analyzed for particle size distribution, pH, soil moisture, total carbon, organic matter, total nitrogen, total phosphorus, cation exchange capacity, exchangeable Ca, Mg, K and Na, and C/N ratio under different cultivation practices including conventional tillage (CT), conservation agriculture (CA), sustainable rice platform (SRP), organic rice farming (ORF), and green manure (GM). For rice grains, quality was assessed based on protein and amylose content using Near Infrared Reflectance Spectroscopy (NIRS). Physical quality was determined by measuring the weight of 100 grains, and yield (t/ha) was recorded after harvest. After that, the correlation between soil properties and rice quality and yield was investigated. Preliminary results indicate that soil properties under different agricultural practices inhibit significant variation in the physico-chemical characteristics of rice grain quality and yield. Understanding these relationships can provide valuable insights for optimizing soil management strategies to improve rice production and grain quality in Cambodia.*

Keywords: Phka Rumdoul rice variety; Soil characteristics; Rice grain quality; Rice agrosystems; Near Infrared Reflectance Spectroscopy

1. INTRODUCTION

Rice (*Oryza sativa* L) is one of the most important crops in the world and is the staple food for nearly half of the world's population [1]. It is cultivated in more than 100 countries across six continents and occupies over 11% of all arable land [2]. About 90% of the global rice-growing areas are located in Asia, especially East Asia (33%). Actually, rice is the staple food in Cambodia and is considered food security. It is the most important field crop today.

[1*] Corresponding author: Chanvorleak Phat
E-mail: phatchanvorleak@itc.edu.kh; Tel: +855- 93418000

A total of 2,226,000 households were estimated to have rice farms, or about 63% of all households in the country (3,553,021 households according to the 2019 Cambodia General Population Census) [3]. Among rice varieties, Phka Rumdoul is one of the most widely cultivated aromatic rice varieties in Cambodia, internationally recognized for its distinctive fragrance and superior grain quality. It is known for its high resilience and high aroma test scores, making it an outstanding variety that has won up to six times the world's best rice award. This rice is a gold mine for Cambodian agriculture. It limits the use of chemical fertilizers and offers great export opportunities. Focusing on Phka Rumdoul rice variety allows for a targeted exploration of the specific

factors, such as agro-system conditions, genotype livelihoods and enhancing Cambodia agricultural exports. Studying this variety help generate insights that are directly applicable to improving its production system, protecting its unique qualities, and ensuring sustainable intensification. However, Soil health is a major concern, considering that soil is essential to rice production [4]. Along with population growth comes an increasing demand for rice production on the soil. Soil serves as the primary medium for nutrient supply, root development, and microbial interactions that directly affect rice growth, grain composition, and yield. This pressure results in nutritional imbalances, biodiversity loss, soil deterioration, and the depletion of soil organic carbon [5]. Rice varies in terms of grain length, color, thickness, stickiness, aroma, growing conditions, and production practices, all impacting its quality and nutritional profile [6]. In this study, rice grains and the corresponding soil samples were collected from fields cultivated with the Phka Rumdoul variety in nine provinces of Cambodia, for a total of 66 samples under different farming practices. The soil samples were subject to analysis for their physico-chemical properties. For rice grains, key quality indicators, including amylose content, protein content, seed weight, and yield were measured. After that, we studied the relationship between soil properties and rice grain quality to determine whether different farming systems have an effect on grain quality and yield. Thus, this study examines the effect of soil properties on the quality and yield of Phka Rumdoul rice in Cambodia.

2. METHODOLOGY

2.1 Sample survey

Phka Rumdoul rice grains were collected around 1kg for each site from the farmer who was working with the HealthyRice Consortium (FEF project). After drying at 14% humidity, the grains were sent to the HealthyRice laboratory at ITC to be stored at 4ºC before analysis. At the same time, a soil sample was collected from each site where the rice variety Phka Rumdoul was grown and stored in a laboratory at room temperature.

Table 1. Survey sample of nine provinces in Phka Rumdoul field across Cambodia

Sample code	Provinces	Agriculture practices
n(1)	Phnom Penh	CA
n(2)	Kampong Speu	CT
n(3)	Kampong Chnang	CT
n(4)	Battambang	CA
n(5)	Siem Reap	SRP
n(6-10)	Preah Vihear	SRP, GM, CT
n(11-12)	Stung Treng	SRP, CT
n(13)	Rattanakiri	SRP
n(14-66)	Kampong Thom	CA, CT, ORF

(CT) Conventional tillage, (CA) Conservation agriculture, (SRP) Sustainable rice platform, (GM) Green manure, (ORF) Organic rice farming, (NA) No data, (DAP) Diammonium hydrogen phosphate, (KCL) Potassium Chloride, (Urea) Nitrogen only, (NPK) Nitrogen, Phosphorus, Potassium

2.2 Sample collection

Phka Rumdoul rice variety and soil samples were collected from nine provinces across Cambodia during the harvest season in November 2024, including Phnom Penh, Kampong Speu, Kampong Chhnang, Kampong Thom, Battambang, Siem Reap, Preah Vihear, Stung Treng, and Ratanakiri. These rice fields were cultivated with different farming practices, fertilizer treatments, generation of rice seeds, day of sowing, and harvest, and so on. The differences in farming practices and environmental conditions in these regions contribute to variations in the quality, characteristics, and nutritional composition of both rice grains and soils.

2.3 Soil physicochemical properties analysis

Soil samples from the different sites (**Table 1)** were collected from nine provinces, totaling 66 samples. In each field, soil was collected from five plots (150 g per plot) at a depth of 10 to 15 cm from the surface. The five soil cores were thoroughly mixed. The composite samples were then air-dried at room temperature on an Agro-food wrapping paper (300 g per sample) to prevent chemical changes caused by direct sunlight. Once dried, the samples were ground with a pestle to break soil clumps and sieved through a 2mm mesh. The prepared samples were analyzed for physical and chemical properties, including, Soil Moisture content % (Oven 105°C and 24hs), Total Carbon, and Organic matter (Black& Walkey Method), Total Nitrogen (Kjeydal sulfuric Method), Total phosphorus (Nitric Digestion), Cation Exchange Capacity (Ammonium acetate pH7.0), Exchangeable Ca and Mg (EDTA titration) Exchangeable K and Na (Flame photometer), pH 1:2.5 (H_2O, KCl), C/N ratio (unit). The analysis was conducted at the Soil Science Laboratory in the Royal University of Agriculture.

2.3 Nutritional and physical quality analysis

After collecting the rice grains, the husks were removed using the Rice Husking Machine (brown rice model P-1) to separate the husk from the seeds, resulting in brown rice. Brown rice was used to analyze nutrition, such as protein and

amylose, by using Near-infrared spectroscopy (Kett AN-920 Near-infrared Grain Tester). To evaluate the physical quality of rice grains, the weight of Phka Rumdoul brown rice of 100 seeds was measured. From each sample, 100 brown rice grains were selected randomly. The selected grains were cleaned to remove any dust or impurities and weighed using an analytical balance with a precision of 0.01g.

2.5 Statistic analysis

Principal component analysis (PCA) was conducted to evaluate the relationships among the different 66 fields of nine provinces of soil physicochemical properties including evaluate the relationship between soil properties and rice grain quality and yield, which were introduced as variables in the PCA using R studio version 4.4.3.

3. RESULTS AND DISCUSSIONS

3.1. Variability of soil physicochemical properties

The physical and chemical properties of soils from 66 fields in nine provinces were analyzed. A principal component analysis (PCA) based on these soil physicochemical properties is presented in **Fig. 1.** The first principal component (Dim1) explained 27.9% of the total variance, while the second component (Dim2) explained 17.3%. Soil organic carbon (SOC), soil organic matter (SOM), C/N ratio, and clay content are clustered together, indicating a strong positive correlation among these variables. In contrast, pH (both H_2O and KCl) and exchangeable bases (Na, K, Mg, and Ca) are closely aligned, suggesting a related pattern or influence. Sand and moisture content appear in the opposite direction to cation exchange capacity (CEC) and clay, suggesting that soils with high sand content tend to have lower CEC and clay content. In addition, electrical conductivity (EC), available phosphorus (P), and total nitrogen (N) also show some correlation, although not as strong as other variables. The black dots represent individual soil samples coded by field. Their distribution in PCA space indicates the diversity of soil properties among the sampled fields.

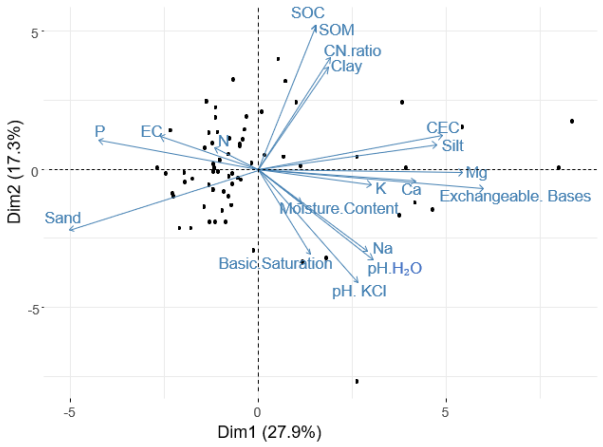


Fig. 1. Principle component analysis (PCA) of soil physico-chemical properties in 66 fields. Each point represents a soil sample. EC µS/cm, CEC: Total Cationic Exchange Capacity; C/N Ratio Carbon/Nitrogen Calcium (Ca), Magnesium (Mg), Potassium (K), Sodium (Na), Exchangeable Bases meq/100g, Basic Saturation, SOM: Soil organic matter, SOC: soil organic carbon, N: total nitrogen, P: Available Phosphorus, Clay, Silt, Sand, Moisture Content.

3.2 Protein content

Protein content from 66 field samples was shown in **Fig. 2**. On average, protein content in rice grains across all samples was approximately 8%. However, two samples one sample from the organic field in Stung Treng province and another from Kampong Thom under conservation agriculture, showed a higher protein content of 11.2% and 11.1%, respectively. Whereas the lower protein content was observed in samples from Preah Vihear under sustainable rice platform (7.1%) and Preah Vihear under the green manure system (7%). Typically, the protein content in brown rice is around 8% [7]. If protein content is higher or lower, several factors can influence including genetic differences, environmental conditions, and management practices [8]. In our result, the higher protein content in the Stung Treng which practices organic rice farming, and kampong Thom plot under conservation agriculture may be related to soil nitrogen levels. High soil nitrogen often leads to increased protein content as nitrogen is a key element in amino acids and protein [9].

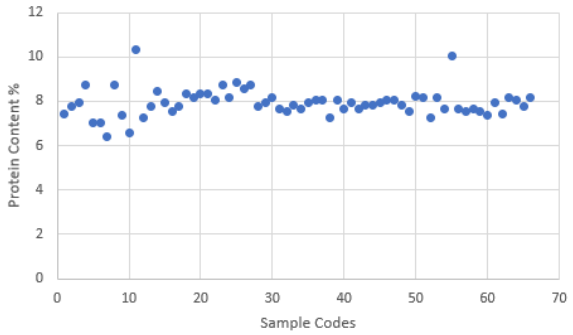


Fig. 2. Percentage of Protein content in Rice grain from 66 fields using NIRS

3.3 Amylose content

Fig. 3 presents of amylose content in rice grain from the 66 fields. Rice is classified based on amylose content as waxy (0-2%), very low (3-9%), low (10-19%), intermediate (20-25%), and high (>25%) [10]. In our result, the average amylose content in rice grains was approximately 17%. However, one sample from Battambang under sustainable rice platform displayed a noticeably higher amylose content of 21.1%, categorizing it closer to the higher amylose range. This suggests a potential variation in starch composition, which environmental conditions, soil properties, or subtle varietal differences in that specific field could influence. Higher amylose rice tends to be less sticky and firmer when cooked.

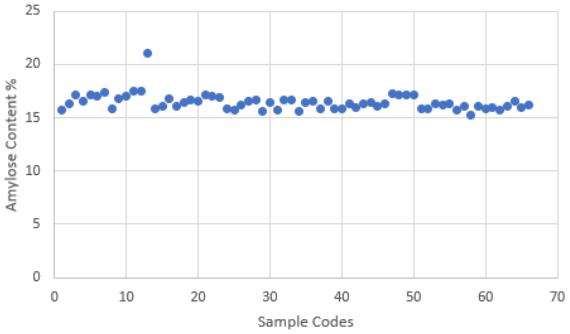


Fig. 3. Percentage of Amylose content in rice grain from 66 fields using NIRS

3.*4 Physical value*

Fig. 4 shows the seed weight of rice grain from 66 fields. On average, the 100-seed wet weight per sample was about 2.0 g, which is within the expected range for the Phka Rumdoul rice variety. The highest values were recorded in samples from Battambang, which practices a sustainable rice platform and Kampong Thom under conservation agriculture with weights of 2.4 g and 2.5 g, respectively. These results suggest that grains from these fields were relatively larger, possibly due to more favorable growing conditions or improved nutrient availability, particularly during the grain-filling stage. In contrast, the lowest 100 seed weight was observed in the sample from Preah Vihear, which practices conventional tillage, indicating smaller grain. This could be due to environmental stress, suboptimal soil conditions, or reduced nutrient uptake during critical stages of reproductive development [11].

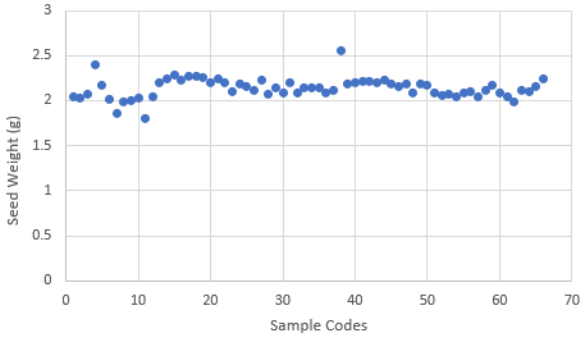


Fig. 4. Measurement of seed weight of 100 Rice Grains from 66 fields

3.5 Yield rice grain

Fig. 5 shows the grain yield (tons per hectare) recorded from 66 rice fields after harvest. On average, the rice yield for all fields was approximately 2.6 t/ha with the standard deviation (0.66). Among these, Siem Reap, which practices under sustainable rice platform had a relatively high yield of 4.47 t/ha. Phka Rumduol typically yields 3.5-5 t/ha under rainfed lowland conditions. This indicates favorable growing conditions and potentially effective nutrient or water management in this area. Fields with higher yields may have benefited from better input management or less stress during critical stages of growth, while fields with lower yields may have been affected by nutrient limitation, drought, or pest pressure.

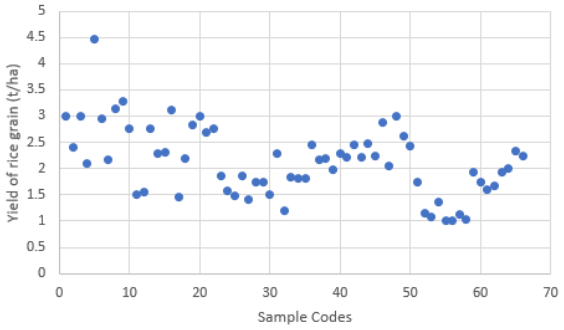


Fig. 5. Yield of rice grain (t/ha) of 66 fields

3.6. Correlation of soil properties on rice grain and yield quality

As previously described in **Fig. 1** soil physicochemical properties including pH, electrical conductivity (EC), organic matter content, soil texture (sand silt, clay), and cation exchange capacity (CEC), were analyzed across 66 fields from nine provinces in Cambodia. These factors play fundamental roles in determining soil fertility, nutrient availability, and overall growing conditions for rice quality and rice production. In **Fig. 6** shows the correlation between various soil properties and rice quality parameters such as protein, amylose content, seed weight, and yield (t/ha). Understanding these relationships provides insight into how soil condition directly influences rice performance and can inform effective soil and nutrient management strategies to improve both grain quality and productivity. The first principal component (Dim1) explains 25.3% of the total variance, while the second principal component (Dim2) explains 15.5%. Yield strongly correlates positively with amylose content, CEC, silt, Mg, Ca, K, and exchangeable bases. Furthermore, strong positive correlations are observed among SOC, SOM, and clay as well as between the two measure of soil pH (H_2O, HCl). Interestingly, rice yield exhibits a negative correlation with seed weight, which may indicate a trade-off between grain number and grain size. This suggests that in certain environments, particularly those with limited resources or that yield more than others, plants may prioritize producing a larger number of smaller grains rather than fewer larger ones. This phenomenon aligns with the resource allocation theory, in which environmental and physiological factors influence assimilate partitioning between seed number and seed size. Additionally, while P and N are crucial for plant growth, their apparent lack of a positive association with yield in this study suggests that they may not have been the main limiting factors in the sampled fields. In some cases, an oversupply or imbalanced nutrient availability can lead to inefficient nutrient uptake or imbalances that do not increase yield. The influence of yield in these fields on soil structure, micronutrients, and other properties such as exchangeable bases and CEC, which enhance nutrient retention and availability, is also a possibility. Conversely, yield was negatively associated with soil pH (H_2O and KCl), sodium (Na), moisture content, and base saturation. This indicates that excessive sodium, moisture, or certain pH conditions may have limited rice productivity. Additionally, sand content was negatively correlated with SOC, SOM, clay, C/N ratio, seed weight, P, N, and protein content, likely reflecting the poor nutrient and water-holding capacity of sandy soils. Sand content also has a negative correlation with amylose content. This implies that soil with higher sand content tends to be associated with lower amylose content levels in rice grain.

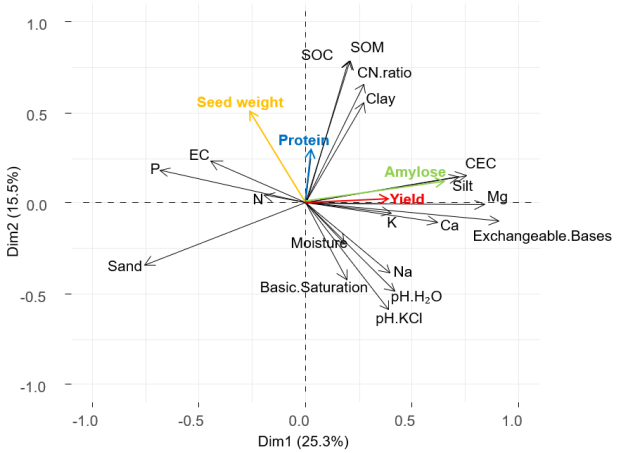


Fig. 6. Principal component analysis (PCA) of soil physicochemical properties correlation with rice quality and yield quality

4. CONCLUSIONS

We measured protein, amylose content, seed weight, and yield of Phka Rumdoul rice grains from the 66 fields in nine provinces of Cambodia. Using PCA, we found that amylose and yield were correlated with CEC, silt, K, Ca, Mg, and exchangeable bases, while protein was correlated with SOC, SOM, C/N ratio and clay. Seed weight was correlated with EC, P, N, and sand. Thus, rice quality and yield may depend on soil properties. In addition, we also found that in the sample from Stung Treng under the practices of organic rice farming cropping system and Kampong Thom under conservation agriculture gave higher protein content, and Battambang under SRP had higher amylose content and seed weight. In the Siem Reap field, under SRP practices, also was a relatively high yield. Based on the results, these higher levels of rice quality and yield are under CA and SRP. Therefore, the cropping systems interaction in the field seems to affect the seed quality. CA and SRP should be further studied to see the correlation between agro-systems with rice quality at the same sites. However, as a term of agro-system factors such as rice genetics, climate change (temperature, weather, humidity), and edaphic factors (e.g., soil properties, soil microbiota) [9], should be further studied to understand which factors affect soil properties and correlate with rice quality and productivity. The analyses were carried out on the Phka Rumdoul rice variety but the seeds were not all sourced from the same origin. As a result, potential genetic divergences within this variety have contributed to the difference observed. Moreover, the repeat experiment with a genetically homogeneous or certified seed stock, such as one provided by CARDI, would help disentangle environment and agronomic effects from genetic factors. Additionally, it should be interesting to see whether the different agro-systems investigated in this study have influenced the aromatic quality of Phka Rumdoul rice.

ACKNOWLEDGMENTS

This research was supported by the Fonds Equipe France (FEF) 2024-14 project, funded by the French Ministry for Europe and Foreign Affairs. Rice and soil samples were provided by the Sansom Mulp Prey Organization, CIRAD, and GDA-DALRM, with the latter two supported by the ASSET project (AFD-EU). The authors also thank farmers from various fields and provinces for access to their samples. Additional support came from the French Research Institute for Sustainable Development (IRD) and ITC through the LMI LEAD project.

REFERENCES

[1] S. Muthayya, J. D. Sugimoto, S. Montgomery, and G. F. Maberly, "An overview of global rice production, supply, trade, and consumption," *Ann. N. Y. Acad. Sci.*, vol. 1324, no. 1, pp. 7–14, 2014, doi: 10.1111/nyas.12540.

[2] J. S. Song and C. C. Kuo, "Farming practice affects rice field animal biota during cultivation but not fallow periods in Taiwan," *Ecosphere*, vol. 13, no. 5, pp. 1–15, 2022, doi: 10.1002/ecs2.4069.

[3] FAO in cambodia, "Cambodia Agriculture Survey 2021 Report now released." [Online]. Available: https://www.fao.org/cambodia/news/add-news/detail-events/es/c/1674999/

[4] J. Lehmann, D. A. Bossio, I. Kögel-Knabner, and M. C. Rillig, "The concept and future prospects of soil health," *Nat. Rev. Earth Environ.*, vol. 1, no. 10, pp. 544–553, 2020, doi: 10.1038/s43017-020-0080-8.

[5] L. Montanarella *et al.*, "World's soils are under threat," *Soil*, vol. 2, no. 1, pp. 79–82, 2016, doi: 10.5194/soil-2-79-2016.

[6] P. P. Behera *et al.*, "Genetic gain and selection of stable genotypes in high zinc rice using AMMI and BLUP based stability methods," *Emirates J. Food Agric.*, vol. 35, no. 12, pp. 1–23, 2023, doi: 10.9755/ejfa.2023.3180.

[7] H. Hoogenkamp, H. Kumagai, and J. P. D. Wanasundara, "Rice Protein and Rice Protein Products," *Sustain. Protein Sources*, pp. 47–65, 2017, doi: 10.1016/B978-0-12-802778-3.00003-2.

[8] Y. J. Shijie Shi, Keqiang Pan, Gaoyu Zhang, Dan Zhao, Hui Zhou, Juan Liu, Cougui Cao, "Differences in grain protein content and regional distribution of 706 rice accessions," *Sci. Food Agric.*, vol. 103, no. 3, pp. 1593–1599, 2023.

[9] Z. Zhou, K. Robards, S. Helliwell, and C. Blanchard, "Composition and functional properties of rice," *Int. J. Food Sci. Technol.*, vol. 37, no. 8, pp. 849–868, 2002, doi: 10.1046/j.1365-2621.2002.00625.x.

[10] I. Kumar, G. S. Khush, and B. O. Juliano, "Genetic analysis of waxy locus in rice (Oryza sativa L.)," *Theor. Appl. Genet.*, vol. 73, no. 4, pp. 481–488, 1987, doi: 10.1007/BF00289183.

[11] S. F. Chao Wu, Kehui Cui, "Heat Stress Decreases Rice Grain Weight: Evidence and Physiological Mechanisms of Heat Effects Prior to Flowering," *Int. J. Mol. Sci.*, vol. 23, no. 18, p. 10922, 2022, doi: 10.3390/ijms231810922.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Analysis of pesticide residues in rice grains using QuEChERS method follow by GC-MS

Aykim Heng [1], Chanvorleak Phat [1,2], Leanghok Bun [2], Mathilde Sester [2,3], Florent Tivet [3], Thanon Nuon [4], Sereyvath Yoeun[1,2*]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3]CIRAD, *UPA Agroecologie et intensification Durable des cultures Annuelles (AIDA)*

[4]*Faculty of Agricultural Science, Royal University of Agriculture, Phnom Penh, Cambodia*

[*]*Corresponding author: svath@itc.edu.kh*

Abstract: *Pesticide residues in rice grains are a significant concern for food safety and public health, especially when rice is a staple food for billions of people worldwide, making its safety a critical concern. Rice is grown in warm and humid climates that attract insect pests and farmers depend largely on the application of agrochemicals such as fertilizers and pesticides to overcome this problem. This study aims to analyze the present of pesticide residues in rice grains collected from Kanghot, Battambang province, during two different cycles: Cycle 1 (end July to end September) and Cycle 2 (end October to early January), sampling was conducted during two distinct cultivation cycles: Cycle 1, from late September to early October, and Cycle 2, from early January to late January. A total 29 samples which 14 samples in Cycle 1 and 15 samples in Cycle 2, the rice grains samples were collected from plot A1, plot A2, plot B1, plot B2 and plot D. Samples were prepared using the QuEChERS (Quick, Easy, Cheap, Effective, Rugged, and Safe) method, followed by gas chromatography-mass spectrometry (GC-MS) analysis. This QuEChERS principle involves a two-step process: first, extraction with acetonitrile, salt, and water; second, cleanup using dispersive solid-phase extraction (d-SPE) with sorbents like magnesium sulfate (MgSO4) to remove excess water and primary secondary amine (PSA) to remove sugars, fatty acids, organic acids, and anthocyanine pigments. The result indicates the presence of Chlorpyrifos (an organophosphate insecticide) was detected in rice grain samples from in A1 and B2 during Cycle 1. In contrast, no residues of Chlorpyrifos were found in Plot A2, Plot B2, or Plot D. During Cycle 2, Isoprothiolane (a systemic fungicide) was detected only in Plot A2, while no residues were detected in Plot A1, Plot B1, Plot B2, or Plot D. This seasonal variation highlights the dynamic nature of pesticide application and residue accumulation. The study underscores the effectiveness of the QuEChERS-GC-MS technique for rapid and accurate pesticide residue analysis and emphasizes the critical need for continuous monitoring to safeguard rice quality and public health.*

Keywords : Pesticide Residue; QuEChERES; GC-MS; Rice Quality.

1. INTRODUCTION

Rice (Oryza sativa) is one of the most consumed cereals and an excellent source of carbohydrates and energy. In the period from 2022 to 2023, approximately 520 million tons of rice were consumed worldwide. Rice is the daily staple for over half of the world's population, particularly in Asia, where it serves as the primary source of energy for millions of people. More than 90% of the world's rice production is from Asia. It is a major grain cereal grown in over 100 countries and consumed by over two billion people in the world [1]. In Cambodia is one of the Asian countries which is agriculture country especially produce rice. Battambang Province, located in the northwest of the country, is known for its high-quality rice production and is a critical area for rice cultivation.

[1*] Corresponding author: Sereyvath Yoeun
E-mail: svath@itc.edu.kh; Tel: +855-10348588

Humans use pesticides to protect their crops and improve the quality of their food[2]. Its production process, however, unavoidably exposes it to pesticides which may detrimentally impact its quality due to residues. To ensure high rice yields and enhance quality, farmers usually rely on pesticides to control pests and diseases to protect the crop from damage and ensure food safety. However, if pesticides are used inappropriately or irrationally, residues are retained in rice and these pesticide residues can enter the human body. Pesticide residues in agricultural products refer to chemical substances that may remain on crops after the use of pesticides to control pests, diseases, etc[3].

It is important to know the presence of pesticide residues in food to guarantee food safety. Pesticides in the food chain is one of major public health concern, where several studies have analysed pesticide residues content in rice grains. The monitoring of pesticide residues is important to ensure not only food safety but also compliance with good manufacturing practices. Thus, several government authorities and international organizations have established maximum residue levels (MRLs) to regulate pesticide concentrations in food products[4].

The objective of this study is to investigate the presence of pesticide residues in rice grains (husked rice) from Kanhot, Battambang Province, was carry out using QuEChERS (quick, easy, effective, rugged and safe) extraction technique base on European equivalent, the EN 15662 method followed by analysis using gas chromatography–mass spectrometry (GC-MS).

2. METHODOLOGY

2.1 Standard and Material

Certified reference materials and analytical-grade reagents were used throughout the study. Acetonitrile was purchase from RCI Labscan Limited and ultrapure water was used as extraction solvents. QuEChERS extraction and clean-up salts were purchased from Agilent Scientific Instruments (USA). The extraction salt packet contained 4 g magnesium sulfate ($MgSO_4$), 1 g sodium chloride (NaCl), 1 g sodium citrate, and 0.5 g disodium citrate sesquihydrate. Clean-up was performed using 15 mL dispersive solid phase extraction (d-SPE) tubes containing 150 mg primary secondary amine (PSA) and 900 mg magnesium sulfate ($MgSO_4$).

A standard pesticide mixture containing 203 compounds (catalog no. 32562) was purchased from RESTEK Corporation (USA) and used for method calibration and quantification. Rice grain samples were collected from Battambang Province, Cambodia, for analysis.

2.2 Study area and Sample Collection

The study was conduct in Kanghot irrigation system located in Battambang province.

The study area consisted of three sites: Site A, Site B, and Site D. Sites A and B were each divided into two plots. Plot A1 and Plot A2 for Site A, and Plot B1 and Plot B2 for Site B while Site D comprised a single plot, designated as Plot D.

Rice sample were collected from Kanghot, Battambang province during two different growing cycles: Cycle 1 collected from end September to early October and Cycle 2 collected from early January to end January from plot A1, plot A2, plot B1, plot B2 and plot D. A total 29 samples were collected using randomize sampling approach to ensure representativeness across fields.

2.3 Sample Preparation

Rice samples were store in refrigerator at -20℃. To prepare the sample the analysis, the harvested rice grains first were dehulled using a rice husking machine. The dehusked grains were then homogenize into fine power using a blender. Approximately 5g of homogenize sample was use for pesticide residue extraction.

The QuEChERS method originally published by Anastassiades et al. (2003) and Lehotay et al. (2005) for the extraction and determination of pesticide residues in produce was followed with modification explained below for the extraction of pesticide residues in rice and validated [1].

Before conducting the experiment, a recovery study of rice grain samples was performed using a standard pesticide mixture containing 203 compounds (catalog no. 32562). The recovery study is a critical component of method validation, as it assesses the accuracy and efficiency of the analytical method. A recovery rate within the range of 70% to 120% is considered acceptable for ensuring the reliability of the analytical process.

In this experiment the rice grains sample were sampling triplicate. The procedure involved weighted accurately 5g of homogenized rice grain sample into 50mL centrifuge tube using an electric balance, followed by the addition of 5 mL of ultrapure water. The mixture was vortexed for 1 minute to ensure proper hydration. Subsequently, 10 mL of acetonitrile (ACN) was added, and the sample was vortexed again for 1 minute. A ceramic homogenizer was placed into the tube, and QuEChERS (quick, easy, effective, rugged and safe) extraction salts (containing: 4g MgSO4, 1g NaCl, 1g NaCitate, 0.5g disodium citrate sesquihydrate) were added. The tube was shaken thoroughly until the salts dissolved completely, then vortexed for an additional 1 minute. The sample was centrifuged at 4,000 rpm and 4 °C for 5 minutes. Following centrifugation, 6 mL of the supernatant was transferred to a clean-up tube containing dispersive solid phase extraction (d-SPE) materials vortexed for 1 minute. The tube was then centrifuged again at 4,000 rpm and 4 °C for 5

minutes. Finally, 1 mL of the cleaned supernatant was filtered through syringe filter and transferred into an autosampler vial with an insert. From this, a 100 µL aliquot of the extract was taken, and 10 µL of internal standard (ISTD) solution was added prior to GC-MS analysis.

3. RESULTS AND DISCUSSION

3.1 Result

In this study, pesticide residue analysis was conducted at three sites: Site A, Site B, and Site D. Sites A and B were managed by local farmers following their usual agricultural practices, while Site D was designated for experimental purposes, where controlled pesticide application was carried out for research analysis. This study investigated pesticide residues in rice grains collected from five agricultural plots A1, A2, B1, B2, and D located in Kanghot, Battambang Province, Cambodia, during two cultivation cycles. The rice grains samples were sampling triplicate, the extraction of pesticide residues was performed using the Quick, Easy, Cheap, Effective, Rugged, and Safe (QuEChERS) method, followed by detection and quantification through Gas Chromatography–Mass Spectrometry (GC-MS). The result of two Cycles was shown as Table 1 and Table 2.

Table 1. The Result of pesticide residue analysis

Site	Code Sample	Detected Compound (Cycle 1)	Concentration (mg/Kg) (Cycle 1)
A	Plot A1	Chlorpyrifos	1.56 ±0.04
	Plot A2	ND	
B	Plot B1	ND	
	Plot B2	Chlorpyrifos	0.22±0.03
D	Plot D	ND	

Table 2. The Result of pesticide residue analysis

Site	Code Sample	Detected Compound (Cycle 2)	Concentration (mg/Kg) (Cycle 2)
A	Plot A1	ND	
	Plot A2	Isoprothiolane*	
B	Plot B1	ND	
	Plot B2	ND	
D	Plot D	ND	

* Isoprothiolane was not included in the list of target compounds; therefore, its quantification could not be performed.

3.2 Discussion

In Cycle 1 (end July to end September), Table 1 shown that there is only one compound which is Chlopyrifos (an organophosphate insecticide) detect in plot A1 with average concentration 1.56 mg/Kg (SD = 0.04) and plot B2 with average concentration 0.22 mg/Kg(SD = 0.03) while another three plots: plot A2, plot B1 and plot D do not detect any compound. These findings indicate that Chlorpyrifos was selectively applied in certain plots, possibly due to varying pest pressure, differing agricultural practices, or farmer preference. Chlorpyrifos is used to control many different kinds of pests, including termites, mosquitoes, and roundworms [5].

In Cycle 2 (end October to early January), Table 2 shown that Isoprothiolane (systemic fungicide) was detected in plot A2. Although quantification was not possible due to this compound being outside of the 203 target compounds standard mix used in the Gas Chromatography–Mass Spectrometry (GC-MS) analysis, its presence was confirmed. There is no detection of any compounds in plot A1, plot B1, plot B2 and plot D.

The absence of pesticide residues in several plots across both cycles may be attributed to several factors: the timing of application relative to sample collection, where residues may have dissipated or degraded before harvest and sample collection, environmental degradation, by the temperature during rice grain processing may influence the persistence of pesticide residues, rainfall, soil type and organic matter content[6]. Chlorpyrifos is known to degrade faster in moist and warm environments, potentially leading to non-detection in some plots, variability in pesticide usage by farmers including differences in pesticide type, frequency, and dosage[7]. Additionally, for residues present at trace levels, limitations in detection sensitivity could also contribute to non-detection in some cases.

The difference in compound detection across both cycles due to seasonal variation in the application pesticides and degradation behavior[8].

The reported half-life of isoprothiolane in rice ranges from 7.5 to 23.5 days [9], While the half-life of chlorpyrifos in rice plants, based on data from 2009-2010, has been found to range between 3.24 to 4.28 days [10].

According to the European Union (EU), the maximum residue limit (MRL) for chlorpyrifos in rice grains is set at 0.01 mg/kg[11]. According to the Environmental Protection Agency (EPA) for Isoprothiolane in husked rice is set at 6 mg/Kg [12]. In Cycle 1 of this study, the concentration of chlorpyrifos detected in Plot A1 was 1.56 mg/kg, which is 156 times higher than the EU MRLs. Similarly, in Plot B1, the concentration was 0.22 mg/kg, which is 22 times higher than the permitted limit.

These results highlight the variability in pesticide usage and residue distribution across smallholder rice fields,

emphasizing the importance of continued monitoring to assess food safety and compliance with maximum residue limits (MRL).

Chlorpyrifos is an organophosphate pesticide that inhibits acetylcholinesterase (AChE), an essential enzyme for nerve function[13]. High levels of exposure can lead to acute cholinergic syndrome, characterized by symptoms such as excessive salivation, muscle twitching, respiratory distress, convulsions, and potentially death due to respiratory failure. Even at lower levels, chronic exposure to chlorpyrifos has been associated with long-term neurological effects, including memory loss, confusion, depression, and sleep disturbances. These symptoms may persist even after exposure has ceased[14].

Pesticide exposure was estimated by multiplying the residue concentration with the average daily per capita intake of cereal grains and flours, based on GEMS/Food Consumption cluster diet data. The resulting intake was expressed as a percentage of the Acceptable Daily Intake (ADI) and the Acute Reference Dose (ARfD) for an individual weighing 60 kg. The Estimated Daily Intake (EDI) of pesticide residues was calculated as grain consumption multiplied by the residue[15].

Long-term risk assessment of intakes was performed by calculating the chronic risk quotient (RQ-c) by dividing the estimated daily intake by the relevant acceptable daily intake[16].

The health risk assessment for chlorpyrifos is based on an Acceptable Daily Intake (ADI) of 0.01 mg/kg body weight per day over a lifetime. Its Acute Reference Dose (ARfD) is set at 0.1 mg/kg body weight, which is considered safe for consumption within a 24-hour period without adverse health effects [17].

In the case of isoprothiolane, International Estimated Daily Intakes across the 17 GEMS/Food cluster diets ranged from 0 to 2% of the maximum ADI of 0.1 mg/kg body weight. The Joint FAO/WHO Meeting on Pesticide Residues (JMPR) determined that establishing an Acute Reference Dose (ARfD) for isoprothiolane was not necessary, concluding that short-term dietary exposure from the evaluated uses is not expected to pose a risk to public health.[18].

4. CONCLUSIONS

Analysis of pesticide residue in rice grains during two different cultivation Cycles was detected Chlorpyrifos in Cycles 1. And Isoprothiolane was detected in Cycles 2 in rice grains samples collected from Kanghot, Battambong province. The presence of these compounds indicates ongoing pesticide application in rice cultivation and highlights the importance of monitoring pesticide residues to ensure food safety and compliance with international standards. The findings emphasize the need for proper pesticide management and awareness among farmers regarding the safe and judicious use of agrochemicals. Analysis the presence of pesticide residue in rice grains plays a crucial role in food safety.

ACKNOWLEDGMENTS

This research was financially supported by the Water Resources Management and Agro-ecology Transition for Cambodia (WAT4CAM) projects.

REFERENCES

[1] R. Harischandra Naik *et al.*, “Determination of 72 Chemical Pesticides and Estimation of Measurement of Uncertainty in Rice Using LC-MS/MS and GC-MS/MS,” *Food Anal Methods*, vol. 14, no. 9, pp. 1788–1805, Sep. 2021, doi: 10.1007/s12161-021-02000-9.

[2] D. Nandan, “An Overview on the Study of Pesticides,” 2018. [Online]. Available: www.jetir.org

[3] T. Jiang, F. Deng, W. Dong, Q. Zhang, and P. Liu, “Rice quality prediction and assessment of pesticide residue changes during storage based on Quatformer,” *Sci Rep*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-59816-8.

[4] L. Pareja, A. R. Fernández-Alba, V. Cesio, and H. Heinzen, “Analytical methods for pesticide residues in rice,” Feb. 2011. doi: 10.1016/j.trac.2010.12.001.

[5] E. Reaves and A. Director, “Chlorpyrifos Proposed Interim Registration Review Decision Case Number 0100,” 2020. [Online]. Available: www.regulations.gov

[6] N. Yigit and Y. S. Velioglu, “Effects of processing and storage on pesticide residues in foods,” Nov. 29, 2020, *Bellwether Publishing, Ltd.* doi: 10.1080/10408398.2019.1702501.

[7] E. Reaves and A. Director, “Chlorpyrifos Proposed Interim Registration Review Decision Case Number 0100,” 2020. [Online]. Available: www.regulations.gov

[8] S. F. Materu, S. Heise, and B. Urban, “Seasonal and Spatial Detection of Pesticide Residues Under Various Weather Conditions of Agricultural Areas of the Kilombero Valley Ramsar Site, Tanzania,” *Front Environ Sci*, vol. 9, Mar. 2021, doi: 10.3389/fenvs.2021.599814.

[9] G. Chen *et al.*, “Dissipation rates, residue distribution and dietary risk assessment of

isoprothiolane and tebuconazole in paddy field using UPLC-MS/MS," *Int J Environ Anal Chem*, pp. 1–13, 2020, doi: 10.1080/03067319.2020.1792894.

[10] X. Zhang, Y. Shen, X. Y. Yu, and X. J. Liu, "Dissipation of chlorpyrifos and residue analysis in rice, soil and water under paddy field conditions," *Ecotoxicol Environ Saf*, vol. 78, pp. 276–280, Apr. 2012, doi: 10.1016/j.ecoenv.2011.11.036.

[11] European Union, "Corrigendum to Commission Regulation (EU) 2020/1085 of 23 July 2020 amending Annexes II and V to Regulation (EC) No 396/2005 of the European Parliament and of the Council as regards maximum residue levels for chlorpyrifos and chlorpyrifos-methyl in or on certain products," 2020.

[12] Environmental Protection Agency, "Isoprothiolane; Pesticide Tolerances," Dec. 2021.

[13] M. G. Robson, S. Jaipieam, P. Visuthismajarn, W. Siriwong, and M. Borjan, "Inhalation exposure of organophosphate pesticides by vegetable growers in the bang-rieng subdistrict in Thailand," *J Environ Public Health*, vol. 2009, 2009, doi: 10.1155/2009/452373.

[14] D. B. Barr and J. Angerer, "Potential uses of biomonitoring data: A case study using the organophosphorus pesticides chlorpyrifos and malathion," *Environ Health Perspect*, vol. 114, no. 11, pp. 1763–1769, Nov. 2006, doi: 10.1289/ehp.9062.

[15] J. Mbabazi, "Principles and Methods for the Risk Assessment of Chemicals in Food," *International Journal of Environmental Studies*, vol. 68, no. 2, pp. 251–252, Apr. 2011, doi: 10.1080/00207233.2010.549617.

[16] B. Lozowicka *et al.*, "Pesticide residues in grain from Kazakhstan and potential health risks associated with exposure to detected pesticides," *Food and Chemical Toxicology*, vol. 64, pp. 238–248, 2014, doi: 10.1016/j.fct.2013.11.038.

[17] Inventory of evaluations performed by the Joint Meeting on Pesticide Residues (JMPR), *Pesticide residues in food, 2004 : report of the joint meeting of the FAO Panel of Experts on Pesticide Residues in Food and the Environment and the WHO Core Assessment Group on Pesticide Residues, Rome, Italy, 120-29 September 2004*. World Health Organization : Food and Agriculture Organization of the United Nations, 2004.

[18] Joint FAO/WHO Meeting on Pesticide Residues (JMPR), "5.23_ISOPROTHIOLANE__299_," Sep. 2017.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessment of Pesticide Residues in Rice and Water from Kanghot Paddy Fields in Battambang Province

Leanghok Bun [1], Chanvorleak Phat [1,2], Mathilde Sester [1,3], Florent Tivet [3], Thanon Nuon [4], Sereyvath Yoeun [1,2*]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *CIRAD, UPR Agroécologie et Intensification Durable des cultures Annuelles (AIDA)*

[4] *Faculty of Agricultural Science, Royal University of Agriculture, Phnom Penh, Cambodia*

[*] *Corresponding author: svath@itc.edu.kh*

Abstract: *This study aims to assess pesticide residues in rice and water samples collected from a Kanghot paddy field. There were two plots (C1 and C2) of paddy field and one canal where points were pointed out for sample collection in three stages of rice growth within two cycles, first and second cycle. At the vegetative and reproductive stages, water samples from the paddy field and canal were collected in triplicates. During the ripening stage, both paddy rice and canal water were sampled in triplicates. Water samples were extracted using solid-phase extraction, while paddy rice was de-husked, ground into fine particles, and extracted using the QuEChERS method. Qualitative and quantitative of pesticide analysis were conducted using GCMS. A total of 18 samples were analyzed, including 14 water sample, 4 rice samples, and 6 control samples spanning six sampling stages. Eleven pesticide compounds were detected in the paddy field water, including isoprothiolane, buprofezin, fipronil, fenoxanil, chlorpyrifos, tebuconazole, chlorfenapyr, hexaconazole, paclobutrazol, azoxystrobin, and pyraclostrobin. Among these, paclobutrazol and isoprothiolane were most found at the highest concentrations in paddy field water. During the second cycle, eight pesticide compounds were detected in paddy field water at plot C1, whereas chlorpyrifos was the only compound found in rice samples. According to the World Health Organization, chlorpyrifos is classified as moderately hazardous to humans (Class II, 1999). Alarmingly, chlorpyrifos was exceeded MRLs in 25% of the total rice samples analyzed. These findings indicate that pesticides applied during rice cultivation by farmers can leave residues in rice. Despite its limited occurrence, the high level poses a potential health risk due to the toxic effects of its metabolite, chlorpyrifos oxon, emphasizing the need for strict monitoring and regulation of pesticide residues in food. The main sources of pesticide contamination appear to be direct application by paddy field owners for crop protection and productivity enhancement, as well as runoff from pesticides used in nearby fields sharing the same irrigation system.*

Keywords: Quick, Easy, Cheap, Effective, Rugged, and Safe (QuEChERS); Solid phase extraction; Gas chromatography-mass spectrometry (GC-MS)

1. INTRODUCTION

Rice, a staple food for more than half the world's population, holds a paramount position in global agriculture and food security, serving not only as a dietary essential but also as a cultural and economic cornerstone in many regions [1]. While crucial for ensuring food security, rice production faces a significant challenge, balancing high yields with environmental protection [2]. While massive amounts of synthetic fertilizers and pesticides have increased rice yields in the past [3], unfortunately, <15 % reaches the target organisms[4]. The remaining pesticides contaminate rice paddy environments, impacting surface water, soil, and groundwater, posing potential risks to ecosystems and human health [5, 6, 7].

Rice is the dominant crop in Cambodian agriculture and a rice-based farming system has existed in the country for more than 2,000 years [8]. Rice-based farming systems have

[1]* Corresponding author: Sereyvath Yoeun
E-mail: svath@itc.edu.kh; Tel: +855-10 348 588

been the backbone of Cambodia's agriculture, with a long history. Cambodian farmers have rich experience in rice production and have developed various rice farming systems such as rainfed lowland rice, rainfed upland rice, deep water rice, and irrigated dry season rice. Moreover, rice is a dominant crop for almost all farmers: more than 80 percent of Cambodian farmers grow rice [9].

Rice paddies, typically distributed around water sources, require flooded conditions during growth but are drained before harvest [10]. This unique environment influences the transport and transformation of substances, including pesticides. The agricultural management and the overuse of chemical fertilizers and pesticides are major contributors to soil erosion, land degradation, and declining soil health [11]. The increasing demand for agricultural products and changing regional climates further exacerbate this issue, leading to increased pesticide consumption. Therefore, the objective of this study was to assess pesticide contaminants in rice and water samples using qualitative and quantitative analysis with GC-MS.

2. METHODOLOGY

2.1 Samples collection

In this study, a total of 18 samples were collected in triplicate (n=3) from the Kanghot paddy field in Battambang province. There are three randomly selected sites: two sites in the paddy field (plots C1 and C2) and one on canal C (Figure 1). Sample collections were done with two cycles in the vegetative, reproductive, and ripening stages of rice growth. At the vegetative and reproductive stages, 12 water samples (1 liter) from plot C1, C2 and canal were collected. During the ripening stage, 4 paddy rice samples (300g) from plot C1 and plot C2 and 2 water samples of canal C were sampled (Table 1).



Fig. 1. water and rice samples collection area in Kanghot.

2.2 Samples preparation for rice and water

Paddy rice was de-husked and ground into fine particles with a dry-grinder machine, then rice powders were packed in plastic bags and stored in a refrigerator at -20ºC until extraction (Figure 2). One liter of water sample was adjusted with 1 mL of phosphate solution (1 mol/L, pH 7) and then pre-filtered to remove suspended solids through a fiberglass filter (Whatman GMF 2UM). The suspended solids on the surface of the filter were eluted with 10 mL of acetone to dissolve compounds that adsorbed on suspended solids and stored in a refrigerator at 4 °C for the next extraction (Figure 3).

Table 1. the dates of water and paddy rice samples in two cycles.

No.	n	Cycle	devStage	Site	codeUnik	Type	Amount	Sampling
1	1	1st	Vegetative	Blank C	K1CBlV1	Water	1L	Aug 5, 2024
2	3	1st	Vegetative	Canal C	K1CCnV1	Water	1L	Aug 5, 2024
3	3	1st	Vegetative	Plot C1	K1CMgV1	Water	1L	Aug 5, 2024
4	3	1st	Vegetative	Plot C2	K1CCtV1	Water	1L	Aug 5, 2024
5	1	1st	Reproductive	Blank C	K1CBlR1	Water	1L	Sep 14, 2024
6	3	1st	Reproductive	Canal C	K1CCnR1	Water	1L	Sep 14, 2024
7	3	1st	Reproductive	Plot C1	K1CMgR1	Water	1L	Sep 14, 2024
8	3	1st	Reproductive	Plot C2	K1CCtR1	Water	1L	Sep 14, 2024
9	1	1st	Ripening	Blank C	K1CBlG1	Water	1L	Oct 7, 2024
10	3	1st	Ripening	Canal C	K1CCnG1	Water	1L	Oct 7, 2024
11	3	1st	Ripening	Plot C1	K1CMgG1	Paddy rice	300g	Sep 28, 2024
12	3	1st	Ripening	Plot C2	K1CCtG1	Paddy rice	300g	Sep 28, 2024
13	1	2nd	Vegetative	Blank C	K2CBlV1	Water	1L	Nov 24, 2024
14	3	2nd	Vegetative	Canal C	K2CCnV1	Water	1L	Nov 24, 2024
15	3	2nd	Vegetative	Plot C1	K2CMgV1	Water	1L	Nov 24, 2024
16	3	2nd	Vegetative	Plot C2	K2CCtV1	Water	1L	Nov 24, 2024
17	1	2nd	Reproductive	Blank C	K2CBlR1	Water	1L	Dec 29, 2024
18	3	2nd	Reproductive	Canal C	K2CCnR1	Water	1L	Dec 29, 2024
19	3	2nd	Reproductive	Plot C1	K2CMgR1	Water	1L	Dec 29, 2024
20	3	2nd	Reproductive	Plot C2	K2CCtR1	Water	1L	Dec 29, 2024
21	1	2nd	Ripening	Blank C	K2CBlG1	Water	1L	Jan 28, 2025
22	3	2nd	Ripening	Canal C	K2CCnG1	Water	1L	Jan 28, 2025
23	3	2nd	Ripening	Plot C1	K2CMgG1	Paddy rice	300g	Jan 28, 2025
24	3	2nd	Ripening	Plot C2	K2CCtG1	Paddy rice	300g	Jan 28, 2025

2.3 Reagent and standard

The reagents that were used in this study included acetonitrile (HPLC grade, Germany), acetone (99.8%, HPLC grade, UK), n-hexane, and dichloromethane (HPLC grade, Korea). Deionized water and ultrapure water were purified using Puric-α ultrapure water technology. QuEChERS extraction pouches (EN method), which contain 4 g of magnesium sulfate, 1 g of sodium chloride, 1 g of sodium citrate, and 0.5 g of sodium hydrogen citrate sesquihydrate (Part No. 5982-5650), and dispersive SPE 15 mL (EN, fatty sample), which presents 149.9 mg of primary secondary amine (PSA), 149.9 mg of C18EC, and 900.2 mg of magnesium sulfate (MgSO4) (Agilent Technologies, United States). PSL3 and AC2. There are 203 external standards of pesticide compounds that exist in 9 groups with purity greater than 95%. (Resteck, United States). To determine the amount of concentration, these external standards were prepared in six various concentrations, such as 0.05, 0.1, 0.25, 0.5, 0.75, and 1 (mg.Kg-1). With these variable concentrations of the external

standards, the unknown concentration will be able to be quantify.

2.4 QuEChERS extraction for rice samples

a. Extraction steps

In the beginning, weigh 5 g of homogenized sample into 50 mL of QuEChERS centrifuge tube (EN method). 5 mL of ultra-pure water and 10 mL of acetonitrile were added to the centrifuge tube and vortexed for 1 min. QuEChERS's salt (Part No. 5982-5650), which includes sorbents such as 1 g sodium chloride, 4 g magnesium sulfate, 1 g sodium citrate, and 0.5 g sodium hydrogencitrate sesquihydrate ($C_6H_9NaO_8$), was added to the extraction tube. Immediately, handshake for 1 min and vortex for 1 min as well. Then the sample was centrifuged at 4000 rounds per minute at 4ºC for 5 min.

b. Clean-up steps

After centrifuging, 6 mL of supernatant was transferred into a 15 mL centrifuge tube (Part No. 5982-5156), which contains 149.9 mg of primary secondary amine (PSA), 149.9 mg of C18EC, and 900.2 mg of magnesium sulfate; then vortex for 1 min and centrifuge at 4000 rpm, 4ºC for 5 min. Lastly, 1 mL of the upper layer was passed through a 0.45 µm syringe filter and injected into a 1.5 mL GC-MS vial (**Figure 2**). Vegetable sample extractions were triplicated preparations for GC-MS analysis, wrapped with Parafilm, and stored in the refrigerator at -20ºC before injection.

2.5 Solid-phase extraction

After sample preparation, pre-filter samples were subjected to an extraction process by using solid-phase extraction method, which was developed by Jinya [12]. This method used two types of cartridges like Polymer Sorbent (PLS3) and Activated Carbon (AC2). First, cartridges were conditioned with 5 mL of dichloromethane followed by 5 mL of acetone then added 10 mL of purified water. Noted conditioning with each solvent was dropped 1 or 2 drops and hold for 2 min on the cartridges before conditioning with new solvent. Then, the high pressure solid-phase extraction instrument was set up. One liter of water sample passed through the cartridge at the speed of 15 mL/min. Later, the cartridge was dried under nitrogen gas stream for 40 min. After drying, PLS3 cartridge was eluted with 2 mL of acetone followed by 3 mL of dichloromethane whereas AC2 was eluted with 3 mL of acetone. Each elution process was followed by a holding time of 1 or 2 min on a cartridge to ensure well-eluted solvent from the cartridge. Solutions passed through PLS3 and AC2 were combined and concentrated to a volume of approximately 1 mL by streaming nitrogen gas. Adding 10 mL of n-hexane into 1 mL concentrated solution and then dehydrated with sodium sulfate salt. After that, concentrated tube was washed with 1 mL hexane five times and pipetted of each washing to pass through 8 g of sodium sulfate salt. The combined eluted was then concentrated again under gentle nitrogen stream until less than 1 mL volume (Figure.3).

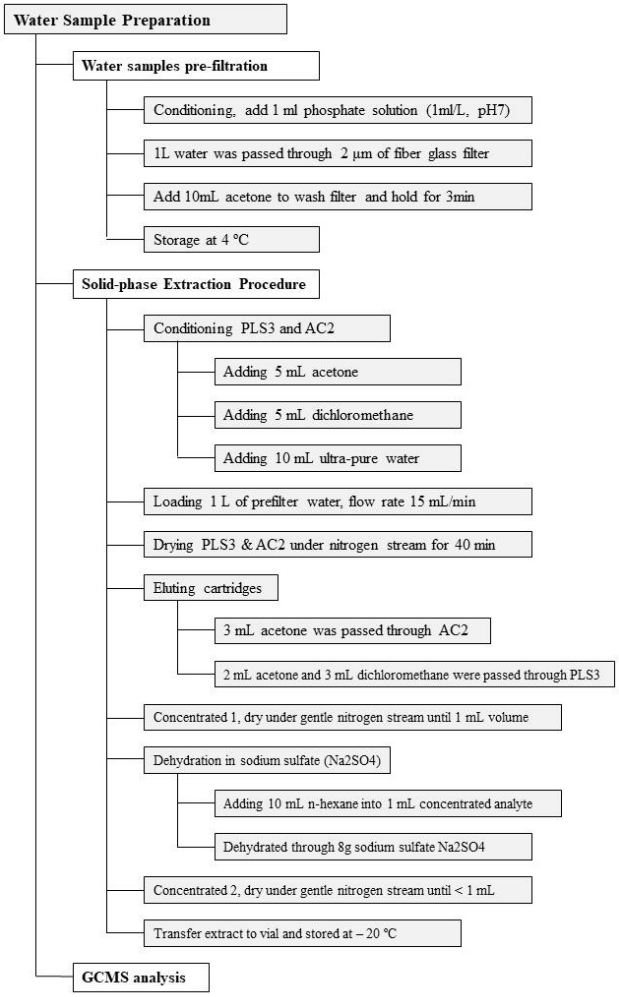


Fig. 3. Solid-phase extraction procedure for water samples.

2.7 Gas chromatography mass spectrometry analysis (GC-MS)

The GC-MS was operated with an DB-5ms capillary column (30 m x 0.25 mm, thickness 0.25 µm) and helium with purity ≥ 99.999 % was selected as the carrier gas (at a total flow of 50 mL/min while the column flow was 1.23 mL/min). The oven temperature was set to be 40 °C for 1 minute, then increased to 310 °C at a rate of 8 °C/min for 4 minutes. The injector and interface temperatures were 250 and 270 degrees Celsius, respectively. The injected mode was splitless mode and the injection volume was 1 µL. The EI for MS was 70 eV,

while the ion source and interface temperatures were 200 °C and 300 °C, respectively. The sample was assessed in full scan mode in the m/z 45-600 range to confirm the pesticide compounds found in the spectral library search. The identification of analysts was done by comparing the retention time and mass spectra, predicting the structural formula of compounds, and comparing them using the MS database. Exactly, all sample matrices and external reference standards in this investigation were analyzed using this GC-MS technique.

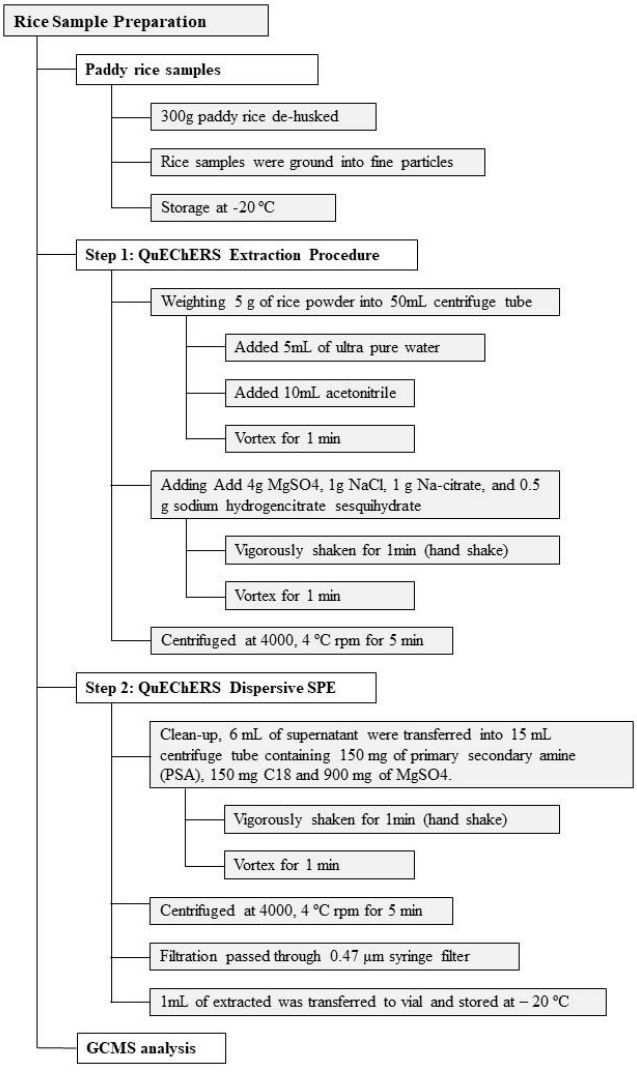


Fig. 3. QuEChERS extraction procedure for rice samples.

3. RESULTS AND DISCUSSION

3.1. Method validation

3.1.1. Limit of detection (LOD) and limit of quantification (LOQ)

To evaluate the capability of analytical instrument and method, the limit of detection (LODs) and limit of quantification (LOQs) were measured by manually injection in GC-MS at the spiking concentration range 0.0025 – 1 µg. L^{-1}. The LOD and LOQ were calculated by considering a value 3 and 10 of signal-to-noise (LOD = 3 S/N and LOQ = 10 S/N), respectively. LOD and LOQ study was prepared and randomly injected in triplication (n = 3) for each matrix and equipment. The LOD and LOQ values in rice matrix were 0.01 – 0.05 µg.g-1 and 0.0333 – 0.167 µg.g^1 for 7 reference pesticide standards. The LODs and LOQs determined in water matrix were 0.005 – 0.033 µg.L^{-1} and 0.0017 – 0.111 µg.L^{-1} for 8 reference standards (Table 2).

Table 2. Limit of detection (LODs) and limit of quantification (LOQs) of 8 reference standard in rice and water matrixes.

No.	Compounds	Rice		Water	
		LOD (µg.g^{-1})	LOQ (µg.g^{-1})	LOD (µg.L^{-1})	LOQ (µg.L^{-1})
1	Azoxystrobin*	0.02	0.067	0.033	0.111
2	Buprofezin*	0.05	0.167	0.025	0.083
3	Chlorfenapyr*	0.01	0.033	0.005	0.017
4	Chlorpyrifos*	0.02	0.067	0.005	0.017
5	Fenoxanil	-	-	-	-
6	Fipronil*	0.015	0.05	0.003	0.008
7	Hexaconazole*	0.01	0.033	0.01	0.033
8	Isoprothiolane	-	-	-	-
9	Paclobutrazol*	0.01	0.033	0.01	0.033
10	Pyraclostrobin	-	-	-	-
11	Tebuconazole*	-	-	0.017	0.05

Noted: "-": not available; "*": reference standards available

3.1.2. Recovery study

The accuracy (recoveries in the range of 70–120%) and precision (relative standard deviation [RSD] < 20%) of the extraction methods were evaluated in triplicate (n = 3) through recovery experiments at a concentration of 1 µg·g^{-1} for rice and 1 µg·L^{-1} for water. The QuEChERS extraction method for paddy rice demonstrated satisfactory accuracy for all reference pesticides, falling within the acceptable range. For the water samples, the recovery yields of buprofezin and chlorpyrifos using the SPE extraction method were below 70%. In terms of precision, most of the reference pesticide compounds showed RSD values below 20% in both rice and water matrices. However, for chlorpyrifos in the water matrix, the RSD exceeded 20% (Table 3).

Table 3. Limit of detection (LODs) and limit of quantification (LOQs) of 8 reference standard in rice and water matrixes.

No.	Componds	Rice		Water	
		Recovery rate %	RSD %	Recovery rate %	RSD %
1	Azoxystrobin*	95.90	11.48	109.36	4.31
2	Buprofezin*	93.86	8.29	38.23	11.15
3	Chlorfenapyr*	88.97	1.11	67.20	7.73
4	Chlorpyrifos*	93.39	6.84	43.48	20.85
5	Fenoxanil	-	-	-	-
6	Fipronil*	74.18	0.62	105.76	5.49
7	Hexaconazole*	91.69	7.16	111.67	3.72
8	Isoprothiolane	-	-	-	-
9	Paclobutrazol*	80.72	2.20	116.22	7.83
10	Pyraclostrobin	-	-	-	-
11	Tebuconazole*	77.82	0.45	117.15	4.35

Noted: "-": not available, "*": reference standards available

3.2. Samples analysis

The assessment of pesticide residues in rice and water were sampled from three points (plots C1, C2 and canal C) that located in Kanghot paddy filed. Sample collections were done with two cycles in the vegetative, reproductive, and ripening stages of rice growth. For the water and rice samples were extracted by following solid-phase extraction and QuEChERS method, respectively. After that, the extracted rice and water samples were analyzed on GC-MS, which is capable to screen 451 pesticide compounds (compared with the instrumental database and the NIST library version 14) and quantify the concentration with 203 target external standard solutions [13]. According to the results were demonstrated that four insecticides, six fungicides, and one plant growth retardants (PGR) [14] active substances were found in paddy field water in two cycles. Moreover, the hazardous class of those active pesticide substances according to World Health Organization (WHO) were also included [14] (Table 4). Insecticides were presented pesticide active substances such as, Chlorfenapyr is used to kill insects that are resistant to organophosphorus insecticides. It works by disrupting the insect's ability to produce energy [15]. It was found in four samples: K2CMgR1, K1CCtR1, K2CCtV1, and K2CCtR1, with concentrations of 0.1723, 0.1589, 0.3079, and 0.5024, respectively. Chlorpyrifos was found in one paddy field water (K1CMgR1), and the concentration was 0.7053 ($\mu g.L^{-1}$). Buprofezin is known as one of the most successful IGRs (Insect Growth Regulator) in the world for control of Homopterous insects such as planthoppers, whiteflies, mealybugs and scale insects, and is widely used in more than 80 countries [16]. Buprofezin was presented in two samples K1CMgR1and K2CMgV1 with concentrations of 0.0658 and 1.0734 ($\mu g.L^{-1}$), respectively. Fipronil is an effective phenylpyrazole insecticide commonly used in agriculture for pest control. In creature, fipronil could be metabolized to several more toxic metabolites. These compounds can contaminate or migrate into foodstuffs and potentially pose danger to human health [17]. There are two samples were found in two samples K1CCtV1 (5.7971 $\mu g.L^{-1}$) and K2CCtR1 (0.0504 $\mu g.L^{-1}$). For fungicide was found with Azoxystrobin is a very broad spectrum fungicide with systemic activity and both protectant and curative actions used on a broad range of crops including fruits, vegetables, small grains, and turf grass [18]. As the result, it was found once in K2CMgR1 (0.5265 $\mu g.L^{-1}$). Hexaconazole is a broad-spectrum systematic fungicide used to treat seed-borne and soil-borne diseases caused by *ascomycetes, basidiomycetes,* and *imperfect fungi*. It is typically applied as a foliar spray to control several diseases in cereal, vegetable, and fruit crops [19]. Sample code K2CMgV1, it was found 0.4081 $\mu g.L^{-1}$. Tebuconazole is a fungicide effective against various foliar diseases in cereals and other field crops [20]. Tebuconazole, developed in 1986 by Bayer, is a typical triazole fungicide and has been widely registered on cereals, vegetables, fruits, medicinal plants, and ornamental flowers because of its high efficiency, low toxicity, broad-spectrum, systemic characteristics, and plant growth regulating properties in Asia, Europe, and North and South America [15,16]. In this study, Tebuconazole was detected in two sample K2CCtR1 (0.6386 $\mu g.L^{-1}$) and K1CCtR1 (36.2787 $\mu g.L^{-1}$). Fenoxanil is a new and protective fungicide with residual effects, belonging to the propionamide chemical -class. Fenoxanil is melanin biosynthesis inhibitor used to control of rice blast caused by the fungus *Pyricularia oryzae*. It was discovered by Shell GmbH, developed jointly by BASF and Nihon Nohyaku Co. Ltd. and launched in 2000 [22]. This compound was presented in one sample K1CMgV1. Isoprothiolane Blast caused by *Pyricularia grisea Sacc*. is one of the major diseases of rice. In situations of sudden occurrence of the disease, use of fungicides is the only effective solution for the immediate suppression of the disease. An attempt was made to test the efficacy of Isoprothiolane for the control of blast [22]. According to the result, it was found in five paddy field waters, such as K1CMgR1, K2CMgR1, K1CCtR1, K2CCtV1, and K2CCtR1. Pyraclostrobin belongs to strobilurin class of agricultural fungicides [23], used around the world, with a preventive mode of action [24]. As a fungicide family, the strobilurins all have a common mode of action. Strobilurin fungicide pyraclostrobin (F-500) initially developed to control fungal pests showed growth promoting and yield enhancing qualities under field conditions [25]. It was detected once in K1CCtR1. For the last three pesticide active substances above, we were not able to quantify the concentration due to an unavailable external standard. Paclobutrazol, a member of triazole plant growth regulator group, is used widely in agriculture. It is a cell elongation and internode extension inhibitor that retards plant growth by inhibition of gibberellins biosynthesis [26]. In this study, paclobutrazole was commonly found in ten samples:

Table 4. qualitative and quantitative analysis of rice and paddy field water samples

No.	n	Sample code	Type	Qualitative and quantitative analysis of water (µg.L^{-1}) and rice (mg.Kg^{-1})										
				Fungicide							Insecticide			PGR
				Fenoxanil (-)	Azoxystrobin* (U)	Hexaconazole* (III)	Isoprothiolane (II)	Pyraclostrobin (-)	Tebuconazole* (II)	Chlorfenapyr* (II)	Chlorpyrifos* (II)	Buprofezin* (III)	Fipronil* (II)	Paclobutrazol* (II)
1	3	K1CMgV1	Water	-	-	-	-	-	-	-	-	-	-	27.8239
2	3	K1CMgR1	Water	+	-	-	+	-	-	-	0.7053	1.0734	-	4.7965
3	3	K1CMgG1	Rice	-	-	-	-	-	-	-	0.6612	-	-	-
4	3	K2CMgV1	Water	-	-	0.4081	-	-	-	-	-	0.0658	-	1.2012
5	3	K2CMgR1	Water	-	0.5265	-	+	-	-	0.1723	-	-	-	0.7986
6	3	K2CMgG1	Rice	-	-	-	-	-	-	-	-	-	-	-
7	3	K1CCtV1	Water	-	-	-	-	-	-	-	-	-	5.7971	2.1706
8	3	K1CCtR1	Water				+	+	36.2787	0.1589	-	-	-	2.4213
9	3	K1CCtG1	Rice	-	-	-	-	-	-	-	-	-	-	-
10	3	K2CCtV1	Water	-	-	-	+	-	-	0.3079	-	-		+
11	3	K2CCtR1	Water	-	-	-	+	-	0.6386	0.5024	-	-	0.0504	3.0938
12	3	K2CCtG1	Rice	-	-	-	-	-	-	-	-	-	-	-
13	3	K1CCnV1	Water	-	-	-	-	-	-	-	-	-	-	-
14	3	K1CCnR1	Water	-	-	-	-	-	-	-	-	-	-	0.1196
15	3	K1CCnG1	Water	-	-	-	-	-	-	-	-	-	-	0.1852
16	3	K2CCnV1	Water	-	-	-	-	-	-	-	-	-	-	-
17	3	K2CCnR1	Water	-	-	-	-	-	-	-	-	-	-	-
18	3	K2CCnG1	Water	-	-	-	-	-	-	-	-	-	-	-

*Noted: "-": Not Detected; +: Detected; *: Available for quantify; II =Moderately hazardous, III = Slightly hazardous; U = Unlikely to present acute hazard in normal use*

K1CMgV1 (27.8239), K1CMgR1 (4.7965), K2CMgV1 (1.2012), K2CMgR1 (0.7986), K1CCtV1 (2.1706), K1CCtR1 (2.4213), K2CCtV1 (No data), K2CCtR1 (3.0938), K1CCnR1 (0.1196), and K1CCnG1 (0.1852) (µg.L^{-1}). At the same time, canal C is supplies water for rice growth irrigate for plots C1 and C2, as the results were shown that only Pacloburazol was found in two samples (K1CCnR1 and K1CCnG1) and the average concentration is 0.1524 µg.L^{-1}. Pacloburazol is moderately hazard by WHO class II [14].

Usage of chlorpyrifos as an insecticide on various crops is very common throughout the world [14, 15]. It was detected in only one of rice sample (K1CMgG1) in first cycle in concentration 0.6612 mg.Kg^{-1} exceeding the MRL value of chlorpyrifos in rice (0.01 mg.Kg^{-1}) showing its potential to pose significant health risk to the consumers [29]. Chlorpyrifos itself is moderately hazard by WHO class II [14], but when it enters in the body via food ingestion, the body breaks it down, and creates a toxic form called chlorpyrifos oxon which has the ability to bind permanently to enzymes which control the messages that travel between nerve cells and results in effecting muscles as muscles do not function correctly. The body then must make more enzymes so that normal nerve function can resume. The body can break down and excrete most of the unbound chlorpyrifos in feces and urine within a few days[30]. However, only one of the four rice samples presented the active substance chlorpyrifos, even though its concentration was moderately higher than MRLs.

4. CONCLUSIONS

These findings indicate that pesticides applied during rice cultivation by farmers can leave residues in rice. Despite its limited occurrence, the high level poses a potential health risk due to the toxic effects of its metabolite, chlorpyrifos oxon, emphasizing the need for strict monitoring and regulation of pesticide residues in food. From an environmental perspective, such elevated residues may indicate overuse or improper application of pesticides, which can contribute to soil and water contamination. The main sources of pesticide contamination appear to be direct application by paddy field owners for crop protection and productivity enhancement, as well as runoff from pesticides used in nearby fields sharing the same irrigation system.

ACKNOWLEDGMENTS

The authors would like to acknowledge Water Resources Management and Agro-ecology Transition for Cambodia

(WAT4CAM) for providing financial support for these research activities.

REFERENCES

[1] J. Harbinson *et al.*, "Designing the crops for the future; the cropbooster program," *Biology (Basel).*, vol. 10, no. 7, pp. 1–8, 2021, doi: 10.3390/biology10070690.

[2] P. Atukunda, W. B. Eide, K. R. Kardel, P. O. Iversen, and A. C. Westerberg, "Unlocking the potential for achievement of the UN Sustainable Development Goal 2 – 'Zero Hunger' – in Africa: targets, strategies, synergies and challenges," *Food Nutr. Res.*, vol. 65, p. 10.29219/fnr.v65.7686, 2021, doi: 10.29219/FNR.V65.7686.

[3] A. Aldas-Vargas, B. A. J. Poursat, and N. B. Sutton, "Potential and limitations for monitoring of pesticide biodegradation at trace concentrations in water and soil," *World J. Microbiol. Biotechnol.*, vol. 38, no. 12, pp. 1–18, 2022, doi: 10.1007/s11274-022-03426-x.

[4] X. Zhang, Y. Shen, X. Y. Yu, and X. J. Liu, "Dissipation of chlorpyrifos and residue analysis in rice, soil and water under paddy field conditions," *Ecotoxicol. Environ. Saf.*, vol. 78, pp. 276–280, Apr. 2012, doi: 10.1016/J.ECOENV.2011.11.036.

[5] N. K. Singh, G. Sanghvi, M. Yadav, H. Padhiyar, J. Christian, and V. Singh, "Fate of pesticides in agricultural runoff treatment systems: Occurrence, impacts and technological progress," *Environ. Res.*, vol. 237, p. 117100, Nov. 2023, doi: 10.1016/J.ENVRES.2023.117100.

[6] J. P. van de Merwe, P. A. Neale, S. D. Melvin, and F. D. L. Leusch, "In vitro bioassays reveal that additives are significant contributors to the toxicity of commercial household pesticides," *Aquat. Toxicol.*, vol. 199, pp. 263–268, Jun. 2018, doi: 10.1016/J.AQUATOX.2018.03.033.

[7] G. Verasoundarapandian *et al.*, "Remediation of Pesticides by Microalgae as Feasible Approach in Agriculture: Bibliometric Strategies," *Agron. 2022, Vol. 12, Page 117*, vol. 12, no. 1, p. 117, Jan. 2022, doi: 10.3390/AGRONOMY12010117.

[8] B. Yu and X. Diao, "Cambodia's Agricultural Strategy: Future Development Options for the Rice Sector," *Policy Discuss. Pap.*, no. March, p. 26, 2011.

[9] "Rice Cultivation in Cambodia: A Comprehensive Guide to its History, Culture, and Economic Impact – Aseanz: We build strong relationships between New Zealand and Southeast Asia." Accessed: Apr. 24, 2025. [Online]. Available: https://aseanz.net/2023/09/rice-cultivation-in-cambodia-a-comprehensive-guide-to-its-history-culture-and-economic-impact/

[10] M. Shahid *et al.*, "Climate resilient rice production system: Natural resources management approach," *Oryza-An Int. J. Rice*, vol. 58, no. Special, pp. 143–167, 2021, doi: 10.35709/ory.2021.58.spl.6.

[11] P. Sharma, P. Sharma, and N. Thakur, "Sustainable farming practices and soil health: a pathway to achieving SDGs and future prospects," *Discov. Sustain.*, vol. 5, no. 1, pp. 1–14, Dec. 2024, doi: 10.1007/S43621-024-00447-4/FIGURES/2.

[12] D. Jinya, "A b s tra c t :," *Shimadzu Corp.*, vol. 2, no. 2, pp. 1–5, 2013.

[13] "GC Multiresidue Pesticide Standards Kit." Accessed: Apr. 25, 2025. [Online]. Available: https://www.restek.com/global/en/p/32562

[14] WHO, *The WHO Recommended Classification of Pesticides by Hazard and Guidelines to Classification*. 2020.

[15] "Chlorfenapyr – a new approach." Accessed: May 02, 2025. [Online]. Available: https://agriculture.basf.com/global/en/business-areas/public-health/our-partnerships/stories_partnership_countries/stories/chlorfenapyr-new-approach

[16] "Buprofezin | IGR | Insect Growth Regulator | Insect Control | Insecticde." Accessed: May 02, 2025. [Online]. Available: https://www.nichino-europe.com/products/buprofezin/

[17] X. Li, W. Ma, H. Li, Q. Zhang, and Z. Ma, "Determination of residual fipronil and its metabolites in food samples: A review," *Trends Food Sci. Technol.*, vol. 97, no. October 2019, pp. 185–195, 2020, doi: 10.1016/j.tifs.2020.01.018.

[18] R. HOLLINGWORTH, "Inhibitors and Uncouplers of Mitochondrial Oxidative Phosphorylation," *Handb. Pestic. Toxicol.*, pp. 1169–1261, 2001, doi: 10.1016/b978-012426260-7/50060-4.

[19] "Hexaconazole — toxicity, side effects, diseases and environmental impacts." Accessed: May 01, 2025. [Online]. Available: https://naturalpedia.com/hexaconazole-toxicity-side-effects-diseases-and-environmental-impacts.html

[20] "Tebuconazole (Ref: HWG 1608)." Accessed: May 01, 2025. [Online]. Available: https://sitem.herts.ac.uk/aeru/ppdb/en/Reports/610.htm

[21] "Substance Information - ECHA." Accessed: May 01, 2025. [Online]. Available: https://echa.europa.eu/it/substance-information/-/substanceinfo/100.060.231

[22] M. Brycht *et al.*, "Ultra trace level determination of fenoxanil by highly sensitive square wave adsorptive stripping voltammetry in real samples with a renewable silver amalgam film electrode," *J. Electroanal. Chem.*, vol. 738, pp. 69–76, Feb. 2015, doi: 10.1016/J.JELECHEM.2014.11.029.

[23] D. W. Bartlett, J. M. Clough, J. R. Godwin, A. A. Hall, M. Hamer, and B. Parr-Dobrzanski, "The strobilurin fungicides," *Pest Manag. Sci.*, vol. 58, no. 7, pp. 649–

662, 2002, doi: 10.1002/ps.520.
[24] K. Grossmann, J. Kwiatkowski, and Gü. Caspar, “Regulation of phytohormone levels, leaf senescence and transpiration by the strobilurin kresoxim-methyl in wheat (Triticum aestivum),” *J. Plant Physiol.*, vol. 154, no. 5–6, pp. 805–808, 1999, doi: 10.1016/S0176-1617(99)80262-4.
[25] M. Kanungo and J. Joshi, “Impact of Pyraclostrobin (F-500) on Crop Plants,” *Plant Sci. Today*, vol. 1, no. 3, pp. 174–178, 2014, doi: 10.14719/pst.2014.1.3.60.
[26] B. Desta and G. Amare, “Paclobutrazol as a plant growth regulator,” *Chem. Biol. Technol. Agric.*, vol. 8, no. 1, pp. 1–15, 2021, doi: 10.1186/s40538-020-00199-z.
[27] M. H. W. M. Cardoso, A. W. de Nóbrega, H. de C. Vital, and S. de M. P. Abrantes, “Preparação de um material de referência certificado para controle de agrotóxicos em hortifrutigranjeiros: Estudo da estabilidade,” *Cienc. e Tecnol. Aliment.*, vol. 30, no. 2, pp. 439–446, 2010, doi: 10.1590/S0101-20612010000200022.
[28] R. Li *et al.*, “Chlorpyrifos residue levels on field crops (rice, maize and soybean) in China and their dietary risks to consumers,” *Food Control*, vol. 51, pp. 212–217, 2015, doi: 10.1016/j.foodcont.2014.11.023.
[29] EU Pesticides Database (v3.3), “EU Pesticides Database - Products - Product details,” European Commission. Accessed: Apr. 28, 2025. [Online]. Available: https://ec.europa.eu/food/plant/pesticides/eu-pesticides-database/start/screen/products/details/237
[30] A. Hamid, G. Yaqub, M. Ayub, and M. Naeem, “Determination of malathion, chlorpyrifos, λ-cyhalothrin and arsenic in rice,” *Food Sci. Technol.*, vol. 41, no. 2, pp. 461–466, Oct. 2020, doi: 10.1590/FST.01020.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessment of Pesticide Residues in Rice and Irrigation Canal: A Study of Agrochemical-Free Farming in Rovieng District, Preah Vihear Province

Rethireach Ly [1], Sereyvath Yoeun [1,2], Leanghok Bun [2], Mathilde Sester [2,3], Florent Tivet [3], Rama Sry [4], Chanvorleak Phat [1,2*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[4] *Cambodian Conservation Agriculture Research for Development Center (CARDEC)*

[*] *Corresponding author: phatchanvorleak@itc.edu.kh*

Abstract: *Pollution by pesticides in agricultural ecosystems raises environmental risks and human health, which requires rigorous monitoring, even in agrochemical-free farming systems. This study analyzed pesticide residues in rice, rice field water, and canal water from two rice fields in Preah Vihear Province, Cambodia: one practicing organic (agrochemical-free) farming and the other employing reduced agrochemical use with a single herbicide application and cover crop management. Each sample was sampled triplicated, where the sample of rice was sampled mainly in the rainy season, and the water sample collection was based on the three stages of the rice growth cycle: early growth, mid growth, and late growth stages. A total of 32 samples were collected, comprising 26 water samples and 6 samples of rice, systematically from Demo 1, Demo 2, and the canal. Rice samples were analyzed using a modified QuEChERS extraction method, while water samples underwent solid phase extraction (SPE) before gas chromatography mass spectrometry analysis (GC-MS). A mixed standard solution of 27 target pesticide compounds was used for method validation and quantification of pesticide residues. The validation of the method confirmed reliability, since recovery studies met the standard analytical criteria. The results revealed that all the samples analyzed of rice and water did not exhibit detectable pesticide residues, indicating the effectiveness of these agricultural practices in preventing pesticide pollution. These findings highlight the success of the studied system in minimizing contamination risks and chemical exposure, thus ensuring food security and environmental protection. While reinforcing the sustainable potential of agrochemical-free rice farming, continuous monitoring is recommended to mitigate potential external contamination from surrounding activities. This study contributes valuable insights into pesticide contamination in sustainable agriculture, supporting policy development for agrochemical-free practices and strengthening environmentally friendly strategies.*

Keywords: Pesticide residues; Agrochemical-free; Rice crops; Canal system; SPE; QuEChERS; GC-MS

1. INTRODUCTION[1]

Pesticide application is an important factor in modern agriculture to improve crop protection and production yield. However, those agrochemical products contain toxic ingredients and are harmful to humans. Thus, pesticide residue in food causes a serious problem for consumers' health and pollutes the local environment, like air, soil, and water sources[1]. Cambodia's rice farming relies heavily on pesticides to protect the crops from pests and diseases. However, the increasing reliance on agrochemicals has led to a 61% annual rise in pesticide imports, causing concerns about pesticide residues, environmental damage, and health risks [2]. Many farmers lack proper training in pesticide management, resulting in excessive application and chemical runoff into local ecosystems. Pest resistance and soil degradation are also pushing production costs higher, making conventional farming less sustainable.

Pesticide residues can cause chronic toxicity issues, neurological problems, and endocrine disruption. Careless application of these chemicals results in environmental pollution, negatively impacting biodiversity and water quality

[*] Corresponding author: Chanvorleak Phat
E-mail: phatchanvorleak@itc.edu.kh; Tel: +855-93-418-800

[3]. The Agrochemical-free rice farming has emerged as a sustainable alternative, replacing synthetic pesticides with natural solutions, biological pest control, organic soil amendments, and ecological farming practices [4]. This method enhances food safety, soil fertility, reduces environmental impact, and meets the increasing demand for organic rice. The Rovieng District in Preah Vihear Province serves as a prime example of agrochemical-free farming, promoting a shift towards sustainable rice production.

2. METHODOLOGY

2.1. Study area

The agrochemical-free paddy field experiment was surrounded by paddy cultivated lands that were investigated in the Water Resources Management and Agro-ecology Transition for Cambodia (WAT4CAM) program with two locations WAT4CAM-CS/04-01 with (13°25'00"N latitude and 105°08'41"E longitude) and the WAT4CAM-CS/04-02 with (13°24'32"N latitude and 105°08'59"E longitude), Preah Vihear province, Cambodia (Fig. 1 (a)). The agrochemical-free paddy field has a total area of 43.1ha, but the study was conducted for 9.5ha (Demo 1 for 4.6ha (c) and Demo 2 (b) for 4.9ha) located in Rikreay commune, Rovieng District, Preah Vihear Province, Cambodia. The irrigated water (total area of 13.01 km) is an earthen dug canal from Okombor Reservoir, which was divided into two main canals, main canal I and main canal II, which supply both Demo and Cherng K'ek canal is a continuous earthen dug canals from main canal II that that distributes and retains water for year-round farming activities. For Demo 1, the farmers practice organic farming, and Demo 2 farmers use a minimum amount of chemicals, restricted to a single application of herbicides during the growth stage alongside cover crop management.

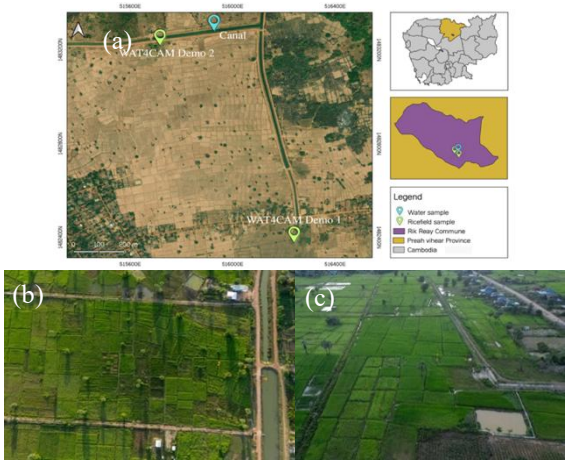


Fig. 1. Mapping of rice and water sample collection: (a) all sample collection; (b) sampling points Demo 2; (c) sampling points Demo 1

2.2. Sample collection

Water samples from paddy fields and irrigation canals were collected in triplicate based on the three stages of the rice growth cycle: early growth, mid growth, and late growth stages in 1 L plastic bottles and additional 1 L plastic bottles containing lab-sourced deionized (DI) water were used to obtain a blank water sample for quality control. All samples were immediately stored in the ice box before being transported to the Laboratory at the Institute of Technology of Cambodia. Rice grains were also collected during harvest season using a technique within the shape of a paddy field, delivered to trampled rice underfoot, and stored in a zip lock before being transported to the laboratory.

2.3. Preparation of standard solution

Twenty-seven references of pesticide (Table 1) were chosen based on their relevance to environmentally induced health effect and their agricultural uses. Twelve analytes were identified as organophosphates and organochlorines due to their resistance to degradation and environmental persistence. These include methyl parathion, mevinphos, malathion, aldrin, dieldrin, heptachlor, γ-hexachlorocyclohexane, isazofos, o,p'-DDT, chloroneb, and fipronil. These pesticides highlight the ecological impact of persistent organic pollutants (POPs). Seven synthetic insecticides of the pyrethroid class are widely used due to their effectiveness and low toxicity to mammals. Eight systemic fungicides, herbicides, and insecticides are of increasing environmental concern. These compounds highlight the need for integrated monitoring systems that incorporate both legacy and contemporary agrochemicals. The compounds Propanil, fipronil, chlorothalonil, and paclobutrazol were dissolved in acetone, while the remaining compounds were dissolved in hexane. A mixed working standard solution containing all 27 chemicals was prepared at 100 ppm in hexane to ensure solvent homogeneity.

Table 1. List of target pesticide standards and their toxicity

Classification	Compounds	Status In Cambodia	Toxicity (WHO)
Insecticide	o,p'-DDT	Banned	I
	Chlorothalonil	--	--
	Heptachlor	Banned	O
	Aldrin	Banned	O
	Dieldrin	Banned	O
	Methyl parathion	Banned	Ia
	Fipronil	--	II
	Mevinphos	Banned	Ia
	Isazofos	Restricted	Ib

	Chlorpyrifos	--	II
	Bifenthrin	Banned	NL
	Lindane	Banned	II
	Chlorfenapyr	--	--
	Malathion	Permitted	III
	Permethrin 1	Banned	Ia
	Biphenyl	--	--
	Cypermethrin (1,2,3&4)	--	II
	Metalaxyl	--	II
Fungicide	Chloroneb	--	O
	Paclobutrazol	--	II
	Pretilachlor	--	U
	Triadimefon	Permitted	U
Herbicide	Propanil	--	II
	Metolachlor	Permitted	III

--: No available information; II: Moderately hazardous; III: Slightly hazardous; Ia: Extremely hazardous; Ib: Highly hazardous; NL: Not listed; O: Obsolete or discontinued for use as pesticides; U: Unlikely to present acute hazard

2.4. Method validation

Method validation in this study was performed based only on accuracy evaluation, given as recovery percentage (%), limit of detection (LOD) and limit of quantification (LOQ). In accordance with extraction procedures unique to each matrix, validation was carried out separately for rice grain, paddy field water, and irrigation canal water matrices.

For water samples, method validation was carried out using solid-phase extraction (SPE) with two different post-filtration rinsing strategies. In Recovery A, the filters were washed with acetone, while Recovery B employed a mixture of acetone and dichloromethane to wash the filters. Blank irrigation water was then spiked with a blended standard solution 10 ml within a concentration of 100 ppm, until it hit roughly 1 ppm, followed by the extraction method described in the extraction section, along with the GC-MS analysis. Extraction efficiency was simply gauged by checking the percentage of each compound that bounced back. The comparison between the two methods came down not only to the average recovery per method but also to tallying up how many compounds landed within the target range of 70–120%. In most cases, the better recovery procedure was approached for further investigation.

For Rice samples, method validation was performed through the QuEChERS extraction to test out the overall approach [5]. In recovery, blank rice samples were spiked at a concentration of 100 ppm with 100µL of external standard to get expected levels of concentration 1 ppm, followed by a one-hour period before extraction to simulate realistic analyte absorption. Next, the spiked rice sample was processed with the QuEChERS, followed by the extraction method described in the extraction section, before being run through GC-MS analysis, and compounds that showed recoveries in the 70%–120% target range were deemed acceptable and helped guide further interpretation of the pesticide residue levels. The recovery percentage (%), known as the ratio of the amount of pesticide recovered after analysis to the amount initially spiked into a sample and expressed as a percentage (%), was calculated by following Eq. 1. LOD is the lowest concentration of a substance in the test sample that the analytical tool can consistently identify. The lowest analyte concentration in a sample that can be reliably measured is known as the limit of quantitation (LOQ). The LOD and LOQ tests were performed in three duplicates. Equations 2 and 3 were used to calculate the LOD and LOQ.

$$\text{Recovery (\%)} = ((P_S - P_B) \times 100)/P_C \quad \text{(Eq.1)}$$

$$\text{LOD} = 3 \times S/N \quad \text{(Eq.2)}$$

$$\text{LOQ} = 10 \times S/N \quad \text{(Eq.3)}$$

where:

- P_S= Peak area of spiking sample
- P_B= Peak area of blank sample
- P_C= Peak area of control sample
- S/N= Signal-to-Noiseratio

2.5. Water sample preparation

Water sample preparation was adapted from Daisuke Jinya (2013). Water samples were pre-filtration before conducting solid phase extraction (SPE), which employed 1 mL of phosphate buffer solution (1 mol/L, pH 7.0) to modify 1L of surface water sample, and then filtered by glass microfiber filters (diameter 47mm) were used to filter out the suspended particles from water samples [6]. To dissolve possibly adsorbed chemicals, 10 mL of acetone was used to elute the suspended particles that the filters had retained. The filtered samples were stored in a refrigerator at 4°C until through an extraction procedure.

2.6. Rice sample preparation

Rice sample preparation was adopted from Agilent Technologies and L.Sun Thermo Scientific (2020) [5][7]. Rice samples were husked by a husker rice machine (NQEK, P-1), which is capable of separating the husk effectively while maintaining the rice grain quality. The husked rice samples were stored in a zip lock bag before transport to laboratory. Before supervising the QuEChERS (Quick, Easy, Cheap, Effective, Rugged and Safe) extraction, the husked rice was ground by a grinder machine to ensuring a homogenous sample and prevent separation of the endosperm and husk, which can lead to inaccurate results also helps in extracting

the desired analytes by increasing the surface area for solvent penetration.

2.7. Solid-phase extraction (SPE)

According to Daisuke Jinya (2013), a 1-liter water sample was extracted by solid phase extraction [5]. This approach utilizes two cartridge types: polymer sorbent (PLS3) and activated carbon (AC2). 5 mL of dichloromethane, 5 mL of acetone, and 10 mL of ultrapure water were used to condition the aforementioned cartridges. The cartridges were then dried for 40 minutes under a nitrogen stream. Following drying, 2 mL of acetone and 5 mL of dichloromethane were used to elute the PLS3 cartridge, while 5 mL of acetone was used to elute the AC2. Under a nitrogen stream, the eluted solution from the PLS3 and AC2 cartridges was mixed and concentrated to a volume of about 1 mL [5]. A sodium sulfate salt filter was then glass pipetted with the hexane, and this combined eluted solution was mixed with washing hexane to create a volume of 1 mL after being concentrated once more with streaming nitrogen gas to a volume of less than 1 mL, and the extract solution was ready for GC-MS Analysis.

2.8. QuEChERS extraction

For the paddy grain, samples weighing 5 g were placed in a 50 mL centrifuge tube, followed by adding 10 mL of acetonitrile and 5 mL of ultra-pure water, and vortexed for 1 minute, respectively. By the EN QuEChERS method, Agilent QuEChERS Extraction Salt Packets were added, and vortexed for 1 minute, then centrifuged for 5 minutes at 4000 rpm with a temperature of 4°C. Following centrifugation, 6 mL of the supernatant was added to a 15 mL QuEChERS dispersive kit centrifuge tube, then centrifuged for 5 minutes at 4000 rpm, 4°C, after being vigorously shaken by vortex for one minute [7]. The 1 mL of extracted sample was transferred to a vial using a nylon syringe filter before GC/MS analysis.

2.9. GC-MS analysis

The pesticide compounds were analyzed by using gas chromatography and mass spectrometry (GC-MS), which was conducted on a model Shimadzu GC-MS-TQ8040 series from Japan equipped with an AOC-20S autosampler, and were obtained using the following expression analytical conditions [8]. The column used was DB-5ms with a length of 30 m, a thickness of 0.25 diameter of 0.25 mm. One microliter of the sample was injected into the GC-MS by an auto injector using splitless mode. After that, the column oven temperature was programmed from 40 to 310°C by holding for 2 minutes at 40°C, increasing the temperature to 310°C at a rate of 8°C/min, and then holding for 5 minutes. The carrier gas was ultra-pure helium at a total flow of 50 mL/min, while the column flow was 1.23 mL/min. The ion source temperature was 200°C, and the temperature of the interface was 300°C. All pesticides were identified by retention time and specific ions, and quantified by the external standard method.

3. RESULTS AND DISCUSSION

3.1. Method validation

A recovery test on the water sample was used to evaluate the solid phase extraction technique with two solvent systems, which included Recovery A, where the filters were washed with acetone, and Recovery B employed with a mixture of acetone and dichloromethane to wash the filters. According to SANTE/11312/2021, the international standard for precision and accuracy should be in ranges of 70-120%, which is an acceptable recovery range [9]. Recovery A, which was washed with Acetone only, performed slightly better than Recovery B, which was washed filter with acetone and dichloromethane. Recovery A had 11 out of 27 compounds within the acceptable recovery range of 70%-120%, while Recovery B had 10 compounds within the target range (Table 2). Recovery A showed fewer overestimated compounds and more consistent values. The overall performance suggests that using only Acetone for filter washing may provide more accurate recovery in pesticide residue analysis. Recovery A, utilizing acetone as a single solvent, yielded equivalent to Recovery B, despite its simpler procedure and reduced solvent complexity. The polarity of acetone probably played the key role in efficiently extracting of moderately polar compounds while reducing matrix interference and variability [10]. This is consistent with the findings from earlier research that highlights the effectiveness of acetone in multi-residue for pesticide analysis. The recovery of target compounds from a rice matrix, using the QuEChERS extraction method, was determined to validate the method's efficiency. Largely as set out by the SANTE/11312/2021 guidelines, out of 27 target compounds in the validation method, 17 compounds (Table 2), roughly 62.96%, were successfully recovered within the acceptable recovery of 70–120% [10]. The remained compounds were nearly the target range, and other compounds tended to stray above or below the target range, which generally can be caused by the significant matrix effect, fast degradation, or poorly retained in GC-MS [11]. The limit of detection (LOD) and limit of quantification (LOQ) were assessed to gauge how sensitive the GC-MS method is for analyzing pesticide residues in rice and water samples. According to SANTE/11312/2021, the majority of compounds had LOQs that were either below or near the legal limit of 0.01 µg/kg, as can be seen from the data in (Table 2). The majority of LOQs in this investigation fell within or around this range, indicating that the approach has sufficient sensitivity for identifying pesticide residues in both matrices.

Table 2. Recovery rates, limit of detection (LOD) and limit of quantification (LOQ) of target pesticide in water and rice

No	Pesticide	Recovery Percentage %		LOD		LOQ	
		Rice	Water	Rice (µg/kg)	Water (µg/L)	Rice (µg/kg)	Water (µg/L)
1	o,p'-DDT	82.60	11.09	0,0075	0,0025	0,015	0,0083
2	Chlorothalonil	0	68.87	0,0150	0,0583	0,033	0,1943
3	Metalaxyl	96.88	99.99	0,025	0,0333	0,050	0,1110
4	Heptachlor	82.27	9.58	0,025	0,0050	0,050	0,0167
5	Aldrin	83.66	11.34	0,025	0,0050	0,050	0,0167
6	Dieldrin	89.58	19.78	0,0075	0,0050	0,015	0,0167
7	Chloroneb	0	99.41	0,0075	0,0050	0,015	0,0167
8	Methyl parathion	0	79.68	0,025	0,0250	0,050	0,0833
9	Propanil	0	108.04	0,0075	0,0250	0,015	0,0833
10	Fipronil	0	107.67	0,0075	0,0025	0,015	0,0083
11	Mevinphos	209.31	68.40	0,05	0,0250	0,100	0,0833
12	Isazofos	90.62	77.40	0,025	0,0050	0,050	0,0167
13	Chlorpyrifos	87.80	25.09	0,01	0,0050	0,020	0,0167
14	Bifenthrin	88.26	8.08	0,025	0,0250	0,050	0,0833
15	Lindane	73.23	0	0,25	0,0250	0,5	0,0833
16	Chlorfenapyr	92.75	35.87	0,005	0,0050	0,010	0,0167
17	Malathion	94.36	87.03	0,05	0,0083	0,10	0,0277
18	Permethrin 1	89.27	44.35	0,005	0,0025	0,010	0,0083
19	Metolachlor	99.53	91.92	0,0075	0,0025	0,015	0,0083
20	Paclobutrazol	92.37	100.61	0,005	0,0100	0,01	0,0333
21	Pretilachlor	97.07	72.23	0,0058	0,0025	0,012	0,0083
22	Cypermethrin 1	61.66	10.37	-	0.051	-	0.154
23	Cypermethrin 2	61.67	10.35	-	0.068	-	0.207
24	Cypermethrin 3	80.95	10.36	-	0.008	-	0.025
25	Cypermethrin 4	94.11	10.58	-	0.039	-	0.118
26	Triadimefon	92.28	16.39	0,025	0,0075	0,050	0,0250
27	Biphenyl	92.28	16.40	0,0066	0,0025	0,013	0,0083

(-): Not available

3.2. Pesticide analysis in water and rice grain

After method validation was completed, the suggested multi-pesticide residue techniques were used to screen for pesticide residues in irrigation canal water, paddy field water, and rice samples that were taken from an agrochemical-free paddy field. By injecting a positive sample and comparing the mass spectrum of each component from the MS detector with the instrumental database and the NIST MS search library version 2.0, to identification of pesticide chemicals. Every pesticide chemical should have a similarity percentage of at least 78% when compared to a spectral library.

The assessment of pesticides in this study was focused on irrigation canal water, paddy field water, and paddy rice. A total of 32 samples were collected, comprising 26 water samples and 6 samples of rice were analyzed. After analysis, the triplicate samples of the canal water, two paddy field waters, and paddy grain samples showed the result with no detection of any pesticide compound. To indicate that the analytical method and instrument performance were consistent, the 6 main compounds of the internal standard were injected with all samples, and the result showed that 5 out of 6 internal standard compounds were detected [12]. According to the study area, Demo 1 is known as the organic field, and the surrounding paddy fields also follow the organic technique with cover crop management, which means that no pesticide or chemical application is made on that sampling side. In the experimental study area of Demo 1, the adjacent paddies were managed by strict organic agriculture, including all cultivation for the ecological balance and the sustainable development of the area. These include management of cover crops, natural pest control, and not using synthetic chemical inputs such as pesticides and chemical fertilizers. The Organic farming systems are meant to nourish soil health and microbial activity, which can further contribute to the

degradation of any external chemical residues that might inadvertently enter the system [13]. Since the organic fields are maintained without the use of synthetic agrochemicals, the absence of detectable pesticide residues in the analytical results is consistent with organic farming principles. This result is in line with research showing that, as compared to conventional farming systems, organic farming systems have noticeably lower or undetectable amounts of pesticide residue [13]. Despite the statements of herbicide use throughout the previous two years, Demo 2 showed no detectable pesticide residues in the study. The study area on Demo 2 followed the management strategy that incorporates cover cropping management and strictly used a single herbicide application. To control weed growth, which is especially common in this study area since farmers using direct sowing have replaced in traditional rice transplanting. This can be attributed to the deviation from traditional rice cultivation practices, in which the farmers used the herbicide product U-T 99, which contains ingredient bispyribac-sodium 20% EC, as a single spray in the early growth stage of rice in July. To promote environmental sustainability and support agroecological practices, field technicians have instructed farmers within the study area and the area around Demo 2 to minimize and reduce chemical consumption by adhering to this single-use application guideline. In accordance with the previous research, our investigation indicated that no detectable pesticide residues were found in the organic rice samples. The investigation found that multiple rice samples from different districts within the Mekong Delta had no discernible pesticide residues include Cau Ngang district (2017), Thoi Binh district (2018), Tri Ton district, Vung Liem district, Nga Nam town, My Xuyen district (2019), and later U Minh Thuong and Tri Ton districts (2021). This absence is greatly due to the observance of effective organic farming techniques and the rice-shrimp aquaculture integration system, which implies less usage of chemicals to maintain shrimp health. Uncompromised compliance with organic guidelines in rice-dominated regions also lessened pollution despite intensive pest pressures. This data showcases the importance of organic agriculture in attaining no-input farming in synergy with environment-friendly production and healthier ecosystem policies [1].

According to the survey from farmers in the study area within the surrounding rice fields, bispyribac-sodium is a common systemic herbicide used in rice farming to control weeds and broadleaf weeds. Among the processes influencing its degradation in the environment are microbial activity, photolysis, and hydrolysis. According to studies of AICRP on Weed Control, bispyribac-sodium has first-order kinetics and half-lives that range from 1.4 to 16 days depending on temperature, pH, and microbial populations [14]. Aline with bispyribac sodium, environmental concentrations in the air and airborne transport are regarded as insignificant because of the poor capacity for volatilization and the estimated rapid photochemical transformation. Bispyribac sodium exhibits moderate to high persistence in anaerobic flooded paddy soils when anaerobic conditions are present. In the aquatic environments, the chemical degrades over 99% in 90 days, demonstrating rapid evaporation [15]. This rapid breakdown often produces metabolites that are not present in traditional analytical libraries, making identification more challenging. Moreover, the compound bispyribac-sodium cannot be identified using GC-MS due to its chemical properties. Bispyribac-sodium is a polar, non-volatile molecule [16] , and this compound was not listed in the GC-MS database and will not be found during analysis. This emphasizes the importance of modernizing the analytical libraries and considering alternative techniques, like LC-MS/MS, which are better suited for identifying polar and less volatile chemicals.

4. CONCLUSIONS

The consistent absence of detectable pesticide residues in both the studied organic farming areas and the associated irrigation canal water strongly suggests the effective implementation of organic farming practices. While natural degradation processes over the considerable period between any potential application and sampling, along with dilution, adsorption to sediments, and the general environmental dissipation, likely contributed to the lack of detection in water samples, it is crucial to acknowledge that the non-detection in both matrices may also be attributed to the inherent limitations of the GC-MS analytical method. Specifically, certain pesticide compounds are not amenable to GC-MS analysis. To achieve a more comprehensive assessment and potentially identify trace levels of a broader range of pesticide residues, future studies should employ a complementary approach using both GC-MS and LC-MS techniques.

ACKNOWLEDGMENTS

This research was financially supported by the Water Resources Management and Agro-ecology Transition for Cambodia (WAT4CAM) projects.

REFERENCES

[1] C. T. Nguyen *et al.*, “Assessment of Pesticide Residues in Organic Rice Production in The Mekong Delta, Vietnam,” *Eur. J. Dev. Stud.*, vol. 2, no. 3, pp. 1–11, 2022, doi: 10.24018/ejdevelop.2022.2.3.95.

[2] P. Schreinemachers *et al.*, “Too much to handle? Pesticide dependence of smallholder vegetable farmers in Southeast Asia,” *Sci. Total Environ.*, vol. 593–594, pp. 470–477, 2017, doi: 10.1016/j.scitotenv.2017.03.181.

[3] O. Tiryaki, “Pestisitlerin Çevredeki Ak ı beti Özet,” no. December, 2016.

[4] H. V. Duong, T. C. Nguyen, X. T. Nguyen, M. Q. Nguyen, P. H. Nguyen, and T. T. Vo, “Evaluating the

Presence of Pesticide Residues in Organic Rice Production in An Giang Province, Vietnam," *J. Sustain. Dev.*, vol. 15, no. 1, p. 49, 2022, doi: 10.5539/jsd.v15n1p49.

[5] Agilent Technologies, "Proven approaches for today's food analysis challenges," *Appl. Noteb.*, vol. 2, pp. 15–20, 2013, [Online]. Available: http://www.chem.agilent.com/Library/brochures/5990-4977EN.pdf

[6] K. Kadokami, S. Pan, D. T. Hanh, X. Li, and T. Miyazaki, "Development of a comprehensive analytical method for semi-volatile organic compounds in sediments by using an automated identification and quantification system with a GC-MS database," *Anal. Sci.*, vol. 28, no. 12, pp. 1183–1189, 2012, doi: 10.2116/analsci.28.1183.

[7] L. Sun, Q. Guo, C. C. Jacob, C. P. B. Martins, and R. Fussell, "Multi-pesticide residues analyses of QuEChERS extracts using an automated online μSPE clean-up coupled to LC-MS/MS," p. 24, 2020.

[8] D. Libraries, "A b s tra c t :," vol. 2, no. 2, pp. 1–5, 2013.

[9] M. Tauseef *et al.*, "Analysis of multiple pesticide residues in rice by LC–MS/MS," *Chem. Pap.*, vol. 75, no. 6, pp. 2871–2879, 2021, doi: 10.1007/s11696-021-01533-x.

[10] A. Lozano, B. Kiedrowska, J. Scholten, A. De Kok, and A. R. Fernández-alba, "NRL - NVWA Improvements in acetone based extraction method for pesticide residues analysis INTRODUCTION," p. 120.

[11] M. Bulai, "Matrix Effect Evaluation in GC / MS-MS Analysis of Multiple," 2023.

[12] T. Yarita, Y. Aoyagi, and T. Otake, "Evaluation of the impact of matrix effect on quantification of pesticides in foods by gas chromatography-mass spectrometry using isotope-labeled internal standards," *J. Chromatogr. A*, vol. 1396, pp. 109–116, 2015, doi: 10.1016/j.chroma.2015.03.075.

[13] C. Benbrook, S. Kegley, and B. Baker, "Organic farming lessens reliance on pesticides and promotes public health by lowering dietary risks," *Agronomy*, vol. 11, no. 7, pp. 1–36, 2021, doi: 10.3390/agronomy11071266.

[14] F. Ahmad, S. Anwar, S. Firdous, Y. Da-Chuan, and S. Iqbal, "Biodegradation of bispyribac sodium by a novel bacterial consortium BDAM: Optimization of degradation conditions using response surface methodology," *J. Hazard. Mater.*, vol. 349, no. 2010, pp. 272–281, 2018, doi: 10.1016/j.jhazmat.2017.12.065.

[15] Qamruzzaman and A. Nasar, "Degradative treatment of bispyribac sodium herbicide from synthetically contaminated water by colloidal MnO2 dioxide in the absence and presence of surfactants," *Environ. Technol. (United Kingdom)*, vol. 40, no. 4, pp. 451–457, 2019, doi: 10.1080/09593330.2017.1396500.

[16] PubChem News, "Bispyribac-sodium." Accessed: Apr. 19, 2025. [Online]. Available: https://www.ncbi.nlm.nih.gov/

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Seasonal Pesticide Residues in Surface Waters of Banteay Meanchey Province, Cambodia

Vivath Tit [1], Leanghok Bun [1,2], Sereyvath Yoeun [1,2], Chanvorleak Phat [1,2*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: phatchanvorleak@itc.edu.kh*

Abstract: *Pesticide residues, even after application to specific fields, can indirectly accumulate in adjacent surface waters through runoff and soil percolation, posing significant risks to non-target aquatic organisms. This study investigated seasonal pesticide contamination in surface waters across Banteay Meanchey Province, northwest Cambodia, an area where modern agriculture significantly impacts canals, reservoirs, and rivers. A total of 40 samples, 20 collected in rainy season and another 20 in dry season from various sites throughout Banteay Meanchey Province, were examined in this study. Solid-phase extraction (SPE) followed by GC-MS was used to analyze pesticide residues samples from various sampling stations during both the rainy and dry seasons, tracing the application and contamination pathways. During the rainy season, chlorfenapyr, fipronil, paclobutrazol, parathion, and atrazine were detected, reflecting diverse pesticide use, particularly of insecticides. Notably, paclobutrazol, a plant growth regulator, was found at multiple stations, indicating widespread application. In the dry season, isoprothiolane, parathion, and atrazine were detected, revealing the use of fungicides, insecticides, and herbicides, respectively. A clear seasonal variation of pesticide residues in surface water was observed, with the mean pesticide concentration during the dry season (0.20 $\mu g/L$) exceeding that of the rainy season (0.10 $\mu g/L$). Alarmingly, parathion, a banned insecticide, was present in both seasons, with a slightly higher concentration recorded during the rainy season (0.20 $\mu g/L$) compared to the dry season (0.19 $\mu g/L$). However, atrazine exhibited a higher concentration during the dry season (0.22 $\mu g/L$) relative to the rainy season (0.21 $\mu g/L$). These findings highlight distinct seasonal pesticide applications and their subsequent transport in surface waters. Consequently, increased awareness of pesticide residues from agricultural runoff is crucial to protect public health and aquatic ecosystems.*

Keywords: Pesticides; Surface water; Runoff; SPE; GC-MS

1. INTRODUCTION

In chemical pest management, pesticides are formulated mixtures of chemical compounds primarily designed to prevent, suppress, or eradicate harmful organisms that threaten agricultural productivity, or to function as a plant regulator [1]. Consequently, pesticide application contributes to higher crop production; however, it also provokes another form of pollution in the environment because of its chemical ingredients [2].

Correspondingly, Cambodia's pesticide demand is around 0.58% of roughly 3 million tons of pesticide used worldwide [3]. In fact, it is estimated that the demand for pesticides in Cambodia increases by 61% on a yearly basis [4]. Considering this trend, inflated pesticide application is evident in the agricultural practices of Cambodian farmers; however, the improper use of these chemicals also occurs spontaneously [5]. Likewise, significant issues related to inadequate knowledge and misconceptions about pesticides, including the miscalculation of risks and illegal chemical use, requires urgent attention and improvement among Cambodian farmers [6]. In this regard, the application of pesticides to specific agricultural fields can result in the indirect accumulation of residues in adjacent surface water resources, either during or after their use [7], and this problem must be intentionally addressed in Cambodia.

The purpose of this study is to analyze pesticide residues in surface water collected during the rainy and dry seasons from Banteay Meanchey province, Cambodia, an agricultural region characterized by numerous water bodies, including canals, reservoirs, lakes and rivers. The occurrences and seasonal variations of pesticide residues in the surface waters were assessed in relation to agricultural application practices and associated transport pathways across the study areas.

[1*] Corresponding author: *Chanvorleak Phat*
E-mail: phatchanvorleak@itc.edu.kh; Tel: +855-93-418-800

2. METHODOLOGY

2.1 Study area

Banteay Meanchey Province is in northwestern Cambodia, bordering Thailand and surrounded by Oddar Meanchey, Siem Reap and Battambang provinces (**Fig. 1**). Its economy relies heavily on agriculture, with key crops such as rice, cassava, banana and soybeans [8]. The province receives an average annual rainfall of approximately 1300 mm, with September and October being the wettest months, averaging 445.9 mm of rainfall. Temperatures in this region range from a minimum of 23℃ to a maximum of 32.1℃, with March and April reaching the highest temperature of 36.2℃ [9]. The majority of people in this province are small-scale farmers, with most relying on rice cultivation as their primary livelihood [10]. Typically, farmers grow crops twice a year: the first runs from March to July and the second from July to November. The cropping season usually begins with the onset of rainfall, as farmers commonly sow their crops after the first rain in the late dry season [11]. Banteay Meanchey is home to two significant rivers, Mongkol Borey and Sisophon, which serve as major tributaries to the Tonle Sap Great Lake, particularly during the rainy season [12].



Fig. 1. Banteay Meanchey Province (Source: Wikipedia)

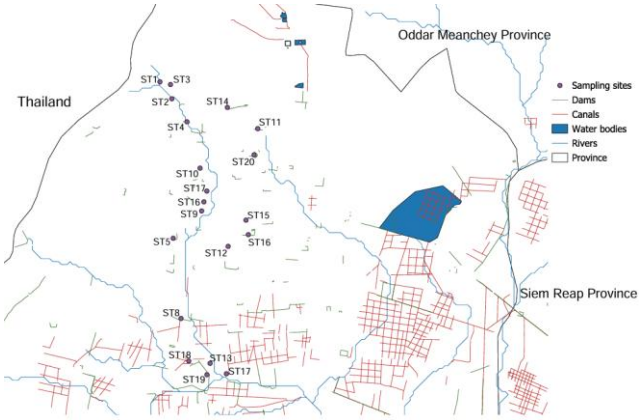


Fig. 2. Sampling station of water samples

2.2 Sample collection

Table 1. Coordination of the sampling sites.

Sampling sites	Latitude	Longitude
ST1	14°00'49"N	102°55'08"E
ST2	13°59'21"N	102°56'10"E
ST3	14°00'35"N	102°56'03"E
ST4	13°57'21"N	102°57'29"E
ST5	13°47'17"N	102°56'17"E
ST6	13°47'38"N	103°02'48"E
ST7	13°51'22"N	102°59'12"E
ST8	13°40'21"N	102°56'58"E
ST9	13°49'39"N	102°58'45"E
ST10	13°53'21"N	102°58'37"E
ST11	13°56'46"N	103°03'37"E
ST12	13°46'35"N	103°01'03"E
ST13	13°36'28"N	102°59'29"E
ST14	13°58'36"N	103°00'59"E
ST15	13°48'51"N	103°02'38"E
ST16	13°50'26"N	102°58'57"E
ST17	13°35'34"N	103°00'54"E
ST18	13°36'40"N	102°57'38"E
ST19	13°35'29"N	102°59'13"E
ST20	13°54'30"N	103°03'18"E

A total of 40 surface water samples were collected; with 20 samples taken during the rainy season (November, 2024) and 20 during the dry season (February, 2025). Sampling stations were identified using Google Earth to pinpoint their locations. Two water samples (ST17 and ST19) were collected from Sisophon River and 8 samples (ST1, ST2, ST7, ST8, ST9, ST10, ST13 and ST16) were taken from the northwestern streams of its branch. Additionally, 8 water samples (ST3, ST4, ST5, ST6, ST11, ST14, ST18 and ST20) were gathered from reservoirs across the province, along with 1 sample (ST12) from a main canal and 1 sample (ST15) from a large pond (see **Fig. 2**). At each station, 1 liter of surface water was collected using the manual grab method. Sampling was conducted approximately 2 to 3 meters from the bank of the water body at a depth of around 30 cm, using a telescopic water sampler. Before collection, the pre-cleaned polyethylene bottle was carefully rinsed with the surface water from the sampling station. The final sample was then

obtained by collecting water from three different subsets within the designated area of each station and combining them into a single 1-liter bottle. After collection, the samples were immediately stored in an ice chest and transported under refrigerated conditions at 4℃ to the laboratory for further extraction of target analytes.

2.3 Sample preparation

To adjust the pH of the water samples, 1 mL of phosphate buffer solution (1 mol/L, pH 7.0) was added. The water sampled were then filtered through a 2 μm fiberglass filter to remove any particulate matter or suspended solids. To eliminate any residual compounds remaining on the filter, 10 mL of acetone were applied. The filtered water samples were subsequently processed using solid-phase extraction.

2.4 Solid-phase extraction

The solid-phase extraction (SPE) procedure was performed following the method described by Jinya (2013) [13]. The SPE cartridge used consisted of a combination of 2 sorbents: PLS3 (styrene divinylbenzene polymer + N-vinyl acetamide) and AC2 (activated carbon). During the extraction process, water samples first passed through the PLS3, followed by the AC2. For cartridge conditioning, 5 mL of dichloromethane, 5 mL of acetone and 10 mL of pure water were sequentially added, allowing each solvent to remain in the cartridge for 2 minutes. After conditioning, 1 L of the water samples was loaded to the cartridge using the vacuum pump. The flow rate was maintained at approximately 15 mL/min by adjusting the applied pressure. Upon completion of sample loading, both sorbents were dried simultaneously under a stream of nitrogen gas for 30 to 40 minutes. For elution, 2 mL of acetone and 3 mL of dichloromethane were added to the PLS3. Separately, 3 mL of acetone was added to the AC2 sorbent. The combined eluates were then concentrated to 1 mL under a gentle stream of nitrogen gas. Prior to dehydration, 5 mL of hexane was pumped into the concentrate, which was then pipetted in a tube containing 8 g of anhydrous sodium sulfate. The original concentrate container was subsequently rinsed five times with 1mL of hexane and each rinse was also transferred to the sodium sulfate tube. The extract was again concentrated under nitrogen gas until the final volume was reduced to less than 1 mL. Finally, 0.01 mL of internal standard solution containing Acenapthene-d10, 1,4-Dichlorobenzene-D4, Napthalene-D8, Perylene-D12, Chrysene-D12, and Phenanthrene-D10 was added prior to instrumental analysis.

2.2 Instrumental analysis

The pesticide residues were analyzed both qualitatively and quantitively using a gas chromatograph coupled with a mass spectrometry (GC/MS) with the GCMS-TQ8040 model (Shimadzu, Japan). A 1 μL sample of the extract was injected in split-less mode, at an injection temperature 250℃. Chromatographic separation was carried out on a DB-5MS capillary column ($30m \times 0.25m, thickness\ 0.25\mu m$). The oven temperature program began at 40℃ (held for a minute), followed by an increase to 310℃ at a rate of 8℃/min, then was held at 310℃ for 4 minutes. Ultra-high purity (99.999%) helium was employed as the carrier gas, with a total flow rate of 50 mL/min and a column flow rate of 1.23 mL/min. The EI for MS was 70 eV, while ion source temperature was set at 200℃, and the interface temperature was maintained at 300℃. Pesticide compounds were identified by comparing their retention times and mass spectra to predict the structural formulas using the MS database. Quantification was performed using the external standards, limited to a selected number of pesticide compounds. Limit of detection (LOD) and limit of quantification (LOQ) were determined using manual injection over a concentration range of 0.0025 to 1 $\mu g/L$. The LOD and LOQ values for the reference compounds can be referenced from the previous study by Thourn et al. (2022) [14].

3. RESULTS AND DISCUSSION

3.1 Occurrence and concentration of pesticide residues

Table 2. Pesticide residue levels in surface water of rainy season.

Compounds	Pesticide classification	Stations	Concentration ($\mu g/L$)
Chlorfenapyr	Insecticide	ST9	≤LOQ
Fipronil	Insecticide	ST8	0.018
Paclobutrazol	Plant growth regulator	ST8, ST12, ST19	0.04, 0.07, 0.11
Parathion	Insecticide	ST14	0.20
Atrazine	Herbicide	ST19	0.21
Average			0.10

≤LOQ: Lower than the limit of quantification

During the rainy season, five pesticide compounds—fipronil, chlorfenapyr, paclobutrazol, parathion, and atrazine—were detected at five out of 20 stations, as shown in **Table 2**. Chlorfenapyr was detected at station 9, though its concentration was below the limit of quantification. Fipronil had the lowest concentration, measured at 0.018 $\mu g/L$ at station 8. Across the study area, paclobutrazol was found at stations 8, 12 and 19, with concentrations of 0.04, 0.07, 0.11 $\mu g/L$, respectively, resulting in a mean concentration of 0.07 $\mu g/L$. The detection of paclobutrazol suggests that it is the predominant single pesticide present during this season. The second highest concentration was observed for parathion,

which was detected at station 14 at 0.20 $\mu g/L$. Atrazine was detected with the concentration of 0.21 $\mu g/L$ at station 19.

Table 3. Pesticide residue levels in surface water of dry season.

Compounds	Pesticide classification	Stations	Concentration ($\mu g/L$)
Isoprothiolane	Fungicide	ST13, ST17, ST19	+
Parathion	Insecticide	ST13	0.19
Atrazine	Herbicide	ST18	0.22
Average			0.20

+: Detected but standard is not available for quantification

During the dry season, three pesticide compounds—isoprothiolane, parathion and atrazine—were detected at four out of 20 stations, as shown in **Table 3**. The concentration of isoprothiolane could not be quantified as it was outside the range of the external standards; however, its presence was confirmed at stations 13, 17 and 19. Significantly, the presence of isoprothiolane at three nearby stations indicated it as the most frequently occurring pesticide during this season. Parathion was detected at station 13, with a concentration of 0.19 $\mu g/L$. Atrazine had the highest concentration of 0.22 $\mu g/L$ at station 18.

Chlorfenapyr is approved for use in Cambodia as an insecticide to control pests, including diamond black worms, defoliators, leaf minors, aphids, white flies, green worm, red mites and grass hoppers [15]. Although chlorfenapyr has the potential to contaminate surface water through spray drift and surface runoff after rainfall, it exhibits a stronger affinity for binding to soil particles than water, and also possesses low water solubility, with average half-life of 0.8 years [16], resulting its detection during the rainy season.

Although its concentration was at the lowest, its presence indicates nearby usage as fipronil was used as an insecticide to eliminate insect pests such as common worms, hoppers and beetles [15]. Fipronil can rapidly dissipate from surface water over time following application and rainfall, primary due to its degradation into byproducts and moderate hydrophobic nature [17]. In this context, the concentration of fipronil residues in this province did not pose a toxicity risk to freshwater vertebrates or aquatic plants. However, it is likely to caused chronic toxicity to freshwater invertebrates, given the EPA's chronic benchmark value of 0.011 $\mu g/L$ [18].

The excessive use of paclobutrazol has been reported across several provinces in Cambodia, where this plant growth regulator is applied to inhibit the growth of various rice varieties, particularly the tall-growing Somali rice variety, and to promote early tillering and flowering [19]. Paclobutrazol is more stable in soil than in water, with a half-life of 182 days in soil and 164 days in water [20]. Despite this, the concentration of paclobutrazol observed in this study is well below levels of concern and is considered safe for aquatic plants, freshwater vertebrates, and invertebrates [18].

Parathion is an organophosphorus pesticide used to control various species of insect pests on crops such as corn, cotton, soybeans, sunflower and wheats in the USA [21]. Parathion is classified as an extremely hazardous pesticide and has been officially banned for use in Cambodia since 2012 [22]. However, the organophosphorus pesticides remain among the most commonly used in developing countries due to their low cost and high effectiveness compared to other classes of pesticides [23]. Due to its high water solubility, parathion is widely found in surface waters, introduced through irrigation runoff, agricultural wastes and rainfall [24].

A range of herbicides is available for use by farmers in Cambodia, with atrazine being one that has been on the market since 2004 [25]. Atrazine is the most frequently detected herbicide in surface water, attributed to its strong hydrophobicity, low solubility and persistence in aquatic environment, with a half-life of about 84 days [26]. The concentration of atrazine detected in this study was significantly lower than the benchmark values for toxicity to freshwater vertebrates (2650 $\mu g/L$) and freshwater vascular plants (4.6 $\mu g/L$) established by the U.S. EPA [18].

Isoprothiolane fungicide has been proposed as a targeted treatment to manage rice blast in rice field caused by fungus called Pyricularia grisea Sacc [27]. With moderate solubility and a half-life of 13 to 16 days, isoprothiolane can persist in paddy field water and reach adjacent irrigation canals through runoff [28]. Therefore, its detected indicates extensive use throughout the dry season.

3.2 Seasonal variation of pesticide residues

Different pesticide compounds exhibited varying application patterns during the rainy and dry seasons, indicating differences in pesticide use for pest control across seasons. The rainy season showed the most frequent detection of pesticide residues in study area compared to the dry season, with five stations detected in the rainy season and four in the dry season. However, the mean concentration of pesticides was lower during the rainy season (0.10 $\mu g/L$) than during the dry season, which had a mean concentration of 0.20 $\mu g/L$. Remarkably, parathion was detected in both seasons, with slight variations in concentration. Atrazine consistently had the highest concentration in both seasons, with a higher concentration observed during the dry season compared to the rainy season. Regarding the location of the study area, station 19 showed variation in pesticide applications across seasons, with paclobutrazol and atrazine detected during the rainy season and isoprothiolane detected during the dry season.

The higher frequency of pesticide residue detection during the rainy season is influenced by factors such as rainfall intensity, watershed characteristics, agricultural practices, as well as the solubility and persistence of the pesticides [29,30]. Thus, the presence of pesticides detected

in this study can attribute to seasonal application patterns and the excessive use of these chemicals, which likely contributed to their frequency detection in nearby surface water, particularly during the rainy season.

The higher pesticide concentration observed during the dry season, characterized by low precipitation, can be attributed to the discharge of residual pesticides into the nearby water sources and the reduced dilution capacity due to lower water levels [31]. Another factor contributing to the lower pesticide concentrations observed during the rainy season, characterized by high precipitation, is the dilution effect caused by heavy rainfall, along with the enhanced degradation of pesticides following their application [32]. Subsequently, the frequency of rainfall, physicochemical properties of the substances and the characteristics of the soil, also influence the pesticide concentration in the surface water [33].

4. CONCLUSIONS

Surface waters near agricultural fields are highly susceptible to pesticide contamination due to intensive and year-round pesticide use. During the rainy season, fipronil, chlorfenapyr, paclobutrazol, parathion and atrazine were detected, while isoprothiolane, parathion, and atrazine were detected during the dry season. The presence of insecticides such as fipronil, chlorfenapyr, and parathion, particularly during the rainy season, suggests frequent use and potential runoff into nearby water sources. Residues of other pesticide types, including plant growth regulator and fungicide, were also detected, reflecting diverse application practices among farmers. Seasonal variation of pesticide concentration was evident, with the dry season exhibiting a higher mean concentration compared to the rainy season, likely due to lower rainfall, agriculture practices, and reduced dilution in the water sources. Key factors influencing pesticide residues in surface water include application frequency, recent inputs, rainfall intensity, surface runoff and pesticide physicochemical properties. To mitigate risks to aquatic ecosystems, it is essential to enhance farmers' awareness of safe pesticide handling and strengthen the enforcement of pesticide regulations in Cambodia.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the facility support provided by the Platform for Aquatic Ecosystem Research (PAER).

REFERENCES

[1] Food and Agriculture Organization of the United Nations (FAO) and World Health Organization (WHO), *The International Code of Conduct on Pesticide Management*. 2014. [Online]. Available: http://www.fao.org/fileadmin/templates/agphome/documents/Pests_Pesticides/Code/CODE_2014Sep_ENG.pdf

[2] W. Aktar, D. Sengupta, and A. Chowdhury, "Impact of pesticides use in agriculture: Their benefits and hazards," *Interdiscip. Toxicol.*, vol. 2, no. 1, pp. 1–12, 2009, doi: 10.2478/v10102-009-0001-7.

[3] FAOSTAT, "Pesticide Use," FAO. Accessed: Apr. 20, 2025. [Online]. Available: https://www.fao.org/faostat/en/#data/RP

[4] P. Schreinemachers, V. Afari-Sefa, C. H. Heng, P. T. M. Dung, S. Praneetvatakul, and R. Srinivasan, "Safe and sustainable crop protection in Southeast Asia: Status, challenges and policy options," *Environ. Sci. Policy*, vol. 54, pp. 357–366, 2015, doi: 10.1016/j.envsci.2015.07.017.

[5] R. J. Flor, H. Maat, B. A. R. Hadi, V. Kumar, and N. Castilla, "Do field-level practices of Cambodian farmers prompt a pesticide lock-in?," *F. Crop. Res.*, vol. 235, no. October 2018, pp. 68–78, 2019, doi: 10.1016/j.fcr.2019.02.019.

[6] C. Chanchao, "Current status of pesticide practices and management approaches toward the safety and health of Cambodia: A review," *Insight Cambodia J. Basic Appl. Res.*, vol. 5, no. 02, pp. 9–21, 2023, doi: 10.61945/cjbar.2023.5.2.03.

[7] C. Khurshid *et al.*, "Pesticide residues in European sediments: A significant concern for the aquatic systems?," *Environ. Res.*, vol. 261, no. May, 2024, doi: 10.1016/j.envres.2024.119754.

[8] S. Phon, "The Impact of Changes in Crop-Livestock Interaction in Banteay Chmar Commune , Thmar Puok District , Banteay Meanchey Province , Cambodia," pp. 2006–2007, 2012.

[9] V. Touch *et al.*, *Adaptation constraints and prospects for future research priorities in lowland rice-based farming systems: learning experiences from Northwest Cambodia*, vol. 26, no. 9. Springer Netherlands, 2024. doi: 10.1007/s10668-023-03566-6.

[10] V. Thangrak, B. Somboonsuke, S. Sdoodee, R. Darnsawasdi, and P. Voe, "The smallholder farmers' perceptions of climate variability impact on rice production and the case of adaptation strategies in Banteay Meanchey, (BMC), Cambodia," *Int. J. Agric. Technol.*, vol. 16, no. 2, pp. 505–516, 2020.

[11] V. Touch, R. J. Martin, F. Scott, A. Cowie, and D. L. Liu, "Climate change impacts on rainfed cropping production systems in the tropics and the case of smallholder farms in North-west Cambodia," *Environ. Dev. Sustain.*, vol. 19, no. 5, pp. 1631–1647, 2017, doi: 10.1007/s10668-016-9818-3.

[12] WEPA, "State of Water: Cambodia." Accessed: Apr.

04, 2025. [Online]. Available: https://wepa-db.net/archive/policies/state/cambodia/river2_2_4.htm?utm

[13] D. Jinya, "Development of Solid-Phase Extraction Method for Simultaneous Analysis of Semi-Volatile Organic Compounds Using a GC-MS Database System," vol. 2, no. 2, pp. 1–5, 2013.

[14] L. Thourn, C. Phat, M. Soung, S. Sieng, S. Heng, and S. Yoeun, "Identification of Pesticide Contamination in Water Sources Surrounding Agrochemical-Free Rice Farming in Battambang Province," vol. 10, no. 2, pp. 63–70, 2022.

[15] K. S. and S. M. Sim Sokcheng, *Sim Sokcheng, Keo Socheat and Sarom Molideth Working Paper Series No. 128 August 2021*, no. 128. 2021.

[16] P. Huang, X. Yan, B. Yu, X. He, L. Lu, and Y. Ren, "A Comprehensive Review of the Current Knowledge of Chlorfenapyr: Synthesis, Mode of Action, Resistance, and Environmental Toxicology," *Molecules*, vol. 28, no. 22, 2023, doi: 10.3390/molecules28227673.

[17] L. Greenberg *et al.*, "Impact of ant control technologies on insecticide runoff and efficacy," *Pest Manag. Sci.*, vol. 66, no. 9, pp. 980–987, 2010, doi: 10.1002/ps.1970.

[18] EPA, "Aquatic Life Benchmarks and Ecological Risk Assessments for Registered Pesticides," United States Environmental Protection Agency. Accessed: Apr. 21, 2025. [Online]. Available: https://www.epa.gov/pesticide-science-and-assessing-pesticide-risks/aquatic-life-benchmarks-and-ecological-risk#ref_1

[19] Phak Seangly, "Farmers warned about overuse of fertilisers," The Phnom Penh Post. Accessed: Apr. 09, 2025. [Online]. Available: https://www.phnompenhpost.com/national/farmers-warned-about-overuse-of-fertilisers

[20] A. Hussain *et al.*, "Waterborne exposure of paclobutrazol at environmental relevant concentration induce locomotion hyperactivity in larvae and anxiolytic exploratory behavior in adult zebrafish," *Int. J. Environ. Res. Public Health*, vol. 17, no. 13, pp. 1–18, 2020, doi: 10.3390/ijerph17134632.

[21] EPA, "Prevention, Pesticides and Toxic Substances (7508C). Fact Sheet for Ethyl Parathion.," *US EPA-Pesticides*, no. December 1986, pp. 1–6, 2000, [Online]. Available: https://www3.epa.gov/pesticides/chem_search/reg_actions/reregistration/fs_PC-057501_1-Sep-00.pdf

[22] MAFF, "List of Pesticide in Cambodia," 2013, [Online]. Available: file:///C:/Users/User/OneDrive/Documents/Project/Cambodia-Banned Pesticide.pdf

[23] S. Y. Wee, T. F. T. Omar, A. Z. Aris, and Y. Lee, "Surface Water Organophosphorus Pesticides Concentration and Distribution in the Langat River, Selangor, Malaysia," *Expo. Heal.*, vol. 8, no. 4, pp. 497–511, 2016, doi: 10.1007/s12403-016-0214-x.

[24] O. T. Ore, A. O. Adeola, A. A. Bayode, D. T. Adedipe, and P. N. Nomngongo, "Organophosphate pesticide residues in environmental and biological matrices: Occurrence, distribution and potential remedial approaches," *Environ. Chem. Ecotoxicol.*, vol. 5, no. October 2022, pp. 9–23, 2023, doi: 10.1016/j.enceco.2022.10.004.

[25] Ministry of Environment Cambodia, "National Profile on Chemicals Management in Cambodia," no. December 2004, p. 162, 2004.

[26] R. M. de Souza, D. Seibert, H. B. Quesada, F. de Jesus Bassetti, M. R. Fagundes-Klen, and R. Bergamasco, "Occurrence, impacts and general aspects of pesticides in surface water: A review," *Process Saf. Environ. Prot.*, vol. 135, pp. 22–37, 2020, doi: 10.1016/j.psep.2019.12.035.

[27] GDA, "Rice Blast-Cambodia: Pyricularia grisea Sacc," PlantwisePlus Knowledge Bank. Accessed: Apr. 17, 2025. [Online]. Available: https://plantwiseplusknowledgebank.org/doi/full/10.1079/PWKB.20147801487

[28] M. R. Abd Rahman, H. Ahmad, and W. Fadhullah, "The residual level of isoprothiolane in paddy field surface water and its acute toxicity level on freshwater prawn, Macrobrachium lanchesteri," *Paddy Water Environ.*, vol. 17, no. 4, pp. 571–579, 2019, doi: 10.1007/s10333-018-00684-0.

[29] I. K. Konstantinou, D. G. Hela, and T. A. Albanis, "The status of pesticide pollution in surface waters (rivers and lakes) of Greece. Part I. Review on occurrence and levels," *Environ. Pollut.*, vol. 141, no. 3, pp. 555–570, 2006, doi: 10.1016/j.envpol.2005.07.024.

[30] A. Claver, P. Ormad, L. Rodríguez, and J. L. Ovelleiro, "Study of the presence of pesticides in surface waters in the Ebro river basin (Spain)," *Chemosphere*, vol. 64, no. 9, pp. 1437–1443, 2006, doi: 10.1016/j.chemosphere.2006.02.034.

[31] T. Nasrabadi, G. N. Bidhendi, A. Karbassi, P. Grathwohl, and N. Mehrdadi, "Impact of major organophosphate pesticides used in agriculture to surface water and sediment quality (Southern Caspian Sea basin, Haraz River)," *Environ. Earth Sci.*, vol. 63, no. 4, pp. 873–883, 2011, doi: 10.1007/s12665-010-0757-2.

[32] G. Wei *et al.*, "Occurrence and risk assessment of currently used organophosphate pesticides in overlying water and surface sediments in Guangzhou urban waterways, China," *Environ. Sci. Pollut. Res.*, vol. 28, no. 35, pp. 48194–48206, 2021, doi: 10.1007/s11356-021-13956-w.

[33] A. A. Thomatou, I. Zacharias, D. Hela, and I. Konstantinou, “Determination and risk assessment of pesticide residues in lake Amvrakia (W. Greece) after agricultural land use changes in the lake’s drainage basin,” *Int. J. Environ. Anal. Chem.*, vol. 93, no. 7, pp. 780–799, 2013, doi: 10.1080/03067319.2012.656099.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Development of Dishwashing Liquid Through Utilization of Various Active Ingredients

Kosal Vaphat [1], Peany Houng [1*], Keakaknika Ly [2], Samphois Nut [2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *LSV Industry Company, Sangkat Phnom Penh Thmey, Khan Sen Sok, 120801 Phnom Penh, Cambodia*

[*] *Corresponding author: peanyhoung@itc.edu.kh*

Abstract: *Dishwashing liquid plays a crucial role in daily life, primarily in eliminating soil, mainly food residue from the kitchenware surface. Dishwashing liquid is formulated with various ingredients, focusing on their specific functions including surfactants, chelating surfactants, pH adjustment reagents, thickeners, water, and other ingredients. This study focuses on developing of dishwashing liquid by investigating the effect of surfactants, natural flavors, and an enzyme on the quality and performance of dishwashing liquid. The formulation was contributed by using various surfactants with a ratio of 10 %w/w, lime, kaffir lime and pineapple as natural flavors, and an enzyme. While quality parameters such as pH, viscosity, stability, density, and foam stability were analyzed. Moreover, the performance of dishwashing liquid was determined by its capacity to clean kitchen soil on two types of kitchenware. Based on the result, surfactant affected the cleaning efficiency. The combination of Sodium Lauryl Sulphate and Cocamidopropyl Betaine provided the highest cleaning capacity. In addition, the type of plate did not affect the cleaning capacity, while only kitchen soil and active ingredients were the main factors affecting the cleaning capacity. The findings of our study enhanced the essential role played by the effect of active ingredients (type of surfactants, natural flavor, and enzyme) on dishwashing liquid. This development serves consumers with the right decision in choosing dishwashing liquid with a priority on cleaning efficiency.*

Keywords: Dishwashing liquid; Surfactant; Kitchen soil; Cleaning capacity; Natural flavors.

1. INTRODUCTION

Dishwashing liquid plays a crucial role in daily life in cleaning the soil, mainly food residue from kitchenware surfaces, including many types of plates such as porcelain plate, stainless-steel dera bowl, etc. To enhance and guarantee the effectiveness of dishwashing liquid, the active ingredients must be considered and determined carefully. In the elimination of kitchen soil, surfactants play a crucial role in the cleaning ability and enhance the quality of dishwashing liquid. Mostly, nonionic surfactants and amphoteric surfactants have the main function in cleaning the food residue or grease on the kitchenware [1]. It also generates the suds or foam, which are the key for consumers to indicate the endpoint on the cleaning power of dishwashing liquid. However, there is limitation dosages of using the surfactants in the dishwashing formulation since there are many ingredients need to be used to produce the dishwashing liquid. To enhance the physiochemical quality, natural flavors such as lime, kaffir lime, pineapple are considered as an additive, which could stand as a natural pH adjuster, natural fragrance and colorant. Moreover, enzyme also strengthen the efficiency by breaking down the greasy oil, which contains protein, fat, and carbohydrate.

The purpose of this study focuses on the investigation of type of active ingredients such as surfactants, natural flavors (lime, kaffir lime, and pineapple), enzyme (Ecoenzyme DET L3) on the physicochemical property and performance of dishwashing liquid to clean kitchenware. Three types of surfactant such as Sodium Lauryl Sulphate (SLS), Sodium Laureth Sulphate (SLES), and Cocamidopropyl Betaine (CAPB) were selected to use in the formulation of dishwashing liquid in this work.

[1*] Corresponding author: Peany Houng
E-mail: peanyhoung@itc.edu.kh; Tel: +855 17 959 492

2. METHODOLOGY

2.1 Formulation of dishwashing liquid

Table 1 shows the concentration of surfactants, natural flavors, and enzyme used in the formulation of dishwashing liquid. Total of 10%w/w of either single surfactant, and surfactant mixtures were used in the formulation containing other fixed ingredients such as Sodium Hydroxide), Ethylenediaminetetraacetic Acid (EDTA), Sodium Chloride, MethylChloroisothiazolinone, Citric acid, Ecoenzyme DET L3, Hydroxypropyl Methylcellulose, and Water.

Dishwashing liquid containing two mixture of surfactants (SLS and CAPB) was used to add natural flavors. 5% w/w of lime, kaffir lime, and pineapple juices were used separately in the formulation as seen in formulation 8 to 10 (F8, F9, F10) (see Table 1). While, enzyme with 0.5% w/w was also used in the dishwashing liquid containing mixture of SLS and CAPB.

Table 1. Type and concentration of active ingredients used in dishwashing liquid formulation

Nº	Formulation	Ingredients	Ratio (% w/w)
SO*	F1	SLS	10
	F2	SLES	10
	F3	CAPB	5:5
	F4	SLS+SLES	5:5
	F5	SLS+CAPB	5:5
	F6	SLES+CAPB	5:5
	F7	SLS+SLES+CAPB	4:3:3
SN**	F8	SLS+CAPB+PA[1]	5:5:5
	F9	SLS+CAPB+L[2]	5:5:5
	F10	SLS+CAPB+KL[3]	5:5:5
SE***	F11	SLS+CAPB+E[4]	5:5:0.5

*OS :Surfactant only,**SN :Surfactant + Natural flavors, ***SE :Surfactant + Enzyme, [1] PA: Pineapple, [2]L :Lime, [3]KL :Kaffir Lime, [4] E: Enzyme.

2.2 Preparation of kitchen soil

All components in **Table 2** were used to prepare kitchen soil.

- All fats were premixed together such as lard or soybean oil, or butter and steric acid by low heating until homogenous condition was formed.
- 10% w/w of all-purpose-flour was added into the solution and mixed it well using a digital mixer (OST 20 digital).
- Hot water was added to fulfill 100% ratio of total ingredients.
- The kitchen soil was cooled down at room temperature without forcing
- It was stored in refrigerator overnight for further use [2].

Table 2. Formulation of kitchen soil

Components	% w/w	
	Soil 1	Soil 2
Soybean oil	9	-
Lard	-	9
Steric acid	1	1
All-purpose flour	10	10
Water	80	80

2.3 Analysis physicochemical property of dishwashing liquid

pH

Dishwashing liquid was diluted with water to create 10% solution of dishwashing liquid. Then, a digital pH meter (AS600) was used to measure the pH of the sample [3].

Viscosity

The viscosity of dishwashing liquid was determined by using a viscometer (NDJ-8S) with different rotors and speeds to ensure the percentage between 10-90%. The highest percentage will be selected for the viscosity value.

Density

10 ml of dishwashing liquid was weighted its mass. Then, the density was calculated through Eq. 1.

$$\rho = \frac{M_f}{V} \qquad \text{(Eq. 1)}$$

ρ denoted as density (g/mL), M_f denoted as a mass of dishwashing liquid (g), and V denoted as the volume of each formulation (mL).

Foam stability

Cylinder shake was a method to measure the foaming stability [4]. 50 ml of 1% dishwashing liquid solution was put into a 250 ml graduated cylinder. The solution was shaken for 10 times. Then, it was left it for 1 min, and the initial height of foam generated was measured by a ruler. After 4 min, the final height of foam was measured. The foam stability was calculated by Eq. 2.

$$\%FS = \frac{h_f}{h_i} \times 100 \qquad \text{(Eq. 2)}$$

FS denoted as foam stability (%), h_f denoted as final height at 4 min (cm), and h_i denoted as initial height at 1 min (cm).

Physical stability

To measure the stability, 2 ml of dishwashing liquid was poured into a mini-tube and it was centrifuged using a minicentrifuge machine (ASONE, HSC 12000) at a rotation speed of 6000 rpm for 10 min. The stability of dishwashing liquid was determined through phase separation. No phase separation after the centrifugation, it indicates that the sample is stable in term of physical stability [5].

2.4 Determination of cleaning capacity of dishwashing liquid

To perform cleaning capability, approximately 5g of prepared kitchen soil (soil 1 and soil 2) were applied on the surface of kitchenware (Porcelain Plate and Stainless-Steel Dera bowl) and maintained them for at least 10 minutes as described. in **Figure 1**.

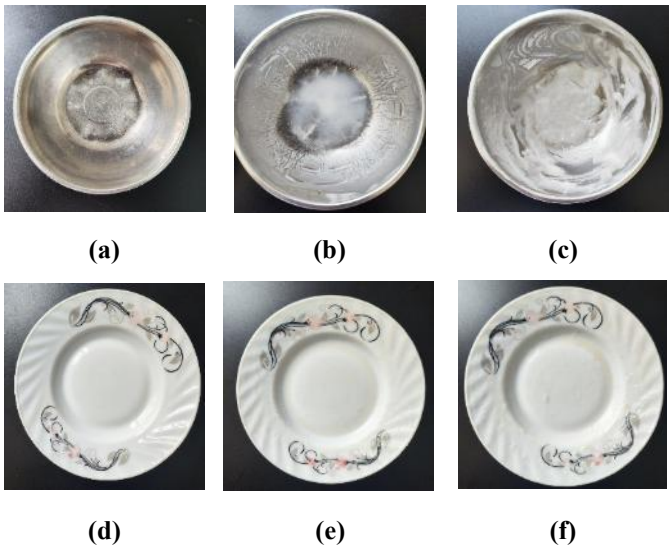


Fig. 1. Preparation of kitchen soil to perform cleaning capacity (a) stainless-steel dera plate, (b) stainless-steel dera plate with soil 1 (c) stainless-steel dera plate with soil 2, (d) porcelain plate, (e) porcelain plate with soil 1, (f) porcelain plate with soil 2

The dishwashing liquid solution was then made in a 1:50% w/w ratio (dishwashing liquid:water). The plate was immersed in a dishwashing detergent solution for a few seconds. To eliminate kitchen soil, six circulates on the front surface and three on the back surface were conducted using a sponge. The cleaned plate was numbered after the foam had decomposed permanently. [2].

3. RESULTS AND DISCUSSION

3.1 Effect of active ingredient on physicochemical property of dishwashing liquid

Table 3 provides the physicochemical property (pH, viscosity, density) of dishwashing liquid affected by various active ingredients. Results showed that pH is high for the dishwashing liquid prepared with mixture of three surfactants such as SLES, SLS, and CAPB. While adding natural flavors into the dishwashing liquid led to lower the pH, which is empowered by lime and kaffir lime. Lime and kaffir lime was known as citrus fruit [6], it had a citric acid that help lower pH. On the other hand, pineapple increase the pH of the dishwashing liquid containing SLS and CAPB due to the presence of an enzyme which is called "Bromalain" in the pineapple [7]. Similarly, using enzyme with concentration of 0.5% w/w also led to increase pH slightly of the dishwashing liquid.

Table 3. Effect of active ingredient on pH, viscosity, density on dishwashing liquid

Nº	pH	Viscosity (mPas)	Density (g/ml)
F1	9.437±0.06	1927±2.83	1.04±0.01
F2	8.787±0.06	257.75±6.01	1.02±0.01
F3	8.253±0.06	10.85±0.21	1.01±0.00
F4	7.707±0.06	449.95±2.90	1.00±0.00
F5	9.127±0.06	11950±98.99	1.11±0.03
F6	9.267±0.06	6315±91.92	1.07±0.01
F7	9.527±0.015	3100±42.43	1.05±0.04
F8	9.267±0.01	8845±7.07	1.05±0.02
F9	5.753±0.015	14540±311.13	1.07±0.02
F10	5.757±0.06	14950±98.99	1.05±0.03
F11	9.233±0.06	7845±21.21	1.05±0.01

Viscosity of dishwashing liquid had wide ranges in between 10 to 14950 mPas. Results showed that surfactant type affected the viscosity of dishwashing liquid. Viscosity value of single SLES, CAPB, and mixture between SLS with SLES tended to be low because SLS and SLES unreacted with each other. Interestingly, the presence of CAPB in combination of surfactants resulted in increased viscosity due to the presence of betaine strcuture. The betaine have positive charge associated to negative charge of anionic surfactant (SLS, SLES) and generate more viscosity [8]. Adding natural flavors (lime and kaffir lime juices) in dishwashing liquid improved viscosity. While adding pineapple juice and enzyme in dishwashing liquid resulted in the decreased viscosity value owning to the fact that in pineapple also contains proteolytic enzyme, which was the key in decreasing viscosity.

Result of density varied in between 1.00 to 1.11 g/ml. This results showed that surfactants type can be the main role affecting on density. Single surfactants (SLS, SLES, CAPB), combination of surfactant (SLS and SLES) tend to be lower the density while the presence of CAPB bind with SLS provided denser in the dishwashing liquid. However, after adding natural flavors and enzyme in the combination of SLS and CAPB, density decrease slightly which is the same as viscosity value of the dishwashing liquid.

Figure 2 provides the effect of active ingredients on foam stability in dishwashing liquid. Single of SLES, CAPB and combination between SLS and SLES were decreased rapidly after 5 min. Combination of SLS and CAPB, and enzyme gave more stable foam which is around 97% to 99% compare to those individual surfactants. In reason that, SLS and CAPB are anionic, amphoteric surfactant, which rich in foam [8]. In contrast, natural flavors (pineapple, lime and kaffir lime) reduced the stability of foam .

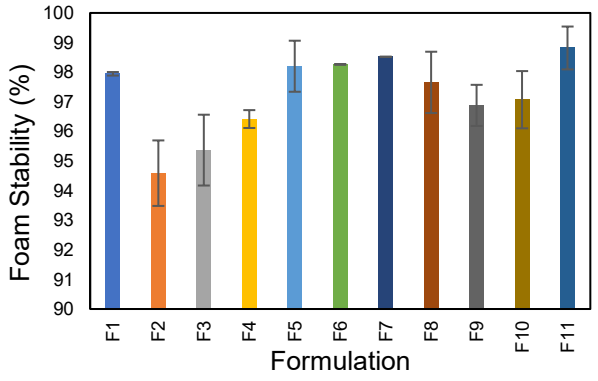


Fig. 2. Effect of active ingredient on foam stability of dishwashing liquid

Figure 3 demonstrate the physical stability of all formulated dishwashing liquid. It show that the active ingredients had no influence on the physical stability of the dishwashing liquid, as each tubes exhibited no phase separation.

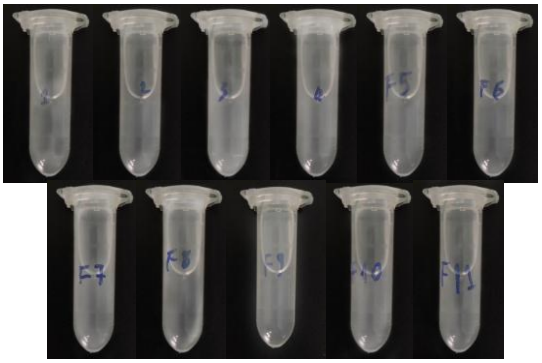


Fig. 3. Effect of active ingredients on physical stability on dishwashing liquid after centrifuged at 6000rpm, 10mins.

3.2 Performance of dishwashing liquid

Figure 4 illustrates the results of cleaning capacity for liquid against two type of soils (soil 1 and soil 2) on two surfaces kitchenware such as Stainless-Steel dera bowl and Porcelain plate. Result showed that surfactants type affected on cleaning capacity. As seen in **Figure 4a** and **Figure 4b**, single SLS, SLES, CAPB, and mixtures of surfactants (SLS:SLES, SLES:CAPB, SLS:SLES:CAPB) showed lower cleaning capacity for the tested plates than the combination of SLS with CAPB. SLS with CAPB was empowered in cleaning performance because the adsorption leads to increase by rolling-up the food residue soil on the plate [9]. Moreover, adding natural flavors were remain unchanged in performance of dishwashing liquid on the tested plates. In contrast, adding enzyme were more effective in cleaning capacity in reason that enzyme contain protease, amylase, lipase, which support breaking down the protein, fatty oil, and carbohydrate in kitchen soil [10].

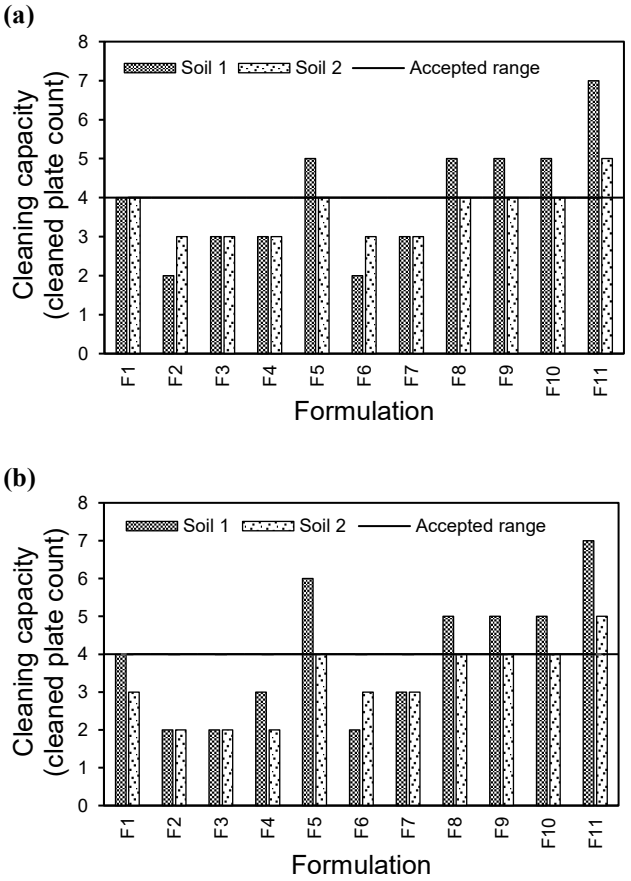


Fig. 4. Effect of active ingredient on cleaning capacity of dishwashing liquid (a) Porcelain plate, (b) Stainless-steel dera bowl

Furthermore, the study found that the type of kitchen soil affected the cleaning ability of dishwashing liquid. Soil 1 and 2 were treated with soybean oil and lard, respectively. Soil 2

had a stronger influence than soil 1 because it included lard, which contains more saturated and unsaturated fat than soybean oil [11]. Therefore, the results proved that different types of kitchenware surfaces (stainless steel dera bowl and porcelain plate) had no effect on dishwashing liquid cleaning capacity.

4. CONCLUSIONS

Surfactants and enzymes are crucial ingredients in formulation of dishwashing liquid because they assisted to break down food residue soil on kitchenware, leading to improve cleaning capacity. While natural flavor has no power on cleaning capacity, it possesses a significant impact on physicochemical qualities, acting as a pH adjuster, viscosity, colorant, and natural smell. Furthermore, the surface of various types of kitchenware has no effect on cleaning capacity; rather, soil is the key factor in minimizing the efficacy of surfactants and enzymes in dishwashing solutions. The results of this study indicated that the formulation of dishwashing liquid could be varied depending on its intended application.

ACKNOWLEDGMENTS

The author wishes to express gratitude to LSV Industry Company for supplying an ingredient and providing laboratory to conduct the experiments on the dishwashing liquid development.

REFERENCES

[1] J. Shi *et al.*, "Basic Building Blocks of Hand Dishwashing Detergents," in *Handbook of Detergents: Part D Formulation*, vol. 128, 1 ed. USA: CRC Press, 2006, ch. Dishwashing Detergents for Household Applications pp. 112-115.

[2] ASSOCASA and INNOVHUB, "Standard protocol for evaluating performances of hand dishwashing detergents," in *HPC Today- Household and Personal Care Today*, vol. 18: tks, 2023, ch. Home Care: Laundry & Cleaning, pp. 44-47.

[3] A. R. Mainkar and C. I. Jolly, "Evaluation of commercial herbal shampoos," *International Journal of Cosmetic Scicence,* vol. 22, no. 5, pp. 385-391, 2000.

[4] K. Klein, "Evaluating Shampoo Foam," *Klein K. (2004). Evaluating Shampoo Foam,* vol. 119, no. 10, pp. 32-35, 2004.

[5] R. Smadi *et al.*, "Liquid Detergent Compostion," *U.S. 6 376 446 B1,* p. 10, 2002.

[6] J. Forsyth and J. Damiani, "Citrus Fruit," *Encyclopedia of Food Sciences and Nutrition,* pp. 1329-1335, 2003.

[7] C. Varilla *et al.*, "Bromelain, a Group of Pineapple Proteolytic Complex Enzymes (Ananas comosus) and Their Possible Therapeutic and Clinical Effects. A Summary," *Foods,* vol. 10, 2021, doi: https://doi.org/10.3390/foods10102249.

[8] K. Klein and I. Palefsky, "Shampoo Formulation," in *Handbook for Cleaning/Decontamination of Surfaces*, vol. 1, I. Johansson and P. Somasundaran Eds., 2007, pp. 277-304.

[9] S. N. Blagojevic´ *et al.*, "Performance and Efficiency of Anionic Dishwashing Liquids with Amphoteric and Nonionic surfactants," *Journal of Surfactants and Detergents,* vol. 19, pp. 363-372, 2016.

[10] A. Naganthran *et al.*, " Improving the Efficiency of New Automatic Dishwashing Detergent Formulation by Addition of Thermostable Lipase, Protease and Amylase," *molecules,* vol. 22, no. 9, p. 1577, 2017, doi: doi:10.3390/molecules22091577.

[11] E. Rendian-Ruedy and B. J. Smith, "Methodological considerations when studying the skeletal response to glucose intolerance using the diet-induced obesity model," *International Bone and Mineral Society,* 2016, doi: 10.1038/bonekey.2016.71.

Mechatronic and Information Technology (MIT)

2 Sub-parallel Sessions

"MIT-1: AI-Driven Digital Transformation"

Chair: Assist. Prof. Dr. VALY Dona, Co-Chair: Assist. Prof. Dr. PHAUK Sokkhey

No.	Topic
1	Uncovering Factors Contributing to Cambodian High School Students' Mathematics Achievement through Machine Learning Algorithms *Authored by: Tong LY, Sokkhey PHAUK, Sothea HAS, Dona, VALY*
2	Harnessing Predictive Analytic for Detecting Students' Learning and Behavioral Engagement in ITC Moodle Platform *Authored by: Chentra PEN, Sokkhey PHAUK, Sothea HAS, Dona VALY*
3	Applying Deep Learning Approaches to Khmer Speech Recognition and Text-to-Speech in Conversational Domains *Authored by: Prakmlis SOENG, Dona VALY, Phutphalla KONG*
4	AI-Based Skill Trend Analysis: Enhancing Job Market Insights and Curriculum Development in Cambodia *Authored by: Sela SONGEAM, Dona VALY, Sothea HAS*
5	Lip-Synchronized Talking Avatar Generation for Khmer Language Using Audio-to-Landmark Modeling *Authored by: Sovanvichea YANN, Dona VALY, Phutphalla KONG*
6	KHR/USD Daily Exchange Rate Short Term Prediction using Deep Learning *Authored by: Hourpheng Hort, Sokkhey Phauk, Pheak Neang, Vesal Khean*
7	A Multilingual Deep Learning Approach Fine-tuning to Khmer News Summarization *Authored by: Soknan KOUM, Kimlong NGIN, Chanthin PHAO, Sane BO, Ousa MA*
8	Handwritten Khmer Digit Recognition Using Sequential RNN Models *Authored by: Kimlong Ngin, Dona Valy, Kimhor Phoeurn , Seyha Va, Bunkhloem Pel, Kana Ty, Sambat Yin, Rotha Pho*
9	Analysis on Machine Learning Models for Imbalanced Data Problem in Payment Fraud Detection *Authored by: Siuphing SEUN, Sokkhey Phauk, Kimlong NGIN, Pakrigna LONG*
10	Interacting Neural Networks: A Comparative Study of Standard and Distributed Architectures on Image Classification *Authored by: Sreyvi UANN, Timoteo CARLETTI, Sokkhey PHAUK, Jarod KETCHA KOUAKEP, Malai NHIM*
11	Enhancing Khmer Language AI: A Privacy-Focused Chatbot with Retrieval-Augmented Generation *Authored by: Nita Li, Seyha Heng, Lay Seng, Jihyun Moon, Sokkhey Phauk, Charmgil Hong*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Uncovering Factors Contributing to Cambodian High School Students' Mathematics Achievement through Machine Learning Algorithms

Tong LY [1*], Sokkhey PHAUK [2], Sothea HAS [3], Dona, VALY [4]

[1*] *Institute of Humanities and Social Sciences, Royal Academy of Cambodia, Russian Federation Blvd., P.O. Box 531, Phnom Penh, Cambodia*

[2,3] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[4] *Department of Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: lytongvps2024@gmail.com*

Abstract: *The educational landscape of Cambodia has witnessed significant transformations over the past decade, yet challenges remain, especially in the teaching and learning of Science, Technology, Engineering and Mathematics (STEM) subjects. The current study aims at detecting important features contributing to Cambodian high school students' mathematics achievement and validating the best machine learning and deep learning models that could accurately identifying at-risk students. Tree-based machine learning and deep learning models including decision tree (DT), random forest (RF), AdaBoost, Bagging, XGBoost, Gradient Boosting Machine (GBM), Super Learner, Artificial Neural Network (ANN), Deep Neural Network (DNN) and ensemble DNN were proposed for experimentations. The dataset contains factors influencing high school students' mathematics learning. Synthetic Minority Over-sampling Technique (SMOTE) was employed to handle imbalance data. The findings from DT, RF, AdaBoost and GBM indicated the significance of attending extra-paid classes, learning engagement and the students' attitudes toward mathematics attainment. The validation of best models using the significant features with all proposed models indicated that RF and Bagging methods outperformed other models with accuracies of 82.0%. ANN and DNN did not perform well in this task with accuracies 70.0% and 76.0% respectively. Super Learner, XGBoost and Ensemble DNN were not yet tested due to the limitation of computer's capacity. The findings from this study not only aim to enhance the theoretical understanding of the determinants of students' mathematical achievement in Cambodia, but also provide implications for the application of machine learning and deep learning techniques for the improvement of mathematics education at high school and beyond.*

Keywords: machine learning; deep learning; feature importance; SMOTE, mathematics learning

1. INTRODUCTION

1.1 Research background

In the era of technology, Science, Technology, Engineering and Mathematics (STEM) education has become the primary focus for educators throughout the globe. Cambodia, without exception, has also focused on STEM education where the mission spans more than a decade [1]. According to Kao [2], the demand for qualified trained workforce in these areas has become one of the priority concerns among educators and industries alike. Despite much focus and efforts have been put in strengthening Cambodian STEM education, the students' performance and their interests in these subject are not satisfactory [3]–[7].

In this study, the focus is on the application of advanced predictive analytics algorithms to understand the students' learning in high school mathematics. Existing literatures have indicated that the students' learning in this subject has become critical concerns; however, there is not much research on the issues. Understanding the characteristics of the students' learning, the underlying factors contributing to their success or failure in mathematics will contribute significantly to the students' learning in this subject. In this regard, Alsariera et al. [8] mentioned that comprehensive understanding about the student learning is vital for enabling educators to identify

[1*] Corresponding author: Tong LY
E-mail: lytongvps2024@gmail.com; Tel: +855-12 250 887

students' weak points and assist them in improving their academic achievement.

Though without much study in the context of Cambodia, in educational setting, numerous studies have highlighted various factors including individual, family, school and teacher factors etc. [9][10] that directly or indirectly contributed to the students' learning, not necessarily mathematics. In the age of technology, digital tools also play crucial role in the students' learning [11][9], especially in the age of online learning and eLearning resources are abundance. Moreover, Al-Alawi et al. [10] indicated that amount of time spent studying at the university and prior performance in secondary school also link to their performance.

1.2 Machine learning in education

Machine learning is emerging nowadays as an important tool for decision support in many areas of research. In the field of education, both educational organizations and students are the target beneficiaries [12]. By harnessing a robust dataset drawn from diverse educational backgrounds, incorporating different factors including academic data, socioeconomic demographics, and individual learning interactions would provide significant insight into students' learning. In addition to this, R. Abdrakhmanov *et al.* [13] argued that, when utilized in education, machine learning models have ability to convert large, complex datasets into actionable insights about student performance. In this context, machine learning models such as linear regression, KNN, NB, have become more popular and have been tested within educational contexts, each offering distinct advantages [14].

Model like Random Forest has ability to provide insight into feature importance, which lead to comprehensive understanding of influential factors affecting the student performance. Neural Networks, on the other hand, could capture non-linear relationships among data. Given the varied nature of educational data, recent studies have advocated for ensemble and hybrid models [15] to improve accuracy of model prediction.

Table 1: related review on related applications of machine learning and deep learning models in educational data

Reference	Data	Algorithms	Accuracy
Phauk et al. [16]	Survey data Sample: 1204	ANN, SMO, LR, 1-NN, Boosted C5.0, Bagged CART, **RF**, DBN	48.5%-96.69%
Huynh-Cam et al. [17]	Survey data Sample: 2407	RF, C5.0, **CART**, MLP	69.02%-80.00%
Jedidi et al.[18]	Survey data Sample: 480 Features: 16	**RF**, LR, XGBoost, MLP, Bagging, Voting	67.5%-77.0%
Lu et al. [19]	Survey data Sample: 1,101	LOG, SVC, RFC, **XGBoost**	85.16%-89.22%

These studies validated the performance of machine learning models, especially ensemble methods in identifying the significant features contributing to students' learning.

1.3 Research problem

Uncovering the factors that hinder or influence the students' learning of mathematics is initial step in enhancing the students' learning. Understanding the characteristics of these features help teachers, educators and policymakers to tailor educational interventions to improve the students' learning in mathematics. In Cambodia, students' learning outcomes were often studied using conventional metrics—standardized testing, classroom participation, and basic performance tracking methodologies with the utilization of basic statistical methods or linear models. Existing literature covered aspects such as motivation, learning strategies, while simple statistical methods, for example, t-test, ANOVA were utilized [20]. There exist some studies in mathematics education which utilized machine learning and deep learning approaches to predict the students' learning in mathematics [18][23]–[27]; however, these studies did not directly detect important features.

1.4 Objectives

This study utilizes the power of machine learning and deep learning to understand the characteristics of students' learning characteristics. The comparison of various tree-based machine learning and deep learning were experimented with the educational dataset collected using survey questionnaire. With this purpose, the study possesses the following two objectives:

- To identify the features that significantly influence Cambodian high school students' mathematics achievement using machine learning techniques.
- To validate machine learning and deep learning models that could accurately predict and classify at-risk students based on the significant features.

2. METHODOLOGY

2.1 Research processes

The study followed steps presented in the following figure. Various experiments of the machine learning models were conducted to identify the features that significantly contribute to the students' mathematics achievement.

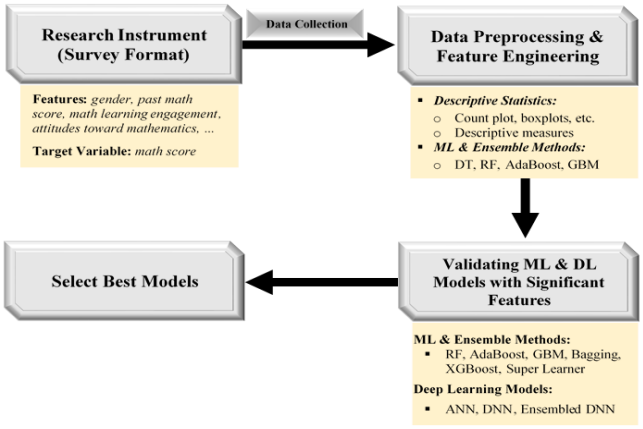


Fig 1: Research processes carried out in the study.

2.2 Research dataset

The dataset in this study was from an empirical study on '*factors influencing the Cambodian high school students' mathematics achievement'*. The features include the variables about the students' demographic information, their learning engagement and their attitudes toward mathematics[2]. All features were categorically measured while the target variable was in numerical value.

2.3 Proposed models

Several tree-based machines learning and deep learning models presented in the following table were experimented.

Table 2: Proposed models for the study

Objective 1	Objective 2
Decision Tree	Random Forest
Random Forest	AdaBoost
AdaBoost	Gradient Boosting Machine
Gradient Boosting Machine	Bagging Method
	XGBoost Model
	Super Learner
	Artificial Neural Network
	Deep Neural Network
	Ensembled DNN

2.4 Model performance metrics

Each model's performance was assessed using accuracy f1-score, recall and precision metrics which based on the following table and equations.

Table 3: Confusion matrix structure

	Predicted Positive	Predicted Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

[2] In this study, the attitudes toward mathematics were restricted to negative emotion or namely the students' math anxiety.

Note: TP: True Positive, TN: True Negative, FP: False Positive and FN: False Negative.

The model performance metrics were calculated using the equations below:

- $Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$ (Eq. 1)
- $Recall = \frac{TP}{TP+FN}$ (Eq. 2)
- $Precision = \frac{TP}{TP+FP}$ (Eq. 3)
- $F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall}$ (Eq. 4)

2.5 Data preprocessing

The dataset was explored using different Exploratory Data Analysis (EDA) methods including data visualization, statistical measures. From the EDA, '*SchType*' was removed from the dataset. '*SelfLearnPerWeek*' was regrouped into '*Less than 3H and More than 3H*', 'Attendance' was regrouped into '*Always, Often and Less Frequent*' due to sample bias among categories these variables (**Fig 2**).

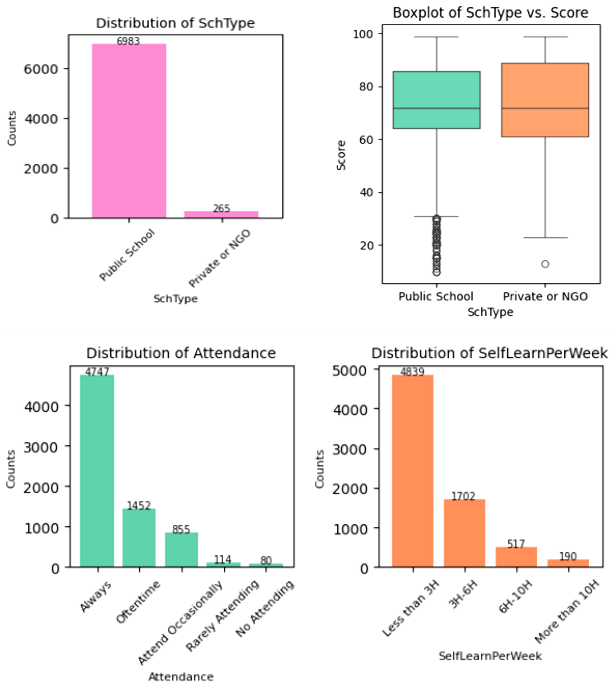


Fig 2: Variable distributions leading to removal and regrouping of the categories.

To apply the machine learning models, the target variable: 'Score' was categorized into '*Poor Learner, At-risk, Satisfactory* and *High Performing*'. There were imbalanced classes among target variable and to addressing the issue, Synthetic Minority Over-sampling Technique (SMOTE) was

implemented to address minority groups. The pre-SMOTE sample size, 7,248 students, had increased to 15,752 after SMOTE was applied.

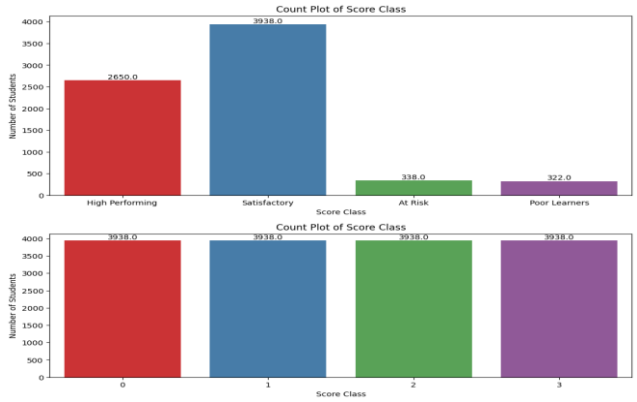


Fig 3: Sample sizes before and after implementing SMOTE

2.6 Model's experiments setup

The dataset was split into train, validation and test sets with 80:10:10 ratio with respective sub-sample sizes 12,601, 1,575 and 1,575.

- *For Objective 1:*

The decision tree was trained to check feature importance with its default methods, MDI and PFI for the basis of feature removal in the main experiments.

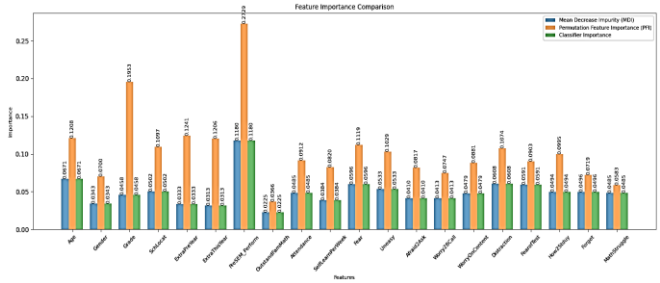


Fig 4: Feature importance from preliminary analysis of decision tree

The main experiments of the four tree-based machine learning models: DT, RF, GBM and AdaBoost were conducted with six sets where certain feature(s) was/were removed based on the contribution to the models presented in Fig 4. Feature(s) with less contribution to the model was/were subjected to be removed for each set presented below:

Table 4: Sets used in experiments to select features contributing to students' mathematics achievement

Feature Set	Feature(s) removed
Set 1	None (all included)
Set 2	OutstandFamMath
Set 3	Gender
Set 4	ExtraPreYear
Set 5	SelfLearnPerWeek
Set 6	Gender, ExtraPreYear, OutstandFamMath, SelfLearnPerWeek

In the experiments to select the best set and identify the significant features, MDI and PFI feature methods were used as the basis for feature selection.

- *Objective 2:*

The proposed machine learning and deep learning models were trained only on the selected set which provided the highest model performance based on Objective 1.

3. RESULTS AND DISCUSSION

The experiments revealed inconsistent accuracies and features importance among the four models (Fig 5). Random Forest (RF) provided the highest accuracy on the six sets.

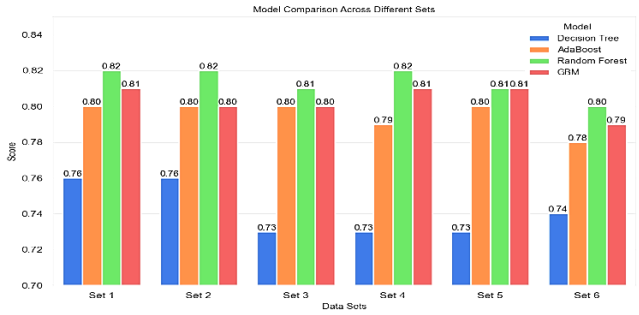


Fig 5. The comparison of the models' accuracies across the six feature sets

The best performing of these four models were on Set 1 with DT (76.0%), AdaBoost (80.0%), RF (82.0%) and GBM (81.0%). The significant factors contributing to the students' mathematics achievement were selected from the feature importance of Random Forest and only feature importance in Set 1 were presented.

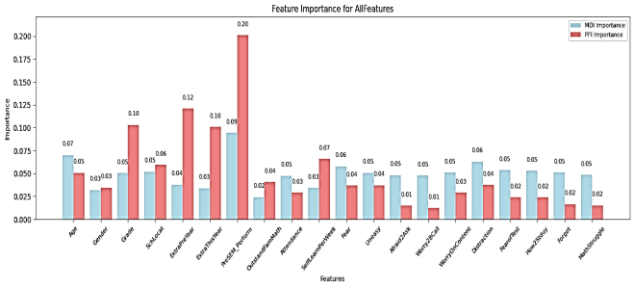


Fig 6. Feature importance from Random Forest for Set 1

From this results, the students' past achievement played the most significant role in their future attainment while their learning engagement in mathematics contributed to their learning outcome. Extra-paid classes emerged as one of the key players in students' math achievement. Compared with other features, factors related to attitudes toward mathematics play less significant role.

Fig 7 and Table 5 present the comparison analyses of the proposed models with the selected feature set (Set 1).

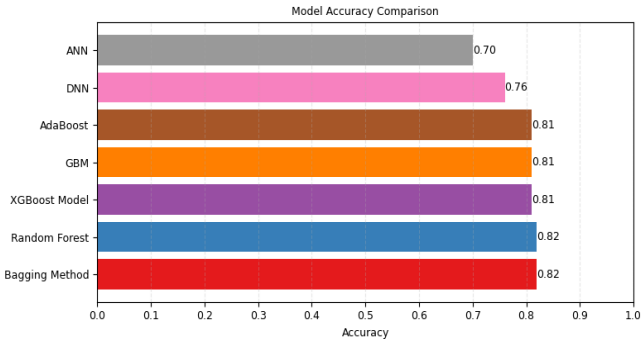


Fig 7. The comparison of the models' accuracies across the six feature sets

Table 5: The comparison of the models' accuracies

Model		F1	Recall	Precision
AdaBoost (acc. 0.81)	Macro average	0.81	0.81	0.81
	Weighted average	0.81	0.81	0.81
Random Forest (acc. 0.82)	Macro average	0.81	0.82	0.81
	Weighted average	0.82	0.82	0.82
GBM (acc. 0.81)	Macro average	0.81	0.81	0.81
	Weighted average	0.81	0.81	0.81
Bagging Method (acc. 0.82)	Macro average	0.82	0.82	0.82
	Weighted average	0.82	0.82	0.82
XGBoost Model (acc. 0.81)	Macro average	0.81	0.81	0.81
	Weighted average	0.81	0.81	0.81
ANN (acc. 0.70)	Macro average	0.70	0.71	0.70
	Weighted average	0.70	0.70	0.70
DNN (acc. 0.76)	Macro average	0.69	0.70	0.70
	Weighted average	0.69	0.70	0.69

Random forest and bagging methods emerged as the best models outperformed other algorithms with the accuracy at 82.0%. This finding was in line with Phauk et al. [16] and Jedidi et al. [18] and indicated the power of RF in educational data mining. In relation to this, Chandel & Arora [26] and ŞEVGİN [27] also validated the power of other ensemble learning such as Bagging, Boosting and Stacking.

4. CONCLUSIONS

The current study utilized the power of machine learning and deep learning to identify the significant features contributing to the students' mathematics achievement and predict their learning outcome and identify at-risk students. Tree-based machine learning models such as Decision Tree, Random Forest, AdaBoost and Gradient Boosting Machine were applied to select features that significantly contributing to students' mathematics achievement. Several features including demographic variables, students' learning engagement and their negative math attitudes toward mathematics were found to have significantly affect their mathematics learning.

Furthermore, the research successfully validates the predictive power of various machine learning and deep learning models in predicting mathematics learning outcome and classifying at-risk students. Random Forest and Bagging methods outperformed other algorithms providing accuracy at 82% while other models provide accuracies ranging from 70% to 81%. Deep learning models, ANN and DNN, did not perform well compared to machine learning in this task. The reason for this underperforming of neural network models may be related to the number of features in the data.

REFERENCES

[1] S. Kao and S. Kinya, "A Review on STEM Enrollment in Higher Education of Cambodia: Current Status, Issues, and Implications of Initiatives," Int. Dev. Coop., vol. 26, no. 1&2, pp. 123–134, 2020, [Online]. Available: https://www.researchgate.net/publication/339815574_A_Review_on_STEM_Enrollment_in_Higher_Education_of_Cambodia_Current_Status_Issues_and_Implications_of_Initiatives

[2] S. Kao, "Family socioeconomic status and students' choice of STEM majors: Evidence from higher education of Cambodia," Int. J. Comp. Educ. Dev., vol. 22, no. 1, pp. 49–65, 2020, doi: 10.1108/IJCED-03-2019-0025.

[3] JICA, "Data Collection Survey on Human Resource Development for Industrialisation in the Education Sector in the Kingdom of Cambodia Final Report," p. 242, 2016.

[4] C. Ban, "The Current State of Cambodia's STEM Education: A Case Study of the Preah Sisowath New Generation School," Phnom Penh, Cambodia, 2020.

[5] L. Phirom, K. Sothy, R. K. Chhem, and G. Smith, "De-framing STEM discourses in Cambodia," no. 127, 2021, [Online]. Available: https://www.cdri.org.kh/storage/pdf/WP 127E_De-framing STEM_1622794628.pdf

[6] L. Phirom, "Student Engagement in STEM Education: Global Trends and Implications for Cambodia," Cambodia Dev. Rev., vol. 20, no. 3, pp. 6–11, 2016.

[7] P. Vann, "STEM Education in Cambodia : Progress and

Challenges," in Innovations and challenges in Cambodian education: Youth's perspectives, 2023, pp. 113–130.

[8] Y. A. Alsariera, Y. Baashar, G. Alkawsi, A. Mustafa, A. A. Alkahtani, and N. Ali, "Assessment and Evaluation of Different Machine Learning Algorithms for Predicting Student Performance," Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/4151487.

[9] A. Melkumyan and M. Sahakyan, "Psychological and educational aspects of studentsadaptation in the context of digitalization," BIO Web Conf., vol. 120, pp. 1–11, 2024, doi: 10.1051/bioconf/202412001063.

[10] L. Al-Alawi, J. Al Shaqsi, A. Tarhini, and A. S. Al-Busaidi, "Using machine learning to predict factors affecting academic performance: the case of college students on academic probation," Educ. Inf. Technol., vol. 28, no. 10, pp. 12407–12432, 2023, doi: 10.1007/s10639-023-11700-0.

[11] T. B. Taha and M. Salim, "The Impact of Technology on Students' Psychological and Educational Performance," vol. 06, no. 01, pp. 91–95, 2023.

[12] P. Balaji, S. Alelyani, A. Qahmash, and M. Mohana, "Contributions of machine learning models towards student academic performance prediction: A systematic review," Appl. Sci., vol. 11, no. 21, 2021, doi: 10.3390/app112110007.

[13] R. Abdrakhmanov et al., "Applying Computer Vision and Machine Learning Techniques in STEM-Education Self-Study," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 1, pp. 819–827, 2024, doi: 10.14569/IJACSA.2024.0150182.

[14] R. Abdrakhmanov, A. Zhaxanova, M. Karatayeva, G. Z. Niyazova, K. Berkimbayev, and A. Tuimebayev, "Development of a Framework for Predicting Students' Academic Performance in STEM Education using Machine Learning Methods," Int. J. Adv. Comput. Sci. Appl., vol. 15, no. 1, pp. 38–46, 2024, doi: 10.14569/IJACSA.2024.0150105.

[15] M. A. Aboamer et al., "Investigation in Analyzing the Food Quality Well-Being for Lung Cancer Using Blockchain through CNN," J. Food Qual., vol. 2022, 2022, doi: https://doi.org/10.1155/2022/5845870.

[16] S. Phauk, N. Sin, T. Ly, and O. Takeo, "Multi-models of educational data mining for predicting student performance in mathematics: A case study on high schools in cambodia," IEIE Trans. Smart Process. Comput., vol. 9, no. 3, pp. 217–229, 2020, doi: 10.5573/IEIESPC.2020.9.3.185.

[17] T. T. Huynh-Cam, L. S. Chen, and H. Le, "Using decision trees and random forest algorithms to predict and determine factors contributing to first-year university students' learning performance," Algorithms, vol. 14, no. 11, 2021, doi: 10.3390/a14110318.

[18] Y. Jedidi et al., "Predicting students' academic performance and modeling using data mining techniques," Math. Model. Comput., vol. 11, no. 3, pp. 814–825, 2024, doi: 10.23939/mmc2024.03.814.

[19] J. Lu, Y. Liu, S. Liu, Z. Yan, and X. Zhao, "Machine learning analysis of factors affecting college students ' academic performance," no. December, pp. 1–14, 2024, doi: 10.3389/fpsyg.2024.1447825.

[20] T. Ly, S. Phauk, and S. Chea, "Relationship between Students' Attitudes toward Mathematics and Their Achievement in Probabilities and Statistics: A Case of Engineering Students at ITC," Cambodian J. Humanit. Soc. Sci., vol. 1, no. 1, pp. 36–53, 2022.

[21] S. Phauk and T. Okazaki, "Integration of Educational Data Mining Models to a Web-Based Support System for Predicting High School Student Performance," Int. J. Comput. …, vol. 15, no. 2, pp. 131–144, 2021, [Online]. Available: https://www.researchgate.net/profile/Sokkhey-Phauk/publication/349226410_Integration_of_Educational_Data_Mining_Models_to_a_Web-Based_Support_System_for_Predicting_High_School_Student_Performance/links/6025c8a5299bf1cc26bcd5e7/Integration-of-Educational-D

[22] S. Phauk and O. Takeo, "Hybrid machine learning algorithms for predicting academic performance," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 1, pp. 32–41, 2020, doi: 10.14569/ijacsa.2020.0110104.

[23] S. Phauk and O. Takeo, "Comparative Study of Prediction Models for High School Student Performance in Mathematics," IEIE Trans. Smart Process. Comput., vol. 8, no. 5, pp. 394–404, 2019, doi: 10.5573/IEIESPC.2019.8.5.394.

[24] S. Phauk and O. Takeo, "Development and optimization of deep belief networks applied for academic performance prediction with larger datasets," IEIE Trans. Smart Process. Comput., vol. 9, no. 4, pp. 298–311, 2020, doi: 10.5573/IEIESPC.2020.9.4.298.

[25] S. Phauk and O. Takeo, "Study on dominant factor for academic performance prediction using feature selection methods," Int. J. Adv. Comput. Sci. Appl., vol. 11, no. 8, pp. 492–502, 2020, doi: 10.14569/IJACSA.2020.0110862.

[26] M. Chandel and H. Arora, "A Review on Ensemble Learning Algorithms in Data Mining," vol. 11, no. 6, 2024.

[27] H. ŞEVGİN, "A comparative study of ensemble methods in the field of education: Bagging and Boosting algorithms," Int. J. Assess. Tools Educ., vol. 10, no. 3, pp. 544–562, 2023, doi: 10.21449/ijate.1167705.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Harnessing Predictive Analytic for Detecting Students' Learning and Behavioral Engagement in Moodle Platform

Chentra PEN [1*], Sokkhey PHAUK [1], Sothea HAS [1], Dona VALY[2]

[1*, 1] *Departement of applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2]*Departement of Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: chentrapen168@gmail.com/pen.chentra@itc.edu.kh*

Abstract: *Understanding and monitoring student engagement in online learning platform is essential to enhancing academic results as blended learning become more significant in higher education. This study utilized predictive analytic on 131 student engagement data from the ITC Moodle Learning Management System (LMS). Discovering aspects of student engagement in Moodle and linking them to their academic achievement is the main goal. Features including absent, assignment, viewing course materials, time spent on tests, midterm score and logging in frequently are all important measures of engagement and link to their achievement. This was conducted the Synthetic Minority Oversampling Technique (SMOTE) to face the imbalance classes and followed by training a variety of machine learning algorithms, such as Gradient Boosting, XGBoost, Random Forest, and Logistic Regression, to forecast students' grade of the first semester using engagement data with evaluating these models based on accuracy, precision, recall, and F1 score. The result Gradient Boosting outperformance with accuracy 81%. Apart from predictions, these models additionally highlighted underlying patterns of student behavior, giving teachers important information about the elements that have the greatest impact on academic achievement. The proposed approach makes it feasible to identify at-risk students early on, giving teachers the opportunity to put timely interventions and individualized support plans into place. This project intends to assist in the building of more successful blended learning environments by improving our understanding of how participation translates into learning outcomes, which will eventually promote better academic achievement and student happiness.*

Keywords: Moodle; blended learning; clustering; predictive analytic; feature importance

1. INTRODUCTION

Ensuring universal equitable access to high-quality, inclusive education is one of the globally priorities defined by UNESCO under Sustainable Development Goal 4 (SDG 4). The aim highlights the importance of offering equitable access to education, encouraging opportunities for lifelong learning, and utilizing innovative technology to enhance academic results. To respond to SDG4 and the Cambodia government, Cambodia Cyber University Network (CCUN) the Cambodian Ministry of Education and Sports has been rapidly advancing online learning programs in July 2024, and the integration of artificial intelligence (AI) into learning settings in line with SDG 4 in February 2025.

Blended learning provides flexible and fascinating educational experiences by combining classroom instruction with online learning environments [1] and utilized the Educational Data Mining (EDM) to improve the students' performance [2]. Accurately forecasting academic performance and knowing student engagement habits are essential elements of successful blended learning environments [3] and an early forecasting to detect the student at risk of failing [4] in a blended learning. Yet, rich interaction data from platforms such Moodle LMS with small records and

[1*] Corresponding author: Chentra PEN
E-mail: chentrapen168@gmail.com; Tel: +855-99 996 939

attributes and rarely used in traditional approaches, in particular in Cambodia.

In order to close this gap, this study uses predictive analytics to analyze Moodle student behavior and establish a direct correlation between it and semester results. In the case, we aim to:

Objective1: Determine the patterns of students' leaning engagement in Moodle through recorded data blended classes.

Objective2: Apply the statistical analytics and machine learning algorithms to predict and analyze the students' academic achievement and determine at-risk students from patterns of engagement in Moodle.

2. METHODOLOGY

2.1 Dataset and features

The study utilized a dataset consisting of 131 students enrolled in a blended learning environment from department of Applied Mathematics and Statistics (AMS) of Institute of Technology of Cambodia (ITC). The dataset only included metrics such as gender, number of absent, score of assignment, number of viewing course materials, time spent on tests in second, midterm score and number of logging in, and final semester results.

2.2 Data Preprocessing and Analysis

The raw data collected from the Moodle LMS and mapped to semester result was first cleaned by removing records in rows 10, 72, 91, 99, 113, 114 with missing values and there were no duplicates, ensuring data integrity for reliable modeling. In [5] "to keep the generality, the target variable was separated into four ordinal categories" and in the researcher the final semester result were categorized into five performance levels (A to F).
In order to understand key patterns and "satisfying forecast of accuracy" [5] within student engagement metrics, we conducted exploratory data analysis (EDA). The methods included descriptive statistics to summarize the statistics, distribution visualization to inspect the normality and variability of engagement variables (e.g., gender, absent, assignment, viewing course materials, time spent on tests, midterm score and logging in, and final semester results.). Correlation analyses (Pearson and Spearman) identified important relationships between engagement behaviors and academic outcomes. Additionally, KMeans clustering combined with Principal Component Analysis (PCA) [6] revealed distinct student engagement profiles, guiding subsequent predictive modeling and interpretation.

2.3 Machine Learning Models

We investigated some reviews on machine learning (ML) models and some technical to improve the performance of models which related to our works.

Table 1. related work on ML models and some technicals

Reference	Dataset	Algorithms	Accuracy
"utilized the DL with wide data to enhance students performance ", Aslam et al. [7].	1044 records	DL	0.93-0.96
"Utilized a Random Forest amount of four ML to predict academic performance " , Sokkhey & Okazaki [5].	1204 records	RF	89.23
"To increase the predictive accuracy of student performance .", Asselman et al. [8]	9407 records	LiR, RF, GrB, Adaboost	0.71-0.78
"Forcasting instructor performance in higher education", , Almasri et al.[9]	5820 records	VoE(Extra Tree, RF, Decision Tree, GB)	0.9161

Secondly, we utilized the stages below in our ML training models

Stage 1: Since the imbalance happened, majority in grade B and C and few in grade F, the *Synthetic Minority Oversampling Technique (SMOTE)*[7].
Stage 2: RobustScaler for handling outliers [10]
Stage 3: We trained and evaluated five machine learning models to predict student academic outcomes based on their engagement behaviors:

1. **Random Forest(RF):** Robust, effective in handling noisy and imbalanced datasets, and provides clear interpretability through feature importance[5].
2. **Gradient Boosting(GB):** Optimized and tuned for this study, demonstrating excellent capability in modeling complex engagement patterns. [8]
3. **Logistic Regression(LR):** Utilized for its simplicity and interpretability, serving as a reliable baseline model.
4. **XGBoost(XGB):** Highly efficient and scalable, optimized with hyperparameter tuning for strong predictive accuracy.
5. **Voting Ensemble(VoE):** Combines the strengths of the best-performing models, enhancing overall predictive stability and accuracy. [9]

Stage 4: Hyperparameter tuning with GridSearchCV.
we used **GridSearchCV** with **Stratified K-Fold cross-validation** to tune hyperparameters for Random Forest, Gradient Boosting, and Logistic Regression. Key parameters like n_estimators, max_depth, learning_rate, and class _weight were optimized to improve accuracy, precision, recall, and F1 score. The best-performing configurations were selected for final evaluation.

Stage 5: Evaluation (accuracy, precision, F1, recall)
To assess the performance of each classification model, we used the following metrics:

1. Accuracy
The proportion of correctly predicted instances among all instances.

$$Accuracy = \frac{TP+TN}{\text{TP+TNTP+TN+FP+FN}} \quad \text{(Eq. 1)}$$

2.Precision

$$Precision = \frac{TP}{TP+FP} \quad \text{(Eq. 2)}$$

3.Recall

$$Precision = \frac{TP}{TP+FN} \quad \text{(Eq. 3)}$$

2.F1 Score is the harmonic mean of Precision and Recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad \text{(Eq. 4)}$$

Where

- TP: True Positives
- TN: True Negatives
- FP: False Positive
- FN= False Negatives

Stage 6: Feature importance plots.

Gradient Boosting, feature importance measures how much each feature contributes to reducing the **loss function** across all boosting iterations. It is typically calculated based on the **total gain** from all splits where the feature is used:

$$Importance_{Gain}(x_i) = \sum_{s \in split\ on\ x_i} \Delta L_s \quad \text{(Eq. 5)}$$

Where ΔL_s is the **reduction in the loss function** (e.g., log loss) achieved by split s involving feature x_i.

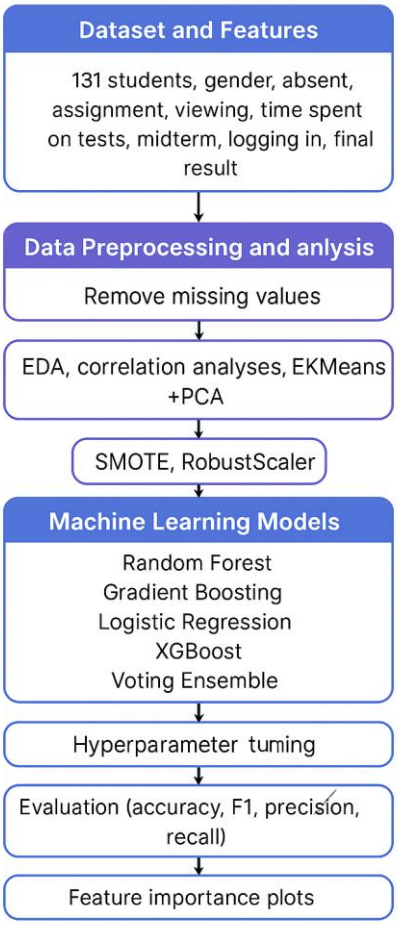


Fig. 1. Flowchart

3. RESULTS AND DISCUSSION

Dataset Imbalance and Preprocessing

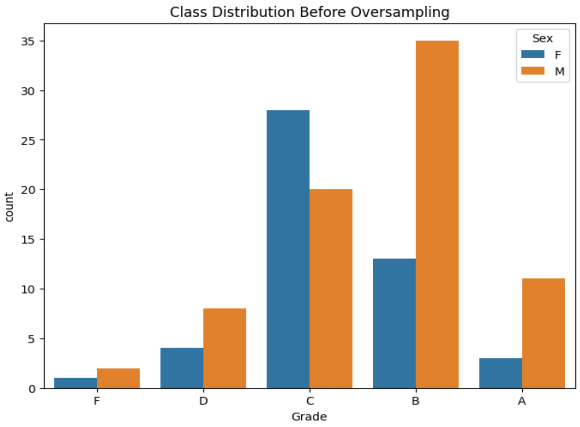


Fig. 2. Before oversampling

The initial analysis revealed a significant class imbalance, with the majority of students falling into grades B and C, while grades A, D, and F were underrepresented (Fig. 2). To address this, the Synthetic Minority Oversampling Technique (SMOTE) was applied, resulting in a balanced dataset suitable for model training (Fig. 3). This step was essential to prevent model bias and ensure fair performance across all grade categories.

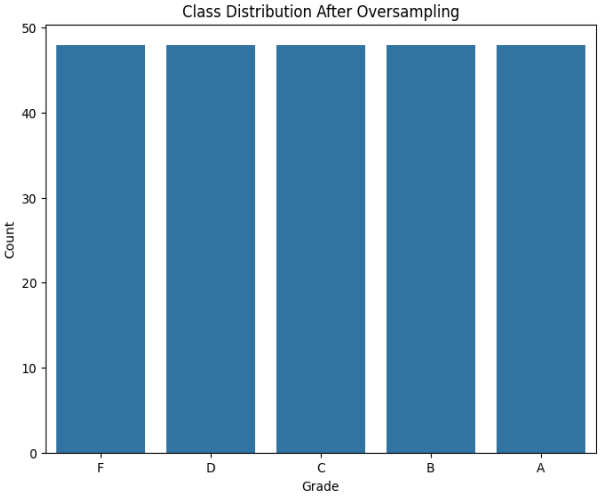


Fig. 3. After oversampling

Correlation Analysis

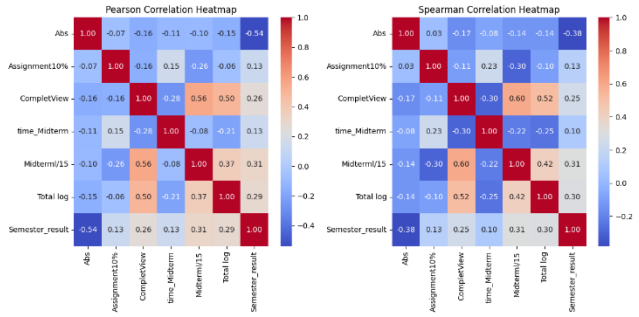


Fig. 4. Result of correlation analysis

Correlation results using Pearson and Spearman methods (Fig. 4) revealed meaningful relationships between engagement features and academic performance. A moderate to strong negative correlation was observed between absenteeism and final semester scores ($r \approx -0.54$), indicating that higher absence rates were associated with lower academic achievement. Conversely, login frequency and course material viewing were positively correlated with performance, emphasizing the importance of consistent participation in online learning environments.

Clustering and Student Engagement Patterns

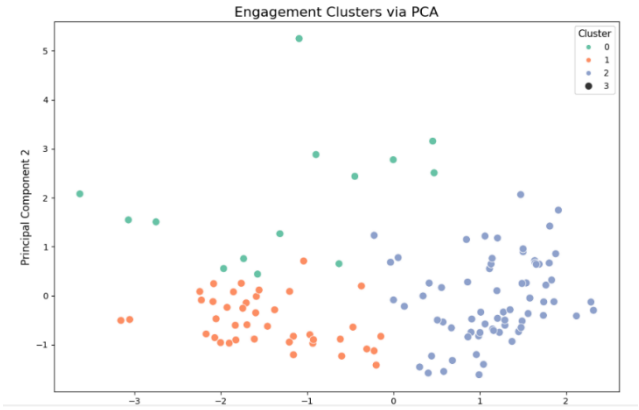


Fig. 5. Clusters Analysis

K-Means clustering combined with Principal Component Analysis (PCA) enabled the identification of distinct engagement profiles (Fig. 5). Students were grouped into three clusters, which can be named as highly engaged, moderately engaged, and less engaged. These clusters provide a valuable framework for targeting interventions based on behavioral engagement levels.

Model Performance and Comparison

Table 2 summarizes the performance metrics for all proposed machine learning models.

Table 2. Model Performance

Model	Accuracy	Precision	Macro F1	Recall
Gradient Boosting	**0.81**	**0.83**	**0.81**	**0.81**
Voting Ensemble	0.79	0.81	0.79	0.79
Random Forest	0.77	0.77	0.77	0.76
XGBoost	0.69	0.68	0.68	0.68
Logistic Regression	0.62	0.61	0.61	0.63

With the highest accuracy (0.81) and F1 score (0.81), Gradient Boosting was the outperforming model among other models, proving its better ability to deal with the data interactions and predict student outcomes. Utilizing several models to improve prediction accuracy, the Voting Ensemble did as well. Random Forest had acceptable results and strong interpretability. While Logistic Regression performed the worst but provided useful baseline comparisons, XGBoost demonstrated an acceptable degree of accuracy.

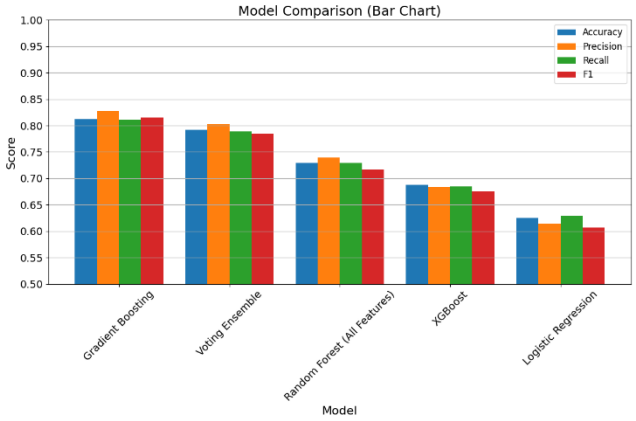


Fig. 6. Comparison bar graphs of Accuracy, Precision, Recall, and F1 score.

The feature importance analysis from the Gradient Boosting model (Fig. 7) reveals which student engagement behaviors had the greatest influence on predicting academic performance. Among all features, **midterm scores** emerged as the most influential predictor, highlighting the strong relationship between formative assessments and final academic outcomes. This was followed by **assignment submission frequency** and **number of absences**, indicating that consistent task completion and regular attendance are key determinants of success in blended learning. Additionally, features such as **login frequency** and **course material views** also contributed meaningfully, underscoring the importance of sustained interaction with the Moodle platform. These insights suggest that both active participation and early academic indicators (e.g., midterm results) play a pivotal role in shaping final student performance. Such information can guide instructors in prioritizing interventions for students showing early signs of disengagement.

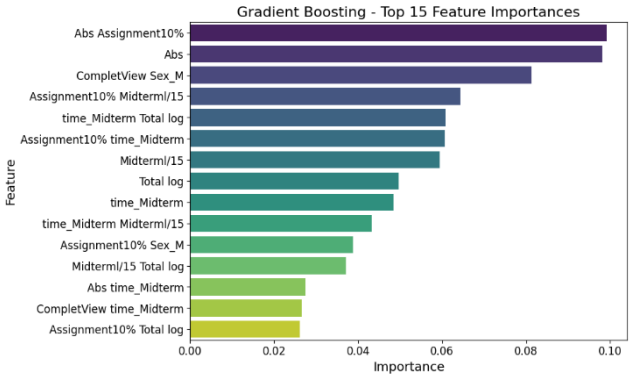


Fig. 7. Captions should be provided as well (Figure format can be varied depending on the type of your data)

Implications for Learning Analytics

The integration of predictive analytics and feature importance ranking in this study offers actionable insights for enhancing teaching and learning strategies in blended environments. The high predictive performance of models such as Gradient Boosting highlights the feasibility of accurately identifying at-risk students based on behavioral engagement data. Furthermore, the feature importance analysis provides clear evidence of which activities—such as midterm performance, assignment completion, and attendance—are most strongly associated with academic success.

These findings suggest that Learning Analytics can be leveraged not only for **early intervention**, but also for **targeted instructional design**. For instance, students with low midterm scores or irregular login patterns can be flagged for academic support or mentoring. Educators can prioritize promoting consistent engagement behaviors, such as timely assignment submission and active interaction with learning materials.

Moreover, institutional stakeholders can use these insights to inform policy and resource allocation—such as reinforcing early feedback mechanisms or investing in automated alert systems within Moodle. By translating data-driven insights into pedagogical actions, learning analytics serves as a powerful tool for personalizing learning pathways, improving student retention, and ultimately contributing to the goals of equitable and quality education as envisioned in SDG 4.

4. CONCLUSIONS

The research conducted here successfully demonstrated how predictive analytics could be utilized to evaluate and enhance student participation in blended learning environments, particularly when utilizing the ITC Moodle platform. Our findings demonstrated the strong predictive capacity of models like Gradient Boosting and Voting Ensemble, indicating the important role that online engagement behaviors—such as attendance, frequency of logins, and content interaction—have on academic achievement.

We recommend increasing the size of the data set for further studies for enhancing resilience and generalizability. Further understanding of student behavior may also be feasible by including additional interaction indicators like forum participation and video interaction. Future studies might investigate these relationships further since every online encounter has the potential to significantly impact academic results. In addition, studying how students perceive and experience blended learning settings may provide insightful qualitative data to support quantitative analytics and, in the end, advance a thorough comprehension of student engagement and academic achievement.

ACKNOWLEDGMENTS

I would like to thank the Institute of Technology of Cambodia for providing resources and support for this research.

I also want to thank my classmates and instructors for their insights and sense of collaboration, which improved my knowledge.

REFERENCES

[1] D. Lalima and K. Lata Dangwal, “Blended Learning: An Innovative Approach,” *Univers. J. Educ. Res.*, vol. 5, no. 1, pp. 129–136, 2017, doi: 10.13189/ujer.2017.050116.

[2] L. C. Liñán and Á. A. J. Pérez, “Educational data mining and learning analytics: Differences, similarities, and time evolution,” *RUSC Univ. Knowl. Soc. J.*, vol. 12, no. 3, pp. 98–112, 2015, doi: 10.7238/rusc.v12i3.2515.

[3] I.-H. Jo, Y. Park, J. Kim, and J. Song, “Analysis of Online Behavior and Prediction of Learning Performance in Blended Learning Environments,” *Educ. Technol. Int.*, vol. 15, no. 2, pp. 71–88, 2014.

[4] P. Sokkhey and T. Okazaki, “Developing web-based support systems for predicting poor-performing students using educational data mining techniques,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 7, pp. 23–32, 2020, doi: 10.14569/IJACSA.2020.0110704.

[5] P. Sokkhey and T. Okazaki, “Hybrid machine learning algorithms for predicting academic performance,” *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 1, pp. 32–41, 2020, doi: 10.14569/ijacsa.2020.0110104.

[6] C. Ding and X. He, “Principal component analysis and effective K-means clustering,” *SIAM Proc. Ser.*, vol. 2, no. 2, pp. 497–501, 2004, doi: 10.1137/1.9781611972740.54.

[7] N. Aslam, I. U. Khan, L. H. Alamri, and R. S. Almuslim, “An Improved Early Student’s Performance Prediction Using Deep Learning,” *Int. J. Emerg. Technol. Learn.*, vol. 16, no. 12, pp. 108–122, 2021, doi: 10.3991/ijet.v16i12.20699.

[8] A. Asselman, M. Khaldi, and S. Aammou, “Enhancing the prediction of student performance based on the machine learning XGBoost algorithm,” *Interact. Learn. Environ.*, vol. 31, no. 6, pp. 3360–3379, 2023, doi: 10.1080/10494820.2021.1928235.

[9] A. R. Almasri, N. A. Yahaya, and S. S. Abu-Naser, “Predicting Instructor Performance in Higher Education Using Stacking and Voting Ensemble Techniques,” *J. Theor. Appl. Inf. Technol.*, vol. 103, no. 2, pp. 740–749, 2025.

[10] C. K. Ting, N. Ibrahim, S. H. Huspi, W. Mohd, and N. Wan, “Multidimensional Context Clustering to Analyse Student Engagement in Online Learning Environment,” vol. 14, no. 2, pp. 89–96, 2024.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Applying Deep Learning Approaches to Khmer Speech Recognition and Text-to-Speech in Conversational Domains

Prakmlis SOENG[1*], Dona VALY [2], Phutphalla KONG [2]

[1] *Graduate student, Master program of Computer Science, Graduate school, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: prakmlis_soeng@gsc.itc.edu.kh*

Abstract: *Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) systems have achieved remarkable success in various digital contexts, enhancing communication, accessibility, and user experiences. However, these innovations primarily benefit high-resource languages, leaving low-resource languages such as Khmer significantly underserved. This research explores the development of ASR and TTS systems for the Khmer language by fine-tuning pre-trained models within an end-to-end deep learning architecture. To address the challenge of limited annotated data, the first objective of this research is to construct a high-quality, domain-specific dataset for both TTS and ASR. This customized dataset will include paired text-audio data for TTS and audio-transcription pairs for ASR, enabling effective training. For TTS, the study employs the Coqui TTS framework, which supports high-performance, multilingual, and modular voice synthesis. It will be adapted to generate expressive, conversational Khmer speech. For ASR, the approach leverages self-supervised learning by fine-tuning pre-trained models such as Wav2Vec2 on Khmer speech data. The research also emphasizes deployment and integration by developing APIs that combine both ASR and TTS components, facilitating seamless integration into real-world applications. These efforts aim to significantly advance Khmer speech technology and provide a replicable framework for other low-resource languages. The findings have the potential to improve accessibility and inclusivity in natural language processing for underrepresented linguistic communities.*

Keywords: Automatic Speech Recognition, Text-to-Speech, End-to-End architecture.

1. INTRODUCTION

Speech technologies such as Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) have become essential components in modern human-computer interaction, powering virtual assistants, customer service bots, accessibility tools, and voice-enabled applications. These systems have seen considerable advancements through deep learning, particularly for high-resource languages like English, Mandarin, and Spanish. However, the benefits of such technologies have not been equally distributed. Low-resource languages, such as Khmer, continue to face critical challenges due to the lack of labeled data, limited linguistic resources, and minimal research attention [1].

Khmer, the official language of Cambodia and spoken by over 16 million people, remains largely unsupported in the global speech technology landscape. Its unique phonetic structure, script complexity, and lack of standardized digital corpora make the development of ASR and TTS systems particularly challenging. To bridge this gap, this research focuses on developing a complete, end-to-end Khmer speech pipeline, guided by four key objectives.

The first objective is to construct a high-quality dataset that can support both ASR and TTS tasks. This involves collecting and curating paired audio-transcription data for ASR and paired text-speech data for TTS. The second objective is to build an expressive, natural-sounding Khmer TTS system using the Coqui TTS framework, which offers modular, multilingual capabilities optimized for real-time and production-ready synthesis. The third objective is to develop a robust Khmer ASR system by fine-tuning self-supervised pre-trained models such as Wav2Vec2.0 on the newly created Khmer dataset, enabling accurate transcription of spoken Khmer. Finally, the fourth objective is to ensure smooth integration of both components through the development of

[1*] Corresponding author: Prakmlis Soeng
E-mail: prakmlis_soeng @gsc.itc.edu.kh; Tel: +855-11 41 92 09

APIs, facilitating deployment in real-world applications such as voice interfaces, chatbots, and digital accessibility tools.

By addressing these four objectives, this research aims to significantly advance Khmer speech technology and demonstrate a practical roadmap for enabling speech processing in other underrepresented languages. The outcome is expected to enhance linguistic inclusivity, improve access to digital services for Khmer speakers, and contribute to the broader field of low-resource language processing.

2. Related Work

2.1 Speech Recognition in Low-Resource Languages

Automatic Speech Recognition (ASR) has advanced considerably with the adoption of deep learning techniques. Self-supervised models such as Wav2vec [2] and its improved variant Wav2vec 2.0 [1] have demonstrated strong performance in low-resource settings by learning useful representations from raw audio without the need for large labeled datasets. These models are particularly advantageous for under-resourced languages like Khmer, where transcription data is limited.

2.2 Evolution of TTS Models

The development of TTS has followed a similar trajectory, evolving from unit-selection methods and parametric approaches [4] to neural end-to-end systems like Tacotron and WaveNet. Tacotron [5] demonstrated the feasibility of synthesizing natural speech from text using a sequence-to-sequence architecture, while WaveNet [6] significantly improved the realism of generated audio through autoregressive waveform modeling. Building on these foundations, the Coqui TTS framework [7] offers a flexible and open-source alternative that is well-suited for multilingual and low-resource scenarios, making it an ideal choice for Khmer TTS development.

2.3 Frameworks for Low-Resource Language Development

Comprehensive studies, such as that by Besacier [1] have identified key strategies for building ASR systems for under-resourced languages, including phoneme transfer, multilingual training, and data augmentation. These techniques inform the methodology of this research, particularly in constructing a usable Khmer speech corpus and adapting pre-trained models through fine-tuning.

2.4 Gaps in Existing Khmer Speech Research

While some prior work has explored Khmer ASR using hybrid or cross-lingual approaches, these efforts are fragmented and lack comprehensive system integration. To date, there is no publicly available Khmer speech pipeline that combines both ASR and TTS with support for conversational applications. This research addresses that gap by developing a complete, integrated system and enabling deployment through modern API services.

3. Proposed Method

This research introduces a novel end-to-end Khmer speech processing pipeline that integrates Automatic Speech Recognition (ASR) and Text-to-Speech (TTS) modules in conversational context. The primary novelty of this work lies in the construction and utilization of a high-quality, conversational Khmer speech dataset, which was custom-collected specifically for this study. Unlike existing Khmer resources, this dataset was designed to capture realistic speech scenarios, including informal expressions, natural intonation, and diverse speaker accents. The conversational nature and speaker diversity make it uniquely suited for training deep learning models capable of handling real-world language variability—an area previously underrepresented in Khmer speech research. The pipeline is divided into four key stages: (1) data preprocessing, (2) ASR model fine-tuning, (3) TTS model training, and (4) integration through application programming interfaces (APIs).

3.1 Data Preprocessing

This research utilizes both publicly available and custom-collected Khmer speech data. In addition to data sourced from Hugging Face and OpenSLR, we constructed a novel conversational Khmer dataset specifically for this study. It includes 24 hours of speech from a large number of diverse speakers, featuring naturally spoken dialogue, informal expressions, and varied regional accents. This conversational focus distinguishes it from previous Khmer datasets, which often consist of scripted or read speech.

All audio files were standardized to a 16 kHz mono-channel WAV format and amplitude-normalized. For ASR, the audio-transcription pairs were manually aligned and segmented into utterances of 5–15 seconds to ensure precise synchronization. The text was cleaned and tokenized into subword units, using either byte pair encoding (BPE) or SentencePiece for character-level representation.

For TTS, a similar preprocessing pipeline was applied. The text was normalized to handle Khmer numerals, punctuation, dates, and non-standard script forms. Paired utterances were trimmed for silence and aligned with the spoken audio. Optional phoneme extraction was considered depending on the model configuration.

This careful preprocessing ensures that both ASR and TTS components are trained on high-quality, domain-specific data reflective of real-world conversational Khmer usage.

3.2 ASR System with Wav2Vec 2.0

The ASR component is built by fine-tuning the Wav2Vec 2.0 architecture, a self-supervised model that learns speech representations from raw audio without requiring large amounts of labeled data. A pre-trained multilingual checkpoint (e.g., XLSR-53) is used as the base model. Fine-tuning is performed on the Khmer audio-transcription dataset using the CTC (Connectionist Temporal Classification) loss function, which is well-suited for speech-to-text alignment without explicit frame labeling.

Training is conducted using the Hugging Face Transformers library. Hyperparameters such as learning rate, batch size, warm-up steps, and number of epochs are selected based on validation performance. The output is a Khmer ASR model capable of transcribing spoken Khmer into UTF-8 encoded text with high accuracy.

3.3 TTS System with Coqui TTS

For the Text-to-Speech (TTS) system, this research adopts the VITS (Variational Inference Text-to-Speech) architecture provided within the Coqui TTS framework. VITS is a state-of-the-art end-to-end TTS model that combines variational autoencoders, normalizing flows, and adversarial training to synthesize high-quality, expressive speech directly from text. Unlike traditional two-stage systems, VITS eliminates the need for separate spectrogram prediction and vocoding modules, allowing fully end-to-end training and inference.

The model is trained on the complete Khmer speech dataset, including the custom-collected recordings to enhance speaker and phonetic diversity. Input text is normalized and tokenized, then passed through the text encoder. The decoder directly generates audio waveforms conditioned on latent variables and linguistic features, effectively modeling prosody and speaker variation.

Training is performed using adversarial loss functions, reconstruction loss, and Kullback–Leibler divergence to balance audio realism and stability. Key hyperparameters such as learning rate, batch size, segment length, and latent dimension size are carefully tuned for Khmer speech. The resulting model is capable of generating fluent, natural, and emotionally expressive Khmer audio, suitable for conversational and interactive applications.

3.4 System Integration and Deployment

To support practical use cases, both ASR and TTS systems are integrated into a modular architecture via RESTful APIs. The backend is implemented in Python using frameworks such as FastAPI or Flask. The ASR API accepts audio input and returns transcriptions, while the TTS API accepts Khmer text and returns generated audio.

This modular integration allows external systems—such as chatbots, mobile applications, or web platforms—to interact seamlessly with the speech pipeline. The system is containerized using Docker for portability and deployed on local servers or cloud environments depending on resource availability. The APIs are secured and optimized for efficient, low-latency processing to support real-time applications.

4. Experiments and Results

4.1 Experimental Setup

All training and evaluation were conducted using Google Colab Pro with access to an NVIDIA Tesla T4 GPU (16GB VRAM), 25GB of RAM, and Python 3.10. The training framework was based on the Hugging Face transformers and datasets libraries, with PyTorch 2.0 backend and CUDA 11.8.

ASR component

For the ASR component, the pre-trained facebook/w2v-bert-2.0 model was fine-tuned using the Wav2Vec2BertForCTC class. Adapter layers were enabled for low-resource adaptation, and Connectionist Temporal Classification (CTC) loss was used for alignment-free training. The model was trained on the prepared Khmer dataset with character-level tokenization.

The key training settings used in this experiment included a batch size of 8, with gradient accumulation steps set to 4 to simulate a larger effective batch size. The model was trained for 10 epochs using a learning rate of 5e-5, with 500 warmup steps to stabilize early training. Evaluation was performed every 300 steps, using Word Error Rate (WER) as the primary metric to monitor performance. To optimize memory usage and training efficiency, both mixed precision (fp16) and gradient checkpointing were enabled.

Step	Training Loss	Validation Loss	WER(%)
300.0	9.2596	0.4418	38.81
600.0	1.0093	0.3614	25.59

Table 1: ASR Model Training and Evaluation Results

TTS component

The TTS model used was VITS (Variational Inference Text-to-Speech), trained using the Coqui TTS framework. The model was trained from scratch on a 22.05 kHz sampled Khmer speech dataset, paired with a carefully curated and normalized character-level transcript. A custom Khmer

character set of 86 unique symbols was used, covering consonants, vowels, modifiers, and punctuation.

The training was conducted using the VITS model architecture with a hidden channel size of 192. The character set included 86 unique symbols, and the training batch size was set to 32, with an evaluation batch size of 16. The model was trained for 100 epochs using the multilingual_cleaners text cleaner. Phoneme usage was disabled (use_phonemes=False), while speaker embedding was enabled to support speaker-specific synthesis. Audio samples used during training ranged from 32×256×4 to 160,000 samples in length. All training outputs, including checkpoints and logs, were saved to the "MyTTSDataset" directory.

Step	Training Loss	Validation Loss	MOS (Predicted)
Epoch 50	2.3512	0.9321	3.81
Epoch 100	1.8749	0.7483	4.12

Table 2: presents the model's training and evaluation progression based on training/validation loss and predicted MOS (Mean Opinion Score).

The results fom both the ASR and TTS components indicate strong potential for Khmer language speech applications, even with limited resources. The ASR model showed significant improvements during training, with the Word Error Rate (WER) dropping from 38.81% at step 300 to 25.59% by step 600, demonstrating effective fine-tuning through adapter-based training. Although the final WER indicates room for improvement, especially in handling rare or complex word structures, the model successfully captured the phonetic and orthographic characteristics of Khmer script.

For the TTS component, the predicted MOS improved from 3.81 at epoch 50 to 4.12 at epoch 100, indicating increasingly natural and intelligible synthesized speech. The use of a custom character-level tokenizer and speaker embeddings likely contributed to improved pronunciation and expressiveness. While occasional mispronunciations or flat intonation were observed in long or out-of-distribution sentences, the model generally produced fluent and rhythmically appropriate output. Overall, both components performed well under low-resource constraints and serve as a solid baseline for future Khmer ASR and TTS research.

4.2 Datasets

This research utilizes three main sources of Khmer speech data: Hugging Face, OpenSLR, and a custom-collected dataset. The Hugging Face dataset contributes approximately 20 hours of transcribed audio for training and evaluation, along with an additional 2.6-hour test set specifically curated to assess model performance on unseen data. To further enhance model diversity and reduce data sparsity, the OpenSLR-42 dataset, containing around 3 hours and 58 minutes of speech, was incorporated. These publicly available resources provided a strong baseline for developing both ASR and TTS components.

To complement these datasets and address limitations in speaker and pronunciation variety, a custom dataset was constructed. The goal was to collect 24 hours of high-quality speech data from 100 Khmer speakers. Each participant read scripted sentences ranging from 1 to 10 seconds in duration, ensuring a diverse representation of speech patterns, accents, and speaking styles. This combined dataset—spanning public and original sources—served as the foundation for training and evaluating deep learning models tailored to Khmer conversational speech.

5. CONCLUSIONS

This research addresses the critical need for speech technologies in low-resource languages by developing an end-to-end pipeline for Khmer Automatic Speech Recognition (ASR) and Text-to-Speech (TTS). The study began with the construction of a high-quality, task-specific dataset tailored for both ASR and TTS applications, providing the foundational resource needed for effective model training.

Leveraging self-supervised learning, the Khmer ASR system was developed by fine-tuning pre-trained Wav2Vec2.0 models to accurately transcribe spoken Khmer. Concurrently, the Khmer TTS system was built using the Coqui TTS framework, enabling the synthesis of natural and expressive Khmer speech suitable for conversational contexts.

To ensure real-world applicability, both systems were integrated through the development of robust APIs, allowing seamless interaction with external applications such as voice interfaces, assistive tools, and chatbots. This integration lays the groundwork for scalable, deployable Khmer speech solutions.

The outcomes of this research make a substantial contribution to the advancement of Khmer language technologies and demonstrate a replicable methodology for similar initiatives in other low-resource languages. Future work will focus on improving prosody modeling, incorporating speaker adaptation, and extending the system for real-time applications and multilingual interoperability.

Acknowledgement

The authors would like to thank the contributors who provided the pre-trained models used in this research. We gratefully acknowledge the Hugging Face and Coqui TTS communities for their open-source tools and resources. This work was carried out using Google Colab Pro for model training and experimentation.

REFERENCES

[1] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: A survey," *Speech Communication*, vol. 56, pp. 85–100, 2014.

[2] S. Schneider, A. Baevski, R. Collobert, and M. Auli, "wav2vec: Unsupervised pre-training for speech recognition," in *Proc. Interspeech*, 2019.

[3] A. Baevski, H. Zhou, A. Mohamed, and M. Auli, "wav2vec 2.0: A framework for self-supervised learning of speech representations," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12449–12460, 2020.

[4] H. Zen, K. Tokuda, and A. W. Black, "Statistical parametric speech synthesis," *Speech Communication*, vol. 51, no. 11, pp. 1039–1064, 2009.

[5] Y. Wang, R. Skerry-Ryan, D. Stanton, Y. Wu, R. J. Weiss, N. Jaitly, Z. Yang, Y. Xiao, Z. Chen, S. Bengio et al., "Tacotron: Towards end-to-end speech synthesis," in *Proc. Interspeech*, pp. 4006–4010, 2017.

[6] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. W. Senior, and K. Kavukcuoglu, "WaveNet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499*, 2016.

[7] Coqui, "Coqui TTS: A deep learning toolkit for Text-to-Speech," GitHub Repository, [Online]. Available: https://github.com/coqui-ai/TTS, 2021.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

AI-Based Skill Trend Analysis: Enhancing Job Market Insights and Curriculum Development in Cambodia

Sela SONGEAM [1*], Dona VALY [2], Sothea HAS [2]

[1] *Graduate student, Master Program of Computer Science, Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: sela_songeam@gsc.itc.edu.kh*

Abstract: *The demand for skills in the jobs market is rapidly changing in response to the new emerging technology and skills that align with the industry's needs. This constant change is a challenge for new job seekers, especially freshly graduated students, to identify and learn trending skills. In this context, universities play an important role in preparing their students to be ready for the job market. Therefore, having the right curriculum that aligns with the demands of stakeholders is the fundamental step to embedding the right skills for students. The objective of this research is to build a data collection tool to gather online job postings, analyze online job postings to identify trending industry skills through utilizing Natural Language Processing (NLP) techniques. Furthermore, the research proposes a skill-based recommendation system designed to assist job seekers in aligning their existing competencies with market needs and to guide universities in updating curricula to bridge identified skill gaps. In this research, we will make use of several methods, including web scraping, NLP approaches such as skill extraction using AI-driven models, keyword matching with standardized taxonomies (ISCO, ESCO), and trend analysis using data mining and data visualization. As a result of this research, we are expected to provide real-time insights into skill demand across industries, allowing job seekers to make informed career decisions, employers to optimize recruitment strategies, and educational institutions to adjust curricula to meet workforce requirements. By leveraging AI-powered labor market analysis, this research contributes to building a future-ready workforce and fostering data-driven decision-making in education and employment sectors.*

Keywords: Job Market, Web Scraping, Data Analysis, Natural Language Processing (NLP)

1. INTRODUCTION

Skills form the foundation of job performance and employability. In today's rapidly evolving landscape, new skills are continually emerging and existing ones are being updated to keep pace with technological advancements and the introduction of new tools. The evolution has become a challenge for job seekers and graduate students, who constantly need to learn and upgrade their knowledge and skills for their career goal. However, the challenge mostly ended up in the skills mismatch between young workers and employer's needs.

According to the survey by EUROCHAM [1], which conducted a Skill Gap Assessment across 106 companies in over 11 sectors in Cambodia, approximately 74% of companies reported difficulties in finding qualified staff. A significant portion of these challenges—about 60%—were attributed to area-specific labor shortages and general skill gaps. The report also highlighted various sector-specific skills that are in high demand. However, despite the value of such assessments, their insights often remain underutilized. The results are presented in static formats and lack visibility among the broader public, especially job seekers and education providers. Notably, EuroCham's data collection methods involved outreach through emails, social media platforms, webinars, phone calls, company visits, and messaging applications, which may limit the representativeness and timeliness of the findings.

[1*] Corresponding author: Sela SONGEAM
E-mail: sela_songeam@gsc.itc.edu.kh; Tel: +855-93 829 726

To address these limitations, our research proposes a data-driven approach to skill trend analysis by leveraging web scraping techniques to collect real-time job posting data from major online job platforms in Cambodia. Unlike traditional surveys that are limited in scope and frequency, job postings provide continuously updated insights into the labor market and reflect the actual demand for skills across various industries. By extracting and analyzing job descriptions using natural language processing (NLP) and skill taxonomies, we aim to uncover emerging skill trends and gaps. This methodology offers a scalable, transparent, and automated way to support curriculum development, inform policy decisions, and empower job seekers with actionable information.

In this work, we seek to contribute to:

- Utilizing web scraping framework of the public available job posting data across sectors in Cambodia.
- Applying Large Language Model (LLM) with a few-shot prompting to extract job information from scraped HTML page, and skill extraction from job description.
- Adapted standardized terminology of job category and skills.
- Building a full automate pipeline for collecting job posting and visualize skill trends.

2. RELATED WORK

In this section, we review related works, in the area of occupational-related datasets, general occupational data mining and analysis, and skill extraction.

Zhang et al. [2] developed SKILLSPAN, a high-quality span-level annotated dataset based on the ESCO taxonomy, explicitly designed for extracting skills from English job postings. Their introduction of domain-specific models, JobBERT and JobSpanBERT, demonstrated significant performance improvements compared to general-purpose models like BERT and SpanBERT, effectively addressing the scarcity of domain-specific annotated datasets.

Khaouja et al. [3] systematically reviewed existing approaches, emphasizing the distinction between structured skill repositories (e.g., ESCO, O*Net, e-CF) and dynamically generated skill lists from external knowledge bases (e.g., Wikipedia, DBpedia). They observed that employing word embeddings and NLP techniques to augment skill bases substantially improved accuracy and adaptability. However, this line of research also highlighted persistent challenges such as handling false positives, refining skill granularity, and achieving comprehensive skill coverage across industries.

Senger, E., et al [4] conducted a thorough review of deep learning approaches, highlighting taxonomy-driven skill extraction and the emerging role of LLMs such as GPT-4. They showed that deep learning models like BERT, RoBERTa, and BiLSTM-CRF have significantly enhanced skill extraction and classification performance compared to traditional statistical methods. However, they identified critical areas requiring further research, including detecting emerging skills and extracting implicit skills.

In this paper, we specifically extend the integration of web scraping with LLM-driven skill extraction methodologies. We propose an end-to-end pipeline that employs web scraping techniques combined with few-shot prompting methods using advanced language models such as GPT to extract structured data directly from unstructured HTML job postings. This methodological approach addresses existing gaps identified in previous studies, particularly the need for dynamically adaptable and accurate skill extraction systems that seamlessly integrate structured taxonomies like ESCO.

3. METHODOLOGY

This section outlines the methodology adopted in this research, detailing the systematic approach we implemented to collect, process, and analyze occupational job data from online postings. Our methodology involves several stages, as illustrated in **Figure 1**, which depicts the overall pipeline of our research process.

3.1 Web Scraping

We first collected data from targeted job-listing websites using automated web scraping techniques. This initial step involved accessing job portals and then we scraped all the list of jobs that posted on the website. Each job entry is processed to eliminate irrelevant content by stripping HTML tags, JavaScript, and CSS using Python BeautifulSoup, retaining only visible text content. Next, we normalize whitespace and remove non-alphabetic tokens. The cleaned textual content is then passed into the LLM (llama3.2) along with a few-shot prompting strategy, enabling the model to output structured job information in JSON format. We provide statistical insights (as summarized in **Table 2**) into the average size of text inputs fed into the LLM to better understand the scope and complexity of the extraction task.

To evaluate the accuracy of our LLM-based extraction, we manually created a ground truth dataset by writing Python scripts to extract job information directly from the HTML using developer tools and tag inspection for comparing against the LLM outputs, which were generated from the cleaned text of the same job postings. We then measured the alignment between LLM-extracted data and manually extracted values using exact match accuracy and field-level precision, recall, and F1-score. This helped assess the reliability of using LLMs for structured information extraction in real-world web scraping tasks.

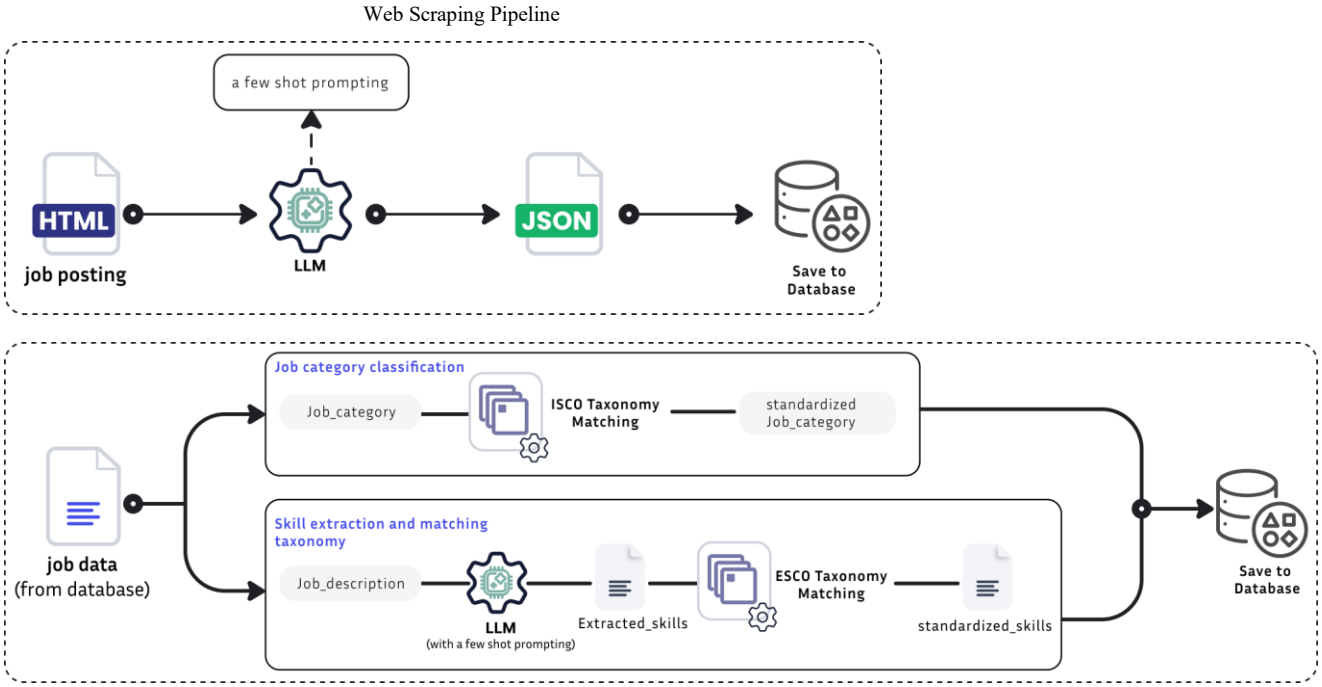


Figure 1: Overview of the Job Posting Extraction and Normalization

Average per input	**Job information card**	**Job detail page**
Characters	1103.5	5997.1
Words	64.8	814.1

Table 2: Statistic of text input to LLM for job information extraction

Field	Type	Precision	Recall	F1-Score	Exact Match Accuracy
job_title	*JC*	0.83	0.83	0.83	83.85%
company_name	*JC*	0.90	0.86	0.88	86.98%
job_location	*JP*	0.85	0.80	0.82	70.56%
job_category	*JP*	0.75	0.7	0.72	68.96%
job_description	*JP*	0.65	0.51	0.63	52.02%
posted_date	*JC*	0.41	0.33	0.36	33.65%
closing_date	*JP*	0.70	0.43	0.53	43.23%
job_url	*JC*	1.0	0.61	0.76	61.90%
company_industry	*JP*	0.78	0.68	0.72	69.23%
Overall (avg.)		**0.76**	**0.63**	**0.69**	**63.37%**

Table 3. Overall LLM Extraction Performance (Precision, Recall, F1). Type JC - is the information from Job Card, and JP- is the information extracted from the job details page

As shown at **Table 3**, the LLM does very well on concise, structured fields (e.g. job_title, company_name, job_url with F1 ≈ 0.83–1.00) because we input a short chuck of each job information card to LLM but much less so on dates and long text (posted_date F1 0.36, job_description F1 0.63). In summary, our Web Scraping Pipeline have collected around 2000 jobs in Cambodia, from various hiring platforms include LinkedIn, BongThom and CamHr. Our data were collected from January to April of 2025.

3.2 Job Category Classification

Occupational data extracted from job postings often includes inconsistencies and ambiguities due to diverse terminologies and varying interpretations of job roles across different industries and countries. This inconsistency poses a significant challenge for accurately categorizing and analyzing labor market trends. To address this problem, we incorporate the International Standard Classification of Occupations (ISCO) into our data pipeline. ISCO provides a structured and globally accepted classification system, allowing us to standardize job category, minimize ambiguity, and facilitate reliable cross-sector and cross-country analyses of occupational data [5].

We utilize sentence embeddings combined with cosine similarity to match job categories extracted from job postings

to their most relevant ISCO descriptions. Each job category is first transformed into a numerical embedding vector, enabling quantitative similarity measurement. The embedding vectors are then compared against those derived from ISCO descriptions, and the best-matching ISCO code and labels are assigned if a similarity threshold of 0.6 or higher is met. This ensures that our job market analysis aligns precisely with internationally recognized occupational standards, enhancing the reliability and comparability of our findings

3.3 Skill extraction and matching taxonomy

Skill Extraction: In this phase of our pipeline, we extract skills systematically from job descriptions using a LLM (llama3.2) enhanced with prompt engineering techniques. First, each job description text is input into the LLM with carefully crafted few-shot prompting, instructing the model to identify and separate hard skills and soft skills explicitly. The output from the LLM is structured into a standardized JSON format, clearly distinguishing between different skill categories. Extracted skill phrases were then matched to a standardized taxonomy to ensure consistency and comparability.

To validate the effectiveness of the skill extraction methodology, we employed the SkillSpan Dataset, a benchmark annotated corpus specifically designed for evaluating skill extraction tasks. SkillSpan contains human-annotated skill spans across various job description texts, categorized into distinct skill types such as hard skills and soft skills. We compared the skill extraction output from our LLM-based method against the SkillSpan annotations using standard evaluation metrics, including precision, recall, F1-score, and exact-match accuracy. Because skill spans often consist of multi-word phrases (e.g. "integration testing", "cloud native applications"), we report both a strict exact-match score and a more forgiving partial-match score. Under the partial-match criterion, we count a predicted span as correct if its normalized text overlaps in any way with a gold span (i.e. one string is a substring of the other). This lets us credit the model even when it splits or trims a long phrase and provide a comprehensive evaluation, demonstrating both the accuracy and practical applicability of our skill extraction framework.

Evaluation metrics: We compared the spans produced by our LLM pipeline against these annotations using standard information-retrieval measures. We retained about the 603 job-description records that contain at least one skill annotation. Each description was processed by the LLM with few-shot prompts that instruct it to return separate lists of hard and soft skills. Predicted spans were token-normalized (lower-cased, punctuation stripped) and aligned to the gold spans under two criteria: Exact match – a predicted span must match the gold span character-for-character, and Partial match – a predicted span is considered correct if either string is a substring of the other, allowing minor boundary mismatches such as "integration testing" vs. "testing integration". The equation (1) was used to calculate the exact-match accuracy for this experiment:

$$Exact\text{-}match\ Accuracy = \frac{1}{N}\sum_{i=1}^{N} f(x) \qquad \text{(Eq.1)}$$

where

$$f(x) = [S_i^{pred} = S_i^{gold}] = \begin{cases} 1, & if\ set\ match, return\ 1 \\ 0, & otherwise \end{cases}$$

N = total number of evaluated documents,
S_i^{pred} *= set of predicted skill spans in document i,*
S_i^{gold} *= set of gold (annotated) skill spans in document i,*

For Partial-match accuracy, we use the same equation but with the different method of counted matching skill set. if the predicted text overlap with the goal span, we counted as matched, which can be describe as:

$$f(x) = [S_i^{pred} \equiv S_i^{gold}] = \begin{cases} 1, & if\ set\ overlab, return\ 1 \\ 0, & otherwise \end{cases}$$

From our experiment, **Table 4** shows that Span-level EXACT-MATCH got a very low score because it demands a perfect, character-for-character alignment between the predicted span and the gold span. In practice, even a tiny mismatch—an extra space, a dropped hyphen, or inclusion/exclusion of a trailing word—counts as wrong. LLMs often get the gist of a skill phrase right but will "trim" it differently (e.g. predicting "project management" vs. "project management skills"), so almost every slight boundary error penalizes the exact-match metric heavily. PARTIAL-MATCH eases this requirement by treating any overlap between the predicted and gold spans as correct. That way, if the model captures the core of the skill mention—even with fuzzy boundaries—it still scores as a hit.

Skill Matching Taxonomy: In our experiment, we use ESCO (the European Skills, Competences, Qualifications and Occupations) as a standard skill taxonomy [6]. it was designed to accommodate both hard and soft skills, with finely grained hierarchy—spanning thousands of narrowly defined skill concepts linked to broader competency areas—enables granular tagging of extracted entities and supports aggregation at varying levels of abstraction (e.g., from "cloud infrastructure management" up to "IT operations") To effectively match these extracted skills, we experimented with two embedding-based matching strategies:

• **OpenAI Ada-002 (1536-Dimensional) Embeddings**: Extracted skill phrases were converted into vector embeddings using OpenAI's text-embedding-ada-002 model, known for its capability to capture deep semantic meanings. These embeddings were stored and indexed in a PostgreSQL database utilizing the vector extension. Similarity searches

Span-level	Precision	Recall	F1-Score	Accuracy
EXACT-MATCH metrics	0.152	0.150	0.151	1.350%
PARTIAL-MATCH metrics	0.593	0.586	0.590	-

Table 4. Span-Level Skill Extraction Performance

Metric	OpenAI Ada-002 (1536D)	FastText (300D)
Dimensionality	1536	300
Embedding Source	text-embedding-ada-002	crawl-300d-2M
Matching Threshold	≥ 0.70	≥ 0.70
Similarity Metric	Cosine Similarity	Cosine Similarity
Match Accuracy (%)	81.3%	73.2%

Table 5. Comparison of Embedding Models for Skill Matching Performance

were performed using cosine similarity, retaining only matches with similarity scores of 0.70 or higher.

• **FastText (300-Dimensional) Embeddings**: As an alternative approach, we employed the FastText crawl-300d-2M pre-trained embeddings. FastText is beneficial for handling noisy text data and out-of-vocabulary terms through sub-word modeling. Similarly, these embeddings were indexed separately and matched using cosine similarity.

The result shows at **Table 5** indicate that Ada-002 embeddings achieved a higher overall match accuracy of 81.3%, outperforming FastText, which yielded 73.2%. Both models were used specifically for matching extracted skill phrases to the ESCO taxonomy. Ada-002 demonstrated better capability in capturing the meaning of skill descriptions.

4. RESULT

This section provides a detailed analysis of the skill distributions and trends extracted from job postings data collected from online platforms in Cambodia. The primary objective of this analysis is to identify both established and emerging skills, prevalent competencies within specific job categories, and derive insights useful for curriculum planning and workforce strategies. After we do all the data collection and pre-processing, job classification, skill extraction, and standardized skill matching, the data was stored in a structured database, facilitating flexible queries and trend analyses across various job categories and sectors. This enables dynamic insights and the ability to monitor real-time shifts in skill demand.

As an illustrative example, focusing specifically on the job category of "Information Technology Training", **Figure 2**, illustrates the monthly trends of the top five most frequently occurring skills. Notably, Microsoft Endpoint Configuration

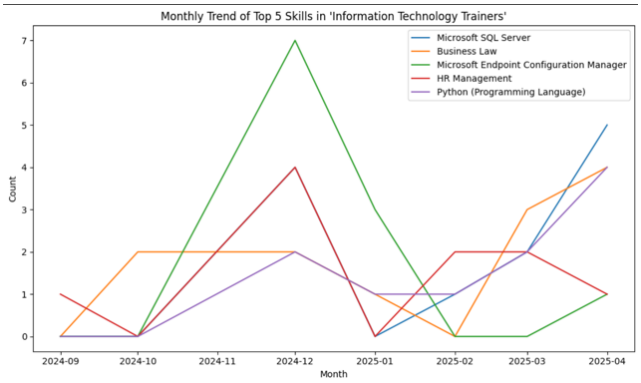


Figure 2. Top 5 skills trend from Sep-2024 to Apr-2025 of the category



Figure 3. Word Cloud of all skills in the category



Figure 4. Top 10 Skills in the category

Manager peaked sharply in December 2024, while Python (Programming Language) and Microsoft SQL gained popularity in early 2025, reflecting a potential shift in technological emphasis within IT training roles. **Figure 3,** presents a word cloud capturing the overall distribution of skill terms. The size of each phrase corresponds to its

frequency across job descriptions, offering a quick visual grasp of dominant skills. Key observations include the prominence of Microsoft SQL Server, Business Law, and Python, along with interdisciplinary competencies such as HR Management and Document Management Systems (DMS). **Figure 4**, displays the top 10 hard skills by frequency in a horizontal bar chart. This provides a clearer quantification of skill occurrences, confirming that Microsoft SQL Server and Business Law are among the most consistently required qualifications for IT trainers in the analyzed period.

The analysis of skill trends within the domain of Information Technology Trainers reveals both persistent and emerging competencies shaping the labor market. Technical skills such as Microsoft SQL Server, Python, and Microsoft Endpoint Configuration Manager consistently appeared across job postings, indicating a sustained demand for infrastructure management and programming expertise. Additionally, the presence of interdisciplinary skills like Business Law and HR Management highlights the growing expectation for IT professionals to possess cross-functional knowledge.

These results provide actionable insights, beneficial to educational institutions and policymakers in designing targeted curricula and workforce development initiatives that effectively respond to evolving labor market requirements.

5. CONCLUSIONS

This research introduces a comprehensive end-to-end framework for AI-driven skill trend analysis, explicitly designed to improve insights into Cambodia's job market and support informed curriculum development. By employing advanced large language models (LLMs) for skill extraction and integrating standardized taxonomies (ESCO and ISCO), we established a robust methodology capable of converting unstructured job posting data into structured, actionable insights. Our embedding-based similarity matching approach, comparing OpenAI Ada-002 and FastText embeddings, demonstrated strong semantic alignment between extracted skills and standardized skill descriptions.

Our analysis highlighted essential domain-specific and interdisciplinary competencies, especially within the Information Technology Training sector. Through clear and intuitive visualizations—including monthly skill trend graphs, word clouds, and frequency-based bar charts—we effectively illustrated both persistent and emerging workforce skill demands. These insights serve as valuable inputs for curriculum designers, educational institutions, policymakers, and employers, helping bridge the skill gap between educational offerings and labor market requirements.

Overall, our study provides a scalable, automated, and data-driven solution for labor market analysis. Future research directions may include expanding the analytical framework to perform regional and international comparisons, integrating labor supply-side data for comprehensive market assessments, and further fine-tuning LLM methodologies for even more precise domain-specific skill extraction and matching.

REFERENCES

[1] European Chamber of Commerce in Cambodia (EuroCham), Skill Gap Assessment Report: Digital, in collaboration with the Swiss Agency for Development and Cooperation (SDC) and the Skills Development Programme (SDP) implemented by Swisscontact, Phnom Penh, Cambodia, 2024. Available: https://www.eurocham-cambodia.org/uploads/782f9-final-report-skill-gap-assessment-report-digital-3.pdf.

[2] M. Zhang, K. N. Jensen, S. D. Sonniks, and B. Plank, "SKILLSPAN: Hard and soft skill extraction from English job postings," arXiv, Apr. 2022.

[3] I. Khaouja, I. Kassou, and M. Ghogho, "A survey on skill identification from online job ads," IEEE Access, vol. 9, pp. 14821-14836, Aug. 2021. doi:10.1109/ACCESS.2021.3106120.

[4] E. Senger, M. Zhang, R. van der Goot, and B. Plank, "Deep learning-based computational job market analysis: A survey on skill extraction and classification from job postings," Proceedings of the First Workshop on Natural Language Processing for Human Resources (NLP4HR 2024), Mar. 2024.

[5] International Labour Organization, The International Standard Classification of Occupations (ISCO-08) companion guide, Version 1.0, Geneva: Int. Labour Office, Sep. 2023. Available: https://www.ilo.org/sites/default/files/wcmsp5/groups/public/@dgreports/@stat/documents/publication/wcms_896661.pdf.

[6] M. le Vrang, A. Papantoniou, E. Pauwels, P. Fannes, D. Vandensteen, and J. De Smedt, “ESCO: Boosting job matching in Europe with semantic interoperability,” Computer, vol. 47, no. 10, pp. 57–64, Oct. 2014.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Lip-Synchronized Talking Avatar Generation for Khmer Language Using Audio-to-Landmark Modeling

Sovanvichea YANN [1*], Dona VALY [2], Phutphalla KONG [2]

[1] *Graduate student, Master program of Computer Science, Graduate school, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: sovanvichea_yann@gsc.itc.edu.kh*

Abstract: *Recent advancements in deep learning have enabled the generation of talking-face videos directly from speech. However, most existing solutions either rely heavily on speaker-specific data or are designed around Western facial datasets, which limits their generalizability especially for underrepresented languages like Khmer. In this paper, we introduce a two-stage, speaker-independent framework focused specifically on lip-synchronized avatar generation for the Khmer language. Our method first predicts lip landmarks from audio using a Transformer-based model, guided by prior facial landmarks from reference images. Then, a rendering module generates lip visuals that reflect the predicted motion while preserving the speaker's identity. To support this work, we utilize the KH-Reading dataset, which includes culturally diverse speech data tailored for Khmer. Evaluation using both visual quality and lip-sync accuracy metrics shows that our approach outperforms previous person-generic models and better preserves visual identity and temporal consistency. This research aims to open the door for more inclusive virtual avatar technologies, especially for low-resource language communities.*

Keywords: Lip synchronization; Audio-to-lip generation; Khmer virtual avatar; Lip landmark prediction; Generative AI

1. INTRODUCTION

The ability to generate talking-face videos from speech has emerged as a powerful tool in applications such as video dubbing, [1], [2], [3], online education [4], and virtual communication. These systems aim to convert spoken audio into synchronized lip movements on a human face, producing visually realistic and expressive animations. Despite recent progress [5], many existing approaches depend on training data from specific speakers or culturally narrow datasets. As a result, they often struggle to generalize across diverse languages and facial characteristics.

This challenge is particularly evident for underrepresented languages like Khmer. Most popular models are trained on Western facial datasets and are not optimized to handle the unique phonetic structure and facial expressions found in Khmer speech. Consequently, Khmer-speaking users frequently encounter avatars that lack authenticity in both articulation and appearance, limiting the usefulness of these systems in local applications.

To address this gap, we propose a new framework for generating lip-synchronized avatars that is both person-independent and culturally adaptive. Our method follows a two-stage design. In the first stage, a Transformer-based model predicts detailed lip landmarks from the input audio. These predictions are enhanced by incorporating static reference images that preserve speaker-specific lip characteristics without requiring speaker-specific training. In the second stage, a rendering module transforms the predicted landmarks into high-quality lip visuals through spatial alignment and texture-aware synthesis.

To validate our method, we created the KH-Reading dataset and used it for our experiments. This culturally tailored corpus of Khmer speech contains rich phonetic and emotional variety. Our results show significant improvements in lip synchronization and visual realism over existing person-generic baselines.

The key contributions of this work are:

- We introduce a two-stage audio-to-lip generation framework optimized for person-generic use in low-resource languages.

[1*] Corresponding author: Sovanvichea Yann
E-mail: sovanvichea_yann@gsc.itc.edu.kh; Tel: +855-17 715 327

- We design a Transformer-based architecture for lip landmark prediction that improves temporal coherence and visual identity.
- We incorporate reference-guided rendering to produce visually consistent lip animations using static images.
- We demonstrate the effectiveness of our model on Khmer speech data, contributing a culturally relevant solution for speech-driven avatars.

2. Related Work

2.1 Audio-Driven Talking Face Generation

Audio-driven talking face generation methods are typically categorized into two groups: person-specific and person-generic approaches. Person-specific models leverage prior video recordings of the target speaker for training or fine-tuning, enabling the synthesis of highly personalized and photorealistic talking-head videos. These methods often utilize 3D Morphable Models (3DMM) [10] or Neural Radiance Fields (NeRF) [11] to predict expression parameters from audio and subsequently render expressive face images. Examples include NVP [12] and FACIAL [13], which map audio to 3DMM parameters before applying neural rendering techniques. Other methods [5],[6] control dynamic NeRF representations using audio inputs to achieve fine-grained animation control. While effective, these methods are limited by their dependence on speaker-specific data, making them impractical in real-world scenarios where such data may not be available.

To overcome this limitation, person-generic models aim to generalize to unseen speakers without requiring additional fine-tuning. Notable examples include Wav2Lip [3], which uses an encoder-decoder architecture guided by a lip synchronization discriminator, and SyncTalkFace [5], which introduces an audio-lip memory module to enrich lip region representations. PC-AVS [16] takes a modular approach by separating speech content, head pose, and identity into distinct latent spaces. These are then combined to generate dynamic talking face videos. While these methods offer generalization across identities, they often fall short in delivering high-fidelity lip synchronization, especially when applied to low-resource languages or culturally specific facial behaviors.

2.2 Landmark-Based Talking Face Generation

Many recent advances in audio-driven talking face generation rely on facial landmarks as an intermediate representation between speech and visual output. Landmark-based approaches offer interpretability and controllability, particularly for lip synchronization tasks, which are critical in speech-driven avatar generation.

One of the early landmark-based systems was proposed by Suwajanakorn et al.[17], who used a Recurrent Neural Network (RNN) to map audio features to mouth landmarks, followed by texture synthesis. Subsequent works such as ATVGnet [6], MakeItTalk [18], and APB2Face [19] expanded on this by incorporating Long Short-Term Memory (LSTM) modules and simple fusion strategies involving pose and blink features alongside audio cues. These models, while effective in modeling temporal sequences, are limited in capturing long-range dependencies and adapting to unseen identities.

MakeItTalk [18] introduced a hybrid of LSTM and attention mechanisms to improve temporal modeling, while Xie et al. [8] proposed a two-stage framework using separate encoders for audio, pose, and reference image features. Their fusion process relies on weighted combinations of embeddings, which may not fully exploit the complex relationships between modalities.

However, most existing landmark-based methods are developed for high-resource languages and lack adaptability to culturally specific articulation styles. These approaches often fail to capture the phonetic and visual nuances of underrepresented languages like Khmer, highlighting the need for a dedicated framework that ensures both identity preservation and accurate lip synchronization in low-resource settings.

3. Proposed Method

To generate realistic and culturally faithful lip movements from speech, we propose a person-generic, lip-only synthesis pipeline. Our framework (Figure 1) is designed to produce temporally synchronized and identity-consistent lip animations using only audio input and a few visual references. The system consists of two primary stages: (1) audio-to-lip landmark prediction and (2) lip rendering via reference-guided synthesis.

3.1 Audio-To-Landmark Generation

The first stage translates raw speech signals into structured lip landmark sequences. Instead of relying on LSTM models as used in many prior works, we introduce a Transformer-based network that excels in modeling long-term dependencies between audio features and dynamic lip positions.

We begin by converting speech into Mel-spectrogram representations of size of 16×80, i.e., $h = 16$ and $w = 80$ [3], capturing essential frequency-temporal patterns. These features are processed through an audio encoder to extract a temporal embedding. Concurrently, reference lip landmarks and facial pose information are encoded to provide structural guidance. Positional encodings and modality indicators are incorporated to retain temporal ordering and modality distinctions.

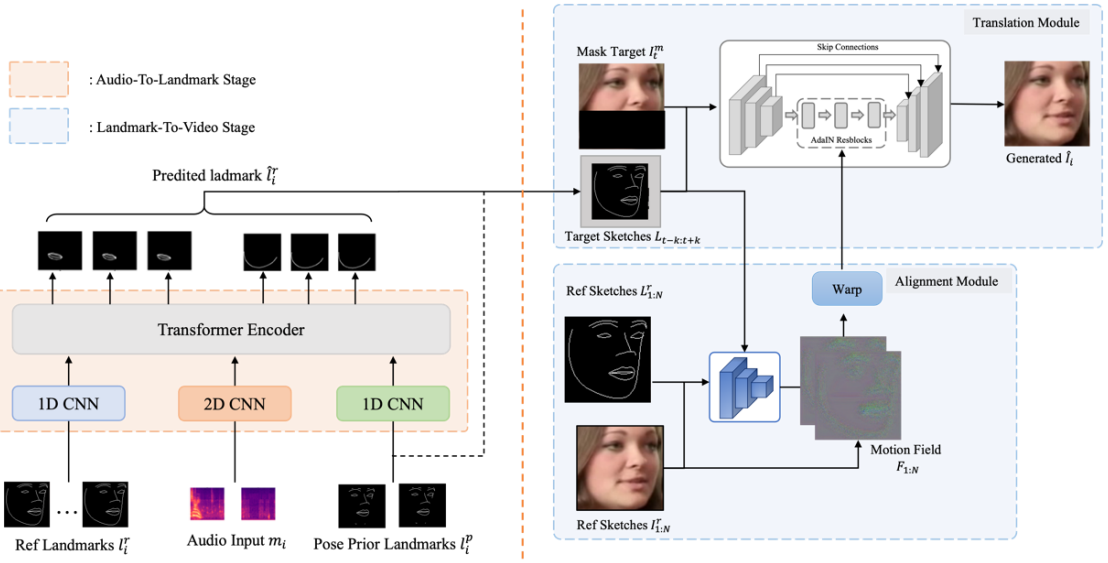


Figure 1: Our framework consists of two main stages. (1) ***Audio-to-Landmark Generation****: In this stage, a Transformer-based landmark generator takes three inputs—audio features, reference lip landmarks, and pose prior landmarks—to predict the dynamic positions of the lips (and optionally the jaw). These predicted landmarks are then combined with pose landmarks to form a complete sketch representing the target mouth configuration. For simplicity, positional and modality encodings used internally in the model are not shown. (2)* ***Landmark-to-Video Rendering****: In this stage, the target sketches guide an alignment module that takes multiple static reference images and their corresponding sketches to compute motion fields. These motion fields warp the reference images and their visual features to match the desired lip pose and expression. Finally, with help from both the warped visual features and the original audio input, a translation module synthesizes the final output image by combining the target sketch with a lower-half masked face image, resulting in a realistic and synchronized lip generation.*

The core of the network is a Transformer encoder that jointly processes these embedding to infer lip landmark $\{\hat{l}_t^l \in \mathbb{R}^{2\times n_l}\}_{t=1}^{T}$ sequences for T = 5 adjacent frames, The predicted sequences are designed to reflect both the rhythm of the speech and the speaker's identity.

3.1.1 Transformer-based Landmark Generator

We extract the audio embedding a_t from the audio Mel-spectrograms with an encoder module E_a similar to [3]. We also utilize 1D convolution layers to construct a pose encoder E_p and a reference encoder E_r which extract pose embedding p_t and reference embedding r_i from pose landmarks l_i^r, respectively. This press is formulated as follows:

$$a_t = E_a(m_t) \quad t = 1, 2, \dots, T \quad (1)$$
$$p_t = E_p(l_t^p) \quad t = 1, 2, \dots, T \quad (2)$$
$$r_i = E_r(l_i^r) \quad i = 1, 2, \dots, N_l \quad (3)$$

where $a_t, p_t, r_i \in \mathbb{R}^d$ (d is the dimension of audio, pose, or reference embedding).

For differentiation embedding calculated from there types of source signals, we introduction three learnable encoding vectors e_a^{type}, e_p^{type}, and $e_r^{type} \in \mathbb{R}^d$, indicating that the embedding is calculated from audio Mel-spectrogram, prior pose landmarks, and reference landmark respectively. The temporal positional encoding of the t-th frame is denoted as $e_t^{pose} \in \mathbb{R}^d$ which is calculated following the sinusoidal positional encoding. These encoding variables are added to the audio, pose and reference embedding as below,

$$\bar{a}_t = a_t + e_t^{pose} + e_a^{type} \quad t = 1, 2, \dots, T \quad (4)$$
$$\bar{p}_t = p_t + e_t^{pose} + e_a^{type} \quad t = 1, 2, \dots, T \quad (5)$$
$$\bar{r}_i = r_i + e_r^{type} \quad i = 1, 2, \dots, N_l \quad (6)$$

3.1.2 Loss Function for Landmark Generation

Reconstruction loss L_1 to constrain the predicted landmarks:

$$L_1 = \frac{1}{T}\sum_{t=1}^{T}\left(\left\|\hat{l}_t^l - l_t^l\right\|_1 + \left\|\hat{l}_t^j - l_t^j\right\|_1\right) \quad (7)$$

Where l_t^l and l_t^j denote the ground-truth landmarks of lip, respectively.

Improve the temporal smoothness, we also adopt the following continuity regularization to constrain the predicted landmarks:

$$L_c = \frac{1}{T-1}\sum_{t=1}^{T-1}\left(\left\|(\hat{l}_{t+1}^l - \hat{l}_t^l) - (l_{t+1}^l - l_t^l)\right\|_2 + \left\|(\hat{l}_{t+1}^j - \hat{l}_t^j) - (l_{t+1}^j - l_t^j)\right\|_2\right) \quad (8)$$

The overall training loss for audio to landmark stage is defined as follows:

$$L = L_1 + \lambda_c L_c \quad (9)$$

where λ_c is a constant and is set to 1.

3.2 Landmark-To-Video Rendering

In the second stage, we transform predicted lip landmarks [20] into photorealistic visuals. Rather than synthesizing entire facial regions, we focus solely on reconstructing the lip region for clarity and precision. We use a set of static reference images of the target identity, which are spatially aligned to match the predicted lip poses using motion-aware transformations. These aligned references supply high-frequency texture details, preserving the subject's identity across frames.

Our rendering model integrates the predicted landmarks, aligned reference features, and audio embeddings to synthesize high-fidelity lip images. The fusion is performed via spatially-adaptive normalization and audio-driven modulation to reflect speech nuances.

3.2.1 Reference Images Warping

For each reference image I_i^r and its sketch L_i^r, the alignment module encodes the channel-wise concatenation of L_i^r and I_i^r with convolution layers to visual features in two spatial resolutions: $h_i^1 \in \mathbb{R}^{c_1 \times h_1 \times w_1}, h_i^2 \in \mathbb{R}^{c_2 \times h_2 \times w_2}$. To guide the prediction of motion field F_i, the channel-wise concatenation of $2k+1$ target sketches $L_{t-k:t+k}$ are injected into the alignment module by SPADE [21] layers, which modulate the visual features according to target sketches. Here, PixelShuffle layers [22] are utilized for up-sampling. Generally, the function of the alignment module can be formulated as:

$$F_i = G_a(L_i^r, I_i^r, L_{t-k:t+k}) \qquad i = 1, 2, \dots, N \quad (10)$$

3.2.2 Sketch-To-Face Translation

In the translation module G_r, we aim to translate the target sketches concatenated with the masked target face I_t^m to the final face image $\hat{I}_t$. This is performed via the assistance of aggregated warped image $\bar{I}^r$ and feature $\bar{h}^1, \bar{h}^2$. Besides, to enhance the synthesized lip synchronization, the auditory feature $a_t \in \mathbb{R}^d$ extracted by an audio encoder similar to [3] in injected into the translation module through AdaIN layers [23].

At last, the generated full face is pasted onto the original frame during inference. However, since the generated face my include small portions of background with artifacts, we composite the generated face with the background of the original frame through a Gaussian-smoothed face mask, as depicted in **Figure 2**.



Figure 2:Post-processing for generated face pasted onto the original frame through a Gaussian-smoothed face mask.

3.2.3 Loss Function for Rendering

Following [24], the alignment module, translation module, and audio encoder are trained jointly on the perceptual loss of [25], between aggregated warped image $\bar{I}^r$ and the ground truth image I_t, is employed to constrain the alignment module for computing an accurate motion field.

$$L_w = \sum_i \|\emptyset_i(\bar{I}^r) - \emptyset_i(I_t)\|_1 \quad (11)$$

where $\emptyset_i$ is the activation output of the i-th layer in the VGG-19 network. The reconstruction loss L_r, which has a similar structure as that of L_w, and the style loss L_s, which computes the statistic error between activation output in the VGG-19 network, are applied to reduce errors between generated face $\bar{I}_t$, and ground truth I_t.

$$L_r = \sum_i \left\|\emptyset_i(\hat{I}_t) - \emptyset_i(I_t)\right\|_1 \quad (12)$$

$$L_s = \sum_i \left\|G_i^{\emptyset}(\hat{I}_t) - G_i^{\emptyset}(I_t)\right\|_1 \quad (13)$$

Where $G_i^{\emptyset}$ is the Gram matrix derived from the aforementioned activation output $\emptyset_i$. To further enhance the photorealism of the generated image, we also utilize patch GAN loss L_g and feature matching loss L_f from pix2pixHD [26]. To summarize, the training loss for the landmark-to-video rendering stage can be formulated as follows:

$$L = \lambda_w L_w + \lambda_r L_r + \lambda_s L_s + \lambda_g L_g + \lambda_f L_f \quad (14)$$

In experiments, we set $\lambda_w = 2.5$, $\lambda_r = 4$, $\lambda_s = 1000$, $\lambda_g = 0.25$, and $\lambda_f = 2.5$ based on cross-validation.

4. Experiments and Results

4.1 Experimental Setup

Implementation Details when generating landmarks from audio, we utilize the mediapipe tool [21] to detect facial landmarks from each video frame, where $n_l = 41$, $n_j = 16$, $n_r = 131$, and $n_p = 74$. Following [3], we calculate Mel-spectrograms from 16kHz audios using a window size of 800 and hop size of 200. N_l is set to 15, and d is 512. To render

videos from landmarks, 128 × 128 face images (i.e., $H = W = 128$) are generated at 25 fps, and k is set to 2. The number of reference images N is set to 3 during training and $\frac{1}{5}$ of the video length during inference. More details about network architectures and hyper-parameters are included in the supplementary document.

4.2 Datasets

Dataset Two audio-visual speech recognition datasets, **BBC LRS2 Dataset** [27] and KH-Reading Dataset (Figure 3), are used in our experiments. The **KH-Reading Dataset** (24h to 48h) ensures cultural authenticity by capturing Khmer-specific facial expressions, head movements, and language-specific lip synchronization. It offers diverse data featuring various speakers, tones, and emotions, making it crucial for creating realistic Khmer virtual avatars. In contrast, the **BBC LRS2 Dataset**, with each sentence up to 100 characters in length, serves as a reliable benchmark for lip synchronization and talking-head synthesis. Its diverse audio-visual pairings are ideal for pre-training, enabling the model to learn general synchronization patterns. Furthermore, its clean test set ensures an accurate and reliable evaluation of the model's performance. By combining these datasets, the model leverages LRS2's technical robustness and the Khmer dataset's cultural specificity, resulting in a system that is both expressive and authentic for real-world applications.



Figure 3: KH-Reading Dataset

4.3 Evaluation Metrics

Visual Quality Metrics. We evaluate the similarity between the generated frames and ground-truth images using Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) [28]. These metrics quantify pixel-level and structural fidelity, respectively. To assess perceptual quality at the feature level, we use Learned Perceptual Image Patch Similarity (LPIPS) [29] and Fréchet Inception Distance (FID) [30], both of which compare the distribution of deep features between generated and real images. To evaluate identity preservation, we compute the Cosine Similarity (CSIM) between identity vectors extracted using the ArcFace model [31].

Lip Synchronization Metrics. To measure synchronization between lip motion and speech, we adopt the Lip Landmark Distance (LipLMD) [7], which computes the normalized distance between predicted and ground-truth lip landmarks. In scenarios where ground-truth frames are unavailable (e.g., voice dubbing), we utilize SyncNet [8] to extract temporal audio-visual features and compute the SyncScore, which reflects the coherence between generated lip movements and input audio.

4.4 Experiment Result

To better understand how each component in our model contributes to its performance, we conducted ablation studies using the LRS2 dataset and KH-Reading dataset. These experiments focused on three main aspects: the type of sequence model (Transformer vs. LSTM), the use of warping and audio features, and the number of reference images used during rendering.

We first evaluated the benefit of using a Transformer encoder in the landmark prediction stage. For comparison, we replaced it with a more traditional bidirectional LSTM. In this variant, reference embeddings were averaged into a single global vector before being concatenated with pose and audio features. As shown in **Table 1**, the Transformer achieved a lower landmark error (3.01) compared to the LSTM-based model (4.32), demonstrating its superior ability to model temporal and multimodal relationships.

We explored how the number of reference images (N) used during inference influences the quality of the generated videos. As shown in **Table 2**, increasing the number of reference images from 1 to 10 consistently improves performance across all evaluation metrics. Higher PSNR and SSIM values indicate that the output frames are clearer and more faithful to the ground truth. At the same time, lower LPIPS and FID scores suggest better visual realism and perceptual quality. The increase in CSIM also shows that the model preserves the speaker's identity more effectively. Overall, using multiple reference images allows the model to better capture identity details and produce more accurate, natural-looking talking avatars.

Method	Landmark Error
w/ LSTM	4.32
w/ Transformer	3.01

Table 1: Experiment Transformer modules.

Reference Num	PSNR↑	SSIM↑	LPIPS↓	FID↓	CSIM↑
N = 1	31.88	0.9282	0.0354	30.61	0.6372
N = 5	32.34	0.9329	0.0338	29.85	0.6448
N = 10	32.61	0.9351	0.0324	29.12	0.6495

Table 2: Experiment on the number of reference images.

5. CONCLUSIONS

In this study, we proposed a two-stage lip animation system designed to create realistic talking avatars for the Khmer language. Unlike earlier models that rely on data from specific speakers, our framework is person-generic, making it more adaptable for broader applications. By combining Transformer-based audio-to-landmark prediction with a rendering process that uses reference images and spatial transformations, our system achieves strong results in both visual quality and lip synchronization. During our experiments with the KH-Reading dataset, we observed improved consistency in lip shapes and better alignment with speech timing, particularly for complex Khmer phonemes. Looking forward, this work contributes a culturally aware solution for speech-driven animation and has the potential to enhance educational tools, virtual assistants, and digital media in low-resource language settings. We believe this approach can be extended to other languages and communities that are currently underserved by mainstream AI systems.

ACKNOWLEDGMENTS

We would like to thank the BBC for providing the LRS2 dataset, which was instrumental to our research.

REFERENCES

[1] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004, doi: 10.1109/TIP.2003.819861.

[2] P. K. R, R. Mukhopadhyay, J. Philip, A. Jha, V. Namboodiri, and C. V. Jawahar, "Towards Automatic Face-to-Face Translation," in *Proceedings of the 27th ACM International Conference on Multimedia*, Oct. 2019, pp. 1428–1436. doi: 10.1145/3343031.3351066.

[3] K. R. Prajwal, R. Mukhopadhyay, V. Namboodiri, and C. V. Jawahar, "A Lip Sync Expert Is All You Need for Speech to Lip Generation In The Wild," in *Proceedings of the 28th ACM International Conference on Multimedia*, Oct. 2020, pp. 484–492. doi: 10.1145/3394171.3413532.

[4] J. Thies, M. Elgharib, A. Tewari, C. Theobalt, and M. Nießner, "Neural Voice Puppetry: Audio-driven Facial Reenactment," Jul. 29, 2020, *arXiv*: arXiv:1912.05566. doi: 10.48550/arXiv.1912.05566.

[5] S. J. Park, M. Kim, J. Hong, J. Choi, and Y. M. Ro, "SyncTalkFace: Talking Face Generation with Precise Lip-Syncing via Audio-Lip Memory," Nov. 03, 2022, *arXiv*: arXiv:2211.00924. doi: 10.48550/arXiv.2211.00924.

[6] L. Chen, R. K. Maddox, Z. Duan, and C. Xu, "Hierarchical Cross-Modal Talking Face Generationwith Dynamic Pixel-Wise Loss," May 09, 2019, *arXiv*: arXiv:1905.03820. doi: 10.48550/arXiv.1905.03820.

[7] "Animating Portrait Line Drawings from a Single Face Photo and a Speech Signal | ACM SIGGRAPH 2022 Conference Proceedings." Accessed: Apr. 25, 2025. [Online]. Available: https://dl.acm.org/doi/10.1145/3528233.3530720

[8] T. Xie *et al.*, "Towards Realistic Visual Dubbing with Heterogeneous Sources," in *Proceedings of the 29th ACM International Conference on Multimedia*, Oct. 2021, pp. 1739–1747. doi: 10.1145/3474085.3475318.

[9] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[10] "morphmod2." Accessed: Apr. 25, 2025. [Online]. Available: https://www.face-rec.org/algorithms/3D_Morph/morphmod2.pdf

[11] S. Ma, J. Fu, C. W. Chen, and T. Mei, "DA-GAN: Instance-level Image Translation by Deep Attention Generative Adversarial Networks (with Supplementary Materials)," Feb. 18, 2018, *arXiv*: arXiv:1802.06454. doi: 10.48550/arXiv.1802.06454.

[12] A. Shrivastava, T. Pfister, O. Tuzel, J. Susskind, W. Wang, and R. Webb, "Learning from Simulated and Unsupervised Images through Adversarial Training," Jul. 19, 2017, *arXiv*: arXiv:1612.07828. doi: 10.48550/arXiv.1612.07828.

[13] C. Zhang *et al.*, "FACIAL: Synthesizing Dynamic Talking Face with Implicit Attribute Learning," Aug. 18, 2021, *arXiv*: arXiv:2108.07938. doi: 10.48550/arXiv.2108.07938.

[14] X. Liu, Y. Xu, Q. Wu, H. Zhou, W. Wu, and B. Zhou, "Semantic-Aware Implicit Neural Audio-Driven Video Portrait Generation," Jan. 19, 2022, *arXiv*: arXiv:2201.07786. doi: 10.48550/arXiv.2201.07786.

[15] S. Shen, W. Li, Z. Zhu, Y. Duan, J. Zhou, and J. Lu, "Learning Dynamic Facial Radiance Fields for Few-Shot Talking Head Synthesis," Jul. 24, 2022, *arXiv*: arXiv:2207.11770. doi: 10.48550/arXiv.2207.11770.

[16] H. Zhou, Y. Sun, W. Wu, C. C. Loy, X. Wang, and Z. Liu, "Pose-Controllable Talking Face Generation by Implicitly Modularized Audio-Visual Representation," Apr. 22, 2021, *arXiv*: arXiv:2104.11116. doi: 10.48550/arXiv.2104.11116.

[17] "Synthesizing Obama: Learning Lip Sync from Audio." Accessed: Apr. 25, 2025. [Online]. Available: https://grail.cs.washington.edu/projects/AudioToObama/siggraph17_obama.pdf

[18] Y. Zhou, X. Han, E. Shechtman, J. Echevarria, E. Kalogerakis, and D. Li, "MakeItTalk: Speaker-Aware Talking-Head Animation," *ACM Trans. Graph.*, vol. 39,

no. 6, pp. 1–15, Dec. 2020, doi: 10.1145/3414685.3417774.

[19] J. Zhang, L. Liu, Z. Xue, and Y. Liu, "APB2Face: Audio-guided face reenactment with auxiliary pose and blink signals," Apr. 30, 2020, *arXiv*: arXiv:2004.14569. doi: 10.48550/arXiv.2004.14569.

[20] M. C. Doukas, E. Ververas, V. Sharmanska, and S. Zafeiriou, "Free-HeadGAN: Neural Talking Head Synthesis with Explicit Gaze Control," Aug. 03, 2022, *arXiv*: arXiv:2208.02210. doi: 10.48550/arXiv.2208.02210.

[21] T. Park, M.-Y. Liu, T.-C. Wang, and J.-Y. Zhu, "Semantic Image Synthesis with Spatially-Adaptive Normalization," Nov. 05, 2019, *arXiv*: arXiv:1903.07291. doi: 10.48550/arXiv.1903.07291.

[22] W. Shi *et al.*, "Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network," presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1874–1883. Accessed: Apr. 25, 2025. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/Shi_Real-Time_Single_Image_CVPR_2016_paper.html

[23] X. Huang and S. Belongie, "Arbitrary Style Transfer in Real-time with Adaptive Instance Normalization," Jul. 30, 2017, *arXiv*: arXiv:1703.06868. doi: 10.48550/arXiv.1703.06868.

[24] Y. Ren, G. Li, Y. Chen, T. H. Li, and S. Liu, "PIRenderer: Controllable Portrait Image Generation via Semantic Neural Rendering," Sep. 17, 2021, *arXiv*: arXiv:2109.08379. doi: 10.48550/arXiv.2109.08379.

[25] J. Johnson, A. Alahi, and L. Fei-Fei, "Perceptual Losses for Real-Time Style Transfer and Super-Resolution," Mar. 27, 2016, *arXiv*: arXiv:1603.08155. doi: 10.48550/arXiv.1603.08155.

[26] T.-C. Wang, M.-Y. Liu, J.-Y. Zhu, A. Tao, J. Kautz, and B. Catanzaro, "High-Resolution Image Synthesis and Semantic Manipulation with Conditional GANs," Aug. 20, 2018, *arXiv*: arXiv:1711.11585. doi: 10.48550/arXiv.1711.11585.

[27] "Lip Reading Sentences 2 (LRS2) dataset." Accessed: Jan. 31, 2025. [Online]. Available: https://www.robots.ox.ac.uk/~vgg/data/lip_reading/lrs2.html

[28] "Image Quality Assessment: From Error Visibility to Structural Similarity." Accessed: Apr. 25, 2025. [Online]. Available: https://www.cns.nyu.edu/pub/lcv/wang03-preprint.pdf

[29] "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric." Accessed: Apr. 25, 2025. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2018/papers/Zhang_The_Unreasonable_Effectiveness_CVPR_2018_paper.pdf

[30] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," Jan. 12, 2018, *arXiv*: arXiv:1706.08500. doi: 10.48550/arXiv.1706.08500.

[31] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive Angular Margin Loss for Deep Face Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 5962–5979, Oct. 2022, doi: 10.1109/TPAMI.2021.3087709.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

KHR/USD Daily Exchange Rate Short Term Prediction using Deep Learning

Hourpheng Hort*, Sokkhey Phauk, Pheak Neang, Vesal Khean

Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

Corresponding author: hourpheng_hort@gsc.itc.edu.kh

Abstract: *Exchange rate prediction is a crucial aspect of financial analysis, particularly for countries with dual currency systems such as Cambodia. The Khmer Riel (KHR) is widely exchanged with the United States Dollar (USD), making accurate forecasting models essential for businesses and policymakers. This study explores deep learning-based approaches to predict the KHR/USD daily exchange rate using historical data collected from the National Bank of Cambodia (NBC).*

The study implemented and compared four deep learning models: Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and Hybrid GRU and LSTM models. The models were trained using a sequence length of 30 days to forecast one day value and evaluated using Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-squared (R^2). All models perform hand-to-hand with one another; however, GRU demonstrated the best predictive performance leading by a slight improvement, achieving a RMSE of 3.99 and an R^2 of 0.9819. The results indicate that simpler recurrent deep learning models, are effective in capturing complex patterns in exchange rate time series, providing valuable insights for financial decision-making.

Keywords: Khmer Riel; Exchange Rate Prediction; Deep Learning; LSTM; GRU

1. INTRODUCTION[1]

Exchange Rate is defined as the measurement of a value of one currency in terms of another or the domestic price of the foreign currency; it is often referenced against the United State dollar (USD) because it is the most used currency in the international market making it the de-facto benchmark for the values of other currency. The Khmer riel or Cambodia's Riel (KHR) is also measured in reference to the U.S. dollar.

The history of the Cambodian Riel, or Khmer Riel (KHR), began with its initial introduction in 1953, coinciding with Cambodia's independence from French colonization. This first iteration of the currency was abolished in 1979 by the Khmer Rouge regime. Following the overthrow of this genocidal regime, the People's Republic of Kampuchea (PRK) reintroduced the Khmer riels into the economy on March 20th, 1980, through the rebuilt of National Bank of Cambodia (NBC) [1]. Initially, the implied value of the Khmer riels in the 1980s was four to one against the United States Dollar. Following political and economic instability, the Khmer riels has experience series events to hyperinflation depreciated its values to 3744.42 riels per U.S in 1997 [2], [3]. Since 2001 the exchange rate became stable fluctuating around 4000 riels per one dollar [1].

The Cambodia's monetary policy is set by the National Bank of Cambodia (NBC) adopting a de-jure managed float regime. The official exchange rate, quoted as riels per dollars (KHR/USD), governs all external transactions made with/by the central government and state enterprises, and serves as accounting rate by the NBC. This official rate is determined daily with adjustment to remain volatile at only one percent of the market rate [3], [4]. Though international foreign exchange market intervention and carefully supply the Khmer riel according to the country's economic condition and its market, the NBC has continued maintaining a stable exchange rate of the currency pair for more than two decades [5].

Despite having a very stable currency exchange rate, Cambodia is known as the most dollarized country in Asia with the dollarization rate calculated by foreign currency deposits to total deposit (RHS) at slightly more than 90% in August 2024 [4]. Consequently, the U.S. dollar remains the preferred currency for large and wholesale transactions, savings, investments, and international trade in Cambodia. The Khmer Riel, on the other hand, is primarily used for smaller, retail transactions. It's common to observe local vendors, market sellers, small businesses, and farmers pricing their goods in Riels but often preferring or using U.S. dollars for purchases and settling larger bills. This creates a significant daily exchange volume between the two

* Corresponding author: Hourpheng HORT
E-mail: hourpheng_hort@gsc.itc.edu.kh; Tel: +855-85 571 071

currencies, particularly in areas of active trade here in Phnom Penh and across the country.

Given the current stability of the KHR/USD exchange rate, it remains vital to thoroughly study the factors influencing the Khmer Riel's value against the U.S. dollar. Furthermore, as the National Bank of Cambodia continues to implement monetary policies aimed at de-dollarizing the Cambodian economy, this research will provide essential insights for policymakers and those with the authority to make informed decisions.

In their investigation of Cambodia's Foreign Exchange Rate dynamics, [6] found that a Markov-Switching Autoregressive model, specifically MS(2)-AR(1) as evaluated by the Akaike Information Criterion (AIC), provided the best fit. This study on the daily exchange rate clearly detected two distinguished regimes for the KHR/USD exchange rate where state 1 is characterized by higher volatility of 18.16, stronger dependent on the AR(1) with coefficient of 18.26, but lower persistent to remain in its state with probability of 70.66% and higher tendency to transit into state 2 with probability of 29.34%. The state 2 showed drastically lower volatility of only 2.39, less dependent on first lag of 2.39, and demonstrated higher probability of remaining in the same state of 84.47% and less tendency to switch state at only 15.53%. Interestingly, the intercepts for both states were very close (4078.2 and 4078.4), indicating similar long-term averages. Thus, the study highlights that the primary distinctions between these two exchange rate regimes are driven by the volatility, the autoregressive component, and the probabilities of transitioning between them.

Research conducted by [7] analyzed the connection between the Khmer Riel (KHR) to U.S. dollar (USD) exchange rate and Cambodia's broad money (M2). The study established that both monthly time series were stationary after first differencing, and the Engle-Granger cointegration test indicated a significant long-term relationship between them. To model this relationship, the author utilized an Autoregressive Integrated Moving Average with eXogenous variables (ARIMAX) model, with the ARIMAX(1,1) specification identified as optimal based on the Schwarz Bayesian Information Criterion (SBIC). The findings revealed a positive correlation: a 100 million KHR increase in M2 is associated with a 41.8 KHR depreciation per USD, while a decrease of the same magnitude leads to an equivalent appreciation, assuming other relevant economic variables in Cambodia are held constant.

Existing research on the KHR/USD exchange rate had been largely relied on statistical models to explore its dynamics and determinants. However, there is a clear need for studies employing machine learning models to further enhance prediction accuracy and provide a more nuanced understanding of this crucial exchange rate. This paper aimed to address this gap by applying advanced deep learning models—namely Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and a variation of their hybrid forms—to the short-term prediction of the KHR/USD monthly exchange rate. The outcomes of this research could offer significant advancements in forecasting capabilities and a more comprehensive analysis of the KHR/USD daily exchange rate behavior, potentially benefiting financial institutions and policymakers.

This study is organized into four distinct chapters. The first chapter serves as an introduction, providing an overview of the research topic. Chapters two and three delineate the research methodology utilized and present the empirical findings, respectively. Lastly, the concluding chapter succinctly summarizes the principal findings and provides insights derived from the study.

2. METHODOLOGY

2.1 Dataset

The original dataset was collected from the National Bank of Cambodia (NBC) website. Data was the daily exchange rate of USD/KHR 1st January, 2009 to 20th March, 2024 accounting for 5558 days. With datetime index corresponds to a specific day of the month of a specific year, the dataset has 8 columns:

a. Market Exchange Rate - Purchase
b. Market Exchange Rate - Sale
c. Market Exchange Rate - Midpoint
d. Official Exchange Rate - Purchase
e. Official Exchange Rate - Sale
f. Official Exchange Rate - Midpoint
g. Percent change of the market buying rate - M. to M. (Month to Month)
h. Percent change of the market buying rate - From the End of Last Year Through Indicated Month

2.2 Feature Engineering and Data Preprocessing

2.2.1 Feature Engineering

The dataset for both the exchange rate was clean, with no missing values or duplicates identified. Furthermore, the Official Exchange Rate Midpoint was used as the primary raw time series for our target variable (KHR/USD). As highlighted in , raw time series data is often a crucial component of the feature vector, particularly in RNN-based models for financial time series forecasting. To incorporate elements similar to statistical models like ARIMA, we extracted the following additional endogenous features directly from the target time series:

a. Lag Features (1 to 31)
b. Rolling Means (2 to 31)
c. Rolling Standard Deviation (2 to 31)
d. Exponential Moving Average (2 to 31)

Following the generation of endogenous input features, we had a total of 121 features. To ensure the relevance of these features for our model, we calculated the Mutual Information (MI) score for each against the target variable (KHR/USD exchange rate). Features scoring below a threshold of 0.1 were subsequently removed to reduce noise and improve model efficiency. Due to the inclusion of lagged components in our feature set, we anticipated a reduction in the number of usable observation points for model training, resulting in a final dataset of 5527 daily observations.

2.2.2 Data Processing

Applying scaling to input features is crucial before training deep learning models. This step ensures better generalization, avoids bias, promotes faster convergence, and is particularly important in preventing vanishing gradients during backpropagation. Following the guidelines in [8] for optimal input transformation—namely centering the average near zero, ensuring similar covariances, and ideally decorrelating inputs—we opted for normalization. Specifically, for our recurrent neural network (RNN) and related models, normalization to the range of (-1, 1) is generally favored over (0, 1). This practical preference, likely stemming from its interaction with tanh activation functions.

2.3 Autoregressive Integrated Moving Average (ARIMA)

ARIMA is a basic statistical model often use to be a baseline model, which combines Auto-Regressive (AR) part and Moving Average (MA) part using Integrated part to handle non-stationary time series by differencing. $ARIMA\ (p, d, q)$ is written to represent the model where p represent the order of AR part, q represent degree of differencing until stationary, and q the order of MA part. $ARIMA\ (p, d, q)$ is defined by Eq.1:

$$y'_t = \mu + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q} + \epsilon_t \quad \text{(Eq.1)}$$

Where,

y'_t represents the d^{th} differenced series at time t.
μ is the intercept
ϵ_t is the error term at time t
$\epsilon_{t-1,\dots,t-q}$ are the past q error terms
$\phi_{1,2,\dots\ p}$ are autocorrelation coefficient at lags $1, 2, \dots p$
$\theta_{1,2,\dots\ q}$ are moving average coefficient at lags $1, 2, \dots p$

And y'_t is defined by Eq.2:

$$y'_t = (1 - L)^d\ y_t \quad \text{(Eq.2)}$$

Where,

y_t is the original time series
L is the lag operator

2.4 Deep Learning Models

A semantic literature review by on deep learning applications in financial time series forecasting (2005-2019) highlighted the prevalence of Recurrent Neural Networks (RNNs), which were utilized in 51% of the surveyed studies. Multilayer Perceptrons (MLPs) and Convolutional Neural Networks (CNNs) were employed in 20.8% and 14.1% of the reviewed literature, respectively. Within the RNN category, Long Short-Term Memory (LSTM) networks emerged as the most popular architecture (60.4%), followed by Vanilla RNNs (29.7%) and Gated Recurrent Units (GRUs) (9.89%).

Given the sequential nature of time series data, Recurrent Neural Networks (RNNs) are a natural choice for analysis, as they can leverage internal memory to process incoming inputs and capture dependencies over longer time horizons [9]. While Long Short-Term Memory (LSTM) networks have been the prevalent RNN architecture in financial time series forecasting research, Gated Recurrent Units (GRUs), which represent a more streamlined version of LSTMs, offer an intriguing alternative. Interestingly, Gated Recurrent Units (GRUs), despite their release in 2014 (later than LSTM's 1998 introduction), were likely underrepresented in the semantic review by [9]. They have shown promising results, often outperforming LSTMs in tasks characterized by lower computational costs, shorter sequence lengths, and limited datasets.

2.4.1 Long Short-Term Memory (LSTM)

Long Short-Term Memory (LSTM) networks are a specialized type of RNN engineered to effectively handle both short-term and long-term dependencies in sequential data [8]. They achieve this through the introduction of memory cells, which can maintain information over extended time intervals. Crucially, the flow of information into and out of these memory cells (c) is carefully managed by three key gates: the input gate (i), the forget gate (f), and the output gate (o). These gates selectively regulate the information, allowing the network to learn and retain relevant values across arbitrary time spans [8]. The equations (Eq.3-Eq.7), detailing the flow of information of LSTM, are explained step-by-step in an online resource [10].

$$f_t = \sigma(W_f . [h_{t-1}, x_t] + b_f) \quad \text{(Eq.3)}$$

$$i_t = \sigma(W_i . [h_{t-1}, x_t] + b_i) \quad \text{(Eq.4)}$$

$$C_t = f_t * C_{t-1} + i_t * tanh(W_c . [h_{t-1}, x_t] + b_c) \quad \text{(Eq.5)}$$

$$o_t = \sigma(W_o . [h_{t-1}, x_t] + b_o) \quad \text{(Eq.6)}$$

$$h_t = o_t * tanh(C_t) \quad \text{(Eq.7)}$$

Where,

x_t is input vector to the LSTM unit
h_t is output vector (hidden state) to the LSTM unit
$C_t$is cell state vector

f_t, i_t, o_t are activation vectors of forget gate, input gate, and output gate respectively

σ is sigmoid function

W is weight matrices that need to be learned and

b is bias vector parameters that need to be learned

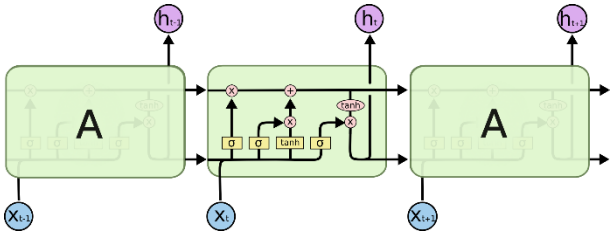


Fig. 1 The Repeating module in an LSTM contains four interacting layers. Source: [10]

Gated Recurrent Units (GRU) emerging as a potent yet alternative to the LSTM is the Gated Recurrent Unit, or GRU, prioritizes the hidden state (h) over a distinct cell state (C) and streamlines the gating mechanism by combining the input (i) and forget gates (f) into an "update gate" (z). This gate determines how much information of the past hidden state should be retained and pass on to the future timestep. Complementing this, the reset gate (r), determines how much of the past information of the past hidden state need to reset or discard [11]. The equations (Eq.8-Eq.11), detailing the flow of information of LSTM, are explained step-by-step in an online resource [10].

$$z_t = \sigma(W_z.[h_{t-1},\ x_t]) \quad \text{(Eq.8)}$$

$$r_t\ = \sigma(W_r.[h_{t-1},\ x_t]) \quad \text{(Eq.9)}$$

$$\tilde{h}_t = tanh(W.[r_t * h_{t-1}, x_t]) \quad \text{(Eq.10)}$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad \text{(Eq11)}$$

where,

x_t is input vector to the GRU unit

h_t is hidden state vector to the GRU unit

z_t, r_t are activation vectors of update gate and reset gate

σ is sigmoid function and

W is weight matrices that need to be learned

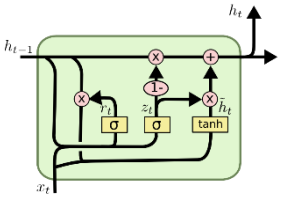


Fig. 2 GRU architecture. Source: [10]

2.4.2 Hybrid Models

Recognizing the potential for enhanced robustness and accuracy by combining the strengths of different modeling approaches, researchers have proposed various hybrid statistical and machine learning models. In our study, however, we specifically focused on the synergistic combination of two sequence-based deep learning models: the Long Short-Term Memory (LSTM) network and its simpler counterpart, the Gated Recurrent Unit (GRU). Our aim is to investigate the benefits of introducing diversity within the deep learning architecture itself, exploring how layering these two distinct sequence models can improve the prediction.

2.5 Evaluation Metrics

Common evaluation metrics for non-classification prediction tasks, relevant to the forecasting of the KHR/USD exchange rate and broad money, include Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), Root Mean Squared Error (RMSE), and the coefficient of determination (R^2) [8]. For this paper, we explore all metrics; however, the R^2 is used to decide the best fit model. The equations (Eq.12-Eq.15) illustrate how all evaluation metrics are calculated.

$$MAE = \frac{1}{n}\Sigma_{i=1}^{n}|y_i - \hat{y}_i| \quad \text{(Eq.12)}$$

$$MAPE = \frac{1}{n}\Sigma_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100\% \quad \text{(Eq.13)}$$

$$RMSE = \sqrt{\frac{1}{n}\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \text{(Eq.14)}$$

$$R^2 = 1 - \frac{\Sigma_{i=1}^{n}(y_i - \hat{y}_i)^2}{\Sigma_{i=1}^{n}(y_i - \bar{y})^2} \quad \text{(Eq.15)}$$

where,

n is the number of data points or observation

y_i is the actual value of the target at i^{th} data point

$\hat{y}_i$ is the predicted value of the target at i^{th} data point

$\bar{y}$ is the mean of actual values of all data points

2.6 Related Works

Predicting financial time series and exchange rates is complex due to various methods, including macroeconomic theories, statistical models, and advanced machine learning techniques like deep learning. In this part, we will discuss statistical models and advanced machine learning models that was used for their work that related to ours.

Statistical models, often called conventional methods by data scientists, rely on the assumption of stationarity (constant mean and variance) in stochastic processes. However, these models often fail to capture the inherent complexity, non-linear relationship present in the time series [12]. For this study, ARIMA will serve as the baseline model. This will allow us to quantify the performance improvement of our proposed advanced model compared to a simpler, less sophisticated approach. As in [13], [14], [15],

the author chose *ARIMA* as a baseline model to compare with the more sophisticated model such that *RNN* and *SVM* in [13], *LSTM* in [14], *LSTM* and hybrid *ARIMA-LSTM* [15]. The comparative analysis revealed that these complex models consistently provided significant improvements in performance over the traditional *ARIMA* model.

Apart from the comparison analysis with the statistical model, there are many works focus solely on comparing deep learning models to one another. [16] has found that LSTM provides better results than Multi-Layer Perceptron (MLP), Simple Recurrent Neural Network (SRNN), and Gated Recurrent Unit (GRU) on their prediction of daily exchange rates of Nepalese Rupee (NPR) against three other currencies like American Dollar (USD), Pound Sterling (GBP) and EURO. Noticeably, recurrent neural networks like LSTM and GRU consistently demonstrated superior performance over simpler architectures such as MLP and SRNN, based on Mean Absolute Error (MAE). Specifically, the LSTM model was most effective for USD/NPR with an MAE of 0.013, and EUR/NPR with an MAE of 0.042. For GBP/NPR, the GRU model emerged as the top performer with an MAE of 0.0177.

[17] and [18] employed and compared LSTM and GRU to predict the daily exchange rate and cryptocurrencies respectively, where [17] studied the daily exchange rate prediction of EUR/USD with different data allocation 70/15/15 and 80/10/10 of train/validation/test set. The results show that GRU model that use the 80/10/10 data allocation was the highest performance among all other variation of the experiment with RMSE value of 0.054, MAPE of 0.037, and R square of 0.97. In the context of cryptocurrency forecasting by [18], GRU model consistently demonstrated superior performance compared to LSTM across various pairs. Specifically, for BTC/USD, the GRU achieved an MAE of 0.0325, outperforming LSTM's 0.0674. Similarly, in ETH/USD, GRU recorded an MAE of 0.0305, which was lower than LSTM's 0.0353. For LTC/USD, GRU again exhibited a slightly better performance with an MAE of 0.0170 compared to LSTM's 0.0179. Consequently, the GRU model was identified as the optimal performer for this prediction task due to its superior evaluation metrics and enhanced generalization. This was evidenced by a more consistent test loss relative to training loss, alongside a faster convergence rate.

3. RESULTS AND DISCUSSION

Our study utilized 5546 daily observations of the primary dataset, the KHR/USD exchange rate and 43 endogenous features derived from the exchange rate itself and go through Mutual Important Score > 0.1. This dataset spans from the 13th January 2009 to the 20th March 2024. Statistical analysis of the target variable revealed an average daily exchange rate of approximately 4072.77 KHR, with a sample standard deviation of 52.95 KHR. The distribution of the target variable exhibited a skewness of 0.95 and a kurtosis of 1.09. A comprehensive summary of this statistical analysis is presented in Table 1.

Table 1. Summary Statistics of KHR/USD daily exchange rate

	KHR/USD daily Exchange Rate
Mean	4072.77 KHR
Median	4065 KHR
Maximum	4241 KHR
Minimum	3980 KHR
Std. Dev.	52.95 KHR
Skewness	0.95
Kurtosis	1.09
Observations	5546 days

3.1 Baseline Model (ARIMA)

In statistical time series analysis, achieving stationarity, often confirmed by unit root tests such as the Augmented Dickey-Fuller (ADF) or Kwiatkowski-Phillips-Schmidt-Shin (KPSS) tests, is a prerequisite for many modeling approaches. The Auto-Regressive Integrated Moving Average (ARIMA) model, recognized for its simplicity, computational efficiency, and interpretability rooted in statistical foundations, is widely adopted as a baseline or "benchmark" model in time series studies.

In this study, *pmdarima* was utilized to automate the selection of ARIMA model components, allowing for a maximum order of 7 for both Auto-Regressive and Moving Average parts, with the degree of differencing automatically determined to achieve stationarity. Applied to the feature-engineered time series for consistent comparison with deep learning models, the optimized ARIMA model, achieving an AIC of 21775.77, demonstrated significant limitations. Its evaluation metrics (MAE: 32.70, MAPE: 0.79%, RMSE: 41.93, and R2: -0.96), particularly the highly negative R2, unequivocally indicated that this linear model was not adequately fitted for the prediction task and failed to generalize effectively. This poor performance was further supported by residual analysis, which suggested the underlying data was essentially white noise, confirming the unsuitability of a linear approach as shown in Figure 3 below.



Fig. 3 Residuals obtained from ARIMA

3.2 Deep Learning Models

Unlike statistical time series analysis, time series analysis using machine learning does not require the time series to be stationary due to their ability to capture complex pattern namely non-linear relationship and furthermore they focus more on accuracy rather than interpretability.

The dataset was trained on training set accounting for 80% of the whole data and among that 80% the last 20% was used as validation set; lastly, the remaining 20% from the training set was used as testing set. After this split, the data was transformed into tensors with input features (X) shaped as (32 batch size, 30 sequence length, 43 input features) and the target variable (y) shaped as (32 batch size, 1 output). We then experimented with three recurrent deep learning models based on the LSTM architecture—LSTM, GRU, and the hybrid GRU-LSTM. These models shared a consistent structure of three hidden layers with 64, 128, and 64 neurons respectively, a learning rate of 0.001, the AdamW optimizer with a weight decay of 0.001, loss function $L1loss()$,and a maximum of 400 epochs. To prevent overfitting and optimize training time, we implemented a dropout rate of 0.3, early stopping with a patience of 20 epochs and gradient clipping set at 1.

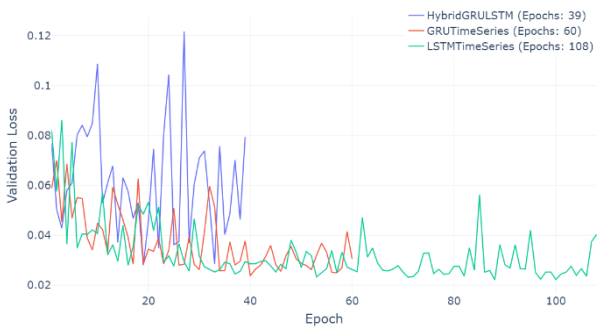


Fig. 4 Validation Loss of Different Deep Learning Models

Fig. 3 illustrates validation loss graphs of all models, where green, red, and blue lines represent LSTM, GRU, and Hybrid GRU-LSTM respectively. The LSTM model demonstrates a trend towards stability, benefiting from extended training up to 108 epochs. Notably, the GRU model achieves comparable validation loss to LSTM but with significantly fewer training epochs, stopping at just 60. In contrast, despite its increased complexity and higher computational demands, the hybrid GRU-LSTM model exhibited instability throughout training, characterized by prominent peaks in validation loss. Consequently, it triggered early stopping at a much earlier stage of only 39 epochs.

Table 2 presents the evaluation metrics for all trained models. As outlined in the methodology, the coefficient of determination, R2, was selected as the primary evaluation measure due to its ability to represent how well the model fits the data by quantifying the proportion of variance in the dependent variable explained by the models. Among all models, the Gated Recurrent Units (GRU) model achieved the highest R2 score at 0.9819, indicating the best fit for the data. Notably, the other models, LSTM and GRU-LSTM, also demonstrated very comparable high R2 scores of 0.9781 and 0.9769, respectively. The GRU model's robustness is evident not only in its highest R^2 score but also in its consistently superior performance across all other key metrics: MAE (2.365), MAPE (5.78%), and RMSE (3.9849). This consistent accuracy and strong fit across various measures solidify the GRU model as the most robust choice for predicting the daily KHR/USD exchange rate.

Table 2. Evaluation Metrics on test set

Metric	LSTM	GRU	GRU-LSTM
MAE	2.6020	2.365	2.8413
MAPE	0.0635 (6.35%)	0.0578 (5.78%)	0.0694 (6.94%)
RMSE	4.3794	3.9849	4.5007
R2	0.9781	0.9819	0.9769

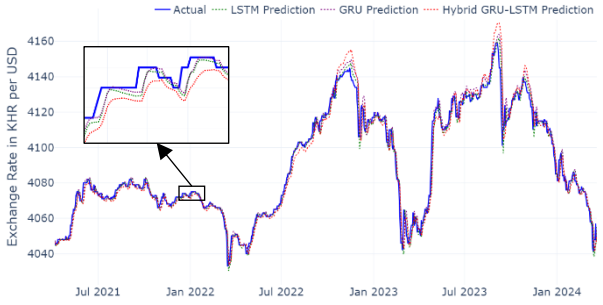


Fig. 5 Actual vs Prediction of Train set

Figure 4 further solidifies the claim that all models effectively capture the overall direction, trend, and major fluctuations within the test set observation period, spanning from 7th March 2021 to 20th March 2024. Observing the predicted graphs, the purple dotted line representing the GRU model's output closely follows the actual data, succeeded by the green dotted line indicating the LSTM predictions. The orange dotted line represents the hybrid GRU-LSTM model, which appears to be the least performant among the three.

In conclusion, the evaluation of our models clearly demonstrates that all three—LSTM, GRU, and the hybrid GRU-LSTM—exhibited remarkably similar performance. However, despite these marginal numerical differences in the evaluation metrics, the consistent outperformance of the GRU model across all indicators suggests it is the preferred choice, even when compared to the more complex LSTM and the hybrid architecture. This underscores an important point: increasing model complexity or employing more advanced methods does not automatically guarantee superior

performance and can potentially lead to overfitting and increased computational costs.

4. LIMITATIONS AND FUTURE RESEARCH

Despite demonstrating significant performance improvements over the ARIMA baseline, and with promising results from the deep learning models, a crucial area for future research lies in enhancing their interpretability. This study focused solely on univariate time series analysis, which, while valuable for understanding the inherent dynamics of the KHR/USD exchange rate, does not incorporate the influence of external factors crucial for real-world interpretation. Future work should therefore explore the inclusion of macroeconomic indicators, sentiment analysis, and other relevant exogenous variables. Furthermore, a thorough investigation of input features, both internal and external, is necessary, specifically considering the unique economic, financial, and monetary contexts of the Khmer Riel and the United States Dollar.

In addition to these efforts, future work could also benefit from exploring other modeling approaches, such as Random Forests or Gradient Boosting, which often offer a strong balance of predictive power and interpretability at a potentially lower computational cost. Additionally, delving into more complex ensemble deep learning models, like stacked LSTMs or attention-based networks, might yield further improvements in forecasting accuracy.

To further address the interpretability challenge inherent in deep learning's ability to capture non-linearities, state-of-the-art methods like Gradient-weighted Class Activation Mapping (Grad-CAM), particularly when applied to the temporal dimension of the data, can highlight the most influential time steps contributing to the model's output. Furthermore, integrating attention mechanisms can provide insights into which parts of the input sequence the model is focusing on when making predictions, thereby improving model transparency and facilitating a more readable and understandable view of the deep learning model's decision-making process [19].

Crucially, this study's data spanned up to March 2024. Given that we are now in late April 2025, over a year later, it would be highly valuable to conduct out-of-sample forecasting using more recent data to ensure the models have truly captured the underlying patterns of the input features and can effectively predict the target variable and scale well to new, unseen data.

5. CONCLUSIONS

In order to predict the KHR/USD daily exchange rate, in this paper, we employed three experimental models—LSTM, GRU, and the hybrid GRU-LSTM-GRU—where each model consisted of three layers with 64, 128, and 64 respectively along with the same training configurations and regularization to prevent overfitting. The empirical experiments results show that the evaluation of our models reveals that while LSTM, GRU, and the hybrid GRU-LSTM achieved comparable overall performance, the GRU model consistently demonstrated superior results across all evaluation metrics with only marginal numerical differences. This consistent outperformance highlights the GRU as the most effective and efficient choice, even when considering the more complex LSTM and hybrid approaches. This reinforces the important insight that increased model complexity or the implementation of sophisticated techniques does not automatically translate to enhanced performance and can potentially lead to overfitting and higher computational overhead.

This proposed study is currently in a relatively early stage of development of a comparative analysis between ensemble learning methods—specifically XGBoost and Random Forest—and recurrent deep learning models as applied in this study. Future work will incorporate various exogenous variables to enhance real-world applicability and implement methods to improve the interpretability of deep learning models. This will allow for a comprehensive comparison of how both ensemble machine learning models and deep learning models differ when applied to financial time series analysis, using the KHR/USD exchange rate as a case study.

ACKNOWLEDGMENTS

My sincere appreciation and deepest gratitude extend to Dr. Sokkhey PHUAK, Dr. Pheak NEANG, and Mr. Vesal KHEAN for their invaluable support throughout this research. Their profound insights, exceptional expertise, and unwavering encouragement with great patience, were instrumental in not only completing this work, but also in fostering a comprehensive understanding of the subject matter. Their supervision has been a great source of direction, and I am honored to have had the opportunity to contribute to this project under their supervision.

REFERENCES

[1] H. Okuda and S. Chea, *Cambodian Dollarization: Its Policy Implications for LDCs' Financial Development*, 1st ed. London: Routledge, 2023. doi: 10.4324/9781003391418.

[2] C. Hay, "Dollarization and macroeconomic performance in Cambodia since the first 1993 general election: a historical perspective," *IJFBS*, vol. 10, no. 2, pp. 27–46, Jun. 2021, doi: 10.20525/ijfbs.v10i2.1179.

[3] H. Chanthol, "Money Demand and Inflation in a Highly Dollarized Economy: Fighting Inflation in Cambodia".

[4] International Monetary Fund. Asia and Pacific Dept, "Cambodia," *IMF Staff Country Reports*, vol. 2025, no. 022, p. 1, Jan. 2025, doi: 10.5089/9798400299711.002.

[5] M. Sean, P. Pastpipatkul, and P. Boonyakunakorn, "Money Supply, Inflation and Exchange Rate Movement: The Case of Cambodia by Bayesian VAR Approach," *JOMEINO*, pp. 63–81, Jan. 2019, doi: 10.31039/jomeino.2019.3.1.5.

[6] S. Lim, "Modelling Cambodia's Foreign Exchange Rate Dynamics: A Markov-Switching Autoregressive Model," *IJEFI*, vol. 15, no. 1, pp. 330–336, Dec. 2024, doi: 10.32479/ijefi.17618.

[7] S. Lim, "Monetary Policy and Foreign Exchange Fluctuations in Cambodia," *14/09/2023*, vol. 2, no. 9, pp. 30–37.

[8] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient BackProp," in *Neural Networks: Tricks of the Trade*, vol. 7700, G. Montavon, G. B. Orr, and K.-R. Müller, Eds., in Lecture Notes in Computer Science, vol. 7700. , Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 9–48. doi: 10.1007/978-3-642-35289-8_3.

[9] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "Financial Time Series Forecasting with Deep Learning : A Systematic Literature Review: 2005-2019," Nov. 29, 2019, *arXiv*: arXiv:1911.13288. doi: 10.48550/arXiv.1911.13288.

[10] C. Olah, "Understanding LSTM Networks -- colah's blog," Understanding LSTM Networks. Accessed: May 27, 2025. [Online]. Available: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

[11] M. S. Islam and E. Hossain, "Foreign exchange currency rate prediction using a GRU-LSTM hybrid network," *Soft Computing Letters*, vol. 3, p. 100009, Dec. 2021, doi: 10.1016/j.socl.2020.100009.

[12] S. Ghahremani and U. T. Nguyen, "Prediction of Foreign Currency Exchange Rates Using an Attention-Based Long Short-Term Memory Network," 2024. doi: 10.2139/ssrn.5038611.

[13] R. Liao, P. Boonyakunakorn, N. Harnpornchai, and S. Sriboonchitta, "Forecasting the Exchange Rate for USD to RMB using RNN and SVM," *J. Phys.: Conf. Ser.*, vol. 1616, no. 1, p. 012050, Aug. 2020, doi: 10.1088/1742-6596/1616/1/012050.

[14] S. S. Namin and A. S. Namin, "FORECASTING ECONOMIC AND FINANCIAL TIME SERIES: ARIMA VS. LSTM".

[15] F. García, F. Guijarro, J. Oliver, and R. Tamošiūnienė, "Foreign Exchange Forecasting Models: ARIMA and LSTM Comparison," in *ITISE 2023*, MDPI, Jul. 2023, p. 81. doi: 10.3390/engproc2023039081.

[16] S. Ranjit, S. Shrestha, S. Subedi, and S. Shakya, "Comparison of algorithms in Foreign Exchange Rate Prediction," in *2018 IEEE 3rd International Conference on Computing, Communication and Security (ICCCS)*, Kathmandu: IEEE, Oct. 2018, pp. 9–13. doi: 10.1109/CCCS.2018.8586826.

[17] M. R. Pahlevi, K. Kusrini, and T. Hidayat, "Comparison of LSTM and GRU Models for Forex Prediction," *SinkrOn*, vol. 8, no. 4, pp. 2254–2263, Oct. 2023, doi: 10.33395/sinkron.v8i4.12709.

[18] R. Kaur *et al.*, "Development of a cryptocurrency price prediction model: leveraging GRU and LSTM for Bitcoin, Litecoin and Ethereum," *PeerJ Computer Science*, vol. 11, p. e2675, Mar. 2025, doi: 10.7717/peerj-cs.2675.

[19] A. Casolaro, V. Capone, G. Iannuzzo, and F. Camastra, "Deep Learning for Time Series Forecasting: Advances and Open Problems," *Information*, vol. 14, no. 11, p. 598, Nov. 2023, doi: 10.3390/info14110598.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

A Multilingual Deep Learning Approach Fine-Tuning to Khmer News Summarization

Soknan KOUM [1*], Kimlong NGIN [2], Chanthin PHAO [1], Sane BO [1], Ousa MA [1]

[1] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.Box 86, Phnom Penh, Cambodia*

[2] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: koumsoknan096@gmail.com*

Abstract: *News summarization is the process of extracting key insights from an original news article in a precise way without spending much time reading the whole text while still obtaining the essence of the information. As newspapers have become less popular and most people now consume news online rather than spending time reading everything from the start, extracting the useful and significant main ideas quickly and effectively is essential. Therefore, it is necessary to develop a high-performance model for the Khmer language, which is one of the most linguistically complex languages due to the absence of spaces between words. The dataset used in this study consists of 2,500 news articles collected from Sabay News and Khmerload through web scraping. It includes two main components: the original article (content) and its corresponding summary, which serves as a reference for evaluation. Since the Khmer language does not contain spaces between words, word segmentation is a critical preprocessing step. To address this challenge, we utilize Khmer-NLP-Tools, a library for advanced Khmer natural language processing in Python. In this paper, two pre-trained models are meticulously fine-tuned to yield the most accurate results: mBART (Multilingual BART), and mT5 (Multilingual T5). To evaluate the quality of summarization, the ROUGE score (ROUGE-1, ROUGE-2, and ROUGE-L) is used to ensure readability and grammatical correctness. Extensive experimental data indicate that the mBART outperforms other models. To sum up, the results of this study hold significant value for applications in various fields such as education, news, government services, and more. Additionally, this study contributes to the development and advancement of Khmer NLP, paving the way for further improvements in the field.*

Keywords: Khmer news summarization; Pre-trained model; Khmer-NLP-Tools; mBART; mT5

1. INTRODUCTION

Over the past several decades, text summarization has grown significantly from simple rule-based [1] systems to powerful neural network-based models, marking a revolution in the machine's process and ability to condense textual information. At that time summarization relied heavily on rule-based approaches, which used manually crafted rules such as keyword matching, sentence position, and cue phrases. These methods were relatively convenient to implement but lack of flexibility and could not adapt well to different types of text.

In the literature on text summarization, the Term Frequency-Inverse Documet Frequency (TF-IDF) was introduced by Kietikul Jearanaitanakij et al. [2], these methods ranked sentences based on features like term frequency inverse document frequency. It was assigned importance to words that appeared frequently in a document but less frequently across a corpus. Also, we can call it automatic extractive summarization, which was found in the work of Yuuki Iwasaki et al, [3], they presented how text is generated. Extractive summarization is selecting key sentences or phrases from the original text without regenerate new words. In contrast, abstractive summarization is to generate entirely new sentences that paraphrase the original article.

In addition, impactful discovery came with Jacob Devlin et al [4] by introduced through a new language representation model called BERT (Bidirectional Encoder Representations from Transformers). Unlike earlier models, transformers can process entire texts in the article and better capture long-range dependencies between words. They discovered that by simply adding a single output layer, the pre-trained BERT model could be fine-tuned to achieve state-of-the-art performance on various tasks, including question answering and language inference without requiring major architectural changes. Similar to abstractive summarization, proposed by Betül Ay and Fatih Ertam and Galip Aydin [5] is T5 (Text-To-Text Transfer Transformer). T5 reframes all NLP tasks into a unified text-to-text format which means both inputs (articles) and outputs (summary) are treated as text strings, making the model flexible and easily adaptable to many tasks. Experimental results showed that T5 achieved strong performance metrics such as ROUGH and BERT score even in low-resource languages. This supports the model's effectiveness in abstractive summarization tasks in the cite of the paper BART model (Bidirectional and Auto-Regressive

[1*] Corresponding author: KOUM Soknan
E-mail: koumsoknan096@gmail.com*; Tel: +855-64 81 702*

Transformers) was found by M. Lewis et al. [6] introduced that this model is a sequence-to-sequence model introduced by Facebook AI (now Meta AI). The achievement was that the model can demonstrate the strong performance in extreme summarization scenarios, where the generated summary is much shorter than the original text. Building on this, Facebook AI later introduced mBART, a multilingual extension of BART, while Google researchers developed mT5, a multilingual version of T5 both aiming to improve summarization performance across multiple languages.

In today's digital age, the amount of information being created and shared is growing rapidly. From traditional news articles to modern social media posts, there is a vast amount of Khmer text available online. However, reading through all these articles can be time-consuming, especially when people are only looking for the important information. This is where Khmer text summarization becomes highly needed.

After seeing this requirement, our goal is to implement an automated Khmer text summarization system that can convert long and complex articles into shorter and simpler summaries. The system is intended to support a wide range of users, including students, professionals, and the public, by helping them grasp the main points of a text without having to read it all.

This study is necessary as it addresses the underrepresented area of Khmer text in natural language processing. Khmer is difficult to study and write as Khmer presents unique challenges, such as non-segmented script, complex word formation, and limited annotated data, which make text summarization especially difficult.

Our current scope focuses on Khmer sports articles, allowing us to develop effective summarization models within a domain with a clear structure and high public interest. Additionally, this study also provides practical benefits for Cambodians, especially students, journalists, and the government, by making information more accessible and faster to consume. Through a careful review of existing summarization research in languages with similar features like writing, structure and character to Khmer, such as Korean, Thai, Japanese, and Hindi. From this review, we identified the most effective techniques and models, including transformer-based models like mBART, and mT5, as well as sequence-to-sequence models and pointer-generator networks. After analyzing the strengths and limitations then we selected and fine-tuned the most suitable models for the Khmer language, aiming to overcome the challenges and produce accurate, readable summaries. In future work, we plan to expand the project to include a wider range of topics, such as education, general news, and other online content.

2. METHODOLOGY

2.1. Research Design

The study aims to develop a Natural Language Processing (NLP) tool using deep learning techniques, with a specific goal of fine-tuning a pre-trained multilingual transformer model for the specific task of abstractive text summarization in the Khmer language.

2.2. Data Collection

Regarding the dataset, we collected a total of 2,500 Khmer news articles from two popular online news platforms: Sabay News [7] and Khmerload [8]. Web scraping techniques, implemented in Python, were used to extract articles. The collected data is stored in CSV format (.csv) and contains six features: title, author, date, content, category, and URL.

Datasets (sources):

- Sabay News: 1,000 articles were collected.
- Khmerload: 1,500 articles were collected.

2.3. Data Cleaning

Regarding the datasets collected from two different popular online news websites mentioned above, some articles were not sufficiently clean or ready for use in reference summarization. Although both datasets contain six features in total but only the content and the reference summary are required for training process. The data cleaning process consists of two main steps. First, we ensure that there is no redundant data in either dataset. If duplicates are found, they are removed. After eliminating duplicated, the next step is to remove unnecessary elements as shown in *Error! Reference source not found.* to prepare the data for reference summarization.

ការប្រកួតកីឡាជាតិ លើកទី៤ ដែលជាព្រឹត្តិការណ៍ធំប្រចាំកម្ពុជា នឹងធ្វើពិធីបើកការប្រកួតជាផ្លូវការហើយនៅល្ងាចថ្ងៃទី៣០ ខែតុលា ឆ្នាំ២០២៤ នេះនៅក្នុងទីលានបាល់ទាត់ នៃពហុកីឡដ្ឋានជាតិ ក្រោមអធិបតីភាព សម្តេចមហាបវរធិបតី ហ៊ុន ម៉ាណែត នាយករដ្ឋមន្ត្រីកម្ពុជា។ក្រោមពីភាគីពាក់ព័ន្ធមានក្រសួងអប់រំ យុវជន និងកីឡា គណៈកម្មាធិការជាតិអូឡាំពិកកម្ពុជា និងសហព័ន្ធកីឡាជាតិ រួមទាំងរាជធានី-ខេត្តទូទាំងប្រទេសបានចំណាយពេលត្រៀមខ្លួនជាច្រើនខែមកនេះឥលូវឈានដល់ថ្ងៃបើកការប្រកួតជាផ្លូវការហើយ។

var gammatag = gammatag || {}; gammatag.cmd = gammatag.cmd || [];
gammatag.cmd.push(function() {
gammatag.defineZone({code:'gax-inpage-async-1700714337',size:[640,1386],params:{siteId:'1700712953',zoneId:'1700714337',zoneType:'Inpage'}});gammatag.sendRequest(); }); ពិធីបើកនៅល្ងាចនេះគឺនឹងមានភាពក្រឹកកក្រេងជាខ្លាំង ព្រោះមានជាការសម្តែងពិសេសៗជាច្រើនផងដែរដូចជាការសម្តែងក្បាច់គុនបុរាណពីផ្នែកនាយទាហាន ការសម្តែងរបាំសិល្បៈបុរាណនានា និងការសម្តែងផ្នែកសិល្បៈសម័យពិសេសវត្តមានរបស់តារាវ៉ៃប វណ្ណដាផងដែរ។ ងាកមកចាប់អារម្មណ៍ចំពោះប្រភេទកីឡាមួយចំនួនក៏បានចាប់ផ្តើមប្រកួតជាបណ្តើរៗផងដែរមុនឈានដល់ការធ្វើពិធីបើកជាផ្លូវការនេះ។ ការប្រកួតកីឡាជាតិឆ្នាំ២០២៤ នេះមាន ៣៦ប្រភេទកីឡាផ្សេងគ្នា និងមានការចូលរួមពី ២៥ រាជធានី-ខេត្ត។

Figure 1: A sample of raw Khmer new article data containing unwanted HTML and advertisement code before cleaning

2.4. Data Preparation

លោក គង់ ចាន់ធីណា ថាខាងម្ចាស់ឧបត្ថម្ភគឺ ISI GROUP បង្ហាញភាពរីករាយដែលបានជួយសហការឧបត្ថម្ភក្នុងការប្រកួតជើងឯកថ្នាក់ជាតិ ឆ្នាំ ២០២៤ កីឡាហាប់គីដូ កន្លងទៅនេះ។ថ្លែងក្នុងឱកាសប្រគល់លិខិតថ្លែងអំណរគុណដល់ក្រុមហ៊ុន ISI GROUP លោក គង់ ចាន់ធីណា អគ្គលេខាធិការសហព័ន្ធខ្មែរកីឡាហាប់គីដូ បានឲ្យដឹងថា ដើម្បីជាការថ្លែងអំណរគុណដល់ក្រុមហ៊ុនដែលបានជួយជាថវិកាមួយចំនួន ដល់ការប្រកួតថ្នាក់ជាតិកន្លងទៅថ្មីៗនេះសហព័ន្ធបានធ្វើលិខិតនិងយកទៅប្រគល់ដល់ក្រុមហ៊ុនផ្ទាល់តែម្តង។ អគ្គលេខាធិការសហព័ន្ធខាងលើនេះបន្ថែមទៀតថា ខាងភាគីតំណាង ISI គឺរីករាយក្នុងការជួយជាថវិកាសម្រួលដល់ដំណើរការនៃការប្រកួតជើងឯកថ្នាក់ជាតិនេះហើយប្រសិនបើមានជាព្រឹត្តិការណ៍ ឬការប្រកួតផ្សេងៗទៀតក្រុមហ៊ុននឹងពិនិត្យសហការតាមជាក់ស្តែងបន្តទៀត។លោក គង់ ចាន់ធីណា បន្តថាការជួយពីបណ្តាក្រុមហ៊ុនឯកជនគឺពិតជាបានសម្រាលបន្ទុកមួយធំរបស់សហព័ន្ធ ព្រោះដូចដឹកស្រាប់ហើយថាសហព័ន្ធខ្មែរកីឡាហាប់គីដូមិនទាន់ទទួលបានកញ្ចប់ថវិកាលើការអភិវឌ្ឍន៍បណ្តុះបណ្តាល ឬរៀបចំការប្រកួតផ្សេងៗនោះទេ។ រាល់ការរៀបចំម្តងៗគឺត្រូវការការប្រមូលថវិកាពីបណ្តាសមាជិកគណៈកម្មាធិការប្រតិបត្តិ ឬថ្នាក់ដឹកនាំសហព័ន្ធដើម។បញ្ជាក់ដែរថា កាលពី ថ្ងៃទី៣០ ខែធ្នូ ឆ្នាំ២០២៤ សហព័ន្ធខ្មែរកីឡាហាប់គីដូ ដឹកនាំដោយលោក គង់ ចាន់ធីណា អគ្គលេខាធិការ រួមជាមួយគ្រូបង្វឹកកីឡាករ កីឡាការិនី បានយកលិខិតថ្លែងអំណរគុណដល់ក្រុហ៊ុន ISI GROUP នៅទីស្នាក់ការក្រុមហ៊ុនង K Mall រ៉ែងស្រែង រាជធានីភ្នំពេញ។ ជាមួយគ្នានោះដែរខាង ISI GROUP មានតំណាងដោយលោក ឡាំ គុយ គួ នាទី អ្នកសម្របសម្រួលគម្រោងនៅក្នុង អង្គការមូលនិធិ ខេ អឹម អេច៕

Figure 2: Original document

លោក គង់ ចាន់ធីណា អគ្គលេខាធិការសហព័ន្ធខ្មែរកីឡាហាប់គីដូ បានថ្លែងអំណរគុណដល់ក្រុមហ៊ុន ISI GROUP ដែលបានឧបត្ថម្ភការប្រកួតជើងឯកថ្នាក់ជាតិឆ្នាំ 2024។ ក្នុងឱកាសប្រគល់លិខិតថ្លែងអំណរគុណ លោកបានលើកឡើងថា ការជួយនេះមានសារៈសំខាន់ក្នុងការចូលរួមសម្រួលដំណើរការប្រកួត។ ISI GROUP បានបង្ហាញភាពរីករាយ និងប្តេជ្ញាចូលរួមសហការទៀតប្រសិនបើមានព្រឹត្តិការណ៍ផ្សេងៗនៅអនាគត។ លោក គង់ ចាន់ធីណា បន្ថែមថា ការជួយពីក្រុមហ៊ុនឯកជនបានសម្រាលបន្ទុកធំបំផុតសម្រាប់សហព័ន្ធ ខណៈដែលសហព័ន្ធមិនទាន់ទទួលបានថវិកាសម្រាប់ការអភិវឌ្ឍន៍ ឬរៀបចំការប្រកួត។ ការរៀបចំត្រូវការថវិកាប្រមូលពីសមាជិកគណៈកម្មាធិការ។ កាលពីថ្ងៃទី 30 ខែធ្នូ ឆ្នាំ 2024 ក្រុមហ៊ុន ISI GROUP បានទទួលលិខិតថ្លែងអំណរគុណនៅទីស្នាក់ការក្រុមហ៊ុន K Mall រ៉ែងស្រែង រាជធានីភ្នំពេញ ។

Figure 2: Reference summary

Regarding both datasets, Sabay News contains 1,000 articles, each with reference summaries manually written by our team. Khmer Load, on the other hand, includes 1,500 articles with reference summaries, as illustrated in ***Figure 3***, which correspond to the original documents shown in ***Figure 2***. Combined, the two datasets consist of 2,500 articles, all with reference summaries. These summaries were segmented into words and sentences using KhmerWordSegmentor [9], a tool specifically designed for the Khmer language. The documents were then divided into three subsets 70% for training, 20% for validation, and 10% for testing.

Table 2.1. The details of dataset after splitting

Dataset	documents	documents		
		Train	Validation	Test
Sabay & Khmerlaod	2,500	1,750	500	250

Table 2.2. The details summarization dataset

Dataset	Total	Avg_content length		Avg_summary length	
		word	sentence	word	sentence
Sabay & Khmerlaod	2,500	1,131.97	6.12	482.92	3.71

2.5. Model Selection

The proposed Khmer news summarization process is illustrated in ***Figure 4***. First, the input text dataset undergoes word segmentation. Since Khmer script does not use spaces to separate words, this segmentation is a crucial step to work on. This process employs the KhmerWordSegmentor tool to divide the text into meaningful units, such as words or phrases.

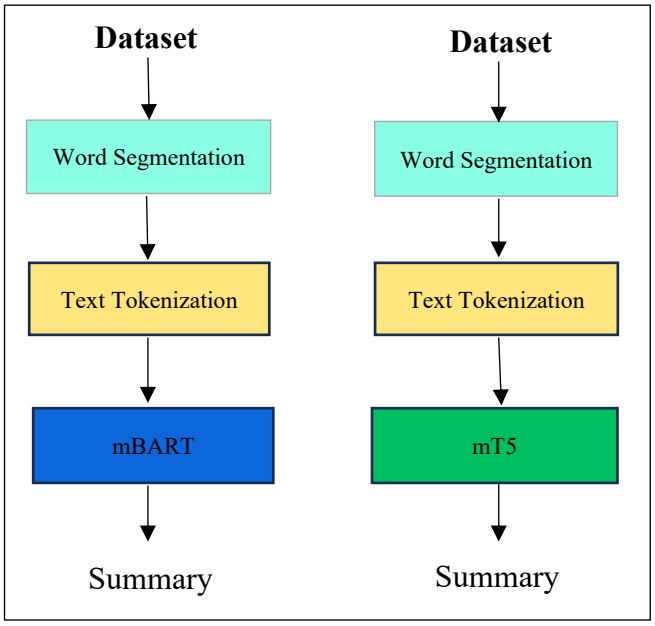


Figure 3: Architecture of Khmer News Summarization Model

Once segmented, the text is tokenized using the pre-trained mBART and mT5 tokenizer, which converts the text into numerical token IDs that the model can process. To maintain a balance between context retention and computational efficiency, we impose length constraints: the input text is limited to 1024 tokens, while summaries are capped at 350 tokens.

Next, the segmented and tokenized data is mapped into a structured format, generating numerical representations for both the content and the summary. These representations are then ready for use in the summarization model. This comprehensive preprocessing step is essential for the success of Khmer text summarization.

Finally, we utilize a pre-trained model called mBART (facebook/bart-base), a multilingual sequence-to-sequence transformer developed by Facebook AI. mBART (multilingual BART) is based on the BART model, originally designed for text translation and generation tasks. It supports 25 languages, including both high-resource and low-resource languages, making it particularly suitable for Khmer-language applications.

2.6. Experiment setup

In this research, we conducted all experiments using Google Colab, which provides a high-performance environment for training deep learning models. Our setup included an NVIDIA A100 GPU (40GB memory), 83.5GB of RAM, and a high-performance CPU, enabling efficient training of large-scale models. For summarization, we compared two state-of-the-art models: mBART and mT5. Both models were fine-tuned on our custom training and validation datasets to ensure a fair evaluation.

Additionally, we tested a proposed preprocessing method, where the input text was first segmented and tokenized before being fed into either mBART or mT5 for summarization. This approach helps the models focus on semantically meaningful units that improve output quality.

Before full training, we performed hyperparameter tuning to optimize key settings, such as the learning rate and batch size, for each model. This step is critical for maximizing performance.

To mitigate overfitting, when fine-tuning large pretrained models on small datasets, we closely monitored the training loss and validation loss. Training was halted if the two metrics diverged significantly, indicating potential overfitting. In our experiments, both models achieved optimal performance after 5 epochs (i.e., five full passes through the training data).

Due to the limited size of our dataset, fine-tuning large pretrained models carried a significant risk of overfitting, where models perform well on test data but generalize poorly to unseen data. To address this, we implemented two key safeguards: First, we employed early stopping by continuously monitoring the divergence between training loss and validation loss, then halting training when this gap began to widen significantly. Second, through empirical evaluation, we determined the optimal training duration to be 5 epochs (five complete iterations through the training dataset), as this provided the best balance between model performance and generalization capability.

Table 2.3. Fine -Tuning Hyperparameter

Hyperparameter	Model	
	mBart	**mT5**
learning_rate	1.00E-05	1.00E-05
weight_decay	0.061	0.044
label_smooth_factor	0.051	0.053
dropout	0.3	0.3

2.7. Evaluation Method

We employed the Recall-Oriented Understudy for Gisting Evaluation (ROUGE) [10] metric which is the standard evaluation measure for text summarization tasks, to assess system performance. The ROUGE score compares system-generated summaries against reference summaries through three distinct approaches: unigram matching (ROUGE-1),

bigram matching (ROUGE-2), and longest common sub-sequence matching (ROUGE-L). These metrics evaluate lexical overlaps at different granularities, providing comprehensive insight into the quality of the generated summaries.

3. RESULTS AND DISCUSSION

After the training process is completed, we notice that mbart50 seems to perform better than mt5 model. The detailed results are shown in **Table 3.1.**

Table 3.1. Results

Model	Rouge-1 (%)	Rouge-2 (%)	Rouge-L (%)
mBART	68.54%	52.91%	64.25%
mT5	51.27%	38.57%	47.98%

Based on the result in table mBART outperforms mt5 across all three ROUGE metrics. The mBART model achieved a ROUGE-1 score of 0.685, which is significantly higher than the 0.513 obtained by mT5. Similar improvements are seen in ROUGE-2 and ROUGE-L, where mBART scored 0.529 and 0.643, respectively, compared to 0.386 and 0.480 by mT5. These results suggest that mBART is better suited for Khmer summarization tasks in this experiment.

4. CONCLUSIONS

This paper successfully developed a Khmer text summarization model using NLP and deep learning techniques for Khmer news articles. The process involved critical steps such as word segmentation with the KhmerWordSegmentor, tokenization, and fine-tuning models based on mBART and mT5. Despite Khmer's linguistic challenges including the lack of word spacing and limited NLP resources the model generated high-quality summaries.

Experimental results demonstrated the model's high accuracy, confirming its effectiveness. The system can efficiently condense lengthy Khmer texts while maintaining correctness. These outcomes highlight the model's practical utility and potential for further enhancement in future work.

ACKNOWLEDGMENTS

We would like to sincerely thank our supervisor, **Mr. NGIN Kimlong**, for his valuable guidance, encouragement, and support throughout this research. His advice helped us overcome many challenges and stay focused.

We are also truly grateful to our team members for their dedication, teamwork, and the time they committed to this project. Without their strong effort and cooperation, this work would not have been possible.

References

[1] P. P. Ghosh, R. Shahariar, and M. A. H. Khan, "A rule based extractive text summarization technique for Bangla news documents," *International Journal of Modern Education and Computer Science,* vol. 11, no. 12, pp. 44-53, 2018.

[2] K. Jearanaitanakij, S. Boonpong, K. Teainnagrm, T. Thonglor, T. Kullawan and C. Yongpiyakul, "Fast hybrid approach for Thai news summarization," *Engineering and Technology Horizons,* vol. 41, no. 4, p. 4, 2024.

[3] Y. Iwasaki, A. Yamashita, Y. Konno and K. Matsubayashi, "Japanese abstractive text summarization using BERT," *ASTES Journal,* vol. 5, no. 6, pp. 1674-1682, 2020.

[4] J. Devlin, . M. W. Chang, K. Lee and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies*, 2019, June.

[5] B. Ay, F. Ertam, G. Fidan and G. Aydin, "Turkish abstractive text document summarization using text-to-text transfer transformer," *Alexandria Engineering Journal,* vol. 68, pp. 1-13, 2023.

[6] M. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov and L. Zettlemoyer, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension," in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, Jul. 2020.

[7] "Sabay News," Latest News in Cambodia, [Online]. Available: https://news.sabay.com.kh. [Accessed 29 Apr 2025].

[8] "Khmerload," News and Entertainment in Cambodia, [Online]. Available: https://www.khmerload.com. [Accessed 29 Apr 2025].

[9] R. Buoy, N. Taing and S. Kor, "Khmer Word Segmentation Using BiLSTM Networks," in *4th Regional Conference on OCR and NLP for ASEAN Languages (ONA 2020)*, Phnom Penh, 2020.

[10] C.-Y. Lin, "ROUGE: A Package for Automatic Evaluation of Summaries," in *Text Summarization Branches Out (WAS 2004)*, Jul, 2004.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Handwritten Khmer Digit Recognition Using Sequential RNN Models

Kimlong Ngin [1], Dona Valy [2], Kimhor Phoeurn [3*], Seyha Va [3], Bunkhloem Pel [3],

Kana Ty [3], Sambat Yin [3], Rotha Pho [3]

[1] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Mechatronics and Information Technology Research Unit, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: e20210823@dtc1.itc.edu.kh*

Abstract: *Handwritten Digit Recognition is one of the crucial tasks for optical character recognition (OCR) with applications of historical document preservation, modern data entry, and accessibility solutions. Unlike capitalized Latin scripts, the complexity of Khmer digits lies in its complex shapes and high variability of handwriting, which poses a challenge in developing robust recognition systems. While most research in handwritten digit recognition relies on image-based data, this paper explores a novel approach using sequential coordinate data for the complex task of recognizing Khmer digits applied on Long Short-Term Memory (LSTM), Bidirectional LSTM (Bi-LSTM), and Gated Recurrent Unit (GRU), incorporating data augmentation techniques to improve performance. The dataset consists of 1583 sequences of handwritten Khmer digit drawings that are collected from a custom-built system. The dataset is split into 60% training set, 20% validation set, and a 20% test set. After applying data augmentation on train set, the dataset expanded to 12,337 sequences. With test set accuracies of 95.27%, 95.58%, and 95.58% for LSTM, Bi-LSTM, and GRU, respectively.*

Keywords: *Khmer Handwriting; Khmer Digit Recognition; RNN Sequential data; Points (x, y)*

1. INTRODUCTION

Handwriting Khmer digit recognition is a fundamental challenge in the field of machine learning and computer vision, with significant applications across education, business, and government sectors in tasks such as automated data entry, identity verification, invoice digitization, financial processing, and more. Consequently, there is an urgent need for a handwritten recognition system to provide efficient and precise recognition into digital formats to facilitate organized storage, effective management, and effortless retrieval.

Over the years, numerous approaches have been proposed for handwritten digit recognition ranging from traditional machine learning methods such as Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) [1] , to deep learning approach such as leveraging Convolutional Neural Networks (CNN) [2] and Recurrent Neural Network (RNN) [3] . Specifically, these deep learning approaches have demonstrated extraordinary performance in handwritten recognition systems.

This project addresses the challenge of recognizing handwritten Khmer digits by shifting from traditional image-based inputs to a sequential data approach that models how digits are drawn. By capturing the temporal dynamics of pen stroke sequences, we aim to develop a recognition system that better reflects the unique characteristics of Khmer handwriting.

2. RELATED WORKS

Khmer Handwritten has attracted growing interest due to its role in digitizing Khmer script. While several studies have investigated Khmer text recognition, specific research handwritten Khmer digit recognition is limited. This section

3* Corresponding author: Kimhor Phoeurn
E-mail: e20210823@dtc1.itc.edu.kh; Tel: +855-61290673

reviews key contributions in Khmer handwritten text recognition and highlights the gap our work aims to fill.

Annanurov et al. [4] introduced a lightweight deep learning architecture, 2+1CNN, for offline Khmer handwritten text recognition. The model comprises two convolutional layers and a fully connected layer. It achieved an average recognition accuracy of 94.9%. Designed for resource-constrained environments, their system adopted a one-against-all classification approach by training 33 distinct binary classifiers for Khmer consonants. The results demonstrated that compact CNNs can effectively process complex scripts like Khmer.

Another relevant study by Annanurov et al. [5] applied Convolutional Neural Networks (CNNs) for recognizing handwritten Khmer syllables. Handwritten samples were preprocessed through grayscale conversion, size normalization to 32×32 pixels, and binary thresholding. A similar one-against-all training strategy was used with an average accuracy of 94.85%.

In contrast to deep learning, Meng et al. [6] proposed a hybrid neural architecture combining Self-Organizing Maps (SOM) and Multilayer Perceptrons (MLP) for Khmer character recognition. Input images (20×20 pixels) were binarized and vectorized into 400-dimensional vectors. SOM was used to cluster inputs, followed by MLP-based classification for each cluster. Although limited by a small dataset of 135 samples derived from printed fonts, this system showcased an effective hierarchical classification pipeline. The authors found that a training loss of 10^{-29} significantly improved performance, increasing the recognition rate from 72% to 94.13% on a 10% noise condition of test set. However, the model showed a drop in accuracy for 65% noised train set and 30% accuracy for noised unseen data.

3. METHODOLOGY

3.1 Data Collection

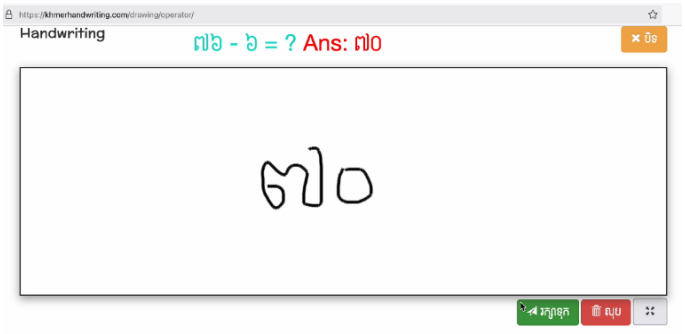


Figure 1: Web Platform for Data Collection

To construct a dataset of Khmer handwritten numbers, a web-based platform as shown in Figure 1 was developed to allow users to draw numbers using their mouse or touchscreen. Each entry includes a digit label, and a stroke trajectory saved in a comma-separated format: $x_1, y_1, x_2, y_2, \ldots$. Multiple strokes are separated by a delimiter (#).

A total of 1,583 samples were collected from Cambodian university students with ages from 18 to 22. Both male and female participants contributed which reflect the handwriting styles typical among young adults in Cambodia.

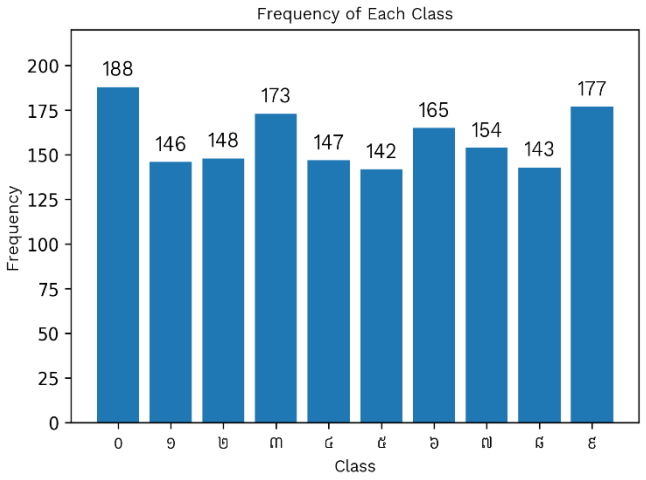


Figure 2: Number of Samples per Digit Class

Figure 2 illustrates the class distribution ranged from 142 to 188 samples per digit, with digit 0 being most frequent and digit 5 the least. While there is some class imbalance, it is relatively small and does not heavily skew the dataset.

3.2 Data Preprocessing

The raw data underwent a multi-stage preprocessing pipeline to prepare it for the models. The pipeline began with cleaning the dataset by removing any entries with missing values. Each digit was then segmented into individual strokes using the '#' delimiter. To ensure uniformity across different drawing scales and positions, all stroke coordinates were normalized to a [0,1] range using the min-max scaling formulas shown in (Eq. 1) and (Eq. 2).

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}} \quad \text{(Eq. 1)}$$

$$y_{scaled} = \frac{y - y_{min}}{y_{max} - y_{min}} \quad \text{(Eq. 2)}$$

Where x and y are the original coordinates, while x_{min} , x_{max}, y_{min}, and y_{max} represent the minimum and maximum coordinate values for that stroke.

Subsequently, each main stroke was divided into uniform substrokes to create fixed-size input sequences for the models. Each substroke was set to contain a fixed length of 8 coordinate pairs. The total number of substrokes for a given main stroke was determined by (Eq. 3).

$$\#\ \text{substrokes} = \frac{\text{total points of main stroke}}{\#\ \text{segments}} \quad \text{(Eq. 3)}$$

Where the number of segments per substroke is 8. If the final substroke contained fewer points than required, it was padded with $(0, 0)$ to ensure a uniform size across all inputs.

3.3 Data Preparation and Augmentation

This dataset was split into a 60% training set, 20% validation set, and a 20% test set. To further enhance dataset diversity and improve generalization, a rotation-based data augmentation technique was applied on the train set. The original stroke coordinates were rotated using angles from -30° to 30°, in 5° increments. This simulates natural variations in handwriting while preserving the character's structure. Mathematically, each point (x, y) in a stroke is rotated about the origin using a standard two-dimensional rotation matrix:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Resulting in:

$$x' = x\cos(\theta) - y\sin(\theta) \quad \text{(Eq. 4)}$$

$$y' = x\sin(\theta) + y\cos(\theta) \quad \text{(Eq. 5)}$$

where θ is the rotation angle in radians. Each stroke with (x, y) coordinates was transformed for every $\theta \in \{-30°, -25°, \dots, 30°\}$. As a result, each original sample was augmented with multiple rotated versions which produce a more robust and representative training set.

3.4 Model Architecture

To model the temporal dynamics inherent in handwritten stroke data, three Recurrent Neural Network (RNN) architectures were evaluated. RNNs are specifically designed for sequential data, making them ideal for interpreting the patterns of how a digit is drawn over time.

3.4.1 Long Short-Term Memory (LSTM)

The primary architecture explored was the Long Short-Term Memory (LSTM) network [7], a variant of RNN. LSTMs are explicitly designed to overcome the vanishing and exploding gradient problems that can affect simpler RNNs, allowing them to effectively learn long-range dependencies within the pen stroke sequences.

An LSTM cell operates as follows:

$$f_t = \sigma(W_f \cdot [h_{t-1}x_t] + b_f) \quad \text{(Eq. 6)}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}x_t] + b_i) \quad \text{(Eq. 7)}$$

$$\tilde{c_t} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad \text{(Eq. 8)}$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad \text{(Eq. 9)}$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad \text{(Eq. 10)}$$

$$h_t = o_t \odot \tanh(c_t) \quad \text{(Eq. 11)}$$

Here, σ is the sigmoid function, $\odot$ is the element-wise multiplication and x_t is the input at time t . The cell state c_t acts as a memory to carry long-term information.

3.4.2 Bidirectional-LSTM (Bi-LSTM)

To capture contextual information from the entire stroke, a Bidirectional LSTM (Bi-LSTM) [8] was also implemented. Unlike a standard LSTM which only processes the sequence in a forward direction, the Bi-LSTM utilizes two separate LSTM layers: one processing the sequence from start to finish, and the other from finish to start. It consists of two LSTMs: one reads the input from $t = 1$ to T, and the other from $t = T$ to 1. The outputs from both directions are concatenated:

$$\overrightarrow{h_t} = LSTM_{fwd}(x_t) \quad \text{(Eq. 12)}$$

$$\overleftarrow{h_t} = LSTM_{bwd}(x_t) \quad \text{(Eq. 13)}$$

$$h_t = [\overrightarrow{h_t}; \overleftarrow{h_t}] \quad \text{(Eq. 14)}$$

3.4.3 Gated Recurrent Unit (GRU)

The Gated Recurrent Unit (GRU) [9], a more modern and computationally efficient variant of the LSTM, was also evaluated. The GRU simplifies the LSTM architecture by combining the forget and input gates into a single update gate and merging the cell state and hidden state. This streamlined design often allows for faster training without a significant loss in performance, making it a strong candidate for this recognition task.

The GRU equations are:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_z) \quad \text{(Eq. 15)}$$

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \quad \text{(Eq. 16)}$$

$$\widetilde{h_t} = \tanh(W_h \cdot [r_t \odot h_{t-1}, x_t] + b_r) \quad \text{(Eq. 17)}$$

$$h_t = (1 - z_t) \odot h_{t-1} + z_t \odot \widetilde{h_t} \quad \text{(Eq. 18)}$$

3.5 Model Training

A set of regularization and optimization techniques was utilized to increase the robustness and generalization capabilities of the models. A dropout rate of 0.3 was applied between the LSTM/GRU layers and fully connected layers to avoid overfitting. The Adam optimizer was used for optimization. Additionally, CrossEntropyLoss function was utilized to handle classification tasks. In addition to these steps, hyperparameters tuning was also done using grid search to find the best hyperparameters. It includes going through various learning rates from 0.01, 0.001, 0.0001 with the number of epochs of 30, 40, and 50 to ensure the final model obtains the best performance.

3.6 Model Evaluation

To evaluate the effectiveness of different recurrent neural network architectures in classifying the stroke point of written Khmer digits, two key evaluation metrics were assessed: Accuracy and F1 score. Since the F1 score is derived from Precision and Recall, these two metrics are also defined below to provide clarity.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad \text{(Eq. 19)}$$

$$Precision = \frac{TP}{TP + FP} \quad \text{(Eq. 20)}$$

$$Recall = \frac{TP}{TP + FN} \quad \text{(Eq. 21)}$$

$$F1\ Score = 2\ \left(\frac{Precision \times Recall}{Precision + Recall}\right) \quad \text{(Eq. 22)}$$

where TP: True Positive, FP: False Positive, FN: False Negative

4. RESULTS

As shown in Table 1, all models demonstrate strong performance, with accuracies exceeding 95% which indicates their effectiveness in capturing sequential patterns in the data.

Table 1: Model Performance Comparison of Test Set

Model	Train Accuracy	Test Accuracy	F1 Score
LSTM	0.9938	0.9527	0.9530
Bi-LSTM	0.9991	0.9558	0.9550
GRU	0.9973	0.9558	0.9561

The LSTM model achieved an accuracy of 95.27% on test set with the F1 score of 95.30%, showing it is a reliable baseline for sequential classification tasks.

The Bi-LSTM model showed a slight improvement over LSTM. This bidirectional context helps capture the full shape and direction of handwritten strokes, leading to an accuracy of 95.58% and the F1 score of 95.50%.

The GRU model achieved the best overall performance, with the same accuracy as Bi-LSTM of 95.58% on test set but surpassing it in F1 score (95.61%). The superior performance

of the GRU in F1 score, despite having the same accuracy as the Bi-LSTM, may be attributed to its simpler architecture.

5. CONCLUSION

In this study, a solution for Khmer handwritten digit recognition system is proposed using LSTM, Bi-LSTM, and GRU. This study aimed to enhance the understanding of the structure of RNN-based models as well as the processing details of the dataset. The results illustrate an exceptional performance of the GRU model compared to the other two with an accuracy of 95.58%.

ACKNOWLEDGEMENTS

The authors would like to express their sincere gratitude to the ViLa Lab for their generous support, insightful feedback, and access to essential resources. Their contributions have played a vital role in enhancing the depth and quality of this research.

REFERENCES

[1] C. Kaensar, "A Comparative Study on Handwriting Digit Recognition Classifier Using Neural Network, Support Vector Machine and K-Nearest Neighbor," in *The 9th International Conference on Computing and InformationTechnology (IC2IT2013)*, P. Meesad, H. Unger, and S. Boonkrong, Eds., Berlin, Heidelberg: Springer, 2013, pp. 155–163. doi: 10.1007/978-3-642-37371-8_19.

[2] S. Ahlawat, A. Choudhary, A. Nayyar, S. Singh, and B. Yoon, "Improved Handwritten Digit Recognition Using Convolutional Neural Networks (CNN)," *Sensors*, vol. 20, no. 12, Art. no. 12, Jan. 2020, doi: 10.3390/s20123344.

[3] I. J. L. Paul, S. Sasirekha, D. R. Vishnu, and K. Surya, "Recognition of handwritten text using long short term memory (LSTM) recurrent neural network (RNN)," *AIP Conf. Proc.*, vol. 2095, no. 1, p. 030011, Apr. 2019, doi: 10.1063/1.5097522.

[4] B. Annanurov and N. M. Noor, "A compact deep learning model for Khmer handwritten text recognition," *IAES Int. J. Artif. Intell. IJ-AI*, vol. 10, no. 3, Art. no. 3, Sep. 2021, doi: 10.11591/ijai.v10.i3.pp584-591.

[5] B. Annanurov and N. M. Noor, "Handwritten Khmer text recognition," in *2016 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*, Pune, India: IEEE, Dec. 2016, pp. 176–179. doi: 10.1109/wiecon-ece.2016.8009112.

[6] H. Meng and D. Morariu, "Khmer character recognition using artificial neural network," in *Signal and Information Processing Association Annual Summit and Conference (APSIPA), 2014 Asia-Pacific*, Dec. 2014, pp. 1–8. doi: 10.1109/APSIPA.2014.7041824.

[7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

[8] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, no. 11, pp. 2673–2681, Nov. 1997, doi: 10.1109/78.650093.

[9] K. Cho *et al.*, "Learning Phrase Representations using RNN Encoder–Decoder for Statistical Machine Translation," in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, A. Moschitti, B. Pang, and W. Daelemans, Eds., Doha, Qatar: Association for Computational Linguistics, Oct. 2014, pp. 1724–1734. doi: 10.3115/v1/D14-1179.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Analysis on Machine Learning Models for Imbalanced Data Problem in Payment Fraud Detection

Siuphing SEUN [1*], Sokkhey Phauk [1], Kimlong NGIN [1], Pakrigna LONG [1]

[1*] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., Phnom Penh, Cambodia*

[1] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., Phnom Penh, Cambodia*

[*] *Corresponding author: siuphing15@gmail.com*

Abstract: *Payment card fraud losses worldwide reached $33.83 billion in 2023 (Nilson Report, 2025). Alarmingly, Deputy Prime Minister and Minister of Interior Sar Sokha (2024) stated that in the first semester of 2024, Cambodians lost nearly $40 million on online fraud and digital fraud. To prevent these significant financial losses, it's crucial to find the predictive models that can better predict the anomalies in transaction. This study investigates what predicting models works best in predicting the anomalies in the payment transaction. The dataset contains types of online transaction, amount of the transactions, names of the sender and receiver, senders' balance of the account before and after the transaction, and receivers' account balance before and after receiving money. Autoencoder, LightGBM, Neural Network, Logistic Regression, Random Forest, and CatBoost were built as prediction models within this research. Each of these algorithms is employed to create prediction models, which are meticulously fine-tuned to yield the most accurate prediction of the fraud payment in the system involving optimizing hyperparameters and selecting the best features to enhance the models' prediction power. Common performance metrics such as Precision, Recall, F1-Score, and AUC-ROC were used to tested each model performance. Extensive experimentation data shows the best performance of CatBoost and Random Forest model, which consistently outperforms other machine learning methods in accurately predicting the payment fraud either training with imbalanced or balanced dataset. These results highlight the model's ability to handle complex, non-linear relationships within the data and its effectiveness in generalizing across different scenarios and conditions. In conclusion, the findings of this study will be beneficial to mostly the bank system as they can apply the models we found into their system to prevent any fraud activities. Spotting the fraud activities in the early stage plays an immense role in preventing the loss.*

Keywords: Payment Card Fraud; Machine Learning Models; Fraud Prediction; Predictive Modeling; Class Imbalance

1. INTRODUCTION

According to the World Payments Report 2025 by Capgemini Research Institute, global non-cash transaction volumes reached 1.411 trillion in 2023 and are projected to rise to 1.65 trillion in 2024. With this frequent transaction will attract malicious actors seeking to exploit vulnerabilities in the payment system.

The term "Fraud" was given the meaning as the crime of getting money by deceiving people (Cambridge, 2023). Payment card fraud losses worldwide reached $33.83 billion in 2023 (Nilson Report, 2025). Alarmingly, Deputy Prime Minister of Interior Sar Sokha (2024) stated that in the first semester of 2024, Cambodians lost nearly $40 million on online fraud and digital fraud. With this enormous amount of money losses due to the fraud payment, it's needed to find the predictive models that can better predict the anomalies in transaction.

Our research applied various machine learning (ML) such as Autoencoder, LightGBM, Logistic Regression, Neural Network, Random Forest, and CatBoost to compare the prediction performance of the abnormalities in the transaction.

Bhattacharyya et al. (2011) compare Logistic Refression and Support Vector machines, and RandomForest to see which one yield the best outcome. Moreover, Oza (2018) used Logistic Regression, Linear Kernel, and Radial Basis Function Kernel to predict fraud payment. Likewise, Oğuz

[1*] Corresponding author: Siuphing SEUN
E-mail: siuphing15@gmail.com; Tel: +855-10 403 311

and Layth (2020) applied RF, NB, SVM, and KNN to study the comparison of the proposed model and found out that RF is the most effective model for classifying credit card fraud transaction as fraudulent or genuine. Similarly, Minjun (2023) proposed 3 ML known as SVM, LR, and Decision tree which got the result in Decision Tree as the best model compare to SVM and LR for its high accuracy. Last but not least, Feng and Kim (2024) selected five machine learning models to perform fraud payment prediction such as Random Forest with adaptive Boosting, Gradient Boosted Decision Tree, K-Nearest Neighbor, Convolutional Neural Network, and Support Vector Machine.

While previous studies have compared the performance of many proposed models in their studies as well as identified factors that lead to fraud payment, there is still a gap in effectively using different approach that can better predict the fraudulent activity with both balanced and imbalanced dataset. For instance, Neural Network that can be a candidate in predicting fraud for its ability to capture complex relationships in data, especially when there's a lot of nonlinear and high-dimensional information. Likewise, Catboost is the ML that shouldn't be left out in applying to train the model to detect for legitimate activity. Although it hasn't been applied by many, Catboost is designed for faster and more accurate results, especially when working with complex datasets. Our study addresses this by applying some ML that haven't been selected to analyze as well as using resampling techniques to balance the data. This research aims validate the effectiveness of ML in predicting fraud activities. The findings of this study will be beneficial to mostly the bank system as they can apply the models we found into their system to prevent any fraud activities. Spotting the fraud activities in the early stage plays an immense role in preventing the loss. Nevertheless, predicting fraudulent payments aligns closely with the goals of SDG9 and SDG16. By leveraging innovative technologies and robust infrastructure (SDG9), we can enhance payment systems to detect and prevent fraud, fostering economic growth and safeguarding resources. Furthermore, ensuring integrity in financial transactions supports peaceful and just societies by combating corruption and fostering trust in institutions (SDG16). Fraud detection not only protects individuals and businesses but also promotes global stability and fairness, making it a valuable tool for advancing the human race.

Table 1 Comparative Analysis of Model Performance

Reference	Title	Dataset	Algorithms	Result
Bhattacharyya et al. (2010)	Data mining for credit card fraud: A comparative study	Internationally operating financial institution	LR, SVM, RF	Random Forest and SVM perform better than Logistic Regression
Oza (2018)	Fraud Detection using Machine Learning	Paysim Dataset	LR, Linear SVM, and SVM with RBF Kernel	The research demonstrates that SVM based methods are more effective
Oğuz and Layth (2020)	Comparative Analysis of Different Distributions Dataset by Using Data Mining Techniques on Credit Card Fraud Detection	European credit card transaction	RF, NB, SVM, and KNN	The study found that RF is the most effective model for classifying credit card fraud transaction as fraudulent or genuine.
Minjun (2023)	Multiple Machine Learning Models on Credit Card Fraud Detection	Paysim Dataset	SVM, LR, and Decision Tree	Decision Tree is chosen to be the best model compare to SVM and LR for its high accuracy score.
Feng and Kim (2024)	Novel Machine Learning Based Credit Card Fraud Detection Systems	European credit card holders dataset	RF with Adaptive Boosting, GB with Decision Tree, KNN, CNN, and SVM	Random Forest with Adaptive Boosting perform the best compare to the other four models.

2. METHODOLOGY

2.1 Dataset

The analysis is based on a large dataset gathered from Kaggle called "Online Fraud Payment Detection" which is a synthesic datasets generated by the PaySim mobile money simulator. The dataset comprises information from 6362620 transactions with 5 different types of payment known as Cash In, Cash Out, Debit, Transfer, and Payment. Each transaction contains 10 features, including unit of time, transaction's type, transaction's amount, sender and receiver's account, balance before and after the transaction of the sender and receiver with the target variables (isFraud). A first examination found that the dataset is imbalanced, with much fewer nonfraud cases than fraud cases. This is a typical issue in fraud payment datasets. Still, it's not really a problem with our study as we want to compare the performance of the selected ML to see which one perform better with both Imbalanced and Balanced dataset used for the training.

To ensure the dataset's reliability and efficiency in training predictive models, thorough cleaning methods were applied. The procedures involved resolving missing values, normalizing continuous attributes, and encoding categorical variables. Each record was thoroughly analyzed for completeness and consistency to ensure that the dataset was reliable and appropriate for testing various predictive algorithms. Additionally, we conducted exploratory data analysis to identify trends and correlations in the dataset, which is essential for developing precise predictive models to improve the prediction of the fraud activity in the payment transaction. The study also addressed class imbalance by utilizing resampling techniques such as oversampling, undersampling, and SMOTE to balance the dataset and enhance model accuracy. While for the case training the selected ML with imbalance, we applied various technique such as sample weight, scale pos weight, etc.

Table 2 Data Description

Variable's Name	Description
step	A unit of time where 1 step equals to 1 hour
type	Type of online transaction
amount	The amount of the transaction
nameOrig	Customer starting the transaction
oldblanceOrg	Balance before the transaction
newbalanceOrig	Balance after the transaction
nameDest	Recipient of the transaction
oldbalanceDest	Initial balance of recipient before the transaction
newbalanceDest	The new balance of recipient after the transaction
isFraud	Indicates whether the transaction is fraud. The 0 value indicates non-fraud while 1 value indicates fraud

2.2 Data Preprocessing

Data Preprocessing is a crucial step in training the model. We follow the 4 steps in order to preprocess the data such as missing value handling, data cleaning, data standardization, and encoding technique.

We also apply Cramer's V to see the association between another variable and target variable (isFraud).

Figure 1 Result of Cramer's V

```
Cramer's V between step and isFraud: 0.6703191487872493
Cramer's V between type and isFraud: 0.05891236771716165
Cramer's V between amount and isFraud: 0.9494821840526401
Cramer's V between nameOrig and isFraud: 0.9991462266201633
Cramer's V between oldbalanceOrg and isFraud: 0.9883278180708011
Cramer's V between newbalanceOrig and isFraud: 0.14284390460836927
Cramer's V between nameDest and isFraud: 0.6706932218607597
Cramer's V between oldbalanceDest and isFraud: 0.5863999685144793
Cramer's V between newbalanceDest and isFraud: 0.6491068887955074
Cramer's V between isFraud and isFraud: 0.9999390422203324
```

Figure 1 shows the result of Cramer's V between every variable compare to isFraud variable so-called target variable to have a better understand of what variable are strongly related to the target variable and what variables are less related to the target variable. According to the result, we can see that the variable "type" is very less related to the target variable which can be drop from the dataset when training model.

2.3 Resampling Methods

To address the dataset's predicted class imbalance, we used three resampling strategies: oversampling, undersampling, and the Synthetic Minority Over-sampling Technique. Each strategy was chosen based on its capacity to improve class balance, hence improving predictive model performance and generalizability.

Oversampling: Using this method, we expanded the size of the minority class ('isFraud = 1') by randomly reproducing instances until the number of fraud cases equaled the number of non-fraud cases. Oversampling ensures that the model is not biased towards the majority class and can learn from the intricacies found in the minority class.

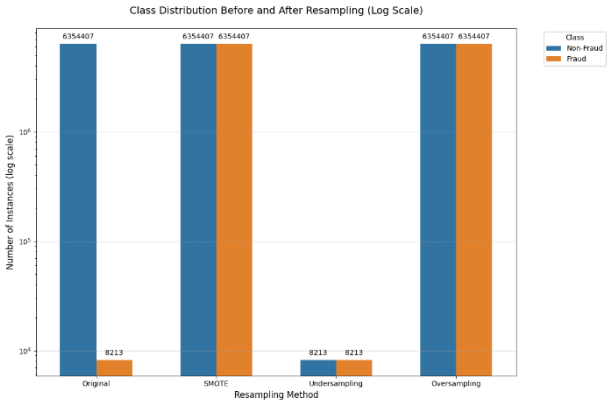
Undersampling: In contrast, undersampling includes removing instances at random from the majority class ('isFraud = 0') in order to equal the number of fraud instances. This technique is one of the several techniques used by data

scientist to extract more accurate information from originally imbalanced dataset. Nevertheless, the drawback is that important information could be lost.

Synthetic Minority Over-sampling Technique (SMOTE): SMOTE generates new synthetic data points for underrepresented classes through interpolation rather than duplicating. This more nuanced representation of the minority class aids models in learning a broader range of features associated with frauds, which can lead to improved prediction of fraud transaction. SMOTE has the advantage of creating more diverse synthetic samples, which can enhance the model's ability to generalize to new, unseen data.

By resolving the class imbalance using these resampling strategies, we saw considerable increases in the performance of our prediction models. Models trained on resampled datasets have higher precision, recall, F1-Score, AUC-ROC, and AUC-PR than those trained on the original imbalanced dataset. This suggests that the models performed better in recognizing both fraud and non-fraud cases, resulting in more trustworthy and generalizable prediction.

Figure 2 Class distribution after resampling technique



2.4 Proposed Models

Selecting the predictive modeling is a crucial phase in the research process as it affects the performance and interpretability of the results. In our study, we purposefully selected a variety of algorithms known as Autoencoder, LightGBM, Neural Network, Logistic Regression, Random Forest, and CatBoost.

2.5 Evaluation Metrics

The efficiency of the predictive modeling is measure by Precision, Recall, F1-Score, AUC-ROC, and AUC-PR in this study as the data that we used to train ML is both balanced and imbalanced data.

In imbalanced datasets, accuracy can be misleading because a model can achieve a high accuracy by simply predicting the majority class for most cases, ignoring the minority class entirely. To be precise, in the case of our dataset where fraud cases make up only less than 1% of the data, predicting "not fraud" for every instance would result in 99% accuracy, but the model would fail to identify any fraudulent transactions.

By focusing on AUC-ROC, F1-Score, Precision, and Recall, we will get a more accurate reflection of how well the model performs, especially in detecting the minority class, which is often the focus in applications like fraud detection.

Table 3 Confusion Matrix

		Actual	
		Positive	**Negative**
Predicted	**Positive**	True positive (TP)	False positive (FP)
	Negative	False negative (FN)	True negative (TN)

Precision measures the correctness of positive predictions by the model, defined as the number of true positives divided by the total number of positive predictions (true and false). The formula is:

$$Precision = \frac{TP}{TP+FP} \quad \text{(Eq. 1)}$$

Recall, also known as the true positive rate, indicates how well the model identifies positive instances. The formula is:

$$Recall = \frac{TP}{TP+FN} \quad \text{(Eq. 2)}$$

The F1-Score is the harmonic mean of precision and recall, providing a balance between the two. The formula is:

$$F1\ Score = \frac{2 \times TP}{2 \times TP+FP+FN} \quad \text{(Eq. 3)}$$

AUC-ROC is called area under the receiver operating characteristic curve which plots the true positive (TP) rate versus the false positive (FP) rate at different classification thresholds. The formula is:

$$AUC - ROC = \int_0^1 TPR(FPR)\, dFPR \quad \text{(Eq. 4)}$$

3. RESULTS AND DISCUSSION

Table 4 Results for fraud prediction using 'Original' dataset

Model Name	AUC-ROC	F1-Score	Precision	Recall
Autoencoder	0.888	0.418	0.975	0.266
Neural Network	0.997	0.798	0.963	0.681
Logistic Regression	0.960	0.258	0.160	0.659
LightGBM	0.998	0.158	0.086	0.992
Random Forest	0.998	0.090	0.047	0.987
CatBoost	0.999	0.689	0.537	0.961

Table 5 Results for fraud prediction applying 'SMOTE'

Model Name	AUC-ROC	F1-Score	Precision	Recall
Autoencoder	0.906	0.04	0.02	0.657
Neural Network	0.998	0.567	0.404	0.95
Logistic Regression	0.961	0.521	0.486	0.563
LightGBM	0.999	0.669	0.512	0.965
Random Forest	0.996	0.760	0.930	0.642
CatBoost	0.999	0.448	0.292	0.962

Table 6 Results for fraud prediction applying 'Undersampling'

Model Name	AUC-ROC	F1-Score	Precision	Recall
Autoencoder	0.775	0.033	0.017	0.335
Neural Network	0.995	0.095	0.05	0.955
Logistic Regression	0.962	0.269	0.169	0.651
LightGBM	0.999	0.24	0.136	0.995
Random Forest	0.994	0.724	0.942	0.588
CatBoost	0.999	0.129	0.069	0.996

Table 7 Results for fraud prediction applying 'Oversampling'

Model Name	AUC-ROC	F1-Score	Precision	Recall
Autoencoder	0.909	0.031	0.015	0.698
Neural Network	0.999	0.325	0.194	0.986
Logistic Regression	0.961	0.503	0.447	0.575
LightGBM	0.999	0.856	0.818	0.897
Random Forest	0.998	0.113	0.060	0.989
CatBoost	0.895	0.006	0.005	0.007

Table 4 shows that Autoencoder achieves the highest Precision of 0.975 indicating it avoids false positives well. However, its Recall (0.266) is the lowest, meaning it misses most fraud cases. Catboost on the other hands stands out with the highest AUC-ROC (0.989) suggesting strong performance despite the imbalance.

Table 5 indicates that Autoencoder achieves a high Recall (0.657) but sacrifices Precision (0.02), leading to many false positives. But, Random Forest and CatBoost excel with almost perfect AUC-ROC (0.999) demonstrating their robustness with balanced data.

We can see in the Table 6 that Autoencoder performs poorly with low Precision (0.017) and Recall (0.335) while Random Forest and Catboost maintain strong AUC-ROC (0.999) making them reliable for fraud prediction even with fewer majority class samples.

According to Table 7, Autoencoder continues to underperform with very low Precision (0.015) but an improved Recall (0.698). Like always, Random Forest and Catboost maintain consistency with high AUC-ROC (0.999) proving their robustness.

4. CONCLUSIONS

The study's investigation into machine learning models for predicting fraud activities in payment transaction using historical of payment transaction to emphasize the exceptional performance of Autoencoder, Light Gradient Boosting Machine (LightGBM), logistic Regression (LR), Neural Network (NN), Random Forest (RF), and CatBoost. Among these, Random Forest and CatBoost are consistently the top-performing models across all techniques. They balance Precision and Recall effectively, achieving near-perfect AUC-ROC scores. On the other hands, Autoencoder, despite its high Precision in the imbalanced dataset, struggling to deliver balanced performance when sampling techniques are applied. This highlights Random Forest and CatBoost's proficiency in capturing the abnormalities in the payment transactions, providing a solid foundation for optimizing and enhancing payment systems.

Fraudulence in payment transactions is a very sensitive issue that should be tackled and prevented in the early stage otherwise the loss will be tremendous to the country's economy and growth. By applying the best model amongst the proposed model in the research to detect the abnormalities in the transaction, fraud activity will be prevented a major percent. Despite challenges, predicting fraudulent payments strongly aligns with the objectives of SDG9 and SDG16. By utilizing advanced technologies and establishing resilient infrastructure (SDG9), payment systems can be improved to identify and prevent fraud, driving economic progress and preserving valuable resources. Moreover, upholding integrity in financial transactions helps build peaceful and just societies by reducing corruption and strengthening trust in institutions (SDG16). Fraud detection safeguards both individuals and businesses while contributing to global stability and fairness,

making it an essential instrument for humanity's advancement.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my advisors, Dr. PHAUK Sokkhey for his expert guidance, unwavering support, and invaluable insights throughout my research. His mentorship has greatly influenced my academic path. Without his support, knowledge, and guidance, this research wouldn't have been possible.

I am also grateful to Mr. NGIN Kimlong, and Mr. LONG Pakrigna for their valuable support and constructive feedback which have greatly contributed to the development of this study. Ther willingness to share their knowledge and provide insightful suggestion has been crucial in refining my work.

Finally, I am deeply grateful to my family and friends for their unwavering support and understanding. Their encouragement and belief in me have been a constant source of motivation.

REFERENCES

Ata, O., & Hazim, L. (2020). Comparative Analysis of Different Distributions Dataset by Using Data Mining Techniques on Credit Card Fraud Detection. *Tehnički Vjesnik*, *27*(2), 618–626. https://doi.org/10.17559/TV-20180427091048

Bhattacharyya, S., Jha, S., Tharakunnel, K., & Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems*, *50*(3), 602–613. https://doi.org/10.1016/j.dss.2010.08.008

Cambridge Dictionary. (2023). FRAUD | Meaning in the Cambridge English Dictionary. Retrieved from dictionary.cambridge.org website: https://dictionary.cambridge.org/dictionary/english/fraud

Capgemini . (2023, September 14). Global non-cash transaction volumes set to reach 1.3 trillion in 2023. Retrieved from Capgemini website: https://www.capgemini.com/news/press-releases/global-non-cash-transaction-volumes-set-to-reach-1-3-trillion-in-2023/

Dai, M. (2023). Multiple Machine Learning Models on Credit Card Fraud Detection. *BCP Business & Management*, *44*, 334–338. https://doi.org/10.54691/bcpbm.v44i.4839

Feng, X., & Kim, S.-K. (2024). Novel Machine Learning Based Credit Card Fraud Detection Systems. *Mathematics*, *12*(12), 1869–1869. https://doi.org/10.3390/math12121869

Horng, P. (2024, August 20). Interior Minister says Cambodians lost $40m to online fraud in H1 2024 - Khmer Times. Retrieved from Khmer Times - Insight into Cambodia website: https://www.khmertimeskh.com/501544975/interior-minister-says-cambodians-lost-40m-to-online-fraud-in-h1-2024/

Nilson. (2023). Card Fraud Losses Worldwide in 2023. Retrieved from Nilson Report website: https://nilsonreport.com/articles/card-fraud-losses-worldwide-in-2023/

Oza, A. (2018). *Fraud Detection using Machine Learning*.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Interacting Neural Networks: A Comparative Study of Standard and Distributed Architectures on Image Classification

Sreyvi UANN [1*], Timoteo CARLETTI [2], Sokkhey PHAUK[1], Jarod KETCHA KOUAKEP [2], Malai NHIM [1]

[1] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Mathematics, University of Namur, Rue de Bruxelles 61, B-5000 Namur, Belgium*

[*] *Corresponding author: sreyvi_uann@gsc.itc.edu.kh*

Abstract: *Recently, Artificial Neural Networks (ANNs) have become to grounded of numerous breakthroughs in computer vision, speech recognition, and natural language processing. Traditional neural networks rely on a centralized, monolithic structure where a single model is trained end-to-end. Nevertheless, as networks grow in size and complexity, centralized models may face limitations in flexibility, interpretability, and robustness. Inspired by the principles of distributed systems and collective intelligence, this research explores an alternative paradigm: Interacting Neural Networks. We present and compare three architectures for image classification using the MNIST dataset: (1) a standard neural network with a single hidden layer, (2) multiple small neural networks combined using majority voting, and (3) multiple small neural networks using a Voter Model with logit blending. While each small neural network operates independently during training, their outputs are combined during evaluation through interaction-based decision-making. In the majority voting model, predictions are made by selecting the class predicted by the majority of models. In the Voter Model, we introduce an interaction loss that encourages each model to align its output with a blended version of the group, fostering a cooperative learning dynamic. The experiments indicate that interacting models can compete or outperform the standard neural network in terms of accuracy and time while maintaining modularity and lower individual model complexity. The Voter Model especially demonstrates consistent performance, benefitting from its shared belief mechanism. Training and evaluation times were also compared, and we present a detailed timing table across different hidden layer sizes. Results show that interacting neural networks outperform the standard model in terms of generalization and stability, especially under the Majority Voting Model, which achieved the highest test accuracy of 99.6%. These findings highlight the benefits of interaction-based architectures for improving neural network performance on image classification tasks.*

Keywords: Distributed Computation; Interacting Neural Networks; Artificial Neural Networks; Feedforward Neural Networks

1. INTRODUCTION

The Artificial Neural Network (ANNs) have been widely adopted for pattern recognition and classification tasks especially for image proceeding [1, 4]. Traditionally, these networks are trained end-to-end in a centralized manner where all layers are optimized simultaneously. Although this approach achieves high accuracy, it has limitations in terms of scalability and interpretability [2].

To address these limitations, distributed or interacting of neural networks have been begun to explore in this study which multiple smaller networks collaborate rather than relying in a single complex model. Inspired by swarm intelligence, multiple small neural networks operate in parallel, each learning independently but sharing decisions or influencing each other during training. In this paper, we propose and evaluate interacting neural networks for the MNIST digit classification problem [5].

We Compare the performance of a single standard neural network with two types of interacting models: majority voting and the voter model. Our main goal is to investigate whether interaction among multiple small models trained neural networks can lead to better or comparable performance with potentially reduced complexity.

2. METHODOLOGY

[1*] Corresponding author: Sreyvi UANN
E-mail: sreyvi_uann@itc.edu.kh; Tel: +855-77 437 155

2.1 Dataset

The MNIST dataset [5] was used which consists of 60,000 training and 10,000 testing grayscale images (0-255) of handwritten digits (0-9). Each image is 28 * 28 pixels and the task is to classify each image into one of the 10-digit classes. The dataset is normalized and flattened into 784-dimensional input vectors.

2.2 Models Setup

All models were implemented using PyTorch and trained using the Adam optimizer with a learning rate of 0.001. The training was conducted under four epoch settings 5, 10, 20 and 50 epochs. The standard neural network was trained with larger hidden sizes of 150, 300, 450 and 600 neurons in a single hidden layer. In contrast, the interacting models consisted of three small neural networks, each trained under the same epoch settings, but using smaller hidden sizes of 50, 100, 150 and 200 neurons.

2.3 Standard Neural Network (Single Model)

The baseline model is a fully connected feedforward neural network with a single hidden layer followed a ReLU activation function. The output layer applied a SoftMax function to generate class probabilities. The model is trained using the standard cross-entropy loss function:

$$\mathcal{L}_{CE} = -\sum_{c=1}^{C} y_c \log \hat{y}_c \quad (1)$$

Where:

- y_c is the true label (one hot encoded)
- $\hat{y}_c$ is the predicted probability for class C.

This loss penalizes the difference between true and predicted class probabilities [1].
To examine the effect of network capacity on performance, four hidden layer sizes were tested with 150, 300, 450 and 600 neurons. For each hidden size, the network was trained using four epoch settings: 5, 10, 20 and 50 epochs. Each configuration was repeated five times to obtain reliable result performance metrics. Evaluation metric included train accuracy, test accuracy and training execution time.

2.4 Interacting Neural Networks: Majority Voting

The Majority Voting architecture consists of three small fully connected neural networks, each trained separately but with interaction incorporated during the training phase. Each model includes one hidden layer with ReLU activation and a Softmax output layer. The hidden layer size was varied across four configurations: 50, 100, 150 and 200 neurons. The training was performed using the Adam optimizer with a learning rate of 0.001 for 5, 10, 20 and 50 epochs and each there was repeated five times for consistency.

During training, interaction is implemented via a custom loss function that encourages models to align their predictions with the majority vote. Each model computes its own cross-entropy loss $\mathcal{L}_i$ and the majority vote $\mathcal{L}_{maj}$ is determined across the three outputs. The interaction loss for each model i is defined as:

$$\mathcal{L}_i^{maj} = \mathcal{L}_i(\mathcal{L}_i + \mathcal{L}_{maj}) \quad (2)$$

The total loss over all model models is:

$$\mathcal{L}_{total} = \frac{1}{k}\sum_{i=1}^{k} \mathcal{L}_i^{maj} \quad (3)$$

During inference, the final prediction is obtained through majority voting: the class predicted by at least two of the three models is selected as the final output. This approach allows for collaborative decision making while maintaining model diversity.

2.5 Interacting Neural Networks: Voter Model

The Voter Model architecture also consists of three small fully connected neural networks, each containing a single hidden layer with ReLU activation and a Softmax output layer. The hidden layer sizes use was 50, 100, 150 and 200 neurons. Training was conducted with 5, 10, 20 and 50 epochs using the Adam optimizer (learning rate = 0.001) and repeated five time for each configuration to ensure consistency.

During training, a single model is randomly selected to serve as the voter model for the batch. The logits from this selected model are used as a reference to guide the training of all three models. Let $\mathcal{L}_i$ denote the standard cross-entropy loss of model i using its own logits and let $\mathcal{L}_{vot}$ be the cross-entropy loss computed using the logits of the selected voter model. The interaction loss for model i is defined as:

$$\mathcal{L}_i^{vot} = \mathcal{L}_i(\mathcal{L}_i + \mathcal{L}_{vot}) \quad (4)$$

The Total loss across all models is averaged:

$$\mathcal{L}_{total} = \frac{1}{k}\sum_{i=1}^{k} \mathcal{L}_i^{vot} \quad (5)$$

During inference, each model outputs its own prediction and one of them is randomly selected to act as the voter. The final prediction for each sample is taken directly from this selected voter model, maintaining consistency with the training interaction logic.

3. RESULTS AND DISCUSSION

This section presents and compares the performance of three neural network architectures: (1) Standard Neural Network, (2) Interacting Neural Networks with Majority Voting and (3) Interacting Neural Networks with the Voter Model. Evaluation is based on accuracy, loss and execution time across varying hidden layer sizes.

Figure 1 and Figure 2 illustrate the test accuracy and training loss of the standard neural network across different hidden sizes (150, 300, 450, 600) and epochs (5 to 50). As shown, test accuracy improves with increasing hidden size and epoch count, reaching around 99.4% at 600 units and 50 epochs. Training loss steadily decreases and stabilizes at values near 0.01, indicating good convergence. The performance is consistent across repeated runs, with smooth and stable curves.

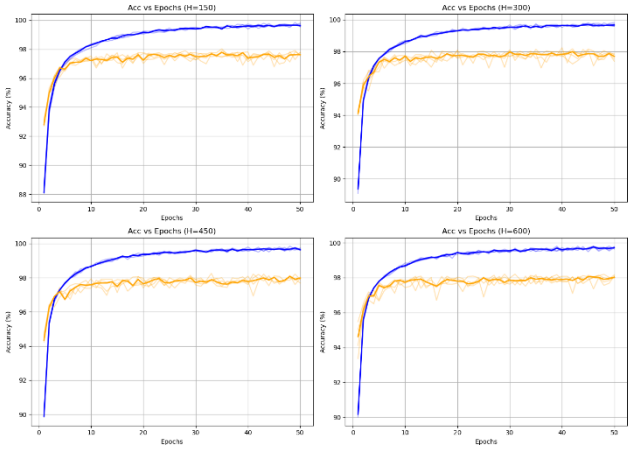


Fig. 1 Accuracy vs Epoch for Standard Neural Network

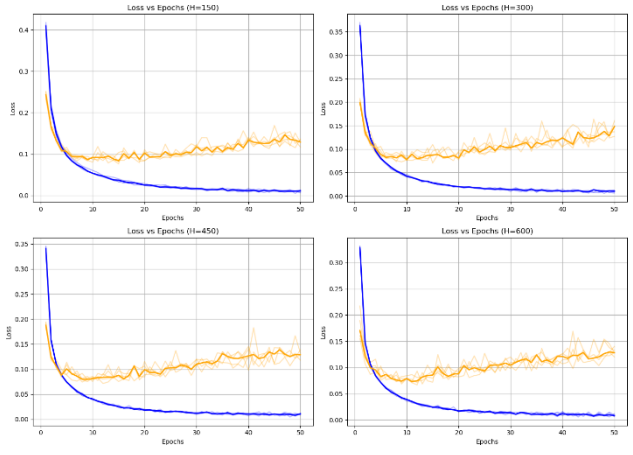


Fig. 2 Loss vs Epochs for Standard Neural Network

Figure 3 and Figure 4 show the test accuracy and loss of the Majority Voting model for hidden sizes 50, 100, 150, and 200. The model demonstrates strong performance, with test accuracy reaching approximately 99.6% at 200 hidden units and 50 epochs comparable to or slightly exceeding the best accuracy of the standard model, despite using fewer parameters. While the loss values are slightly higher than those of the standard model, the test accuracy curves remain stable.

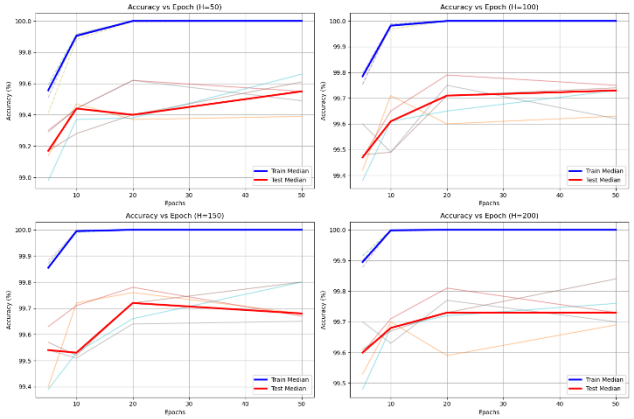


Fig. 3 Accuracy vs Epochs for Interacting Neural Network using Majority Voting

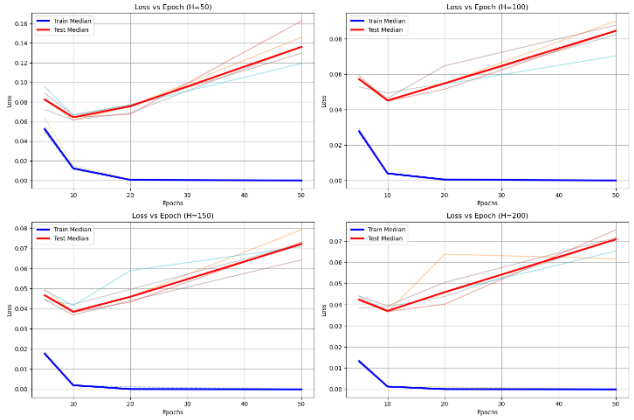


Fig. 4 Loss vs Epoch for Interacting Neural Networks using Majority Voting

For the Voter Model, Figure 5 presents the test accuracy of the Voter Model across different hidden sizes and epoch settings. Accuracy steadily increases with more epochs, reaching around 97% at 200 hidden units and 50 epochs. While it does not surpass the Majority Voting model, it consistently maintains strong generalization and low variance across configurations. Figure 6 shows the training and test loss values. The training loss converges quickly and remains low, while test loss increases slightly with more epochs, which may indicate mild overfitting. Overall, the Voter Model demonstrates stable learning and competitive accuracy under cooperative training with randomly assigned voter guidance.

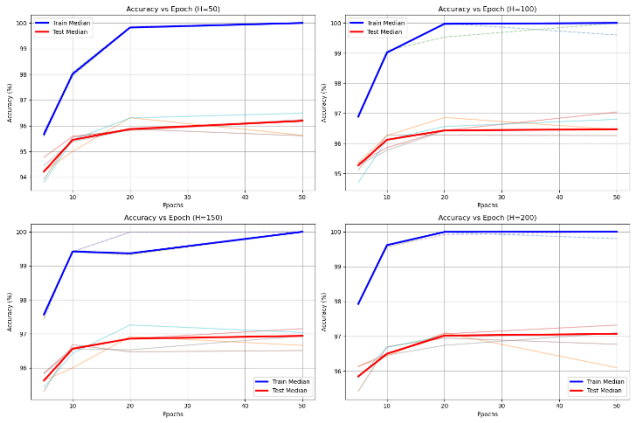


Fig. 5 Accuracy vs Epochs for Interacting Neural Networks using Voter Model

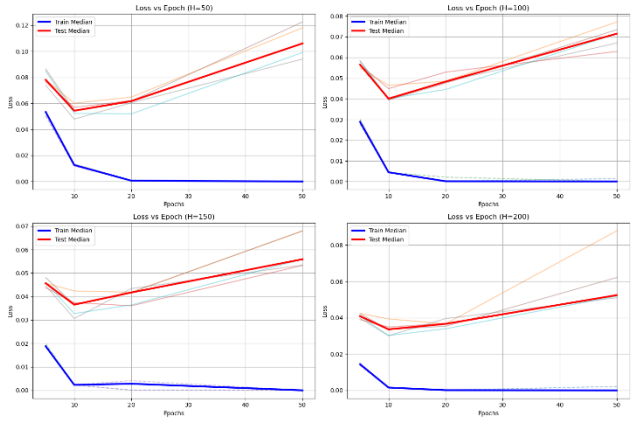


Fig. 6 Loss vs Epochs for Interacting Neural Networks using Voter Model

Execution time was measured the model take how much time for training across different hidden layer sizes to assess the trade-off between performance and processing cost. As shown in Table 1, the Standard Neural Network (Stand NN) recorded the longest execution times overall, especially at larger hidden sizes, due to its monolithic structure. The Majority Voting consistently achieved the shortest execution time while the Voter Model involved more computation.

Tabel 1 Execution Time Comparison

Hidden Size	Stand NN (s)	Majority Voting (s)	Voter Model (s)
50	52.04	29.04	47.48
100	58.70	29.54	63.35
150	73.31	30.03	65.21
200	72.09	29.82	54.14

4. CONCLUSIONS

This paper explored and compared three neural network architectures for image classification on the MNIST dataset: a standard neural network, a system of multiple small neural networks using Majority Voting and another using the Voter Model. The goal was to examine whether interacting architectures can compete with or outperform the standard approach in terms of accuracy, loss and execution time.

Our results show that interacting neural networks principally the Majority Voting Model offer notable advantages. The Majority Voting Model consistently achieved higher test accuracy and lower loss across different hidden layer sizes, indicating better generalization.

Moreover, the execution times were observed for interacting models which were significantly greater than the standard model and some cases, they were even lower, which highlights their practical efficiency. The inclusion of interaction losses in both Majority and Voter approaches permit distributed models to learn collaboratively and improve predictive consensus.

These finding validate the potential of decentralized neural architectures for classification tasks. In future work, these models could be extended to more complex datasets and deeper architectures or real-world scenarios.

ACKNOWLEDGMENTS

The author would like to express sincere gratitude to Prof. Timoteo Carletti for his guidance and support to this research. Special thanks are also extended to Dr. Sokkey Phauk, Mr. Jarod Ketcha Kouakep, and Mr. Malai Nhim for their valuable insights and assistance throughout this project. The author is also thankful to the Erasmus+ program for the opportunity to carry out this research as part of an exchange semester at the University of Namur, Belgium. Appreciation is further extended to the Institute of Technology of Cambodia for providing the academic foundation and support that enabled this experience.

REFERENCES

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” Proceedings of the IEEE, 1998.
[2] G. Hinton, O. Vinyals, and J. Dean, “Distilling the knowledge in a neural network,” arXiv preprint arXiv:1503.02531, 2015.
[3] T. Carletti and C. Martinet, “Collective dynamics of interacting neural networks,” Journal of Complex Systems, 2020.
[4] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” arXiv preprint arXiv:1409.1556, 2014.
[5] Y. LeCun, “The MNIST Database of Handwritten Digits.”

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Enhancing Khmer Language AI: A Privacy-Focused Chatbot with Retrieval-Augmented Generation

Nita Li [1*], Seyha Heng [1], Lay Seng[1], Jihyun Moon [2], Sokkhey Phauk [1], Charmgil Hong[2]

[1] *ReDA Lab, Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia*
[2] *Handong Global University, Pohang, Gyeongsangbuk-do, South Korea*

[*]*Corresponding author: nitali960@gmail.com*

Abstract: *Khmer language processing faces significant challenges due to limited annotated datasets, complicated word segmentation issues, and a scarcity of pre-trained models tailored to the Khmer language. These limitations make it difficult for processing and understanding Khmer text effectively using AI. This research focuses on developing a privacy-focused Khmer chatbot using a Retrieval-Augmented Generation (RAG) framework to improve response accuracy and knowledge retrieval. The chatbot is intended to assist private organizations, including government ministries, where secure access to internal information is essential. The solution provides a means for authorized people the ability to access and retrieve internal documents at anytime and anywhere, without the need for paper-based records or additional staff. This not only reduces operational costs and improves organizational effectiveness but also saves time and ensures data privacy. We compared five different models from Hugging Face—SeaLLM-v3-7b-chat, aisingapore/gemma2-9b-cpt-sea-lionv3-instruct, kimleang123/fine-tuned-KQA-gemma-7B-QLora-64-128, kimleang123/fine-tuned-KQA-mistral-7B-v0.3-lora-128-256, and kimleang123/fine-tuned-KQA-qwen2-7B-QLoRA. This comparison helped us select the most effective model for the chatbot's language understanding capabilities. To find the best-performing model, we evaluate using qualitative self-evaluation and quantitative benchmarking such as perplexity, ChrF++, and latency. Once the best model is selected, we then integrated it into a RAG framework to further enhance knowledge retrieval. Furthermore, we compared two embedding models namely sentence-transformers/all-MiniLM-L6-v2 and intfloat/multilingual-e5-large to see which one performs better in the context of the Khmer language. Besides improving Khmer Natural Language Processing (NLP), this technique addresses real-world challenges including secure document access and digital transformation in organizations. This study's solution promotes technological innovation, efficient resource management, and increase transparency in both public and private sector operations, thereby contributing to the achievement of the United Nations Sustainable Development Goals (SDG 9 and SDG16). The AI-powered chatbot sets the foundation for sustained improvements in productivity, document management, and access to information in both Cambodia's government and private sectors.*

Keywords: Khmer Language Processing; Retrieval-Augmented Generation (RAG); Natural Language Processing (NLP); AI Chatbot

1. INTRODUCTION

Over the past decade, the development of artificial intelligence (AI) and natural language processing (NLP) has significantly improved human-computer interaction. Large language models (LLMs) such as ChatGPT, Gemini, and DeepSeek have transformed how people access information, automate tasks, and communicate with machines by generating human-like responses to a wide range of queries [1,2]. However, despite their impressive capabilities, these general-purpose chatbots often struggle with low-resource languages like Khmer, particularly when answering domain-specific or localized questions that require cultural, contextual, or institutional understanding [3].

Khmer, the official language of Cambodia, is still underrepresented in the AI and NLP research community. This is because there is not much availability of digitized corpora, annotated datasets, or high-quality pretrained language models. As a result, current chatbots usually fall short in providing good service to Khmer-speaking users, particularly in high-stakes situations like ministries or banks where internal documents are not publicly available, and responses must reflect specific institutional knowledge. In such environments, leaders often rely on human assistants to retrieve necessary information which cause delays, higher staffing costs, and limited access to real-time data, especially outside office hours.

Recent studies have shown that Retrieval-Augmented Generation (RAG) is a powerful approach to improve chatbot

[1*] Corresponding author: Nita Li
E-mail: nitali960@gmail.com; Tel: +855-97 995 7535

performance by combining large language models (LLMs) with real-time access to documents [4,5]. This setup architecture allows models to pull information directly from internal or private data sources, making its responses more accurate and relevant to the user's question without exposing sensitive data to public. The RAG framework is especially useful in areas where privacy is important, such as government or business settings, because it can provide answers without exposing confidential data to public.

Building on this concept, the present study introduces a privacy-focused Khmer language chatbot that leverages the RAG architecture to enable secure and appropriate responses from internal document stores. The system is deployed using Gradio [6] to provide an interactive and user-friendly interface. To identify the most effective configuration for Khmer question answering in closed environments, we evaluate and compare five pre-trained LLMs from Hugging Face [7] and two embedding models for retrieval tasks. The proposed solution addresses real-world situations, especially in places like government offices and banks, where it's important to quickly and securely find accurate information. In addition, this work supports Cambodia's move toward digital technology and aligns with two of the United Nations Sustainable Development Goals (SDGs). It supports SDG 9 (Industry, Innovation, and Infrastructure) by encouraging innovation and building local AI tools, and supports SDG 16 (Peace, Justice, and Strong Institutions) by helping create more transparent, efficient, and accountable institutions.

2. METHODOLOGY

2.1 Dataset

Dataset used in this study was collected from two different sources namely, OSCAR (Open Super-large Crawled Aggregated Corpus) [9,10] and Wikipedia. While Wikipedia is an online encyclopedia, OSCAR is a multilingual text corpus derived from web data, designed for training large-scale language models. We obtained 146 text files from OSCAR and 2,482 text files from Wikipedia.

2.2 Data Cleaning and Preprocessing

To prepare the dataset for training, we first applied word segmentation using the Conditional Random Fields (CRF) [11] method. Since the Khmer language, unlike English, does not use spaces between words, word segmentation is a critical step. CRF helps predicting the correct word boundaries based on the surrounding context. This makes it easier for the model to better understand and generate meaningful responses.

For efficiency, especially given the time-consuming nature of running large language models, we selected 10 representative files from the OSCAR corpus for initial experimentation rather than using the full dataset.

2.3 Implementation

To build an effective Khmer chatbot, we started by testing five pre-trained language models from Hugging Face to see which one best understood and responded to Khmer input. Here are the five pre-trained models:

- `SeaLLM-v3-7b-chat`: Based on `Qwen2-7B-chat`'s structure and it is pre-trained with South East Asian languages [12].
- `aisingapore/gemma2-9b-cpt-sea-lionv3-instruct`: Based on Gemma2, this model is part of SEA-LION (Southeast Asian Languages In One Network), which is designed for Southeast Asian languages [13,14]
- `kimleang123/fine-tuned-KQA-gemma-7B-QLora-64-128`: A Khmer fine-tuned model built from Gemma 7B [15]
- `kimleang123/fine-tuned-KQA-mistral-7B-v0.3-lora-128-256`: A Khmer fine-tuned version of Mistral 7B v0.3 [15]
- `kimleang123/fine-tuned-KQA-qwen2-7B-QLoRA:` A Khmer fine-tuned model based on Qwen2 7B [15]

After evaluating all five, we chose the best-performing model and integrated it into a Retrieval-Augmented Generation (RAG) framework. This help the chatbot to generate responses that were both accurate and better aligned with the user's queries by pulling information from a document database. During this process, we also evaluated two different embedding models to determine which one provided better retrieval performance. Here are the embedding models:

- `intfloat/multilingual-e5-large` [16]
- `sentence-transformers/all-MiniLM-L6-v2` [17,18]

Once the RAG setup was complete, we deployed the chatbot using Gradio, providing a user-friendly web interface. All development and testing were conducted on Google Colab using Python programing, which allowed us to access the necessary computing resources efficiently.

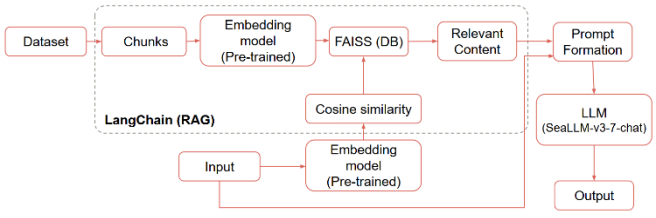


Fig. 1. Diagram illustrating the architecture of the Khmer chatbot system.

Figure 1 provides an overview of the LangChain-based Retrieval-Augmented Generation (RAG) framework used in our Khmer language chatbot. The system uses a pre-trained embedding model to convert both the user's query and chunks of a document dataset into vector form. After converting the queries and document chunks into vectors, these vectors are indexed and queried using FAISS, a specialized tool designed for efficient similarity searches in vector spaces. The system then identifies the most relevant information based on the user's query, which is integrated into a prompt. Once the system forms the prompt, the best-performing model among the five (`SeaLLM-v3-7B-chat`) uses this prompt to generate a response that fits the user's question with accuracy and relevance. Each part of the framework plays a vital role in making this work smoothly:

- *Chunks*: The original dataset is first broken down into smaller, meaningful parts to make it easier to find the most relevant information.
- *Embedding Model:* These chunks are then converted into numerical vector using a pre-trained embedding model. This allows the system to understand the underlying meaning of the text.
- *FAISS (Facebook AI Similarity Search):* The vectorized data is stored in FAISS, a specialized tool designed for quickly finding similar vectors within large datasets.
- *Cosine similarity*: This is used to measure how closely related the chunks are to the user's query. This technique compares the angle between vectors to assess their similarity.
- *Relevant content*: Based on the similarity scores, the system selects the chunks that best match the user's query.
- *Prompt formation*: these relevant chunks are combined with the user's original input to create a comprehensive prompt, which guides the model in generating a well-informed and context-aware response.

2.2 Evaluation Metrics

To assess the performance of our Khmer language chatbot, we used three evaluation metrics: Perplexity, ChrF++, and Latency.

Perplexity measures how well the model predicts a sample. A lower perplexity score indicates better language understanding and generation [19].

$$PP(w) = P(w_1 w_2 \dots w_n)^{-\frac{1}{N}} \quad \text{(Eq. 1)}$$

ChrF++is a character n-gram F-score metric, suitable for evaluating languages with complex word boundaries like Khmer. It provides a reliable measure of text generation quality at the character level [20,21,22].

$$\text{CHRF}\beta = (1+\beta^2)\frac{\text{CHRP.CHRR}}{\beta^2.\text{CHRP} + \text{CHRR}} \quad \text{(Eq. 2)}$$

CHRP = Precision
CHRR = Recall

Latency refers to the response time of the chatbot. We measured how quickly the model returns answers, which is essential for a smooth user experience.

$$Latency = Response\ End\ Time - Request\ Start\ Time \quad \text{(Eq. 3)}$$

These metrics help us understand the chatbot's accuracy, fluency, and responsiveness in real-world use.

3. RESULTS AND DISCUSSION

In this section, we present the key findings of the study. It covers the process of choosing the best pre-trained model for the Khmer chatbot, identifying the most effective embedding model, and evaluating how well the chatbot performed overall.

3.1 Comparison of the Pretrained LLM

Table 1. Evaluation scores based on selected metrics using the input example: សួស្តី! នេះគឺជាការធ្វើតេស្ត។

Model\ Evaluation Scores	Perplexity ↓	ChrF++ ↑	Latency ↓
`Aisingapore`	24	85.33	40.14
`SeaLLMs/SeaLLms-v3-7B-chat`	4.05	87.78	14.41
`Kimleang123/fine-tuned-KQA-gemma-7B-QLora-64-128`	30.76	85.33	21.23
`Kimleang123/fine-tuned-KQA-mistral-7B-v0.3-lora-128-256`	6.84	89.06	7.15
`Kimleang123/fine-tuned-KQA-qwen2-7B-QLoRA`	4.95	6.42	1370.37

Table 1. illustrates the evaluation scores of the five pre-trained models using the input example: សួស្តី! នេះគឺជាការធ្វើតេស្ត។ The table depicts 2 outperforming models namely *`SeaLLMs/SeaLLms-v3-7B-chat`* and *`Kimleang123/fine-tuned-KQA-mistral-7B-v0.3-lora-128-256`*. Now let's look at **Table 2.**

Table 2. Responses generated by five models using the input example: សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៖

Model	Answer

Aisingapore	1. **បរិស្ថានគឺជាផ្នែកសំខាន់មួយនៃជីវិតមនុស្ស ដែលយើងត្រូវរក្សានិងរក្សាសុខភាពវា ដើម្បីសុខសប្បាយនិងអភិវឌ្ឍន៍របស់សង្គមយើង។** 2. **ការប្រើប្រាស់ធនធានធម្មជាតិយ៉ាងមិនប្រសើរនឹងបង្កើតបញ្ហាអាក្រក់ទៅលើបរិស្ថាន ដូចជាការផ្លាស់ប្តូរអាកាសធាតុ និងការអនុគ្រោះរបស់សត្វ ដែលត្រូវតែយើងប្រព្រឹត្តការថែរក្សានិងប្រើប្រាស់យ៉ាងសម្តែង។**
SeaLLMs/SeaLLms-v3-7B-chat	1. បរិស្ថានជុំវិញយើងកំពុងត្រូវបានបំផ្លាញដោយការចំណាយថាមពល និងការរំលោភបំពានធនធានធម្មជាតិ។ 2. ប្រជាជនទាំងអស់គ្នាត្រូវតែមានការយល់ដឹងពីសារៈសំខាន់នៃការការពារបរិស្ថាន ហើយចូលរួមក្នុងការអភិរក្សធនធានធម្មជាតិ។
Kimleang123/fine-tuned-KQA-gemma-7B-QLora-64-128	ប្រយោគទាំងពីរគឺ បរិស្ថានជារឿងរបស់មនុស្ស និងជារឿងរបស់រាល់សត្វលោក បរិស្ថានជារឿងរបស់មនុស្ស គឺជាអ្វីៗដែលនៅជុំវិញមនុស្ស មានបររបស់មនុស្ស និងបររបស់រាល់សត្វលោក គឺជាអ្វីៗដែលនៅជុំវិញសត្វ មានបររបស់មនុស្ស និងបររបស់រាល់សត្វលោក គឺជាបររបស់មនុស្ស និងបររបស់រាល់សត្វលោក បរិស្ថានជារឿងរបស់មនុស្ស និងបររបស់រាល់សត្វ
Kimleang123/fine-tuned-KQA-mistral-7B-v0.3-lora-128-256	ខ្ញុំសរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៖ សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន។ សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន២ សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៣ សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៨ សរសេរប្
Kimleang123/fine-tuned-KQA-qwen2-7B-QLoRA	ប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៖ បរិស្ថានជាអ្វីដែលនៅជុំវិញខ្លួនយើង និងជាប្រភពនៃការរស់រានមានជីវិត បរិស្ថានជាអ្វីដែលនៅជុំវិញខ្លួនយើង និងជាប្រភពនៃការរស់រានមានជីវិត បរិស្ថានជាអ្វីដែលនៅជុំវិញខ្លួនយើង និងជាប្រភពនៃការរស់រានមានជីវិត បរិស្ថានជាអ្វីដែលនៅ

Table 2. presents the generated responses from five pre-trained models when given the prompt: `សរសេរប្រយោគពីរជាភាសាខ្មែរ អំពីបរិស្ថាន៖`. Based on self-judgement, we decided that the responses produced by the *`SeaLLMs/SeaLLms-v3-7B-chat`* model makes more sense compared to the others. Thus, *`SeaLLMs/SeaLLms-v3-7B-chat`* model is the optimal model.

3.2 Embedding Models Comparison

After selecting the optimal pre-trained model, we further enhanced the chatbot's accuracy by adding an embedding model. In this process, we compared two embedding models: *`intfloat/multilingual-e5-large`* and *`sentence-transformers/all-MiniLM-L6-v2`*.

Table 3. Responses generated using two different embedding models for the input example: ហេង ពិទូជានរណា?

Model	Answer
intfloat/multilingual-e5-large	ហេង ពិទូជាតារាចម្រៀងដែលមានឈ្មោះល្បីល្បាញក្នុងស្រទាប់យុវវ័យ
sentence-transformers/all-MiniLM-L6-v2	ហេង ពិទូជា សុ

Table 3. shows the comparison between two embedding models using the question: `ហេង ពិទូ ជានរណា?` The `intfloat/multilingual-e5-large` model outperformed the other by generating the most accurate and relevant answer.

3.3 Chatbot Comparison

To evaluate the effectiveness of our Khmer chatbot, we compared its responses with two other versions: ChatGPT, the original SeaLLM model without the RAG framework (`SeaLLMs/SeaLLms-v3-7B-chat`), and our chatbot using SeaLLM with the RAG framework (`SeaLLMs/SeaLLms-v3-7B-chat` with `intfloat/multilingual-e5-large` as embedding model). A few sample questions were asked, and the responses are shown in Table 4 for comparison.

Table 4. Comparison of chatbot responses between ChatGPT, SeaLLM without RAG, and SeaLLM with RAG

Question	ChatGPT	Original SeaLLM	RAG with SeaLLM
ហេង ពិទូជានរណា?	ហេង ពិទូជាកីឡាករនៅក្នុងក្រុមបាល់ទាត់ប្រទេសកម្ពុជា	សូមអភ័យទោសខ្ញុំមិនអាចផ្តល់ព័ត៌មានបន្ថែមបានទេ។ ប្រសិនបើអ្នកមានសំណួរផ្សេងៗ សូមការផ្តល់ព័ត៌មានបន្ថែម ដើម្បីខ្ញុំអាចជួយអ្នកបានច្បាស់។	ហេង ពិទូជាតារាចម្រៀងដែលមានឈ្មោះល្បីល្បាញក្នុងស្រទាប់យុវវ័យ។
នរណាជារដ្ឋមន្ត្រីក្រសួងមហាផ្ទៃរបស់ប្រទេសអ៊ីតាលី?	រដ្ឋមន្ត្រីក្រសួងមហាផ្ទៃរបស់ប្រទេសអ៊ីតាលីគឺ François Hollande។	នរណាម្នាក់ឈ្មោះ Aymeric Chauprade ជារដ្ឋមន្ត្រីក្រសួងមហាផ្ទៃរបស់ប្រទេសអ៊ីតាលី។ គាត់គឺជាអ្នកនយោបាយដែលមានចំណង់ប្រជាធិបតេយ្យនិងការពារសិទ្ធិមនុស្ស។ ក្រុមហ៊ុនស្ថាបន៍ស៊ីវិលឯកជនគ្រប់គ្រងដោយក្រុមហ៊ុនអូវ៉ា (AVE), ក្រុមហ៊ុនកសិកម្មហាក្សានៅកំណើត, និងក្រុមហ៊ុនវិនិយោគក្រឡាចារ្យ	រដ្ឋមន្ត្រីក្រសួងមហាផ្ទៃរបស់ប្រទេសអ៊ីតាលីគឺលោក Matteo Salvini ។

លោក ប្រធានាធិបតី បារ៉ាក់អូបាម៉ា បំពាក់ មេដាយសេរីភាពដល់ លោក Bob Dylan ពេលណា?	លោក ប្រធានាធិបតី បារ៉ាក់ អូបាម៉ា បានបំពាក់ មេដាយសេរីភាពដល់ លោក Bob Dylan នៅឆ្នាំ 2005។	លោកប្រធានាធិបតីបារ៉ាក់អូបាម៉ា បានបំពាក់មេដាយសេរីភាពដល់ លោក Bob Dylan នៅឆ្នាំ 1990។ នេះគឺជាការផ្ទេរម្ហូបិតប្រឆាំងនឹងការចោទប្រកាន់ពីគណបក្សប្រឆាំង (Alliance of Artists and Composers, Inc.) ដែលបានចោទប្រកាន់ថា Dylan គួរតែប្តូរទៅជាសិល្បករប្រដាប់ខ្លួន (independent artist) ព្រោះគាត់បានបញ្ចេញស្នាដៃដោយស្ម័យប្រវត្តិជាមួយក្រ	លោក ប្រធានាធិបតី បារ៉ាក់អូបាម៉ា បំពាក់មេដាយ សេរីភាពដល់ លោក Bob Dylan នៅថ្ងៃទី ២៩ ខែ ឧសភា ឆ្នាំ ២០១២

As shown in **Table 4**, the chatbot powered by SeaLLM and enhanced with the RAG framework provided more accurate and relevant answers compared to the other two models.

4. CONCLUSIONS

This study successfully developed a privacy-focused Khmer chatbot using the `SeaLLMs-v3-7B-Chat` model integrated with a Retrieval-Augmented Generation (RAG) framework. The system used the `intfloat/multilingual-e5-large` embedding model to generate vector representations of the documents, which were then stored in a database for efficient search. This helped the chatbot to retrieve accurate and meaningful answers, especially when dealing with domain-specific information.

By combining the power of large language models with document retrieval, the chatbot delivers more accurate, context-aware, and domain-specific responses, especially important for low-resource language settings like Khmer. Among the models tested `SeaLLMs-v3-7B-Chat` with RAG outperformed both the original `SeaLLMs-v3-7B-Chat` and general-purpose tools like ChatGPT. Additionally, the embedding model `intfloat/multilingual-e5-large` proved to be the most effective for retrieving relevant information, further improving the system's performance.

Deployed with Gradio, the chatbot provides a user-friendly interface and supports real-world applications in environments where data privacy and fast information access are critical, such as government ministries and banks. The solution improves operational efficiency, reduces the need for human intermediaries, and lowers costs. Importantly, this work contributes to Cambodia's digital innovation goals by aligning with SDG 9 (Industry, Innovation, and Infrastructure) and SDG 16 (Peace, Justice, and Strong Institutions), supporting the development of secure, inclusive, and efficient digital systems.

As a next step, future research will explore generating training data by extracting text from image files to further enrich the chatbot's knowledge and improve performance in specialized domains.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to the UNESCO UNITWIN program, organized by the Handong Global University team, for their teaching, guidance, and support throughout the Data Science Camp from 2023 to 2025. Special thanks to Prof. Charmgil Hong for his invaluable supervision during this project, and to Asst. Prof. Dr. Phauk Sokkhey for his continuous guidance and support in this research.

REFERENCES

[1] T. Brown et al., "Language Models are Few-Shot Learners," *Advances in Neural Information Processing Systems*, vol. 33, pp. 1877–1901, 2020. [Online]. Available: https://arxiv.org/abs/2005.14165

[2] Google DeepMind, *Gemini: Our Largest and Most Capable AI Model*, 2023. [Online]. Available: https://deepmind.google/technologies/gemini/

[3] D. Nekoto et al., "Participatory Research for Low-Resourced Machine Translation: A Case Study in African Languages," in *Findings of the Association for Computational Linguistics: EMNLP 2020*, pp. 1896–1908, 2020. [Online]. Available: https://aclanthology.org/2020.findings-emnlp.195/

[4] P. Lewis et al., "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks," *Advances in Neural Information Processing Systems*, vol. 33, pp. 9459–9474, 2020. [Online]. Available: https://arxiv.org/abs/2005.11401

[5] G. Izacard and E. Grave, "Leveraging Passage Retrieval with Generative Models for Open Domain Question Answering," in *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, pp. 8749–8759, 2020. [Online]. Available: https://arxiv.org/abs/2007.01282

[6] G. Team, "Quickstart." https://www.gradio.app/guides/quickstart

[7] "Hugging Face – The AI community building the future." https://huggingface.co/

[8] "THE 17 GOALS | Sustainable Development." https://sdgs.un.org/goals

[9] P. O. Suarez, "OSCAR," *OSCAR*. https://oscar-project.org/

[10] "oscar-corpus (OSCAR)," Jul. 26, 2023. https://huggingface.co/oscar-corpus

[11] P. Tum, "Word segmentation of Khmer text using conditional random fields," *Medium*, Sep. 28, 2022. [Online]. Available: https://medium.com/@phylypo/segmentation-of-khmer-text-using-conditional-random-fields-3a2d4d73956a

[12] X.-P. Nguyen et al., "SeaLLMs - Large Language Models for Southeast Asia," DAMO Academy, Alibaba Group, pp. 294–304, Jan. 2024, doi: 10.18653/v1/2024.acl-demos.28.

[13] Atoh, "Southeast Asian Languages in One Network Data (SEALD) - AI Singapore," *AI Singapore*, Mar. 18, 2024. https://aisingapore.org/aiproducts/southeast-asian-languages-in-one-network-data-seald/

[14] Aisingapore, "GitHub - aisingapore/sealion: South-East Asia Large Language Models," *GitHub*. https://github.com/aisingapore/sealion

[15]"kimleang123 (Bread)," Oct. 11, 2024. https://huggingface.co/kimleang123

[16]L. Wang, N. Yang, X. Huang, L. Yang, R. Majumder, and F. Wei, "Multilingual E5 Text Embeddings: A Technical Report," Microsoft Corporation, Feb. 2024. [Online]. Available: https://huggingface.co/intfloat/multilingual-e5-large

[17] "sentence-transformers/all-MiniLM-L6-v2 · Hugging Face," Jan. 05, 2024. https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

[18] "SentenceTransformers Documentation — Sentence Transformers documentation." https://www.sbert.net/

[19] A. Kim, "Perplexity Intuition (and its derivation) - TDS Archive - Medium," *Medium*, Sep. 19, 2022. [Online]. Available: https://medium.com/data-science/perplexity-intuition-and-derivation-105dd481c8f3

[20] M. Popović, "chrF++: words helping character n-grams," *Proceedings of the Second Conference on Machine Translation*, pp. 612–618, Jan. 2017, doi: 10.18653/v1/w17-4770.

[21] S. Aycock, D. Stap, D. Wu, C. Monz, and K. Sima'an, "Can LLMs Really Learn to Translate a Low-Resource Language from One Grammar Book?," *arXiv (Cornell University)*, Sep. 2024, doi: 10.48550/arxiv.2409.19151.

[22] M. Popovic, "CHRF: Character N-Gram F-Score for Automatic MT Evaluation," *Proceedings of the Tenth Workshop on Statistical Machine Translation*, pp. 392–395, Sep. 2015, [Online]. Available: https://github.com/kweonwooj/papers/issues/17

Mechatronic and Information Technology (MIT)

2 Sub-parallel Sessions

"MIT-2: Mechatronics Innovations and Control Systems"

Chair: Assist. Prof. Dr. SRANG Sarot, Co-Chair: Assist. Prof. Dr. PEC Rothna

No.	Topic
1	Balloon-observed Gravity Wave Momentum Flux Reconstruction using Transfer Learning Bayesian Neural Network *Authored by: Sothea HAS, Riwal Plougonven, Aurélie Fischer, Raj Rani, Francois Lott, Albert Hertzog, Aurélien Podglajen*
2	Comparing Time Series with Machine Learning and Deep Learning Models for Paddy Rice Price Forecasting in Battambang Province, Cambodia *Authored by: Phirum LENG, Sokkhey PHAUK, Malai NHIM, Chantara TANN*
3	The Forecasting for Cambodia's Rice Production Using Multiple Linear Regression and Tree-Based Ensemble Learning Methods *Authored by: Sithach THONG, Sokkhey PHAUK, Sothea HAS, Ponna PHOK*
4	Model-Based PMSM Control in the dq Frame Using Full-State Feedback with Integral Action and EKF for State Estimation under Stochastic Dynamics *Authored by: Rotanak CHHEAN, Sarot SRANG, Sokserey SREY*
5	Continuously Uploading ECG Waveform Data to Google Sheet using ESP32 *Authored by: Phalla I Ning LEANG, Lichhean LY, Sreypich PHORN, Suchhing LY, Rithea NGETH*
6	Toward a Case Study on Event Detection in Non-Intrusive Load Monitoring Using Geometric Features of the Cumulative Sum *Authored by: Song SOK, Korsorl THOURN, Kimtho PO*
7	Observation of the Effect of Cutting Parameters on Surface Roughness of High Purity Copper in Dry Turning Process *Authored by: Sela Soy, Vongchivorn Chhour, Phearom San, Sophal Hai*
8	Investigating Efficiency of PMSM Using Full State Feedback Control with Integral Action and Dynamic Compensation *Authored by: Bunsrung TAING, Sarot SRANG, Sokserey SREY*
9	Predictive Modeling of Garment Trade Flows in Cambodia: A Data-Driven Approach to Forecasting Export Trends *Authored by: Dara Em, Vesal KHEAN, Sokkhey PHAUK*
10	Reconstructing Higher-Order Interaction in Algal Communities *Authored by: Danut CHHORN, Sokkey PHAUK, Timoteo CARLITTI, Malia LIN*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Balloon-observed Gravity Wave Momentum Flux Reconstruction using Transfer Learning Bayesian Neural Network

Sothea HAS[1,5*], Riwal Plougonven[2], Aurélie Fischer[3], Raj Rani[4], Francois Lott[4], Albert Hertzog[5], Aurélien Podglajen[5]

[1]*Laboratoire de Météorologie Dynamique, Ecole Normale Supérieure, IPSL, Paris, France*
[2]*CNRS/Laboratoire de Probabilités, Statistique et Modélisation (LPSM), Université Paris Cité, France*
[3]*Laboratoire de Météorologie Dynamique (LMD)/IPSL, PSL Research Institute, Paris, France*
[4]*LMD/IPSL, Sorbonne Université, Paris, France*
[5]*Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: has.sothea@itc.edu.kh*

Abstract: *Global atmospheric models rely on parameterizations to capture the effects of gravity waves (GWs) on middle atmosphere circulation. As they propagate upwards from the troposphere, the momentum fluxes associated with these waves represent a crucial yet insufficiently constrained component. The present study employs* ***Bayesian Neural Networks*** *to probe the relationship between large-scale flow and small-scale GWs within the tropical lower stratosphere. The measurements collected by eight superpressure balloons from two Strateole 2 campaigns, comprising a cumulative observation period of 1304 days (680 days in 2019 and 624 days in 2021), provide valuable estimates of the gravity wave momentum fluxes (GWMFs). Multiple explanatory variables, including total precipitation, wind, and temperature, were interpolated from the ERA5 reanalysis at each balloon's location. Given the presence of uncertainty in both large-scale flow inputs and the target, Bayesian Neural Networks were selected as the primary model. These networks were trained on data from the second campaign (2021) and subsequently used to estimate three types of reference GWMFs—eastward, westward, and total momentum—for the first campaign conducted in 2019. We observed that parts of the GW signal are successfully reconstructed, with correlations typically around 0.52 and maximum correlation around 0.77 for certain balloons. The models show significantly different performances from one balloon to another and able to capture high-peak events in GWMFs better than tree-based models. This study also discusses potential limitations, including the intermittent nature of GWMFs and data scarcity, providing insights into the challenges and opportunities for advancing our understanding of these atmospheric phenomena.*

Keywords: Gravity Waves, Momentum Fluxes, Bayesian Neural Networks, Transfer Learning, Superpressure Balloons

1. INTRODUCTION

As computational power advances, climate models and Numerical Weather Prediction (NWP) systems can simulate atmospheric phenomena—such as wind, temperature, and pressure—in greater detail with increasing spatial resolution. However, many small-scale processes, including cloud dynamics, turbulence, and wave interactions, remain unresolved, necessitating continuous refinement. Among these, internal gravity waves significantly influence the dynamics and variability of the middle atmosphere [5]. Gravity waves are waves in a fluid medium that result from the restoring force of gravity or buoyancy [10]. These waves occur at the interface between two media or within a fluid itself, propagating as the displaced fluid seeks to return to equilibrium. The main sources occur in the troposphere, such as flow over mountains and cloud development. Their three-dimensional propagation induces major aggregated impacts in the stratosphere and mesosphere, forcing key aspects of circulation. Enhancing stratosphere modeling is driven not only by long-term climate concerns (e.g., [12] and [13]) but also by the need for better short-term forecasts ([3] and [4]).

Gravity waves (GWs) occur on scales ranging from a few to several hundred kilometers. An important effect stems from

[1]* Corresponding author: Sothea HAS
E-mail: heng_sounean@itc.edu.kh; Tel: +855-86 723 327

their vertical propagation: gravity waves are responsible for vertical transfers of momentum from lower layers (troposphere: denser and with more gravity wave sources) to upper layers (stratosphere and beyond), where they constitute an essential driver of the overall circulation [5]. A significant part of the spectrum of gravity waves has been and remains unresolved in global models, requiring these effects to be represented by parameterizations [7]. Models display sensitivity to these, calling for coordinated efforts to better constrain these parameterizations from both observations and high-resolution modeling [1].

This study utilizes data from the Strateole 2 high-altitude balloon campaigns, conducted from October 2019 to February 2020 and October to January 2021. The first campaign deployed eight superpressure balloons, while the second involved seventeen balloons, circled the Earth at altitudes of 18–20 km within tropical regions. These balloons provided in situ observations of atmospheric variables, including wind speed, enabling precise estimates of gravity waves along their trajectories. These estimates serve as ground truth for reconstruction using machine learning (ML) models and inputs of resolved large-scale flow data derived from the ERA5 dataset.

2. DATA AND METHODOLOGY

2.1 Data

The two Strateole 2 campaigns accumulate a total observation of approximately 1304 days (680 days in 2019 and 624 days in 2021). In this study, three different types of gravity wave momentum fluxes (GWMFs) are considered including absolute (amplitude), westward and eastward. These target GWMFs are derived from observations of the balloons while the large-scale flow input variables are extracted from ERA5 dataset. The large-scale flow inputs include zonal wind velocity (u), meridional wind velocity (v), temperature (temp), total precipitation (tp) and logarithm of surface pressure (lnsp). These fields are retrieved for each balloon, from fields at a resolution of 1° × 1°, at the grid point closest to the balloon position. Additionally, the same input variables have been retrieved in the vicinity of a 5 by 5 horizontal square centered on the grid point closest to the balloon; in the present study, only total precipitation in this extended area around the balloon will be used. In the vertical, the ECMWF model comprises a total of 137 levels and only 4 surrounding levels of the balloons' level are considered. Figure 1 below shows the trajectories of all balloons circling Earth during the first campaign of Strateole 2 in 2019.

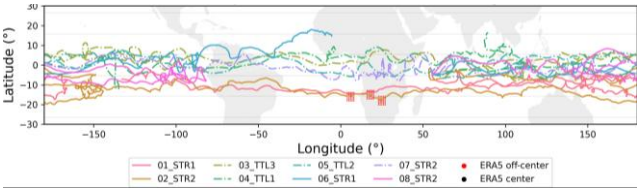


Figure 1. Balloons' trajectories in the first campaign of Strateole 2.

2.2 Methodology

Bayesian Neural Networks (BNNs) are neural network-based probabilistic models [14]. Unlike classical neural networks (NNs), where the weights and biases are deterministic values [8], BNNs treat the parameters as random variables (as illustrated in Figure 2 below).

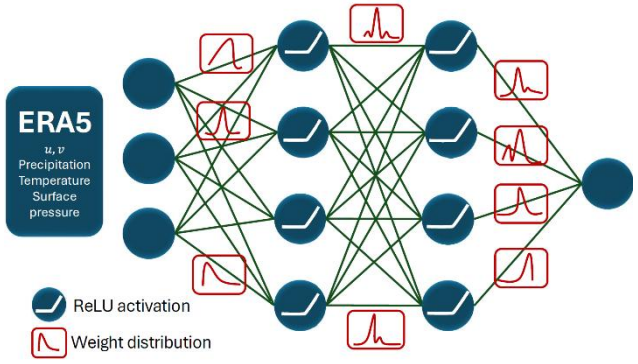


Figure 2. Diagram illustrating the framework of our BNN models.

Training a BNN involves estimating the posterior distribution of parameters W (weights and biases) given the training input-output data denoted as $p(W|D)$, starting from a prior distribution $p_0(W)$. There are two main approaches for training BNNs. First, the "Variational Inference (VI)" approach where the posterior $p(W|D)$ is approximated by some inference distribution $q(W|D)$. In this case, optimizing a BNN is equivalent to nudging the inferential posterior $q(W|D)$ towards the true posterior distribution $p(W|D)$. Second, the "Sampling Method and Monte Carlo" approach, which relies on Monte Carlo approximation and proves efficient for small networks [2]. In this study, we use Gaussian prior $p_0(W) = N(0, \sigma^2 I)$ to sample the initial parameters, and the posterior distribution is estimated using Markov Monte Carlo (MCMC) method [7]. Suppose we obtained the estimated posterior $\hat{q}(W|D)$ after the training phase, then the predicted GWMF $\hat{y}(x)$ of any ERA5 input x is computed by

$$w_i \sim \hat{q}(W|D) \text{ for } i = 1,2,\dots,N,$$
$$\hat{y}(x) = \frac{1}{N}\sum_{i=1}^{N} \mathrm{NN}(x; w_i).$$

In the above computation, the network weight w_i are first sampled from the approximated posterior distribution $\hat{q}(W\,|D)$, then then final prediction $\hat{y}(x)$ is the average over all the predictions $\mathrm{NN}(x; w_i)$ by different networks of weight w_i.

[9] demonstrated, using data from the first Strateole 2 campaign, that gravity wave momentum fluxes (GWMFs) can be partially reconstructed despite their intermittency and the limited availability of observations. However, several high-peak events remain inadequately captured due to their intermittent nature, observational uncertainties, and variability in large-scale flow dynamics. Building on this, additional data from the second campaign will allow us to quantify the uncertainties in GWMF reconstruction, which are inherent in both target variables and inputs. To achieve this, Bayesian Neural Networks will be trained within a transfer learning framework, enabling the effective reuse of pretrained models for predicting different types of the target GWMFs. Figure 3 below provides the overall process of training BNNs models for predicting different types of GWMFs using Transfer Learning framework.

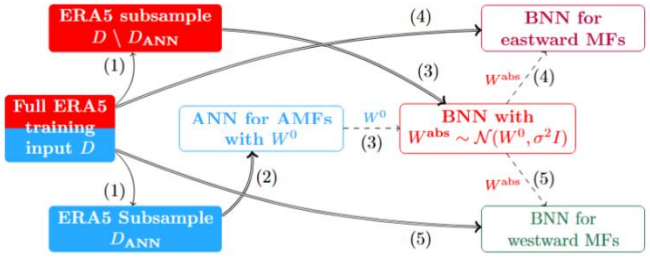


Figure 3. The diagram of Transfer Learning Process of training BNN models. In this diagram, single arrows (→) denote data splitting, double arrows (⇒) indicate that the data on the left is used to train the model on the right, and dashed arrows (⇢) indicate that the parameter on the left model is transferred to the model at the tip of the arrow. First in (1), the full training data D of the first campaign is randomly split into D_{ANN} and $D \setminus D_{ANN}$. In (2), we train a three-hidden-layer ANN on D_{ANN} to predict the absolute momentum fluxes (AMFs) of the first campaign. Then in (3), we initialize a BNN model according to the weight W_0 of the previously trained ANN, then train it using the remaining subsample $D \setminus D_{ANN}$ to predict the same target AMFs. In (4) and (5), two separate BNN models are trained using the resulting weight of BNN obtained in (3), to predict eastward and westward MFs, respectively. These last two BNN models are designed by augmenting the previous BNN model with an additional hidden layer and then training only the last two layers.

The BNN models can be trained efficiently by fine-tuning only a small subset of the transferred parameters, using a limited portion of the full training dataset. In this study, the training data consists of observations from the second balloon campaign, while the first campaign serves as the testing dataset.

3. RESULTS AND DISCUSSION

As BNN is a probabilistic method, the credible interval (CI) of the predictions, which is the equivalent of confident interval in Bayesian method, can be constructed from the posterior distribution of the parameters of the resulting networks. Moreover, Pearson correlation (Corr) between predicted values and the true GWMFs, and the following Continuous Ranked Probability Score (CRPS) will be used as the performance metric of the prediction.

$$\mathrm{CRPS}(F, y_i) = \int_{-\infty}^{\infty} (F(t) - \mathbb{1}_{\{t \geq y_i\}})^2 dt,$$

where F is the prediction range of the true target y_i and 1 denotes the indicator function, i.e.,

$$1_{\{P\}} = 1 \text{ if } P \text{ is true, and } 1_{\{P\}} = 0 \text{ otherwise.}$$

Unlike Pearson correlation, small CRPS indicates accurate prediction around the true target value. Figure 4 below highlights the quality of the predictions of our final models on balloon 1 and 2 from the testing campaign.

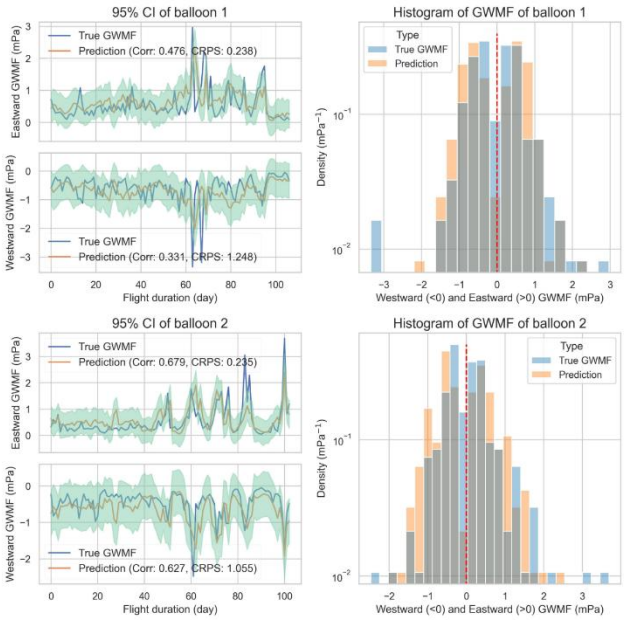


Figure 4. The predictions of eastward and westward GWMFs with their 95% CI on the 1st and 2nd balloon of the first campaign. The positive values correspond to the eastward GWMFs.

Additionally, Figure 5 below highlights the overall prediction of BNN on the whole 1st campaign (eight balloons combined). Moreover, the performance metric of our models on absolute GWMFs of each balloon is also provided in Table 1 below.

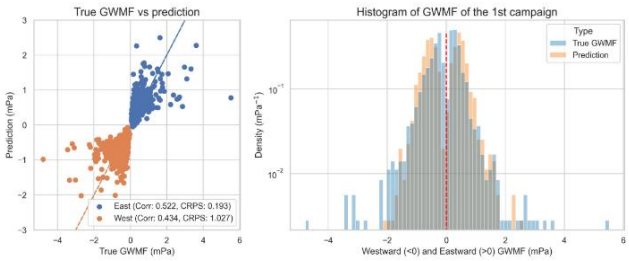


Figure 5. Predictions vs true target GWMFs on the 1st campaign.

Balloon	CRPS	Corr	Corr in [9]
1	0.441	0.436	**0.58**
2	0.378	0.700	**0.74**
3	0.431	0.474	**0.49**
4	0.317	0.464	**0.47**
5	0.357	0.380	**0.55**
6	0.292	0.633	**0.75**
7	0.381	0.468	**0.48**
8	0.361	**0.768**	0.72

Table 1. The comparison between the proposed BNN models (CRPS and Corr) and the correlation of tree-based machine learning models (random forest, extra-trees and adaboost) from [9] (Corr in [9]) in predicting absolute GWMFs from the eight balloons.

According to the correlations in Table 1 above, [9] seems to outperform the result obtained by BNN models proposed in this study. However, the numerical experiments suggested that the BNN models are better at capturing the high-peak events of GWMFs than tree-based models of [9]. Figure 6 below illustrates the predictions of tree-based methods of [9] on same balloons 1 and 2. Compared to Figure 4, BNN models perform better than tree-based models in estimating the high-peak events with a slight decrease in correlation.

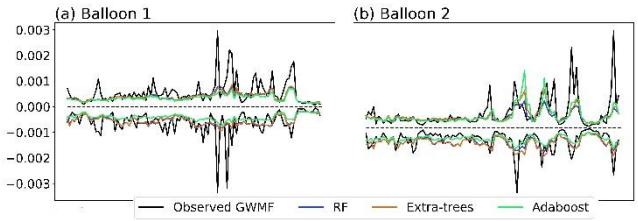


Figure 6. Predictions from tree-based ML models of [9] vs the true target GWMFs on balloon 1 and 2.

The intermittent nature of gravity wave momentum fluxes (GWMFs) presents challenges for accurate predictions; however, certain components can be effectively reconstructed. Given that uncertainties persist in both the input data and the target variables, the performance obtained is rather encouraging. Notably, the results differ across individual balloon flights, suggesting that each balloon encounters distinct atmospheric conditions during its trajectory.

4. CONCLUSIONS

In this study, we trained networks using data from the Strateole 2 second campaign (2021) to estimate three types of reference gravity wave momentum fluxes (GWMFs)—eastward, westward, and total momentum—for the first campaign conducted in 2019. Large scale flow variables including wind, precipitation, temperature are extracted from ERA5 reanalysis data at the nearest position of the balloons and are treated as the inputs for BNN models. The Transfer Learning approach enhances both the accuracy and efficiency of training BNN models, as the knowledge used to predict absolute GWMFs can be transferred and fine-tuned to estimate other types of momentum fluxes. As a result, portions of the gravity wave signal are successfully reconstructed, with correlations typically around 0.52, reaching a maximum of 0.77 for certain balloon flights. Notably, model performance varies significantly between balloons and the 95% credible intervals are relatively large indicating strong uncertainty within the predictions. Furthermore, this study highlights key limitations, including the intermittent nature of GWMFs and data scarcity, underscoring both the challenges and opportunities in refining our understanding of these atmospheric processes.

With additional observations, such as satellite images, model performance is expected to improve, along with enhanced uncertainty quantification.

ACKNOWLEDGMENTS

This research project is supported by Schmidt Science on Data wave project. This work has also received support from the ANR project BOOST3R (ANR-17-CE01-0016-01) and the French-American project Strateole 2 (CNES).

REFERENCES

[1] Alexander, M., Geller, M., McLandress, C., Polavarapu, S., Preusse, P., Sassi, F., et al. (2010). Recent developments in gravity-wave effects in climate models and the global distribution of gravity-wave momentum flux from observations and models. *Quarterly Journal of the Royal Meteorological Society*, **136**(650), 1103–1124. https://doi.org/10.1002/qj.637

[2] Bishop, C. (2006). Pattern recognition and machine learning. Springer.

[3] Butchart, N. (2022). The stratosphere: A review of the dynamics and variability. *Weather and Climate Dynamics*, **3**(4), 1237–1272. **https://doi.org/10.5194/wcd-3-1237-2022**

[4] F. Vitart & A. W. Robertson (Eds.) (2018). *Sub-seasonal to seasonal prediction*. Elsevier.

[5] Fritts, D., & Alexander, M. (2003). Gravity wave dynamics and effects in the middle atmosphere. Reviews of Geophysics, 41(1), 1003. https://doi.org/10.1029/2001RG000106

[6] Fritts, D., & Alexander, M. (2003). Gravity wave dynamics and effects in the middle atmosphere. *Reviews of Geophysics*, **41**(1), 1003. **https://doi.org/10.1029/2001RG000106**

[7] Goan, E., & Fookes, C. (2020). Bayesian neural networks: An introduction and survey. In K. L. Mengersen, P. Pudlo, & C. P. Robert (Eds.), Case studies in applied Bayesian data science: Cirm jean-morlet chair, fall 2018 (pp. 45–87). Cham: Springer International Publishing, doi: 10.1007/978-3-030-42553-1-3.

[8] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press. (http://www.deeplearningbook.org).

[9] Has, S., Plougonven, R., Fischer, A., Rani, R., Lott, F., Hertzog, A., et al. (2024). Reconstructing balloon-observed gravity wave momentum fluxes using machine learning and input from ERA5. *Journal of Geophysical Research: Atmospheres*, 129, e2023JD040281. **https://doi.org/10.1029/2023JD040281**

[10] Hines, C. O. (1960), Internal atmospheric gravity waves at ionospheric heights, *Can. J. Phys.*, **38**, 1441–1481.

[11] Kim, Y.-J., Eckermann, S., & Chun, H.-Y. (2003). An overview of the past, present and future of gravity-wave drag parametrization for numerical climate and weather prediction models. *Atmosphere-Ocean*, **41**(1), 65–98. **https://doi.org/10.3137/ao.410105**

[12] Kremser, S., Thomason, L. W., von Hobe, M., Hermann, M., Deshler, T., Timmreck, C., et al. (2016). Stratospheric aerosol—Observations, processes, and impact on climate. Reviews of Geophysics, 54(2), 278–335. https://doi.org/10.1002/2015rg000511

[13] Solomon, S., Rosenlof, K., Portmann, R., Daniel, J., Davis, S., Sanford, T., & Plattner, G.-K. (2010). Contributions of stratospheric water vapor to decadal changes in the rate of global warming. *Science*, **327**(5970), 1219–1223. **https://doi.org/10.1126/science.118248**

[14] Tishby, Levin, & Solla. (1989). Consistent inference of probabilities in layered networks: predictions and generalizations. In International 1989 joint conference on neural networks (p. 403-409 vol.2). doi: 10.1109/IJCNN.1989.118274.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Comparing Time Series, Machine Learning, and Deep Learning Models for Paddy Rice Price Forecasting in Battambang Province, Cambodia

Phirum Leng*, Sokkhey Phauk, Malai Nhim, Chantara Tann

Applied Mathematics and Statistics, Institute of Technology of Cambodia, Phnom Penh, Cambodia

*Corresponding author: lphirum4@gmail.com

Abstract: *In Cambodia, paddy rice is a staple food and a key income source, with Battambang Province being one of the leading rice-producing regions. Accurate paddy rice price forecasts are essential for farmers, traders, policymakers, and financial institutions to make informed decisions, mitigate risks, and stabilize markets. Several factors influence price fluctuations, including domestic production, global market trends, input costs, and macroeconomic variables like crude oil prices and exchange rates. Developing forecasting models that capture these dynamics is crucial for agricultural sustainability and economic resilience. This study compares traditional time series models (ARIMAX) with machine learning (XGBoost) and deep learning (LSTM) models to evaluate forecasting accuracy. The results show that XGBoost outperforms both ARIMAX and LSTM, achieving the lowest RMSE (48.78), MAE (40.41), and MAPE (2.79%). In contrast, ARIMAX performs moderately, while LSTM exhibits higher error rates. This study offers stakeholders a data-driven approach to agricultural decision-making, supports sustainable farming practices, and enhances market predictability, contributing to economic stability in Cambodia's rice industry. The findings also inform policy recommendations to improve trade efficiency and food security.*

Keywords: Rice Price Forecasting, Time Series Models, Machine Learning, Deep Learning, ARIMAX, XGBoost, LSTM, Market Stability, Forecasting Accuracy

1. INTRODUCTION

As a staple crop and a significant export, rice is essential to Cambodia's economy, and Battambang is a major center for production (FAO, 2024). However, changes in rice prices pose difficulties for farmers, traders, and legislators, affecting economic stability and food security (Sokchea, H., Sovannarith, S., & Phalla, C., 2023). While machine learning models, such as LSTM and XGBoost, have shown promise in increasing accuracy, conventional forecasting techniques like ARIMA struggle to capture intricate pricing patterns. However, few studies have combined the two methods to improve forecasts (Nichani, M., Khairuddin, A. S., & Rahman, A. A., 2024). This study evaluates XGBoost, LSTM, and time series models (ARIMA) for predicting rice prices in Battambang. It aims to determine the most suitable strategy by examining key variables, including crude oil prices and the USD exchange rate. The results will help with economic planning, market stability, and decision-making in the rice industry in Cambodia. Since the agriculture industry faces both market volatility and external shocks, accurate price forecasting becomes a crucial tool for ensuring resilience and informed policymaking. By contrasting machine learning and deep learning techniques with traditional models, this study seeks to improve forecast accuracy while illuminating the seasonal and nonlinear patterns of rice prices (MacLachlan, M. J., Chelius, C. A., & Short, G., 2019). Finally, the study applied a comparative analysis of the models to support supply chain and rice stock management, focusing on their advantages and disadvantages.

Additionally, this study supports the Global Sustainable Development Agenda, particularly the SDGs related to hunger, poverty, and economic growth. By providing accurate price forecasts, the study aims to contribute to SDG 1 (No Poverty), SDG 2 (Zero Hunger), and SDG 8 (Decent Work and Economic Growth) (UNESCAP, 2021), helping policymakers, stakeholders, and farmers make informed decisions that promote long-term agricultural sustainability and economic resilience.

2. METHODOLOGY

This section presents an overview of the steps undertaken in this project, which aims to compare the models and

forecasting. The process involves data acquisition, algorithm modeling, and performance evaluation.

2.1. Dataset and Description

The dataset used in this study comprised historical rice prices from January 2007 to June 2022, totaling 183 monthly records obtained from the World Food Program's website. Additional factors, such as crude oil prices and USD exchange rates from World Bank Commodity Prices, were incorporated to capture macroeconomic influences on paddy rice prices.

2.2. Autoregressive Integrated Moving Average with Exogenous Regressors

The ARIMAX model extends the classical ARIMA model by incorporating external (exogenous) variables that may influence the target variable (Shilpa, S. & Sheshadri, H. S., 2019). The ARIMAX model is useful when external factors, such as oil prices and exchange rates, are expected to impact the variable of interest, which is the rice price in this study. The general form of the ARIMAX model is:

$$y_t = \alpha + \sum_{i=1}^{p} \emptyset_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \epsilon_{t-j} + \sum_{k=1}^{K} \beta_k X_{k,t} + \varepsilon_t \quad (1)$$

Where:

y_t = Target variable at time (t)
α = Intercept (constant term)
$\emptyset_i$ = Autoregressive coefficients (AR part)
θ_j = Moving average coefficients (MA part)
ε_t = Error terms (white noise)
β_k = Coefficients for exogenous variables
$X_{k,t}$ = Exogenous variable(s) at time t
ρ, d, q =Orders of the AR, differencing, and MA, respectively
K = Number of exogenous variables

2.3. Extreme Gradient Boosting

XGBoost is a machine learning algorithm that captures non-linear relationships and complex feature interactions. It minimizes a regularized objective function to improve prediction accuracy and reduce overfitting (Chen, T. & Guestrin, C., 2016). The objective function in XGBoost is defined as:

$$L(\emptyset) = \sum_{i=1}^{n} l(y_i, \hat{y}_i(t)) + \sum_{k=1}^{t} \sigma(f_k) \quad (2)$$

Where:

$L(\emptyset)$ = loss function to be minimized
y_i = Actual value of the target variable at time i
$\hat{y}_i$(t) = Predicted value of the target variable at time i
$l(y_i, \hat{y}_i(t))$= loss function
n = Number of data points in the time series
T = Number of regularization terms
(f_k) = The set of features or model parameters for the regularization term
$\sigma(f_k)$ = Regularization term for feature selection or complexity control (e.g., L1/L2 norm)

2.4. Long Short-Term Memory

LSTM is a specialized type of Recurrent Neural Network (RNN) designed to learn long-term dependencies in sequential data. Unlike traditional RNNs, which suffer from vanishing or exploding gradients when modeling long sequences, LSTM addresses this limitation by introducing gated mechanisms that regulate the flow of information (Ghojogh, B. & Ghodsi, A., 2023), which is expressed by the formula below:

$$f_t = \sigma(W_{fh} h_{t-1} + W_{fx} x_t + b_f) \quad (3)$$

$$i_t = \sigma(W_{ih} h_{t-1} + W_{ix} x_t + b_i) \quad (4)$$

$$\tilde{c}_t = tanh(W_{\tilde{c}_t h} h_{t-1} + W_{\tilde{c}_t x} x_t + b_{\tilde{c}}) \quad (5)$$

$$c_t = f_t \cdot c_{t-1} + i_t \cdot \tilde{c}_t \quad (6)$$

$$o_t = \sigma(W_{oh} h_{t-1} + W_{ox} x_t + b_o) \quad (7)$$

$$h_t = o_t \cdot tanh(c_t) \quad (8)$$

Where:

f_t = Forget gate – filters out irrelevant past information
i_t = Input gate – controls which new data is added to memory
$\tilde{c}_t$ = Candidate state – proposed values for memory update
c_t = Cell state – updated memory combining past and new information
o_t = Output gate – selects which memory to output
h_t = Hidden state – final output of the LSTM at time t

These equations represent two components of the loss function: (1) the sum of the losses between the actual and predicted values, (2) the sum of regularization terms, which help prevent overfitting or enforce specific constraints.

2.5. Parameter optimization and Evaluation metrics

To fine-tune the ARIMAX model, the values of p, d, and q were selected based on the **Akaike Information Criterion (AIC)**, which measures the trade-off between model fit and complexity. The data was first made stationary using differencing, and the ideal value of d was then found using the Augmented Dickey-Fuller (ADF) test. The model with the

lowest AIC value was then picked to determine the values of p and q. The AIC is calculated as:

$$AIC = 2k - \ln(L) \quad (9)$$

Where:

k = Number of parameters estimated
L = Maximum likelihood of the model

After fitting the ARIMAX models, their predictive performance was evaluated using a validation dataset (index 146–183). Three widely used error metrics were applied:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (10)$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad (11)$$

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right| \times 100 \quad (12)$$

Where:

n = Number of observations
A_t, y_i = Actual values
$F_t, \hat{y}_i$ = Forecasted (predicted) values

These metrics quantify the average deviation between predicted and actual values, with lower scores indicating better model performance.

3. IMPLEMENTATION

3.1. Data Preparation and Preprocessing

The work begins with data preparation and preprocessing to ensure its suitability for analysis. Data profiling is performed to assess quality and detect anomalies, followed by data cleaning to address errors, inconsistencies, and missing or duplicate values. The Date column is transformed into a datetime format for more informative analysis. Visualizations are used to understand the data distribution and characteristics.

The significance of this processing is that stable values are replaced for periods longer than 6 months (Apr 2017 to Oct 2017) and (Aug 2020 to Mar 2021). The figure below shows the data after the stable values were replaced.

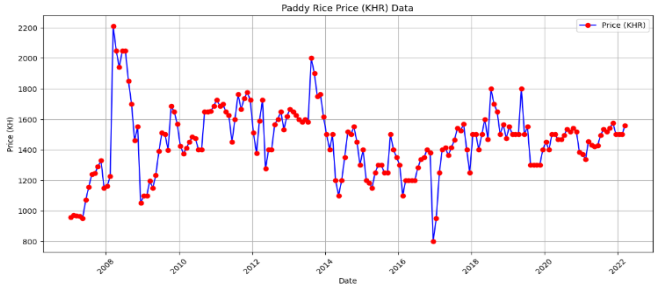


Fig.1. Time plot of Paddy Rice Price (KHR)

3.2. Correlation Matrix

A statistical metric of correlation measures how strongly and in which direction two variables have a linear relationship. The plot below indicates the relationship of variables in this research.

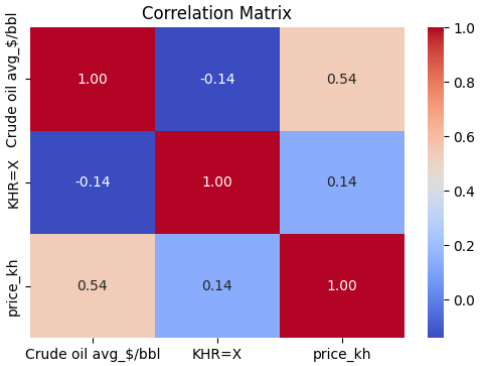


Fig.2. Correlation Matrix of variables

This correlation matrix illustrates the relationships between crude oil price, exchange rate, and rice prices in Battambang, Cambodia. Crude oil shows a moderate positive correlation (0.54) with rice prices, suggesting that rising oil costs may increase rice prices. The exchange rate weakly correlates with crude oil (-0.14) and rice prices (0.14), indicating limited influence. The color gradient reflects the strength and direction of each relationship. Overall, crude oil prices appear to have the most significant impact on rice prices.

3.3. Feature Scaling

To ensure a proper contribution during feature selection, feature scaling standardizes the range of independent variables, and to stop features with higher values from dominating, all features are scaled to a 0-1 range using MinMaxScaler().

3.4. Data Splitting

To assess model performance and prevent overfitting, the data is divided into training and testing sets. The model is trained on the first 146 instances (index 0-145) and evaluated on the remaining 38 unseen instances (index 146-183) to measure its accuracy.

4. RESULTS AND DISCUSSION

4.1. ARIMAX model

The first step is to test for stationarity using the ADF and KPSS tests. Among the three variables, only Crude Oil Price is non-stationary, with p-values of 0.080 (ADF) and 0.022 (KPSS), indicating the need for transformation or differencing. Then, after applying the 1st differencing, the crude oil appears to be stable.

The next step was to identify the p and q using the autocorrelation and partial autocorrelation functions (ACF and PACF) (Holmes, E., 2021) as shown in Fig.3 and 4. The order of p and q with the minimum AIC (Akaike Information Criterion) and BIC (Bayesian Information Criterion) was selected, and the models were ARIMAX (1,0,0), ARIMAX (1, 1, 1), and ARIMAX (1, 1, 2). Among the above models, ARIMAX (1, 1, 2) is the most suitable model based lowest AIC and BIC.

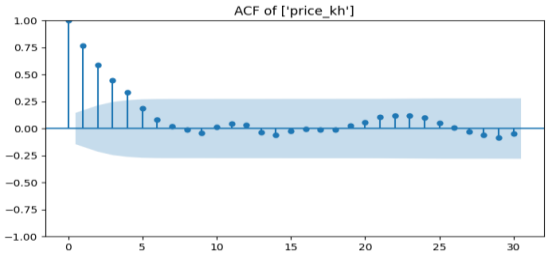


Fig.3. ACF plot of paddy rice price

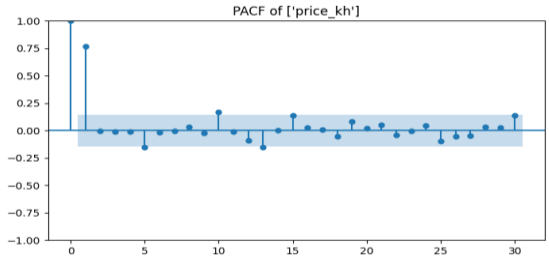


Fig.4. PACF plot of paddy rice price

Refer to the Fig.3 and 4, the PACF cuts off after lag 1, indicating an AR (1) component, and the ACF plot displayed a progressive decrease. The choice of ARIMAX (1,1,2) was influenced by this observation as well as the minimum AIC/BIC.

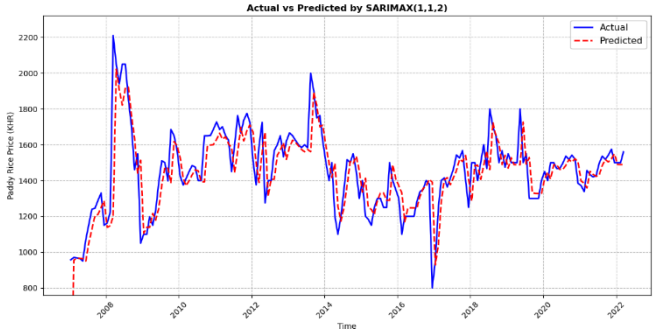


Fig.5. Actual Vs Predicted by ARIMAX (1,1,2)

4.2. XGBoost model

The XGBoost model was trained using a rolling forecast approach with a one-month step size to predict the 12 months following the initial 183-month dataset. Lag features from the previous 12 months were used to capture temporal dependencies. Hyperparameters were tuned using RandomizedSearchCV with TimeSeriesSplit for cross-validation. At each rolling step, the model was trained on all available data up to the prediction point, and the resulting forecasts were stored and plotted to visualize performance over the forecast horizon.

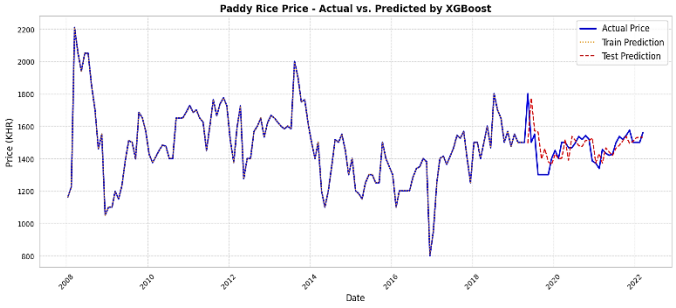


Fig.6. Actual vs Predicted by XGBoost

4.3. LSTM model

The LSTM model was implemented to capture long-term dependencies and sequential patterns in rice price movements. All features were normalized using MinMaxScaler. A sequence length of 12 months was selected, allowing the model to learn from one year of historical data to predict rice prices for the following month.

Input sequences were generated using a sliding window approach, where each sequence consisted of 12 months of past data (crude oil prices, exchange rates, and rice prices), and the target was the rice price in the 13th month. The dataset was split into 80% training and 20% testing sets.

After training, predictions were made on the test set and inverse-transformed to retrieve actual paddy rice price values. Model performance was evaluated using standard metrics: MAE, RMSE, and MAPE.

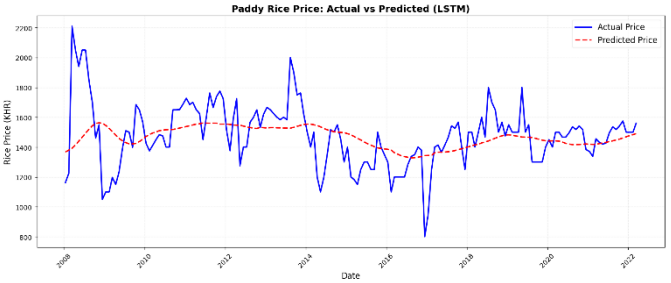


Fig.7. Actual vs Predicted Rice Prices using LSTM

The figure above illustrates the comparison between the actual and predicted rice prices using LSTM. The model successfully followed the overall trend but showed some lags, resulting in larger error gaps in response to rapid price changes.

4.4. Comparison of the models

The performance of the three forecasting models: ARIMAX, XGBoost, and LSTM, was evaluated using standard accuracy metrics, which are summarized in the table below:

Table 1. Performance metrics for the models:

Model	RMSE	MAE	MAPE
ARIMAX (1,1,2)	55.62	45.89	3.85%
XGBoost	48.78	40.41	2.79%
LSTM	98.59	76.15	5.16%

Among the models, XGBoost demonstrated the best performance with the lowest RMSE, MAE, and MAPE, indicating superior accuracy in forecasting paddy rice prices in Battambang Province. The ARIMAX model produced moderate results, reflecting its effectiveness in capturing linear patterns but its limitation in modeling more complex structures. The LSTM model, while suitable for capturing long-term dependencies, yielded the highest error rates in this study, potentially due to the limited dataset size or suboptimal model tuning.

4.5. The paddy rice price predicted using XGBoost

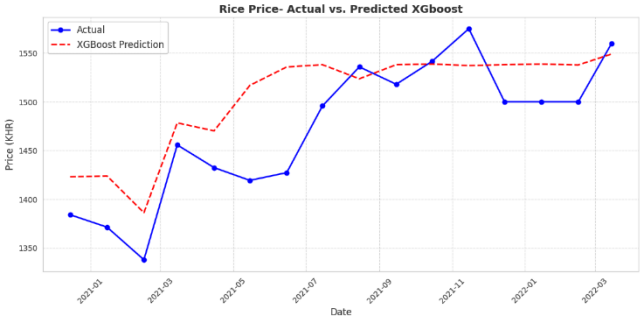


Fig.8. Actual vs. Predicted Rice Prices by XGBoost

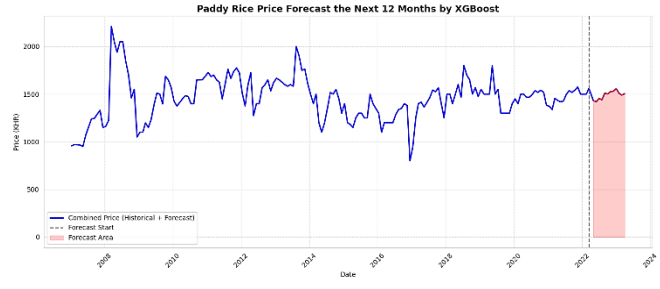


Fig.9. Forecast Paddy Rice Prices by XGBoost

As illustrated in Figure 8, the XGBoost model closely aligns with actual rice price movements, successfully tracking seasonal fluctuations and short-term variations. The deviations between predicted and actual values are minimal, especially during the major harvesting seasons. These findings underscore the capability of advanced machine learning models like XGBoost to capture complex price dynamics that traditional statistical models may overlook.

4.6. Relationship to SDG Implications

The forecast models and Battambang's rice price data show trends that have a direct impact on both income stability (SDG 1) and food security (SDG 2). This study offers useful insights that might assist policymakers in developing methods that stabilize rice prices and provide fair access to food resources throughout the region by precisely forecasting price variations. These strategies should be taken into consideration with additional research and exploration (UNDP, 2021).

5. CONCLUSIONS

This study evaluates ARIMAX, XGBoost, and LSTM models for forecasting paddy rice prices in Battambang, Cambodia. The results highlight that XGBoost achieves the highest forecasting accuracy with the lowest RMSE, MAE, and MAPE. While ARIMAX provides moderate accuracy, LSTM struggles with higher error rates. XGBoost reduced error by 12.3% to 27.5% compared to ARIMAX and 16.2% to 30.6% compared to LSTM across RMSE, MAE, and MAPE metrics. These results highlight XGBoost's superior accuracy and robustness in capturing complex price dynamics.

These findings emphasize the effectiveness of machine learning in capturing complex price dynamics and suggest that policymakers and stakeholders should consider data-driven forecasting methods for agricultural planning.

While the current work focuses on comparative forecasting performance, future research should expand these models by incorporating additional socioeconomic and environmental factors, exploring hybrid models combining time series analysis with deep learning for improved accuracy, and advancing sustainable agricultural development in Cambodia and beyond.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the World Food Program (WFP) for providing access to the valuable rice price data. We also acknowledge the ongoing support, consultation, and thorough review from our

academic advisors and co-advisors in the Master's Program in Data Science (M-DAS), Department of Applied Mathematics and Statistics (AMS), at the Institute of Technology of Cambodia (ITC). Their expertise and dedication significantly contributed to the success of this study.

REFERENCES

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785–794). ACM (Association for Computing Machinery).

FAO. (2024). *GIEWS Country Brief: The Kingdom of Cambodia.* Cambodia: Food and Agriculture Organization of the United Nations (FAO), Global Information and Early Warning System (GIEWS).

Ghojogh, B., & Ghodsi, A. (2023, April). *Recurrent neural networks and long short-term memory networks: Tutorial and survey.* Retrieved from arXiv: https://arxiv.org/abs/2304.11461

Holmes, E. (2021). *Fitting and Selecting ARIMA Models.*

MacLachlan, M. J., Chelius, C. A., & Short, G. (2019). *Time-Series Methods for Forecasting and Modeling Uncertainty in the Food Price Outlook.* United States Department of Agriculture (USDA), Economic Research Service. Retrieved from https://ers.usda.gov/publications/pub-details/?pubid=104490

Nichani, M., Khairuddin, A. S., & Rahman, A. A. (2024). Optimizing financial time series predictions with hybrid ARIMA, LSTM, and XGBoost models. *Studies in Engineering and Exact Sciences*. Retrieved from https://doi.org/10.54021/seesv5n2-582

Shilpa, S., & Sheshadri, H. S. (2019). ARIMAX model for short-term electrical load forecasting. *International Journal of Recent Technology and Engineering*. Retrieved from https://doi.org/10.35940/ijrte.D7950.118419

Sokchea, H., Sovannarith, S., & Phalla, C. (2023). Factors influencing paddy producers' profitability and market sales in Battambang province, Cambodia. *Cogent Food & Agriculture*. Retrieved from https://doi.org/10.1080/23311932.2023.2193311

UNDP. (2021). *Sustainable Development Goals: Progress and Insights in Cambodia.* United Nations Development Programme (UNDP). Retrieved from https://www.kh.undp.org/

UNESCAP. (2021). *Goal Profile: SDG 8 – Decent Work and Economic Growth.* United Nations Economic and Social Commission for Asia and the Pacific.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

The Forecasting for Cambodia's Rice Production Using Multiple Linear Regression and Tree-Based Ensemble Learning Methods

Sithach Thong [1*], Sokkhey Phauk [2], Sothea Has [2] Ponna Phok [2]

[1] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: sithach_thong@gsc.itc.edu.kh*

Abstract: *Cambodia's rice sector has undergone a transformative ative shift in recent decades, evolving from a net importer to a vital pillar of the national economy. This transformation has been fueled by the expansion of cultivated land, improved farming techniques, and technological adaptation. However, despite this progress, the sector remains vulnerable to key challenges such as erratic climate patterns, inefficient resource management, and limited integration of data-driven predictive analytics. Addressing these gaps is crucial for enhancing national food security, guiding policy development, and strengthening Cambodia's position in regional rice markets. This study presents a novel application of advance machine learning models, especially tree-based ensemble methods alongside traditional Multiple Linear Regression to forecast Cambodia's rice production. The study exclusively utilizes a newly dataset sourced directly from Cambodia, integrating data from the Food and Agriculture Organization (FAO), and the World Bank Database. The dataset incorporates variables relevant to Cambodia's unique agricultural context, such as precipitation, cultivated land area, and local surface temperature, providing a tailored and context-aware forecasting framework. The results indicate that Random Forest outperformed all other models, achieving a highest R^2 score (93.34%), with the lowest MAE (14.66 %) and MAPE (0.09%). This research marks one of the first comprehensive applications of ensemble learning to Cambodia's rice production forecasting, using specific dataset and the development of a predictive system designed to inform smart agriculture practices, data-driven approach to sustainable agricultural planning, supporting Cambodia's long-term food security and economic stability of the country.*

Keywords: Rice Production Forecasting; Multiple Linear Regression; Ensemble Learning Methods; Predictive Analytics; Cambodia Agriculture; Sustainable Development Goals; Machine Learning

1. INTRODUCTION

Rice has been recognized as the nation's most significant agricultural product for Cambodia. Its production plays a pivotal role in economy and food security for the country. Rice serves as a source of livelihood for millions of farmers, contributing around 60% to the agricultural GDP, and represent as a major export commodity [1]. Between 2004 and 2012, Cambodia's agricultural gross production grew by 8.7%, largely due to increased rice output, which significantly reduced the national poverty rate from 53% in 2004 to 18% in 2012. The Roylal Government of Cambodia (RGC) has developed a comprehensive policy to enhance rice production and expand future export capacity [2], [3]. This policy underlines the government's commitment to modernizing the sector, ensuring sustainability, enhance food security, strengthening its global compititveness, and supports Sustainable Development Goals related to food security (SDG2).

Over the last decades, many studies have explored the use of statistical and machine learning techniques to predict rice production or yield in various regions. Traditional methods such as time series analysis and linear regression have been commonly applied due to their simplicity interpretability [4]. More recently, machine learning approaches like artificial neural networks, support vector machine, and decision tree-based algorithms have gained attention for their ability to capture nonlinear relatioships and improve predictive accuracy [5], [6], [7]. Additionally, other recent studies also tried to compare the performance between multiple linear regression with advance models such as artificial neural networks [8].

Despite the increasing adoption of predictive analytics globally, Cambodia has lacked context-spcific, data-driven tools, and limited research of predictive analytics applied in

[1*] Corresponding author: Sithach Thong
E-mail: sithach_thong@gsc.itc.edu.kh ; Tel: +855-61575529

agricultural fields. This study addresses these gaps by introducing the use of localized climate and agricultural data with the state-of-the-art ensemble models for the first time to forecast rice production, aiming to enhance policy development for climate-resilient agriculture (SDG 13), and contribute to innovation and sustainable infrastructure in the agricultural sector (SDG 9).

2. METHODOLOGY

2.1 Dataset

Dataset from 1961-2022 were collected from Food and Agricultural Organization (FAO) and the World Bank, supplemented with Cambodia-specific datasets. There are 248 existing observations were gathered, incorporating climate variables, harvested area, and production volumes, forming one of the first localized datasets for machine learning applications in Cambodia's agriculture.

Table 1. Statistical summary of the variables

Variable	Mean	Std	Min	Max
Production (Ton)	4396868.10	3347409.38	538000.00	12207000.00
Harvested Area (Ha)	2074386.60	726660.87	555000.00	3498000.00
Mean Precipitation (mm)	1846.01	156.02	1477.24	2215.91
Mean Surface Temperature (°C)	27.13	0.404	26.3	27.99

2.2 Data Preprocessing

Rice production was used as the target variable, with several inputs variables including harvested area, mean precipitation, and mean surface temperature.

Before building the models and to ensuring data's usability, data preprocessing was implemented in this study. Normalization and transformation were applied to prepare the data and improve its characteristics, make it suitable for modeling. This study employed different transformation techniques such as logarithmetic and square transformation.

$$y' = log\ (x) \qquad \text{(Eq. 1)}$$

Where:

y' = the transformed values after applying natural logarithm
x = the original vavues before applying natural logarithm

$$y' = x^2 \qquad \text{(Eq. 2)}$$

Where:

y' = the transformed values after applying square transformation
x = the original values before applying square transformation

2.3 Machine Learning Models

Multiple linear regression and tree-based ensemble learning (Gradient Boosting, XGBoost, and Random Forest) were implemented. Each model's theoretical foundation and application for regression tasks were outlined.

2.3.1 Multiple Linear Regression

Linear regression ia a statistical model, also known as a traditional approach that can solve the task by considering the historical dataset or parameter. There are two types of linear regression. When the analysis use only one independent variable to predict the dependent variable, it is considered as the simple linear regression (SLR). Otherwise, when the analysis use two or more independent variables, it is considered as the multiple linear regression (MLR). The mathematical representation of the MLR is written as ;

$$\hat{y} = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n \qquad \text{(Eq. 3)}$$

Where:

$\hat{y}$ = dependent variable
β_0 = constant or $y_{intercept}$
$\beta_1, \beta_2, \ldots, \beta_n$ = slope coefficient
$X_1, X_2, \ldots, X_3$ = dependent variable

Multiple linear regreesion is capable of analyzing one variable based on others various variables effectively. The slope shows the magnitude and direction of relationship between an independent variable and the dependent variable. It tells how much the dependent variable Y is expected to change when a specific independent variable change in one unit. Similarly, correlation coefficients are used to measure the degree of linear association between two variables, they show how strongly and in what direction the variables move together [9]. The correlation coefficient formula is written as;

$$r = \frac{\Sigma(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\Sigma(x_i - \bar{x})^2 \Sigma(y_i - \bar{y})^2}} \qquad \text{(Eq. 4)}$$

The coefficient measures how well data can be represented in the regression line and can defines strength and direction of relationship between the dependent and independent variables.

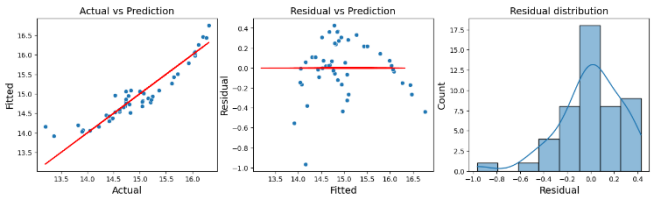


Fig. 1. MLR model diagnostics

2.3.2 Gradient Boosting

Gradient boosting is one of tree-based learning model, widely used for regression and classification work [10]. Gradient boosting regression builds multiple weak learners sequentially, optimizing model performance by minimizing a given loss function, such as mean squred error (MSE). At each step, the model tries to minimize the error (rediduals) of the previous model by fitting a new model to the gradient and improve model accuracy by learning from previous mistakes. Let y be the true value, and we want to predict y using a function $F(x)$. The initial prediction model is written as;

$$F_0(x) = argmin\gamma \sum_{i=1}^{n} L(y_i, \gamma) \qquad \text{(Eq. 5)}$$

Where:

$F_0(x) =$ the initial model
$L(y_i, \gamma) =$ the loss function (mean squared error)
$y_i =$ the actual target value

Compute the negative gradient (residuals) for each iteration $m = 1$ to M;

$$r_i^m = -\left[\frac{\partial L(y_i, F(x_i))}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} \qquad \text{(Eq. 6)}$$

Fit a weak learner function $h_m(X)$, and find the best gradient decent step-size γ_m

$$\gamma_m = argmin_{\{\gamma\}} \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)) \qquad \text{(Eq. 7)}$$

Update the model;

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \qquad \text{(Eq. 8)}$$

Where:

$F_m(x) =$ the updated model (output)
$h_m(x) =$ the weak learner (decision tree)
$\gamma_m =$ the learning rate or step size $(0 < \gamma_m \leq 1)$

The final model after M iterations can be written as;

$$F_M(x) = \sum_{m=1}^{M} \gamma_m h_m(x) \qquad \text{(Eq. 9)}$$

2.3.3 Extreme Gradient Boosting

XGBoost (Extreme Gradient Boosting) is an optimized implementation of Gradient Boosting algorithm, initially developed by Tianqi Chen as part of Distributed Machine Learning Community. Boosting tree algorithms based on the decision tree, which is known as classification and regression (CART). For regression tasks, CART divides the dataset into two subsets at each level according to the boundary for one variable until reaching the tree's maximun depth set by users [11]. It aims at extremely fast, scalable, and portable. XGBoost is used for supervised learning problems, where we use the training data (with multiple features) x_i to predict a target variable y_i. XGBoost extends traditional gradient boosting by including regularization elements in the objective function, improve generalization and prevents overfitting. The mathematical equation of the model is defined as;

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \qquad \text{(Eq. 10)}$$

Where:

$\hat{y}_i =$ the final predicted value for i^{th} data point
$K =$ the number of trees in the ensemble
$f_k(x_i) =$ the prediction of K^{th} tree for the i^{th} data point

The objective function in XGBoost consists of two parts, a loss function and a regularization term. The loss function measures how well the model fits the data and the regularization term simplify complex trees. The general form of loss function is written as;

$$obj(\Theta) = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \qquad \text{(Eq. 11)}$$

Where:

$l(y_i, \hat{y}_i) =$ the loss function which computes the difference between the true value
$\Omega(f_k) =$ the regularization term which discourages overly complex trees

By minimizing the model iteratively (instead of fitting the model all at once), starting with an initial prediction $\hat{y}_i^{\,0} = 0$ and add a new tree in each step to improve the model, then the updated predictions after adding the i^{th} tree can be written as;

$$\hat{y}_i^{\,t} = \hat{y}_i^{\,t-1} + f_t(x_i) \qquad \text{(Eq. 12)}$$

Where:

$\hat{y}_i^{\,t-1} =$ the prediction from the previous iteration
$f_t(x_i) =$ the prediction of the i^{th} tree for the i^{th} data point

The regularization term simplifies complex trees by penalizing the number of leaves in the tree and the size of leaf. It can be defined as;

$$\Omega(f_t) = \gamma T + \frac{1}{2}\lambda \sum_{j=1}^{T} w_j^2 \qquad \text{(Eq. 13)}$$

Where:

$T =$ the number of leaves in the tree
$\gamma =$ the regularization parameter that controls the complexity of the tree

$\lambda =$ the parameter that penalizes the squared weight of the leaves w_j

After splitting nodes in the trees, the informative gain for every possible split is computed. By calculating the information gain for every possible split at each node, XGBoost selects the split that results in the largest gain which effectively reduce the errors and improves the model's performance. The equation for information gain calculation is denoted as;

$$Gain = \frac{1}{2}\left[\frac{G_L^2}{H_L+\lambda} + \frac{G_R^2}{H_R+\lambda} + \frac{(G_L+G_R)^2}{H_L+H_R+\lambda}\right] - \gamma \qquad \text{(Eq. 14)}$$

Where:

$G_L, G_R =$ the sums of gradients in the left and right child nodes
$H_L, H_R =$ the sums of Hessians in the left and right child nodes

2.3.4 Random Forest

Random forests are a supervised learning algorithm, also known as an ensemble learning method that is widely used in regression and classification problems. A random forest is essentially a collection of decision trees, where each tree is slightly different from the others [12]. It provides an improvement over bagged trees by way of a small tweak that decorrelates the trees. The number of decision trees are built on bootstrapped training samples and when building these decision trees, each time a split in a tree is considered, a random sample of m predictors is chosen as split candidates from the full set of p predictors. A random forest built using quantitative variables, give output values also in quantitative values. It is assumed training data is independently selected from the original dataset [13]. A random forest uses Bootstrap Aggregation (Bagging) of a given dataset D with N observations.

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \qquad \text{(Eq. 15)}$$

Where:
$x_i =$ the feature vector of the i^{th} sample
$y_i =$ the target variable

By generating B bootstrap samples D_b by randomly selecting N with replacement $D_1, D_2, D_3, \dots, D_B$. Each dataset D_b trains an individual regression tree denoted by $f_b(x)$. Each decision tree is trained independently on a bootstrap sample D_b. The trees are bult using the CART (Classification and regression Trees) algorithm. When building these decision trees, each time a split in a tree is considered, a random sample of m predictors is chosen as split candidates from the full set of p predictors. Each node in a decision tree split based on a feature x_j that minimizes a loss function. For regression, the best split is chosen by minimizing the Mean Squared Error (MSE).

$$MSE = \frac{1}{n}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \qquad \text{(Eq. 16)}$$

For a given feature x_j and split point s, we partition the data into two groups,

- Left node: $R_L = \{(x_i, y_i) | x_{ij} \le s\}$
- Right node: $R_R = \{(x_i, y_i) | x_{ij} > s\}$

The optimal split is determined by minimizing the weighted sum of MSEs;

$$argmin_{j,s}\left(\frac{|R_L|}{|R|}\sum_{i\epsilon R_L}(y_i - \bar{y}_L)^2 + \frac{|R_R|}{|R|}\sum_{i\epsilon R_R}(y_i - \bar{y}_R)^2\right) \qquad \text{(Eq. 17)}$$

$\bar{y}_L$ and $\bar{y}_R$ are the mean target values in the left and right nodes, respectively. Once all B decision trees are trained, each tree makes an individual prediction for a new input x as;

$$\hat{y}_b = f_b(x), for\ b = 1,2,3, \dots, B \qquad \text{(Eq. 18)}$$

Therefore, the final prediction is the average of all tree inputs. The equation is written as;

$$\hat{y} = \frac{1}{B}\sum_{b=1}^{B} f_b(x) \qquad \text{(Eq. 19)}$$

Instead of using all features at each node, the random forest is built using $m = \sqrt{p}$ by selecting a random subset of features (known as m features out of total p features).

2.4 Evaluation Metrics

In order to determine the forecast errors of the yearly rice production, different between the actual values and those predicted by the models were used. Determining the accuracy of the forecasts was indicated by calculating the values of the forecasting properties of the models. Coefficient of determination or known as R^2 score is the most used metric to measure the variance in the dependent variable that explained by the model, the higher R^2 score the better model performance [14].

$$R^2 = 1 - \frac{\Sigma(y_i - \hat{y}_i)^2}{\Sigma(y_i - \bar{y})^2} \qquad \text{(Eq. 20)}$$

Mean absolute error (MAE) is regularly employed in model evaluation studies [15]. The metric measure the average magnitude of errors between predicted and actual values, providing a straightforward interpretation of prediction accuracy.

$$MAE = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \qquad \text{(Eq. 21)}$$

Mean absolute percentage error (MAPE) is a well known metric has been used for regression model [16]. It measures

the accuracy of forecasting and regression models by expressing prediction errors as a percentage, offering an intuitive measure of how cose forecasts are to actual outcomes.

$$MAPE = \frac{100}{m}\sum_{i=1}^{m}\left|\frac{(y_i-\hat{y}_i}{y_i}\right| \quad \text{(Eq. 22)}$$

3. RESULTS AND DISCUSSION

Random Forest outperformed all others, achieving the highest R^2 score (0.9334) and the lowest error metrics MAE(0.1466) and MAPE (0.0009), indicating the best prediction accuracy. XGBoost and Gradient Boosting also showed good performance, with XGBoost slightly outperforming Gradient Boosting in both MAE and MAPE. Although MLR had a relatively high R^2 score (0.9185), it recorded the highest MAE (0.2086) and MAPE (0.0137), making it the least accurate overall.

Table 2. Models Performance

Models	R^2_{score}	MAE	MAPE
MLR	0.9185	0.2086	0.0137
Gradient Boosting	0.8926	0.1736	0.0118
XGBoost	0.9015	0.1694	0.0115
Random Forest	0.9334	0.1466	0.0009

This study makes a breakthrough in Cambodia's agricultural forecasting by combining machine learning models with context-specific data. The findings support integration of ensemble models in agricultural predictive analysis.

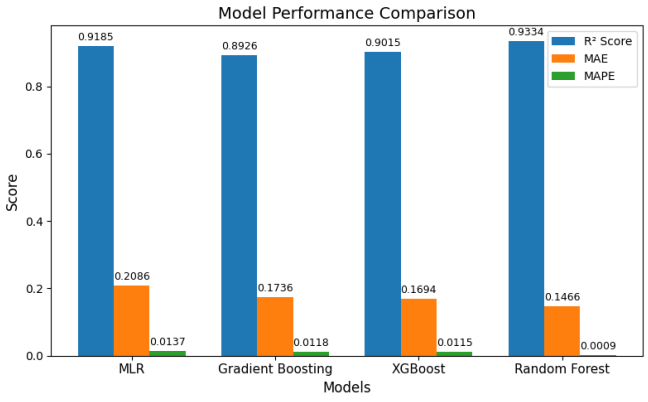


Fig. 2. Comparison of model performance

4. CONCLUSIONS

This study validates the utility of Radom Forest for rice production forecasting in Cambodia and introduces a localized, scalable framework for data-driven agricultural decision-making. It stands as one of the first studies to integrate ensemble learning models with Cambodia-specific datasets. The resuts provide actionable insight for policymakers and stakeholders. The research contributes to multiple Sustain Development Goals, ensuring food security and improve nutrition (SDG 2), fostering innovation and building resilient agricultural infrastructure (SDG 9), advancing climate adaptation and resilience stratagies (SDG 13), and supporting effective, transparent institutions through evidence-based decision-making (SDG 16). Future work should focus on real-time data integration, farmer-level applications, and apply predictive analysis in different rice-cultivating seasons.

ACKNOWLEDGMENTS

On behalf of this acheivement, I extend my deepest gratitude to my advisor Asst, Prof, Dr. PHAUK Sokkhey, for his expert guidance, unwavering support, and invaluable insights throungout my academic. I would also like to pay my gratefull thanks to my co-advisor Dr. HAS Sothea and Mr. PHOK Ponna for his invaluable assistance and advice along my research implemtations. His mentorship has been instrumental in shaping my academic journey and this work would not have been possible without his encouragement and expertise.

REFERENCES

[1] S. Piseth, Y. Monyoudom, and H. Tynarath, "Cambodia's Agri-Food Trade: Structure, New Emerging Potentials, Challenges & Impacts of Covid-19".

[2] "Policy Paper on The Promotion of Paddy Production and Rice Export," Jul. 25, 2010. [Online]. Available: https://faolex.fao.org/docs/pdf/cam189808.pdf

[3] Royal Government of Cambodia, *Cambodia Industrial Development Policy*. Royal Government of Cambodia, 2015. [Online]. Available: https://cdc.gov.kh/wp-content/uploads/2022/04/IDP-English.pdf

[4] P. Ly, L. S. Jensen, T. B. Bruun, and A. De Neergaard, "Factors explaining variability in rice yields in a rain-fed lowland rice ecosystem in Southern Cambodia," *NJAS: Wageningen Journal of Life Sciences*, vol. 78, no. 1, pp. 129–137, Sep. 2016, doi: 10.1016/j.njas.2016.05.003.

[5] Erlin, A. Yunianta, L. A. Wulandhari, Y. Desnelita, N. Nasution, and Junadhi, "Enhancing Rice Production Prediction in Indonesia Using Advanced Machine Learning Models," *IEEE Access*, vol. 12, pp. 151161–151177, 2024, doi: 10.1109/ACCESS.2024.3478738.

[6] N. Gandhi, L. J. Armstrong, O. Petkar, and A. K. Tripathy, "Rice crop yield prediction in India using

support vector machines," in *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)*, Khon Kaen: IEEE, Jul. 2016, pp. 1–5. doi: 10.1109/JCSSE.2016.7748856.

[7] N. Gandhi, O. Petkar, and L. J. Armstrong, "Rice crop yield prediction using artificial neural networks," in *2016 IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, Chennai, India: IEEE, Jul. 2016, pp. 105–110. doi: 10.1109/TIAR.2016.7801222.

[8] M. Piekutowska *et al.*, "The Application of Multiple Linear Regression and Artificial Neural Network Models for Yield Prediction of Very Early Potato Cultivars before Harvest," *Agronomy*, vol. 11, no. 5, p. 885, Apr. 2021, doi: 10.3390/agronomy11050885.

[9] E. Sreehari and S. Srivastava, "Prediction of Climate Variable using Multiple Linear Regression," in *2018 4th International Conference on Computing Communication and Automation (ICCCA)*, Greater Noida, India: IEEE, Dec. 2018, pp. 1–4. doi: 10.1109/CCAA.2018.8777452.

[10] M. Swamynathan, *Mastering Machine Learning with Python in Six Steps*. Berkeley, CA: Apress, 2017. doi: 10.1007/978-1-4842-2866-1.

[11] W. Dong, Y. Huang, B. Lehane, and G. Ma, "XGBoost algorithm-based prediction of concrete electrical resistivity for structural health monitoring," *Automation in Construction*, vol. 114, p. 103155, Jun. 2020, doi: 10.1016/j.autcon.2020.103155.

[12] A. C. Mu, "Introduction to Machine Learning with Python".

[13] J. K. Jaiswal and R. Samikannu, "Application of Random Forest Algorithm on Feature Subset Selection and Classification and Regression," in *2017 World Congress on Computing and Communication Technologies (WCCCT)*, Tiruchirappalli, Tamil Nadu, India: IEEE, Feb. 2017, pp. 65–68. doi: 10.1109/WCCCT.2016.25.

[14] R. L. Sapra, "Using R2 with caution," *Current Medicine Research and Practice*, vol. 4, no. 3, pp. 130–134, May 2014, doi: 10.1016/j.cmrp.2014.06.002.

[15] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature," *Geosci. Model Dev.*, vol. 7, no. 3, pp. 1247–1250, Jun. 2014, doi: 10.5194/gmd-7-1247-2014.

[16] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation," *PeerJ Computer Science*, vol. 7, p. e623, Jul. 2021, doi: 10.7717/peerj-cs.623.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Model-Based PMSM Control in the dq Frame Using Full-State Feedback with Integral Action and EKF for State Estimation under Stochastic Dynamics

Rotanak CHHEAN [1*], Sarot SRANG [2], Sokserey SREY [2]

[1] *Faculty of Mechatronics, Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: rotanak_chhean@gsc.itc.edu.kh*

Abstract: *This paper presents a robust model-based control strategy for a Permanent Magnet Synchronous Motor (PMSM) in the dq reference frame, combining Full-State Feedback with Integral Action (FSFB&I) and an Extended Kalman Filter (EKF) for state estimation under stochastic dynamics. The proposed controller controls the motor's angular speed (ω_m) while enforcing zero d-axis current ($i_d = 0$) to maximize efficiency, despite disturbances such as Coulomb friction as load torque, and measurement noise. A stochastic PMSM model is derived, accounting for electrical and mechanical uncertainties, and an EKF is designed to estimate key states (i_d, i_q, ω_m) and the external load torque (T_L) in real time. Simulation results demonstrate that the EKF provides accurate and fast state estimation, even with significant measurement noise. Without estimation, the system exhibits minor speed tracking delays and substantial i_d fluctuations (±1.5A), indicating energy inefficiency. However, integrating the EKF with FSFB&I significantly enhances performance—smoothing speed tracking and reducing i_d fluctuations to near-zero (±0.003A), thereby improving efficiency. The proposed FSFB&I + EKF framework proves highly effective for precise speed control and energy-optimal control under stochastic disturbances, making it suitable for high-performance PMSM applications. Future work will focus on experimental validation and extending the controller to handle fluctuating loads (e.g., periodic torque variations) with adaptive compensation, further improving robustness in real-world applications.*

Keywords: Permanent Magnet Synchronous Motor (PMSM); Full-State Feedback with Integral Action; Extended Kalman Filter (EKF); Stochastic Control; Load Torque Estimation.

1. INTRODUCTION

Permanent Magnet Synchronous Motors (PMSMs) have been increasingly used in various industrial applications due to their high efficiency, high torque density, and superior dynamic performance. PMSMs employ permanent magnets on the rotor, removing the need for external excitation currents, which leads to reduced losses and improved overall efficiency [1]. The modeling of PMSMs in the direct-quadrature (d-q) rotating reference frame simplifies the machine dynamics by transforming the three-phase time-varying system into a two-axis time-invariant system, which is essential for the implementation of advanced control strategies like Field-Oriented Control (FOC). The desired d-axis current (i_d) is set to zero while the q-axis current (i_q) is adjusting according to the desired torque or speed, thus the stator current vector becomes perpendicular (90 degrees) to the rotor flux angle which produces maximum torque [2].

Despite the advantages of PMSMs, challenges persist in achieving robust performance under uncertainties such as parameter variations, external disturbances, and limited sensor availability. Direct measurement of rotor speed (ω_m) and currents (i_d, i_q) typically requires costly or fragile sensors, which can reduce system reliability. Furthermore, variations in load torque (T_L) are often unknown but significantly impact system performance [3]. Traditional controllers designed without considering stochastic disturbances or using fixed-gain observers may lead to poor estimation and degraded control, especially under noisy conditions. Therefore, developing control strategies that incorporate state estimation and are robust to uncertainties is necessary.

[1*] Corresponding author: Rotanak CHHEAN
E-mail: rotanak_chhean@gsc.itc.edu.kh; Tel: +855-98 711 217

Full-state feedback with integral action improves control system performance by enhancing precision and robustness. Chowdhury et al. [3] showed that, compared to PID control, it yields better transient response, lower overshoot and faster settling in DC motor position control. Similarly, Love et al. [4] found it boosts responsiveness, stability, and disturbance rejection. These advantages make it ideal for systems needing fast dynamics and accurate steady-state performance.

The objective of this paper is to design a robust PMSM controller that utilizes full state feedback with integral action (FSFB&I), incorporating a stochastic model to enhance performance under uncertainty. The proposed controller aims to effectively control the angular speed (ω_m) and the d-axis current ($i_d = 0$) while estimating key states such as i_d, i_q, ω_m, and the external load torque T_L using an Extended Kalman Filter (EKF).

This paper presents the modeling of a surface-mounted PMSM in the d-q frame, the design of a full-state feedback controller with integral action, and the development of an EKF for real-time state and disturbance estimation. The system is assumed to experience coulomb friction as load (bearing friction and/or brake) and measurement noise, representing practical operating conditions. Simulation results are presented to validate the effectiveness of the proposed method. Hardware implementation and experimental validation are considered future work and are not included in the present study.

2. METHODOLOGY

This section presents a structured methodology for modeling, control design, state estimation, and simulation of a PMSM in the dq reference frame. The complete workflow includes: (1) system modeling, (2) controller design using Full-State Feedback with Integral Action, (3) modeling stochastic disturbances, (4) state estimation using EKF, and (5) simulation and validation in MATLAB/Simulink.

2.1 PMSM Mathematical Modeling in dq-Frame

Yusivar et al. [5] and Yesilbag et al. [6] show the mechanical equation of PMSM dynamic model as

$$\frac{d\omega_m}{dt} = \frac{1}{J}(T_e - T_L - B_v\omega_m) \qquad \text{(Eq. 1)}$$

where T_L is the load torque, B_v is the viscous friction coefficient, ω_m is the mechanical angular speeds, J is the inertia of the rotor, and $T_e = \frac{3}{2}\frac{P}{2}[\lambda_m i_q + (L_d - L_q) i_d i_q]$ is the developed electromagnetic torque, P is the number of poles, L_d and L_q are the inductances on dq axis, i_d and i_q are the stator current on dq-Frame, and λ_m is the permanent magnet flux linkage. In most PMSM, such as Surface-Mounted PMSM, the permanent magnets are mounted on the surface of the rotor. This geometry results in a uniform magnetic air gap, which leads to equal d-axis and q-axis inductances $L_d = L_q = L$, Yesilbag et al. [6]. Then, the electrical and mechanical equations of PMSM dynamic model in d-q Frame could be written as

$$\frac{d}{dt}i_d = \frac{1}{L}v_d - \frac{R}{L}i_d + \frac{P}{2}\omega_m i_q \qquad \text{(Eq. 2)}$$

$$\frac{d}{dt}i_q = \frac{1}{L}v_q - \frac{R}{L}i_q - \frac{P\omega_m\lambda_m}{2L} - \frac{P}{2}\omega_m i_d \qquad \text{(Eq. 3)}$$

$$\frac{d\omega_m}{dt} = \frac{3P\lambda_m}{4J}i_q - \frac{1}{J}B_v\omega_m - \frac{1}{J}T_L \qquad \text{(Eq. 4)}$$

where v_d and v_q are the stator voltages on dq axis, R is stator resistance.

2.2 Full State Feedback Controller with Integral Action (FSFB&I) of PMSM

In order to apply FSFB&I, the PMSM dynamic model needs to be compensated the non-linear terms by defining new variables $v_{d0} = -\frac{P}{2}\omega_m L i_q$, $v_{q0} = \frac{P}{2}\omega_m L i_d$, $V_d = v_d - v_{d0}$, $V_q = v_q - v_{q0}$, and let $\dot{i_d} = \frac{d}{dt}i_d$, $\dot{i_q} = \frac{d}{dt}i_q$, and $\dot{\omega}_m = \frac{d}{dt}\omega_m$. Then, the equations of PMSM dynamic model could be written as

$$\dot{i_d} = \frac{1}{L}V_d - \frac{R}{L}i_d \qquad \text{(Eq. 5)}$$

$$\dot{i_q} = \frac{1}{L}V_q - \frac{R}{L}i_q - \frac{P\omega_m\lambda_m}{2L} \qquad \text{(Eq. 6)}$$

$$\dot{\omega}_m = \frac{3P\lambda_m}{4J}i_q - \frac{1}{J}B_v\omega_m - \frac{1}{J}T_L \qquad \text{(Eq. 7)}$$

Then, the state space equation of PMSM dynamic model could be written as

$$\dot{X} = AX + Bu + Ed \qquad \text{(Eq. 8)}$$

where $X = \begin{bmatrix} i_d \\ i_q \\ \omega_m \end{bmatrix}$, $A = \begin{bmatrix} -\frac{R}{L} & 0 & 0 \\ 0 & -\frac{R}{L} & -\frac{P\lambda_m}{2L} \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B_v}{J} \end{bmatrix}$, $B = \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \\ 0 & 0 \end{bmatrix}$, $u = \begin{bmatrix} V_d \\ V_q \end{bmatrix}$, $E = \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{J} \end{bmatrix}$, and $d = T_L$ represents the disturbance to the system.

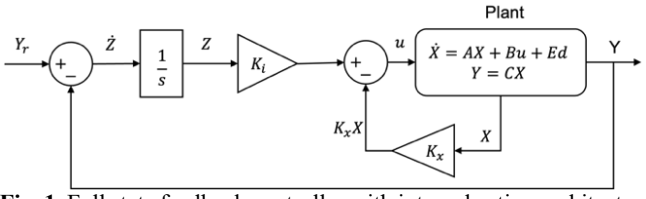


Fig. 1. Full state feedback controller with integral action architecture of PMSM

The **Fig. 1.** illustrates the architecture of FSFB&I of PMSM where the measurement equation is

$$Y = CX \quad \text{(Eq. 9)}$$

where $C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ and the state feedback controller is

$$u = -K_x X - K_i Z \quad \text{(Eq. 10)}$$

where, K_x is the state feedback gain matrix and K_i is the feedforward gain that ensures the tracking of the reference signal $Y_r = [i_d, \omega_m]^T$. To eliminate steady-state error, integral states are introduced for the tracking errors of i_d and ω_m. Let

$$Z = \int (Y_r - Y)dt \quad \text{(Eq. 11)}$$

or,

$$\dot{Z} = Y_r - CX \quad \text{(Eq. 12)}$$

Then, the augmented state vector is written as

$$\begin{bmatrix} \dot{X} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} Ed \\ Y_r \end{bmatrix} \quad \text{(Eq. 13)}$$

or,

$$\dot{X}_i = A_i X_i + B_i u + E_i \quad \text{(Eq. 14)}$$

where $\dot{X}_i = \begin{bmatrix} \dot{X} \\ \dot{Z} \end{bmatrix}$, $A_i = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix}$, $B_i = \begin{bmatrix} B \\ 0 \end{bmatrix}$, and $E_i = \begin{bmatrix} Ed \\ Y_r \end{bmatrix}$. Let $K = [K_x \quad -K_i]$, then $u = -KX_i$. Substituting to the equation **(Eq. 14)**, then

$$\dot{X}_i = (A_i - B_i K)X_i + E_i \quad \text{(Eq. 15)}$$

To compute the state feedback gain K, the pole placement method is applied by assigning a set of desired closed-loop poles $\lambda = \{\lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5\}$ lying in the left-half complex plane to ensure stability. The gain matrix K is then calculated such that the eigenvalues of the matrix $A_i - B_i K$ match these desired poles. This is done by solving the characteristic equation:

$$det\ (A_i - B_i K - \lambda I) = 0 \quad \text{(Eq. 16)}$$

The solution is computed using MATLAB's $place()$ function, which implements the Kautsky–Nichols–Van Dooren (KNVD) algorithm to solve for K numerically. To enhance control performance, feedforward compensation terms v_{d0} and v_{q0} are also added to counteract nonlinear effects in the PMSM model.

2.3 Stochastic Modeling of PMSM for the FSFB&I

In real-world applications, PMSMs are subject to stochastic disturbances such as measurement noise, process noise, and unmodeled dynamics. By using the equations **(Eq. 2)**, **(Eq. 3)**, and **(Eq. 4)**, the dynamic model of PMSM could be written in deterministic form as

$$\begin{bmatrix} \dot{i_d} \\ \dot{i_q} \\ \dot{\omega}_m \end{bmatrix} = \begin{bmatrix} -\frac{R}{L} & 0 & 0 \\ 0 & -\frac{R}{L} & -\frac{P\lambda_m}{2L} \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B_v}{J} \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ \omega_m \end{bmatrix} + \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_d \\ v_q \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{J} \end{bmatrix} T_L + \begin{bmatrix} \frac{P}{2}\omega_m i_q \\ -\frac{P}{2}\omega_m i_d \\ 0 \end{bmatrix} \quad \text{(Eq. 17)}$$

or, it could be written in state space form as

$$\dot{X} = AX + Bu_s + Ed + N \quad \text{(Eq. 18)}$$

where $N = \begin{bmatrix} \frac{P}{2}\omega_m i_q \\ -\frac{P}{2}\omega_m i_d \\ 0 \end{bmatrix}$ is the non-linear term, and $u_s = \begin{bmatrix} v_d \\ v_q \end{bmatrix}$.

To accurately represent the real-world uncertainties in the PMSM system such as sensor noise, load torque variation, and model inaccuracies, we model the PMSM using a Stochastic Differential Equation (SDE) in continuous-time domain as

$$\dot{X} = AX + Bu_s + Ed + N + \sqrt{Q}W \quad \text{(Eq. 19)}$$

where W is the Wiener process (standard Brownian motion), $\sqrt{Q}$ is the Cholesky decomposition or square root of the process noise covariance Q. To implement this model numerically in a digital system (such as Simulink), we discretize the SDE using the Euler–Maruyama method, which is $\dot{X} \approx \frac{X_{k+1} - X_k}{T_s}$ and $W \approx \frac{W_{k+1} - W_k}{T_s}$ where $\Delta W_k = W_{k+1} - W_k \sim \mathcal{N}(0, T_s I)$ and $\Delta W_k = \sqrt{T_s} w_k$ with $w_k \sim \mathcal{N}(0, I)$, T_s is the time step, the subscript k denotes the discrete time step at

k^{th}, and I is the identity matrix. Then, an SDE of the process equation in the discrete-time domain could be written as

$$X_{k+1} = (I + T_s A)X_k + T_s B u_{s,k} + T_s E d_k + \sqrt{T_s Q} w_k \quad \text{(Eq. 20)}$$

and the measurement equation in the discrete-time domain could be written as

$$Y_k = CX_k + \sqrt{R} v_k \quad \text{(Eq. 21)}$$

where $v_k \sim \mathcal{N}(0,\ I)$ and R the measurement noise covariance.

The process equation **(Eq. 20)** and the measurement equation **(Eq. 21)** are used, instead of the deterministic continuous-time dynamic model, to simulate the real-world uncertainties in the PMSM system.

2.4 State Estimation Using Extended Kalman Filter (EKF)

To handle disturbances and noises, EKF is employed to estimate the nonlinear and time-varying internal states of the PMSM, namely d-axis current (i_d), q-axis current (i_q), rotor speed (ω_m), and load torque (T_L). EKF is suitable for nonlinear systems with Gaussian noise characteristics and is applied to the discretized stochastic model. Over short time intervals, we treat T_L as slowly varying or constant, then

$$\dot{T}_L = 0 \quad \text{(Eq. 22)}$$

Then, the deterministic PMSM dynamic model in the continuous-time domain could be written as

$$\begin{bmatrix} \dot{i_d} \\ \dot{i_q} \\ \dot{\omega_m} \\ \dot{T_L} \end{bmatrix} = \begin{bmatrix} -\frac{R}{L} & \frac{P}{2}\omega_m & 0 & 0 \\ -\frac{P}{2}\omega_m & -\frac{R}{L} & -\frac{P\lambda_m}{2L} & 0 \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B_v}{J} & -\frac{1}{J} \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_d \\ i_q \\ \omega_m \\ T_L \end{bmatrix} + \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_d \\ v_q \end{bmatrix} \quad \text{(Eq. 23)}$$

Let the state vector $X_e = [i_d, i_q, \omega_m, T_L]^T$, then the equation **(Eq. 23)** could be written in state space form as

$$\dot{X}_e = A_e X_e + B_e u_s \quad \text{(Eq. 24)}$$

where $A_e = \begin{bmatrix} -\frac{R}{L} & \frac{P}{2}\omega_m & 0 & 0 \\ -\frac{P}{2}\omega_m & -\frac{R}{L} & -\frac{P\lambda_m}{2L} & 0 \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B_v}{J} & -\frac{1}{J} \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $B_e = \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

By using an SDE, the equation **(Eq. 24)** becomes

$$\dot{X}_e = A_e X_e + B_e u_s + \sqrt{Q_e} W_e \quad \text{(Eq. 25)}$$

where W_e is the Wiener process and Q_e is the process noise covariance. Discretize the equation **(Eq. 25)** using the Euler–Maruyama method, we obtain

$$X_{e,k+1} = (I + T_s A_e)X_{e,k} + T_s B_e u_{s,k} + \sqrt{T_s Q_e} w_{e,k} \quad \text{(Eq. 26)}$$

where $w_{ek} \sim \mathcal{N}(0,\ I)$. And the measurement equation is

$$Y_{e,k} = HX_{e,k} + \sqrt{R_e} v_{e,k} \quad \text{(Eq. 27)}$$

where $H = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$, R_e is the square root of the measurement noise covariance, $v_{e,k} \sim \mathcal{N}(0, I)$, and the subscript k denotes the discrete time step at k^{th}.

Welch et al. [7] and Kulikov et al. [8] demonstrated that EKF algorithm operates recursively, updating its state estimates over time as new measurements become available. Each cycle of the EKF consists of two main steps: time update (prediction) and measurement update (correction). The prediction step of PMSM is formulated in discrete time as:

$$\hat{X}_{e,k}^{-} = f\left(\hat{X}_{e,k}, u_{s,k}, 0\right) \quad \text{(Eq. 28)}$$

$$P_k^{-} = \left(I + T_s A_{j,k}\right) P_k \left(I + T_s A_{j,k}\right)^T + T_s Q_e \quad \text{(Eq. 29)}$$

and the correction step is formulated in discrete time as

$$K_k = P_k H_{j,k}^T \left(H_{j,k} P_k H_{j,k}^T + R_e\right)^{-1} \quad \text{(Eq. 30)}$$

$$\hat{X}_{e,k} = \hat{X}_{e,k}^{-} + K_k \left(z_k - h\left(\hat{X}_{e,k}^{-}, 0\right)\right) \quad \text{(Eq. 31)}$$

$$P_k = \left(I - K_k H_{j,k}\right) P_k^{-} \quad \text{(Eq. 32)}$$

where $f\left(\hat{X}_{e,k}, u_{s,k}, 0\right) = (I + \mathrm{T_s} A_e)\hat{X}_{e,k} + \mathrm{T_s} B_e u_{s,k}$, P_k^{-} and P_k are the predicted and corrected error covariance, the Jacobian matrix of the measurement model is $H_{j,k} = H$, z_k is the measurement at step k, $h\left(\hat{X}_{e,k+1}, 0\right) = H_{j,k} X_{e,k}^{-}$, and the Jacobian matrix of the system model at step k is

$$A_{j,k} = \begin{bmatrix} -\frac{R}{L} & \frac{P}{2}\omega_m & \frac{P}{2}i_q & 0 \\ -\frac{P}{2}\omega_m & -\frac{R}{L} & -\frac{P}{2}\left(\frac{\lambda_m}{L} + i_d\right) & 0 \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B_v}{J} & -\frac{1}{J} \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

2.5 MATLAB/Simulink Implementation

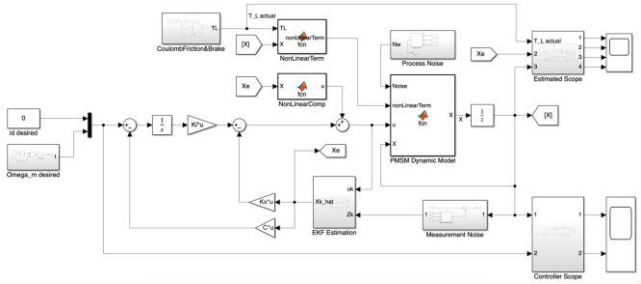


Fig. 2. i_d and ω_m control using Full state feedback with integral action incorporated stochastic PMSM model integrated with the state estimation using EKF

Fig. 2. illustrates the MATLAB/Simulink schematic of controlling i_d and ω_m using FSFB&I incorporated stochastic PMSM model and integrated with the state estimation using EKF. This simulation uses a small PMSM's parameters from Teknic's industrial-grade NEMA 23 modeled M-2310P-LN-04K. The motor parameters are listed in **Table 1.**. In addition, the values used in the Simulink model are $Q_e = 10^{-4} \times diag([0.5, 0.5, 5, 0.5])$, $R_e = 10^{-2} \times diag([0.5, 0.5, 5])$, and $\lambda = \{-90, -43.9823, -43.9823, -100, -100\}$.

Table 1. Parameters of PMSM Model M-2310P-LN-04K

Parameters	Value	Unit
Phase to Phase Resistance	0.74	Ω
Phase to Phase Inductance	0.0004	H
Continue Torque	0.2754	Nm
Rated Current	7.1	A
Back EMF	0.04429	Vpeak/(rad/s)
Electrical Time Constant	0.00056	s
Number of Poles	8	

3. RESULTS AND DISCUSSION

Fig. 5. demonstrates a good performance of the EKF in estimating the states i_d, i_q, ω_m, and T_L. The estimated values closely match the actual ones and respond quickly, even in the presence of the slightly elevated measurement noise shown in **Fig. 4.**. **Fig. 3.** shows that, without estimation, the actual ω_m tracks the desired speed well, with only a slight response delay of approximately 0.07 seconds, even though the desired speed varies sinusoidally. However, small fluctuations in the actual ω_m suggest it is somewhat affected by noise or unmodeled dynamics. The actual i_d, on the other hand, attempts to follow the desired value of $0A$ but exhibits high fluctuations, mostly ranging between $\pm 1.5A$. This indicates that it is significantly influenced by noise or unmodeled dynamics, resulting in increased energy waste and potential efficiency loss in high-performance applications. When EKF estimation is integrated into the control system, the performance of both i_d and ω_m is noticeably improved. As shown in **Fig. 6.**, the speed ω_m becomes smoother, and the fluctuations in i_d are significantly reduced to approximately $\pm 0.003A$ at steady state. This implies a substantial enhancement in efficiency, approaching the maximum when i_d remains near zero. In conclusion, the FSFB&I controller based on the stochastic PMSM dynamic model with EKF achieves excellent control of both i_d and ω_m, resulting in efficiency levels that are close to optimal at the desired operating speed.

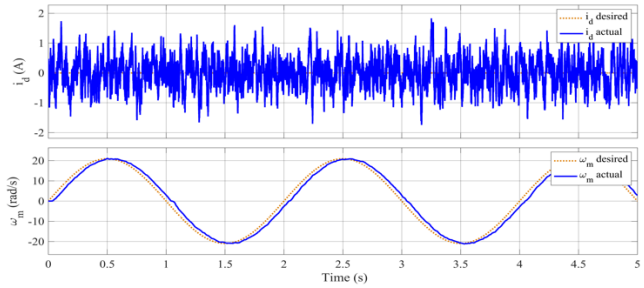


Fig. 3. i_d and ω_m controlled using FSFB&I of stochastic PMSM dynamic model without estimation

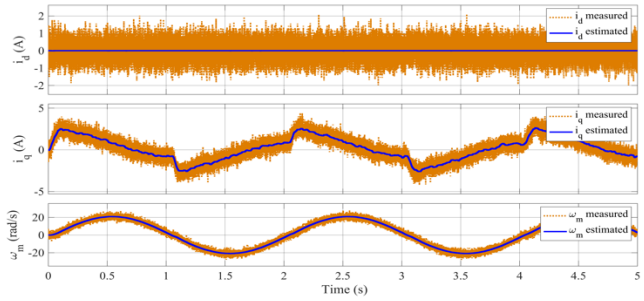


Fig. 4. i_d, i_q, and ω_m estimated using EKF with measurement noise and stochastic PMSM dynamic model

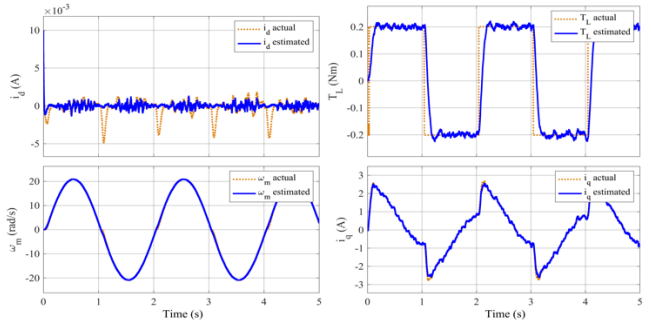


Fig. 5. i_d, i_q, ω_m, and T_L estimated using EKF comparing to actual states of stochastic PMSM dynamic model

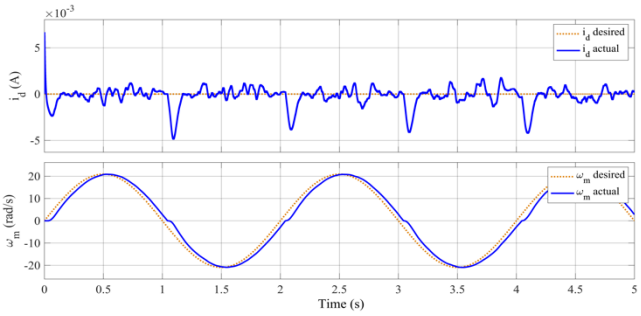


Fig. 6. i_d and ω_m control using FSFB&I of stochastic PMSM dynamic model with state estimation using EKF

4. CONCLUSIONS

This paper presented a robust FSFB&I for a PMSM Motor operating under stochastic dynamics, incorporating an EKF for real-time state and disturbance estimation. The proposed control scheme effectively controlled the motor's angular speed (ω_m) while maintaining the d-axis current ($i_d = 0$) to maximize efficiency, even in the presence of measurement noise, Coulomb friction, and load torque disturbances. The EKF demonstrated excellent performance in estimating key states (i_d, i_q, ω_m) and the external load torque (T_L), closely tracking actual values with minimal delay despite significant measurement noise. Without estimation, the system exhibited fluctuations in i_d and minor speed tracking deviations due to noise and unmodeled dynamics. However, integrating the EKF significantly improved control performance, smoothing the speed response and reducing i_d fluctuations to nearly zero ($\pm 0.003A$ in steady state), thereby enhancing energy efficiency. The simulation results validated the effectiveness of the FSFB&I + EKF approach in achieving precise speed control and near-optimal efficiency under stochastic disturbances. Future work will focus on experimental validation and extending the controller to handle fluctuating loads (e.g., periodic torque variations) with adaptive compensation, further improving robustness in real-world applications.

ACKNOWLEDGMENTS

I gratefully acknowledge the Dynamics and Control Laboratory (DCLab) and the University of Sydney for their financial support. My sincere thanks to my advisor, Dr. SRANG Sarot, and co-advisor, Mr. SREY Sokserey, for their invaluable guidance and support. I also appreciate the support from DCLab colleagues for their valuable discussions.

REFERENCES

[1] M. Y. Bdewi, A. M. Mohammed, and M. M. Ezzaldean, "Design and Performance Analysis of Permanent Magnet Synchronous Motor for Electric Vehicles Application," *ETJ*, vol. 39, no. 3A, pp. 394–406, Mar. 2021, doi: 10.30684/etj.v39i3A.1765.

[2] D. Citharthan, M. Varatharaj, M. A. Babu, and S. Kannan, "Field oriented control of permanent magnet synchronous motor with d-q axis decoupling using feed forward compensation," presented at the INTERNATIONAL CONFERENCE ON SMART TECHNOLOGIES AND APPLICATIONS (ICSTA 2022), Rajapalayam, India, 2023, p. 040008. doi: 10.1063/5.0164857.

[3] A. Chowdhury and D. Debnath, "Performance Comparison between PID Controller and State-Feedback Controller with Integral Action in Position Control of DC Motor," *AMM*, vol. 367, pp. 188–193, Aug. 2013, doi: 10.4028/www.scientific.net/AMM.367.188.

[4] K. Love, C. Kriger, and N. Tshemese-Mvandaba, "Design and Simulation of Full State Feedback Controller for DC Motor," *IJEEAS*, vol. 7, no. 1, Oct. 2024, doi: 10.54554/ijeeas.2024.7.01.008.

[5] F. Yusivar, N. Hidayat, R. Gunawan, and A. Halim, "Implementation of field oriented control for permanent magnet synchronous motor," in *2014 International Conference on Electrical Engineering and Computer Science (ICEECS)*, Kuta, Bali, Indonesia: IEEE, Nov. 2014, pp. 359–362. doi: 10.1109/ICEECS.2014.7045278.

[6] E. Yesilbag and L. T. Ergene, "Field oriented control of permanent magnet synchronous motors used in washers," in *2014 16th International Power Electronics and Motion Control Conference and Exposition*, Antalya, Turkey: IEEE, Sep. 2014, pp. 1259–1264. doi: 10.1109/EPEPEMC.2014.6980685.

[7] G. Welch and G. Bishop, "An Introduction to the Kalman Filter," 2006.

[8] G. Yu. Kulikov and M. V. Kulikova, "Accurate Numerical Implementation of the Continuous-Discrete Extended Kalman Filter," *IEEE Trans. Automat. Contr.*, vol. 59, no. 1, pp. 273–279, Jan. 2014, doi: 10.1109/TAC.2013.2272136.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Continuously Uploading ECG Waveform Data to Google Sheet using ESP32

Phalla I Ning LEANG [1*], Lichhean LY [1], Sreypich PHORN [1], Suchhing LY [1], Rithea NGETH [1]

[1] *Department of Telecommunication and Network Engineering, Institute of Technology of Cambodia, Russian Federation*

[*] *Corresponding author: e20211447@dtc1.itc.edu.kh*

Abstract: *This paper presents an Electrocardiogram (ECG) data acquisition system using the ESP32 as microcontroller, the AD8232 ECG sensor, Google Sheets for cloud storage and MATLAB for designing the user interface (UI). The purpose is to design an affordable ECG remote monitoring system. The first part of work study compares two approaches between the Hypertext Transfer Protocol (HTTP) get request and HTTP post request methods for uploading ECG to Google Sheets, examining their respective advantages and performance. The second part of this work is designing a MATLAB based UI that includes signal processing tools such as digital filter to handle and process the raw data obtained from microcontrollers. The designed UI can display captured ECG waveform and provides an option to send it to the user through email. Although the acquired ECG waveform demonstrates acceptable quality, comprehensive validation through comparative analysis with a certified medical-grade measurement device is essential to ensure clinical reliability.*

Keywords: Continuously uploading; ESP32; AD8232; Google Sheet; User Interface.

1. INTRODUCTION

Cardiovascular diseases (CVDs) are the leading cause of death globally. According to the World Health Organization (WHO) [1], an estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths. Of these deaths, 85% were due to heart attack and stroke. Over three quarters of CVD deaths take place in low and middle-income countries. Out of the 17 million premature deaths (under the age of 70) due to noncommunicable diseases in 2019, 38% were cause by CVDs. One of the main risk factors for CVD is abnormal heart rhythm, also known as arrhythmia, which can lead to serious complications such as stroke, heart failure and sudden cardiac death [1]. Therefore, early detection and diagnosis of arrhythmia is crucial for preventing and treating CVD. This paper considers the electrocardiogram (ECG) data continuously that can serve to provide the detection and prevention for any change of pathosis by hazard.

Previous works have extensively addressed the collection of ECG datasets. In work of [2], a system is developed where ECG data was acquired and wirelessly transmitted via Bluetooth to a smartphone, which then relayed the information to a server and physician over the internet. On the other hand, ECG data were uploaded to an internet of thing (IoT) cloud with storage, Hypertext Transfer Protocol (HTTP) and MQ Telemetry Transport (MQTT) servers in the work of [3]. In the work of [4], data were collected by using analog-to-digital converter MCP3008 ADC, Raspberry Pi and application "plot.ly" for viewing the graph. ECG data acquired by ESP32 on single-lead ECG sensor, AD8232, are uploaded to a paid cloud, Ubidots was presented in the work of [5].

In summary, the cloud-based storage can be open-source storage and paid-storage. This paper aims to propose a remote monitoring system that continuously uploading ECG waveform data to Google Sheet which measure the heart rhythm and store the data for diagnosis by specialists in heart disease and to support the healthcare system. The reason Google Sheet is chosen because it is a popular open-source and is easy to share with the other. In addition, based on our knowledge, it is a challenge to transfer a huge data such as ECG data to Google Sheet. This work found out that uploading data in "string" type can be a solution. In this system, there is also a user interface (UI) designed in MATLAB that makes it easy for users to test and view their ECG data as a graph. ESP32 and AD8232 are employed as microcontroller and ECG sensor, respectively. With MATLAB, the raw signal can be processed with build-in tools

[1*] Corresponding author: Phalla I Ning LEANG
E-mail: e20211447@dtc1.itc.edu.kh; Tel: +855-96 45 00 514

such as digital filter to eliminate the un-wanted signal. Moreover, after testing, the users can send their ECG report as a pdf file to their email.

The rest of paper is organized by presenting the methodology in Section 2. The result and discussion are described in Section 3. Finally, the conclusion is given in the last section.

2. METHODOLOGY

2.1 System Architecture

The proposed system architecture, as shown as in **Fig. 1**, is designed to acquire cardiac electrical activity data through an ECG sensor, and retrieve store data from Google Sheet for MATLAB based user interface generation. To enable cloud data storage, the system incorporates an ESP32 microcontroller module for establishing internet connectivity.

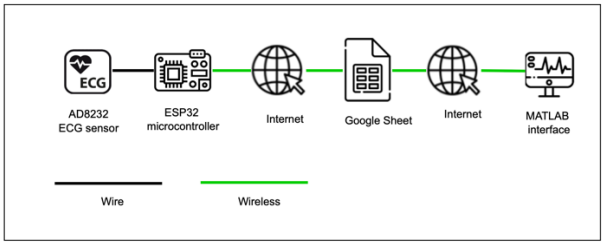


Fig. 1. System architecture of proposed system

2.2 Hardware

2.2.1 Microcontroller (ESP32)

The ESP 32 is a powerful and cost-effective Wi-Fi, Bluetooth, microchip, designed for Internet of Things (IoT) application. It includes a microprocessor, memory, wireless connectivity, TCP/IP, stack, Analog Input (used with analog-to-digital converter), peripheral interface and rich set of peripherals all in one chip. There is an analog input pin with analog to digital converter 12-bit with voltage range 0 to 3 Volts (V). It is used for reading the ECG signal that is a continuous-time analog signal.

2.2.2 AD8232 ECG Sensor

The AD8232 ECG sensor is a compact, low-power sensor module that can record ECG waveforms and monitor the electrical signals of the heart using on three electrodes. It is an integrated signal conditioning block for ECG and other biopotential measurement application. It is designed to convert the tiny, noisy signal from the electrodes into a large, filtered signal that can be easily read by a low resolution [6]. As shown in **Fig. 2** and based on [5], the input is 3.5mm jack for biomedical pad connection with three electrodes. It operates with 3.3 V of voltage.



Fig. 2. AD8232 ECG Sensor

2.3 Software

2.3.1 Arduino IDE

The Arduino IDE (Integrated Development Environment) is an open-source electronics platform based on easy-to-use hardware and software used to write, compile, and upload code to board. Board manager support for other boards like ESP32, ESP8266, STM32, etc.

2.3.2 Google Sheet

Google Sheet is a web-based spreadsheet application developed by Google. It allows you to create, edit, and share spreadsheets online, with support for real-time collaboration, data analysis, and cloud storage. On the other hand, the data can be updated or inserted via uniform resource locator (URL) link. The user can response the access via URL link with "Apps Script" where the callback functions can be created.

2.3.3 MATLAB Interface

MATLAB (Matrix Laboratory) interface allows users to build self-contained apps that handle data input, processing, and result display within a single interface. It features a drag-and-drop GUI layout tool and integrated code editor, making it easy to design and program app behavior. All interactions, from data input to final result display, are handled within the app, making it a powerful solution for creating custom MATLAB-based tools (MathWorks).

2.4 System Design

2.4.1 Reading ECG signal by using ESP32 with AD8232

ECG reading data can be done with the process as shown in **Fig. 3** by using ESP32. *ECGReadDataTask* is the task that continuously reads data from ECG sensor and push it into a local data buffer. An index variable tracks the position in the buffer. On the other hand, *ECGSendDataTask* is the task that triggered when data is ready for upload. It uses HTTP GET or POST requests to send the buffered ECG data to a linked

Google Sheet. The codes are written and compiles by using Arduino IDE. The reading process is done in loop.

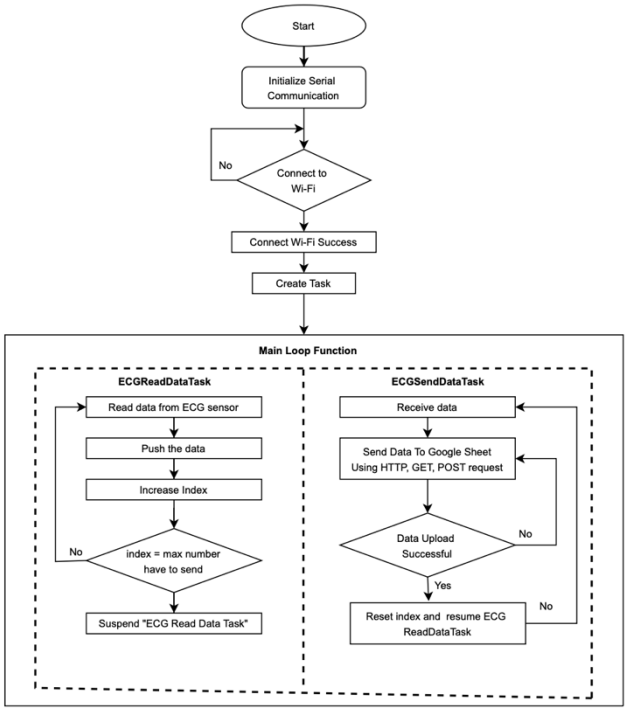


Fig. 3. Process of reading ECG Signal by using ESP32 with AD8232 sensor.

2.4.2 Callback from Google Sheet

A callback function is created for every event that is executed on access through a URL. This is based on the specific use case for the Google Sheets. The sample data demonstrates process concerning Google Sheets, and it can be accessed in **Fig. 4**. Information saved in Google Sheet contains the date, time and the corresponding information as a string.

2.4.3 MATLAB Interface

The MATLAB interface is central to this system as it reads the data from Google Sheets in the real-time, uses a digital filter to process the data and outputs are visualized on a graph within the interface. The digital filters used in this UI are a lowpass filter and bandstop filter to remove noise in ECG, in particular with a Blackman window, which is used to remove the interference of the 50 Hz frequency and its harmonics. The process of MATLAB is shown in **Fig. 5**.

2.5 HTTP Method Implementation

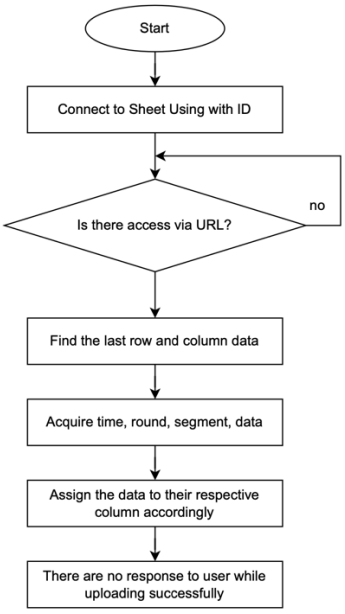


Fig. 4. Flowchart callback function of Google Sheet

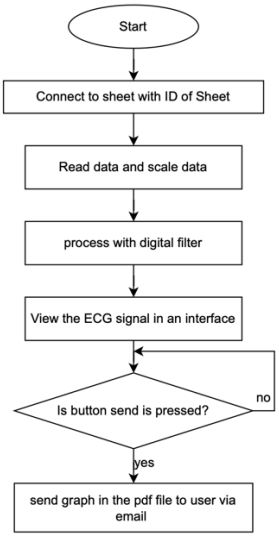


Fig. 5. Process of MATLAB interface

The system utilizes two HTTP upload processes for ECG data that are implemented as HTTP GET Request and HTTP POST Request in the ESP32 microcontroller and performance is evaluated on the basis of reliability, data size limitations and compatibility with Google Sheets Apps Script.

2.5.1 HTTP GET Request

An HTTP GET method requests data from a server without altering its state. It appends parameters to the URL, making it suitable for retrieving non-sensitive data. It is commonly used for viewing content, GET is ideal for requests that don't involve data modification [7].

2.5.2 HTTP POST Request

The HTTP POST method sends data from the client to the server to create or update resources, storing data in the request body. It's suitable for secure data transfer, like images or documents, with security relying on encryption (HTTPS), authentication, and validation [7].

3. RESULTS AND DISCUSSION

3.1 Comparison between HTTP Get Request and Post Request

3.1.1 Experiment Set Up

In order to the comparison between HTTP get request and post request with fairness, the two sets of identical simple data uploading system. The list of components in each set is given in **Table 1**.

Table 1. List of Components in Each Simple Data Uploading System

Component name	Purpose
ESP32 Devkit V1	Microcontroller to generate and upload data to Google Sheet via Wi-Fi
Prolink DL-7202 4G LTE Mobile Wi-Fi modem	Access point to generate Wi-Fi for internet connection with cellular communication
Metfone sim card	Sim card for cellular network
RuiDeng AT34 Cost-effective USB tester	Equipment to measure the energy consumption of ESP32 in data uploading
Anker PowerCore Lite 10000 Power bank	Power bank to serve as power supply for ESP32
Micro-USB USB cable	Cable connected between USB tester and ESP32

From **Table 1**, the internet access of each ESP32 does not interfere with each other because each ESP32 connects to Wi-Fi with different access point. **Fig. 6**. shows the experiment set up for this comparison experiment.

The data is generated with sampling time $T_s = 4$ ms (sampling rate $F_s = 250$ samples/s). The data are grouped into generation or round, while each round consists of 5040

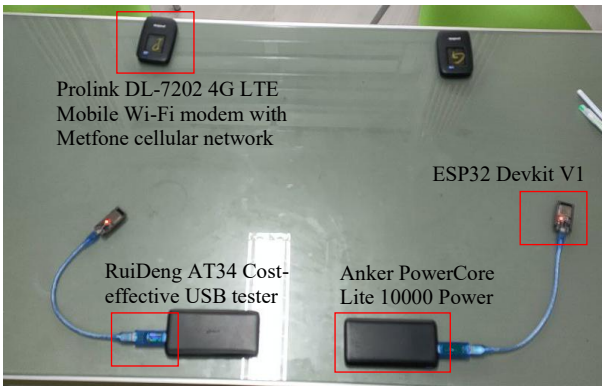
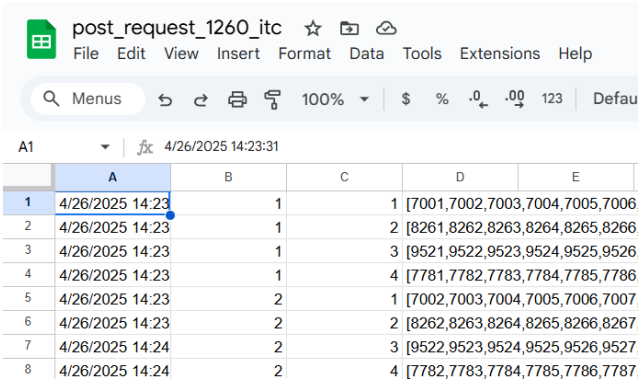


Fig. 6. Experiment set up for comparison between HTTP get request and post request

data with integer value starting from 7001. Thus, the duration of each round should be 5040×4 ms $= 20.16$ s. 5040 data of each round are divided into segments with the same amount of data. The comparison checks the performants of both request methods with different numbers of data per segments that are: 720, 840, 1008 and 1260, while they are correspondent to number of segments per round: 7, 6, 5 and 4, respectively. Each experiment is done for at least 100 rounds. The purpose of this comparison is to provide the information in making decision which method to be used. The received data on Google Sheet are shown in **Fig. 7**.


post_request_1260_itc

File Edit View Insert Format Data Tools Extensions Help

Menus 100% $ % 123 Defau

A1 4/26/2025 14:23:31

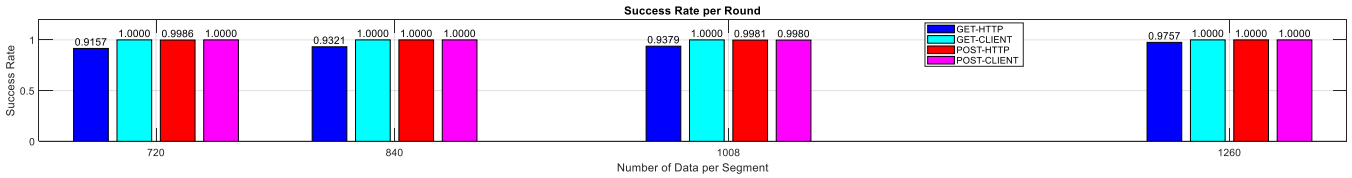
	A	B	C	D	E
1	4/26/2025 14:23	1	1	[7001,7002,7003,7004,7005,7006	
2	4/26/2025 14:23	1	2	[8261,8262,8263,8264,8265,8266	
3	4/26/2025 14:23	1	3	[9521,9522,9523,9524,9525,9526	
4	4/26/2025 14:23	1	4	[7781,7782,7783,7784,7785,7786	
5	4/26/2025 14:23	2	1	[7002,7003,7004,7005,7006,7007	
6	4/26/2025 14:23	2	2	[8262,8263,8264,8265,8266,8267	
7	4/26/2025 14:24	2	3	[9522,9523,9524,9525,9526,9527	
8	4/26/2025 14:24	2	4	[7782,7783,7784,7785,7786,7787	

Fig. 7. Received data on Google Sheet

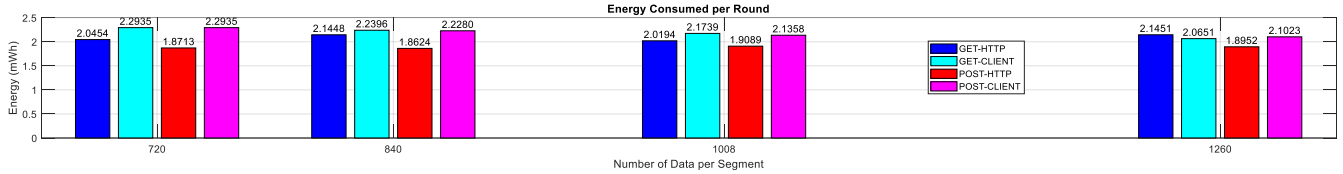
3.1.2 Performance metrics

The performance metrics to be checked are success rate per round, energy consumption per round and taken duration per round in seconds. Success rate per round is the mean of the ratio of number of received segments per round to number of segments per round.

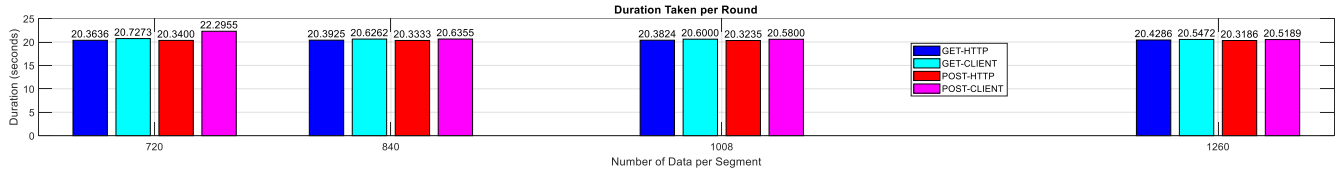
The success rate per round is the mean value of the ratios of the number of received segments per round to the number of segments per round. The success rate might vary with the number of segments per round, if the number of segments per



(a) Result of success rate per round



(b) Result of energy consumption per round



(c) Result of duration taken per round

Fig. 8. Results of comparison

round is larger, then the taken duration to upload one segment is shorter, then the probability of successfully uploading is lower. In addition, if the reference signal received power (RSRP) of modem is not high enough, for example, lower than dBm, the success rate might become lower.

The energy consumption is measured from AT34 cost-effective USB tester in mWh. The energy consumption per round is the ratio of the measured value of energy consumption to the total number of rounds for the experiment test. This metric can compare the overhead of both request methods. The heavy overhead will take high energy consumption.

The taken duration per round is the mean value of the duration taken per round, that is the difference between the timestamp of the reception of the last segment in a round and the timestamp of the reception of the last segment in its previous round. This duration should be close to 20.16 s, if there is difference, the more difference reflects more errors in time domain, i.e., the actual sampling time is higher than the set value 4 ms.

3.1.3 Results and Discussion

The results shown in Fig. 8 are obtained in the case where RSRP ≈ -77 dBm that is a good condition and can be served as reference for stable internet condition. HTTP get request using *WiFiClientSecure* library and *HTTPClient* library are denoted by GET-CLIENT and GET-HTTP, respectively. In addition, HTTP post request using *WiFiClientSecure* library and *HTTPClient* library are denoted by POST-CLIENT and POST-HTTP, respectively.

From the result in **Fig. 8(a)**, the success rates per round approximately equal to one except GET-HTTP. The success rate increases with the number of data per segment, i.e., the time lag between times to upload two successive segments. GET-HTTP (as well as POST-HTTP) needs to wait for the response from Google Sheets while the response code equal to 200 refers for success, and negative for failure. This waiting might make ESP32 miss the chance to upload the data if the waiting time is longer than the time lag between times to upload two successive segments.

From **Fig. 8(b)**, the energy consumption of POST-HTTP is lowest, that can indicate that the protocol overhead of POST-HTTP is the lightest. However, POST-HTTP might face the cyber security issue because of its simple protocol. In this comparison, GET-CLIENT and POST-HTTP do not employ security setting (*setInsecure* is used) to keep the fairness of comparison.

From **Fig. 8(c)**, the duration taken per round of POST-HTTP is shortest, maybe because also its simple protocol. For POST-CLIENT and GET-CLIENT, the order of received segments is not in order, some segments might arrive at Google Sheet after the segments of the other round have arrived. It is because they do not need to wait for response from Google Sheet. Thus, the error in time domain might need to be investigated in the other way.

On the other hand, regarding to data security and privacy will be considered as future work.

3.2 ECG Waveform Visualization and Sharing

Normally, there is an interference of AC power, i.e., 50 Hz sinusoidal interference and white noise. The interference can be eliminated by a bandstop filter, and the white noise can be reduced by a lowpass filter with cut-off frequency of 45 Hz. The resulted ECG signal is shown in **Fig. 9**.

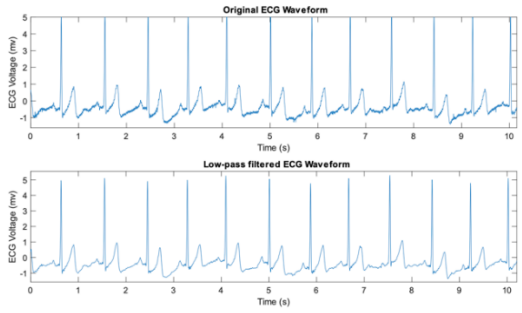


Fig. 9. Comparison of original and filtered ECG waveform in the time domain

The captured figure of MATLAB interface is shown in **Fig. 10**. The user needs need to press the "load" button to the signal. Then, the user needs to enter their email address if he/she would like received their ECG waveform via email. The email reception is shown in **Fig. 11**.

With the resulted ECG waveform, it is understandable by a medicine doctor after consulting with him. However, the ECG waveform might need to compare with medical-grade device to receive better comparison.

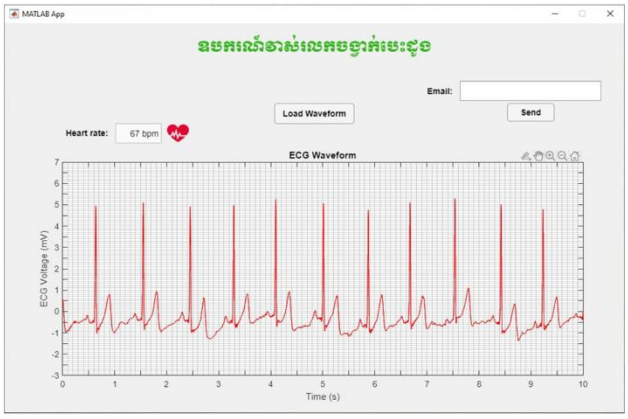


Fig. 10. UI of the MATLAB app

4. CONCLUSIONS

This paper presents a comparison of two approaches in uploading data: HTTP get request and HTTP post request with two used libraries. The results can serve as a reference in selecting the best option in upload data. In addition, an ECG waveform acquisition with a UI for display and sharing ECG

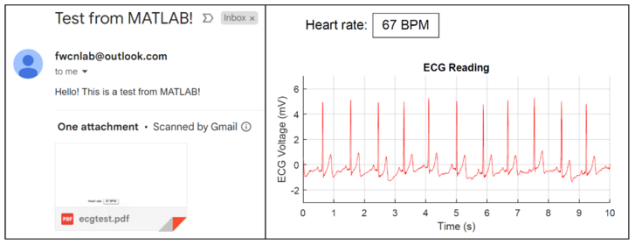


Fig. 11. Email sent from MATLAB app and the contents of the attached PDF

is described. The system can work properly with single-lead sensor AD8232.

ACKNOWLEDGMENTS

This work was supported by JICA-LBE Batch 4 (April 2022-March 2023) and Takahashi Industrial and Economic Research Foundation (April 2024-March 2025).

REFERENCES

[1] World Health Organization. Cardiovascular diseases. https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds)

[2] Chai, J. (2013, November). The design of mobile ECG monitoring system. In 2013 IEEE 4th International Conference on Electronics Information and Emergency Communication (pp. 148-151). IEEE.

[3] Yang, Z., Zhou, Q., Lei, L., Zheng, K., & Xiang, W. (2016). An IoT-cloud based wearable ECG monitoring system for smart healthcare. Journal of medical systems, 40, 1-11.

[4] Patil, P., & Bhole, K. (2018, February). Real time ECG on internet using Raspberry Pi. In 2018 International Conference on Communication, Computing and Internet of Things (IC3IoT) (pp. 267-270). IEEE.

[5] ResearchGate IoT Based ECG Using AD8232 and ESP32 https://www.researchgate.net/publication/378501123_IoT_Based_ECG_Using_AD8232_and_ESP32.pdf

[6] Analog Devices. (2012). Single-Lead, Heart-Rate Monitor Analog Front End AD8232. https://www.analog.com/media/en/technical-documentation/data-sheets/ad8232.pdf

[7] Difference between HTTP GET and POST Methods, 16 September 2024. https://www.geeksforgeeks.org/difference-between-http-get-and-post-methods/

Toward a Case Study on Event Detection in Non-Intrusive Load Monitoring Using Geometric Features of the Cumulative Sum

Song SOK [1]*, Korsorl THOURN [2], Kimtho PO [2]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Department of Telecommunication and Network Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Senior Researcher, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

Corresponding author: soksong168@gmail.com

Abstract: *Non-intrusive load monitoring (NILM) has become more common among load monitoring methods because of its low hardware needs—just one sensor at the main power inlet—while maintaining the current circuit architecture. NILM allows for the identification of single devices and power consumption calculation. Event detection—the identification of changes in appliance states to assess their operational status—is a fundamental component of NILM. Among several techniques, the Cumulative Sum (CUSUM) method, a statistical probability-based approach, is often employed for event detection analysis during appliance activation. Serving as a basis for both aggregated and disaggregated appliance usage analysis, this paper offers a case study on event identification for single-state (ON/OFF) residential appliances employing the CUSUM technique. The work uses a privately owned dataset for assessment. The findings show that CUSUM has great accuracy and works well in event detection. This paper sets the stage for more studies on household appliance monitoring and NILM growth.*

Keywords: Energy Disaggregation; Energy Saving; Event Detection; NILM; Geometric Features of the CUSUM;

1. INTRODUCTION

In this digital era, the standard of living for humans is increasingly defined by energy utilization, as nearly everything relies on electricity and the Internet. At the same time, the demand for electrical energy in residential properties is rising daily [1]. In [2], the Royal Government of Cambodia has enacted the 'National Policy on Energy Efficiency 2022–2030' to improve energy efficiency and support a strong, competitive, and sustainable economy. The numerous strategies for energy saving had been addressed such as intrusive load monitoring (ILM) and non-intrusive load monitoring (NILM). The NILM is a popular approach for load monitoring due to its low hardware requirements and straightforward integration into existing electrical systems [3], [4], [5]. Typically, this technique requires only one sensor at the main electrical supply intake, significantly reducing installation complexity and costs compared to ILM is an invasive system that require sensors for each individual appliance. The most critical aspect of NILM systems is event detection, which involves identifying specific transitions in the operational states of household appliances, such as ON/OFF status changes, by analyzing the transition change.

The foundation of NILM systems is event detection, which consists of gathering power data and applying it to identify when certain changes in the operational state of household appliances, such turning them on or off, take place [6], [7]. Efficient event detection enables a realistic consumption profile at the appliance level, hence supporting general energy reduction initiatives and improving energy management.

This paper shows how to utilize the CUSUM approach with geometric characteristics from [8] on a privately held dataset to find events connected to an appliance that an end user uses in different situations. Because these methods are so

[1] * Corresponding author: Song SOK
E-mail: soksong168@gmail.com; Mobile: +855-968988008

complicated and need so much computing power, they are not commonly used in real-time homes. This case study is the first step in developing a real-time monitoring approach based on NILM. It doesn't use many ways yet, but we are working on changing the features of the methods to make them more accurate and good speed processing.

2. METHODOLOGY

To solve the event detection process, the CUSUM method combined with geometric features, as proposed in algorithm [8], is applied to the private owned dataset as a case study. This approach is intended to facilitate further research aimed at addressing real-time monitoring of individual appliance consumption. As shown in Fig. 1, the flow of work utilized a current case study involving a privately owned dataset and employed the CUSUM method for simulating appliance detection.

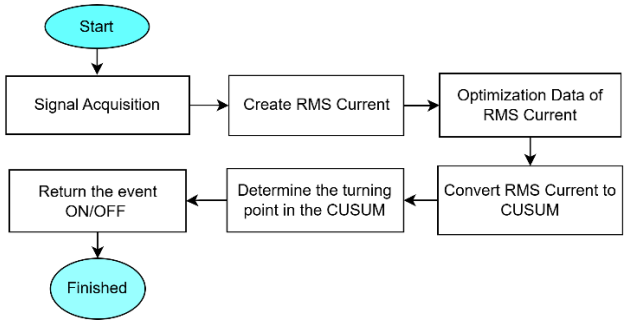


Figure 1. The workflow of the event detection development

2.1. DATASET

The numerous datasets are accessible to the scientific community for experimental purposes, including Tracebase, PLAID, REDD, AMPds, BLUED, GREEND, UK-DALE, iAWE, BERDS, and ACS-F [9], [10], [11]**.** These datasets are not ideal for real-time applications; however, they can be applied in case studies.

A part of the newly owned private dataset, shown in Table 1, was gathered from a Phnom Penh City residential home to present a new case study using the existing CUSUM method for event identification.

This dataset has used the specific components, such as the Acquisition Board Series 2024B, Differential Probe, and Current Clamp T189, that were used in the data collection. Both voltage and current primary signals were recorded at a high sampling frequency of 10 kHz. A low-pass filter with a cutoff frequency of 2 kHz and scope processing set up at 100 milliseconds per division were features of the digital signal processing (DSP) arrangement. The data collection was coordinated using the PicoScope™ software, which ran on a Windows 11 system equipped with 16GB of RAM and a Core i7 CPU. This dataset includes various household electrical appliances such as air conditioners, washing machines, irons, kettles, blenders, ovens, vacuum cleaners, and fans.

The newly acquired dataset has been applied in the current case study to illustrate event detection techniques and to facilitate future work in real-time detection or monitoring of residential electrical appliances.

Table 1. A part of private owned dataset

NO	APPLIANCE TYPE	QTY
1	Blender	8
2	Air Conditioner	8
3	Washing Machine	10
4	Refrigerator	10
5	Iron	6
6	Rice Cooker	11
7	Oven	6
8	Kettle	10
9	Television (TV)	11
10	Fan	41
11	Lighting (Lamp, LED, …)	3
12	Laptop	59
13	Desktop	3
14	PC Monitor	37
15	PC all in one	57

2.1.1. The variety of Appliances

As shown in Fig. 2 [4], individual appliances, along with their working conditions and associated power usage, may each function in various states depending on their purpose or operation. The previous researchers categorized appliances into four types:

- **Type I**: Single-state (on/off) appliances, which refer to the basics of home appliances such as light bulbs, toasters, and monitors. During the ON state, it responds to the specific amount of power.
- **Type II**: Finite state machine or multiple-state appliance, which refers to appliances that exist in more than a single state (ON/OFF), such as washing machines, fridges, air conditioners, etc.
- **Type III**: Infinite state or continuously variable appliance. If it does not have a finite set of states, then it is difficult to use the NILM algorithm for modeling or identification. These appliances are battery chargers, laptops, phones, etc.
- **Type IV**: Always on. This appliance type refers to single or multiple states that have a constant source of energy consumption, such as a fridge and landline phone.

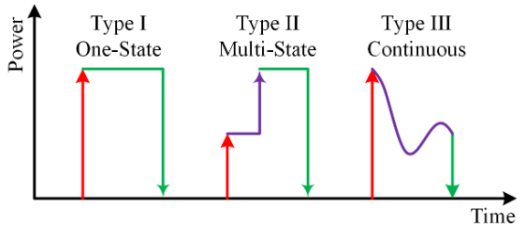
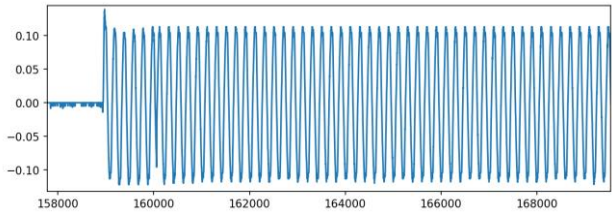


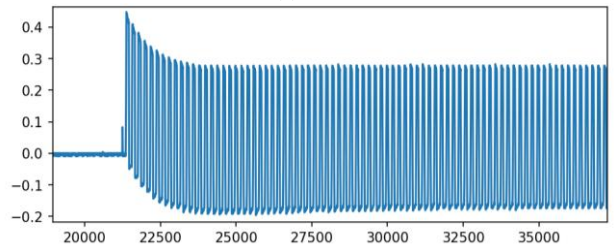
Figure 2: The different types of appliances category by its states

2.1.2. the variety current signal of appliances

As shown in Fig. 3, the data is loaded according to the current data points during both transient and steady states to identify the unique features of individual appliances. The appliance discussed here is solely a sample from the local, privately owned dataset.



(a) Fan



(b) Blender

Figure 3. The Raw Current Signal of Electrical (a) Fan and (b) Blender

2.2. RMS (Root Mean Square) Current

As shown in Fig. 4, the raw current signal is validated error and converted to RMS current. The RMS current value represents a signal segment that is processed within each window. Each signal segment $s_i = \{x_{i1}, x_{i2}, \dots, x_{iN}\}$, is defined by the step size, which indicates the number of segments, as explained in the equation below.

$$K = \left[\frac{T-N}{S}\right] + 1 \quad (1)$$

$$RMS_i = \begin{cases} 0, \ if \ \frac{1}{N}\sum_{j=1}^{N} x_{ij} \leq 0 \ or \ NaN \\ \sqrt{\frac{1}{N}\ \sum_{j=1}^{N} x_{[i*S+j]}^2}, \ Otherwise \end{cases} \quad (2)$$

Where Eq. (1) $N = cycle\ of\ samples$ and S step size, K is the number of segments, and T is the length of signal and Eq. (2) x_{ij} is the j^{th} sample of the i^{th}segment, N is the $window\ size\ =\ cycle\ of\ samples$, and S is the step-size

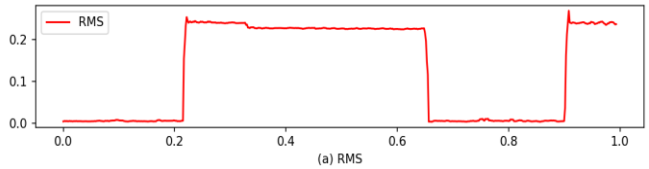


Figure 4. The raw current signal of fan converted to RMS Current

2.3. Convert RMS Current to CUSUM

As shown in Fig. 5, the RMS current has been converted into CUSUM. However, traditional CUSUM does not significantly improve event detection accuracy. To enhance its effectiveness, it requires updates that incorporate additional features, such as geometry, as referenced in [8]. This section concentrates solely on the previously proposed CUSUM for event detection. The following section will examine appliance sample plots. Below is the mathematical representation of the CUSUM process.

$$\mu_0 = \frac{\sum_{i=1}^{N} x_i}{N} \quad (3)$$

$$C_n = \sum_{i=1}^{n}(x_i - \mu_0) = (x_n - \mu_0) + C_{n-1} \quad (4)$$

Where μ_0 is mean of the sample, N is the total number of samples, x_i is sample of RMS, and C_i is the number of CUSUM.

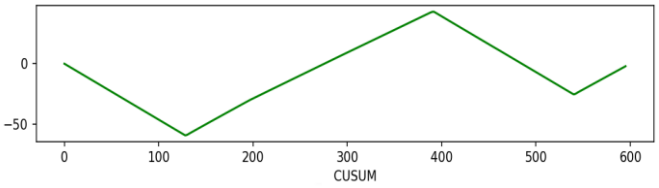


Figure 5. Fan, The converting RMS Current to CUSUM

2.4. Event Detection Based CUSUM

As shown in Fig. 6, this section employs equations (5) and (6) to demonstrate the significant geometric characteristic of RMS current. This characteristic is then applied in conjunction with CUSUM to identify ON/OFF events during the load identification process of actual appliances, taking into account the aggregated energy consumption. We implement the following protocols to recognize and engage the devices.

$$(C_i < C_{i-m}) \cap (C_i < C_{i+m}),\ where\ m \in N \quad (5)$$

To identify the appliance, and turn it off, the following rule is used

$$(C_i > C_{i+m}) \cap (C_i > C_{i-m}),\ where\ m \in N \quad (6)$$

Where C_i is the turning point at i^{th} CUSUM current based, m is a phase combined with the period of C_i

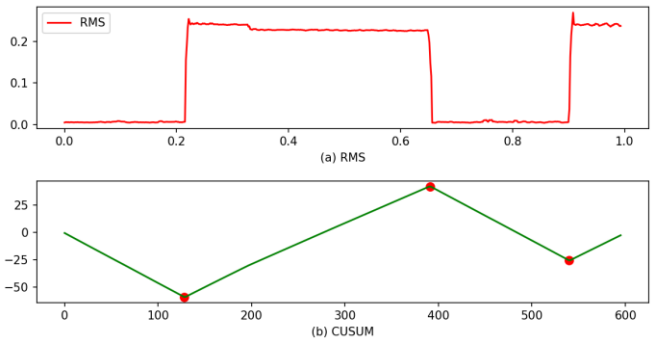


Figure 6. The Fan Hatari includes (a) RMS, (b) CUSUM, and Event Detection (ON/OFF) features for the appliance.

2.5. Evaluate Matrix

In the Ref. [8], [12] have described, the evaluation of the accuracy of the system performance, there are four main indicators used Accuracy, Precision, Recall, and F1 Score. Each indicator has its formula and definition such as below:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$precision = \frac{TP}{TP + FP} \quad (8)$$

$$Recall = \frac{TP}{TP + FN} \quad (9)$$

$$F1 = \frac{2 \times Recall \times precision}{Recall + precision} \quad (10)$$

Where the *Accuracy* presents the ability of the model to judge the authenticity in all situations. *Precision* is the accuracy rate used to present the ability of a model to distinguish real events. *Recall* is the recall rate used to present the ability of a model to detect real events. *F1* is the comprehensive evaluation index used to present the comprehensive ability of a model to detect and identify a real event. These indicators are used after event detection.

Table 2. Evaluation Matrix

NO	APPLIANCE TYPE	QTY	TP	TN	FP	FN	Precision	Recall	F1	ACC
1	Blender	**8**	14	13	0	1	100%	93%	96.6%	96%
2	Air Conditioner	**8**	5	7	4	0	56%	100%	71.4%	75%
3	Washing Machine	**10**	10	7	12	5	45%	67%	54.1%	50%
4	Refrigerator	**10**	9	8	0	1	100%	90%	94.7%	94%
5	Iron	**6**	9	11	3	1	75%	90%	81.8%	83%
6	Rice Cooker	**11**	13	14	1	3	93%	81%	86.7%	87%
7	Oven	**6**	7	7	1	0	88%	100%	93.3%	93%
8	Kettle	**10**	19	13	0	1	100%	95%	97.4%	97%
9	Television (TV)	**11**	11	10	21	2	34%	85%	48.9%	48%
10	Fan	**41**	45	33	5	9	90%	83%	86.5%	85%
11	Lighting (Lamp, LED,etc)	**3**	3	2	0	0	100%	100%	100.0%	100%
12	Laptop	**59**	42	36	131	10	24%	81%	37.3%	36%
13	Desktop	**3**	1	0	20	0	5%	100%	9.1%	5%
14	PC Monitor	**37**	62	32	1	27	98%	70%	81.6%	77%
15	PC all in one	**57**	58	8	73	25	44%	70%	54.2%	40%
Total:		**280**	308	201	272	85	53%	78%	63.3%	59%

3. RESULT AND DISCUSSION

In this case study, only appliances in category type I with a single state (ON/OFF) from the privately owned dataset are utilized to simulate event detection using the geometry feature of CUSUM, resulting in satisfactory outcomes. The formula for determining turning points for events that turn ON and OFF is introduced and implemented in CUSUM. While the experiment with the manual individual appliance achieves high accuracy, complications arise with similar appliances due to overlap.

Looking at Table 2**.** Evaluation Matrix, the high precision and F1 on simple, peak-dominant loads, all of the following get at least 94% F1: blender, kettle, refrigerator, and lighting. These appliances have distinct, high-amplitude transients that CUSUM and geometry can easily pick up (high signal-to-noise ratio). Via Perfect Recall on Many Loads, the air conditioner, oven, and desktop, 100% of ON/OFF events were recorded, even if there were many false alarms. On the other hand, the robust accuracy on the fan and rice cooker shows that appliances with moderate complexity (with recurring cycles) can still keep their accuracy at 85% or more and maybe even 80% to 90% F1. There are a lot of problems with the weakness, like low precision on low-power or complex loads, high false positives on electronics, and class imbalances & small sample effects. First, the low precision on low-power or complex loads has laptops (24% precision), TVs (34%), and desktops (5%); these have low step-changes or very variable idle draws, so CUSUM flags numerous false events. Utilizing the second, there are many false positives on electronics. The laptop has 131 false positives and only 42 true positives. There are also 73 FPs in "PC all-in-one," which means it is 40% accurate. The third issue is class imbalance and the effects of small sample sizes. Appliances with few samples (such as desktops and

lighting) have unstable TN/FP counts; thus, a single mistake might change the percentage a lot.

4. CONCLUSIONS AND FUTURE WORK

Event detection is crucial for identifying each appliance used in a household. By understanding the events occurring during their operation, we can observe state changes in the appliances and their characteristics. This case study, which utilizes a privately owned dataset and the CUSUM method, demonstrates successful results only for single-state appliances; other types of appliances do not perform well with this method. Totally, the low accuracy is achieved compared to all appliance types but it is the main pilot step for research improvement.

The advanced denoising techniques can enhance signal cleaning and address slow baseline shifts in CUSUM. The expanded feature set comprises frequency-domain descriptors and time-frequency patterns. Automated searches can optimize CUSUM parameters tailored to each device type. Hybrid detection and classification can be accomplished by inputting candidate events into a compact CNN. In-field validation requires the collection of continuous, whole-house power measurements from actual homes to confirm that the NILM solution operates effectively in a non-intrusive and multi-appliance configuration.

The other phases will concentrate on classification, energy disaggregation, and software development for managing individual appliances' energy consumption. The results will provide feedback to end consumers, ultimately aiding in improved family economic growth and promoting sustainable environmental protection.

ACKNOWLEDGEMENT

The author sincerely appreciates the valuable contributions and support received from existing online resources, colleagues, family members, professors, advisors, and particularly the HEIPs Project for their significant role in enhancing research capabilities during the ITC's events, such as the 14th Scientific Day and other academic activities.

REFERENCES

[1] A. Faustine, N. H. Mvungi, S. Kaijage, and K. Michael, "A Survey on Non-Intrusive Load Monitoring Methodies and Techniques for Energy Disaggregation Problem," Mar. 10, 2017, *arXiv*: arXiv:1703.00785. doi: 10.48550/arXiv.1703.00785.

[2] "National Energy Efficiency Policy 2022-2030."

[3] T. Ono, A. Hagishima, and J. Tanimoto, "Non-Intrusive Detection of Occupants' On/Off Behaviours of Residential Air Conditioning," *Sustainability*, vol. 14, no. 22, p. 14863, Nov. 2022, doi: 10.3390/su142214863.

[4] M. I. Azad, R. Rajabi, and A. Estebsari, "Non-Intrusive Load Monitoring (NILM) using Deep Neural Networks: A Review," Jun. 08, 2023, *arXiv*: arXiv:2306.05017. doi: 10.48550/arXiv.2306.05017.

[5] A. Zoha, A. Gluhak, M. Imran, and S. Rajasegarar, "Non-Intrusive Load Monitoring Approaches for Disaggregated Energy Sensing: A Survey," *Sensors*, vol. 12, no. 12, pp. 16838–16866, Dec. 2012, doi: 10.3390/s121216838.

[6] Z. Zhu, Z. Wei, B. Yin, S. Zhang, and X. Wang, "A novel approach for event detection in non-intrusive load monitoring," in *2017 IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing: IEEE, Nov. 2017, pp. 1–5. doi: 10.1109/EI2.2017.8245715.

[7] "Adaptive Event Detection for Representative Load Signature Extraction."

[8] M.-S. Tsai and Y.-K. Lin, "Applying the Geometric Features of Cumulative Sums to the Development of Event Detection," *Energies*, vol. 16, no. 20, p. 7207, Oct. 2023, doi: 10.3390/en16207207.

[9] A. Ridi, C. Gisler, and J. Hennebert, "Processing smart plug signals using machine learning," in *2015 IEEE Wireless Communications and Networking Conference Workshops (WCNCW)*, New Orleans, LA, USA: IEEE, Mar. 2015, pp. 75–80. doi: 10.1109/WCNCW.2015.7122532.

[10] A. Ridi, C. Gisler, and J. Hennebert, "ACS-F2 — A new database of appliance consumption signatures," in *2014 6th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, Tunis, Tunisia: IEEE, Aug. 2014, pp. 145–150. doi: 10.1109/SOCPAR.2014.7007996.

[11] H. Rafiq, P. Manandhar, E. Rodriguez-Ubinas, O. Ahmed Qureshi, and T. Palpanas, "A review of current methods and challenges of advanced deep learning-based non-intrusive load monitoring (NILM) in residential context," *Energy and Buildings*, vol. 305, p. 113890, Feb. 2024, doi: 10.1016/j.enbuild.2024.113890.

[12] M. Irani Azad, R. Rajabi, and A. Estebsari, "Nonintrusive Load Monitoring (NILM) Using a Deep Learning Model with a Transformer-Based Attention Mechanism and Temporal Pooling," *Electronics*, vol. 13, no. 2, p. 407, Jan. 2024, doi: 10.3390/electronics13020407.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Observation of the Effect of Cutting Parameters on Surface Roughness of High Purity Copper in Dry Turning Process

Sela Soy [1], Vongchivorn Chhour [2], Phearom San [3], Sophal Hai [2*]

[1] *Faculty of Graduate School, National Polytechnic Institute of Cambodia, Phnom Penh, 12409, Cambodia*
[2] *Faculty of Mechanical of Engineering, National Polytechnic Institute of Cambodia, Phnom Penh, 12409, Cambodia*
[3] *Faculty of Civil Engineering and Architecture, National Polytechnic Institute of Cambodia, Phnom Penh, 12409, Cambodia*

[*]*Corresponding author: khaisophal@gmail.com*

Abstract: *The main objective of the study is to analyze the various cutting parameters, such as spindle speed, feed rate, and depth of cut, to investigate the surface quality of high-purity copper (UNS C10700) in the dry turning process with an uncoated carbide insert of nose radius 1.2 mm. The surface roughness (R_a) of the turning process was measured by a surface roughness tester (Surf test SJ-210, Mitutoyo, Japan). The statistical methods of an analysis of variance (ANOVA) and signal-to-noise ratio (S/N) are applied to investigate the effects of the turning input parameters on surface roughness. The result of this study indicates that the feed rate and depth of cut significantly affected surface roughness (Ra), but the spindle speed was insignificant. It found that the cutting parameters for optimum surface quality are a spindle speed of 580 rpm, a feed rate of 0.039 mm/rev, and a depth of cut of 0.35 mm.*

Keywords: Dry turning process; Surface roughness; ANOVA; signal-to-noise ratio (S/N).

1. INTRODUCTION

Copper and its alloys are widely used in electrical and mechanical applications, due to their high conductivity, corrosion resistance, and machinability. In addition, copper is also a soft and ductile material. One of the applications of copper machining is the production of motor commutators, which are cylindrical devices that transfer electric current between stationary and rotating parts of an electric motor.

The UNS C10700 copper is most commonly used in electrical conductors, which are operated at high temperatures. The surface roughness of the commutator is an important factor that affects the performance of the motor [1]. The surface quality of motor commutators is critical for their performance and reliability, as it affects the contact resistance, friction, and wear between the commutator and the brushes. The commutation is satisfactory good if brushes do not sparkle. In order to maintain cylindrical shape of the active commutator surface for different rotation speeds and temperatures, the same forces should act on each of the commutator bar [2]. Surface roughness is the most significant surface quality measure during the turning process. It is often induced by several machining parameters, such as spindle speed, feed rate, and depth of cut. Torres et al. [3] study the influence of different machining parameters on surface roughness has been analyzed. All these parameters have been studied in terms of depth of cut (D_c), feed rate (f), cutting speed (V_c) and tool radius (r). Moreover, Design of Experiments (DOE) techniques have been used to predict the surface quality in advance and to select the optimal turning conditions. The experimental results revealed that for dry turning operations and for the amplitude parameters, the most significant factor was the interaction effect between the depth of cut and the feed rate. Dry turning operation is one of the most popular conventional metal cutting methods for manufacturing industries. This is due to its high performance in terms of dimension accuracy, time saved, cost saved and a high-quality surface. Dry turning is characterized by the absence of any cutting fluid, and unlike minimum quantity lubrication (MQL) and cryogenic turning does not require any additional delivery system. Hence, from the environmental perspective, dry turning is ecologically desirable; and from an economic perspective, it decreases manufacturing costs by 16 to 20% [5].

1* Corresponding author: Sophal Hai
E-mail: khaisophal@gmail.com; Tel: +855-966803424

At present, dry machining is possible without considerable tool wear; as such, research work has been focused on the surface finish aspect of dry turning. This research has focused on experimental study and investigate of the cutting parameters spindle speed, depth of cut, and feed rate, for best surface roughness for motor commutator, during dry turning of silver bearing copper using uncoated carbides tool. ANOVA was used to analyze the effects of the machining factors, spindle speed, feed rate, and depth of cut, on the responses.

2. METHODOLOGY

2.1 Design of experiment process

The experimental conditions for dry turning process of UNS C10700 copper carried out in this research are shown in Table 1. It was determined in three levels and cutting parameters such as spindle speed (*N*), feed rate (*f*) and depth of cut (D_c) with dry turning environment. Figure 1 illustrates the schematic of the research process. The dry turning process was conducted on the Lath after defining the material of workpiece and experiment conditions. The turning of workpieces was cut in 27 pieces of specimens for different experiment conditions in order to measure the surface roughness by surface roughness tester. The surface roughness data were observed and analyzed using statistical methods for optimizing the experimental conditions for surface roughness.

Table 1. Experimental conditions for dry turning process.

Levels	Spindle Speed *N* (rpm)	Feed Rate *f* (mm/rev)	Depth of Cut D_c (mm)	Tool Nose Radius *r* (mm)	Environment
1	350	0.039	0.2		
2	580	0.060	0.35	1.2mm	Dry Turning
3	920	0.108	0.5		

In the present experimental work, silver bearing of copper (UNS C10700) having dimension of diameter Ø 30mm and length 500mm were use as a work material. The reason for selecting this material is that, UNS C10700 copper is a type of lightly alloyed grade of copper, has good corrosion resistance and high thermal, and a high electrical and thermal conductivity. The UNS C10700 copper is most commonly used in electrical conductors, which are operated at high temperatures. The chemical composition, and mechanical properties of workpieces were shown in Table 2, and Table 3 [10].

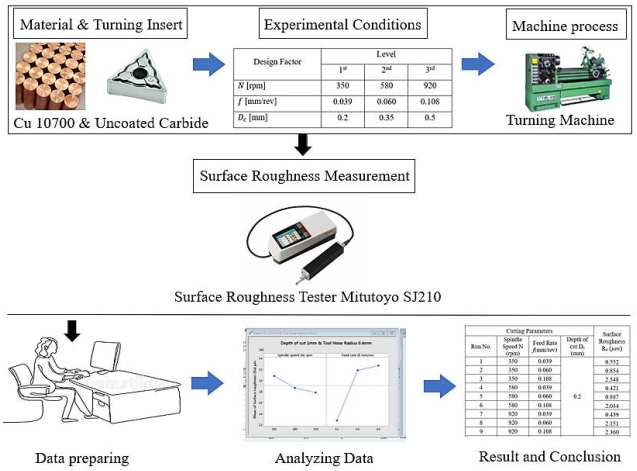


Fig. 1. Schematic of the research process.

2.2 Materials and cutting tools

In the present experimental work, silver bearing of copper (UNS C10700) having dimension of diameter Ø 30mm and length 500mm were use as a work material. The chemical composition, physical and mechanical properties of workpieces were shown in Table 2,Table 3 [10].

Table 2. Chemical composition of the UNS C10700 copper.

Element (UNS C10700 copper)	Content (%)
Cu	99.95
Ag	0.07813

Table 3. Mechanical properties of the UNS C10700 copper

Properties	Metric
Tensile strength	22-455 MPa
Tensile strength, yield (depending on temperature)	69.0-365 MPa
Elongation at break (in 101.6 mm)	55%
Modulus of elasticity	115 GPa
Poisson's ratio	0.300
Shear modulus	44.0 GPa

Cutting insert: as cutting insert of hard metal with uncoated carbides were selected this present experiment. The selected radius for cutting insert was *r*=1.2mm type TNMG160404-VP2-H01 combine with the tool holder type WTJNR2020-K16 of Korloy company.

2.3 Machining process

The machine operation was performed by using turning machine of TIPL-410 (HL460×1000GN, TONGIL, CHANGWON, KOREA) with a spindle power of 7.5kW and maximum spindle speed of 1500rpm by dry turning process. The surface roughness was measured at three different locations by using a surface roughness tester (Surftest SJ-210). To ensure accurate measurement values for surface roughness, measurements were taken in different zones. The average surface roughness value, Ra, was determined using the formula $R_a = (R_{a1}+R_{a2}+R_{a3})/3$.

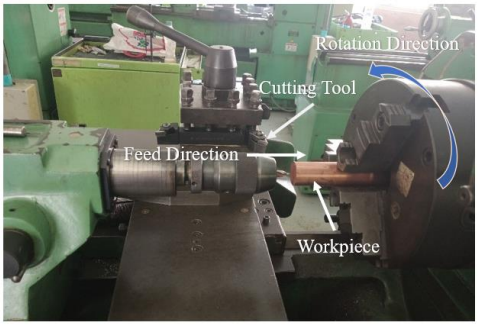


Fig. 2. Experimental set up for turning process

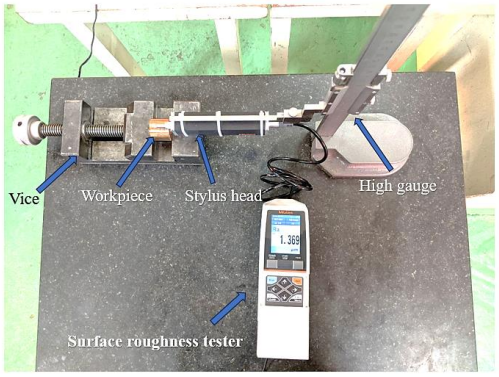


Fig. 3. Measuring surface roughness with surface roughness tester.

2.4 ANOVA And S/N Ratio Analysis

Statistical analysis was employed to examine and optimize the experimental conditions for surface roughness. In this study, Minitab statistical software was used for the analysis. ANOVA was conducted to determine which cutting parameters significantly impact responses such as surface roughness, tool wear, temperature, and cutting force [23-25]. As detailed in Section 3.2, degrees of freedom (DF) are used to calculate the adjusted mean squares (Adj MS); the adjusted sum of squares (Adj SS) measures the contribution of additional cutting parameters to the response; Adj MS is the Adj SS divided by the DF; the F-value is the Adj MS divided by the error value; and the P-value indicates which cutting parameters significantly affect the responses. A P-value of 0.05 or less signifies a significant effect of the cutting parameter on the response [26, 27].

The S/N ratio, developed using Taguchi methodology, is a metric aimed at optimizing the robustness of a product or process. The response values for surface roughness (Ra) were converted into S/N ratios. These ratios are categorized into three types: larger-is-better, smaller-is-better, and nominal-is-best. The appropriate S/N ratio type is chosen based on the study's objective to analyze the experimental data. In this research, the smaller-is-better S/N ratio was used to analyze the cutting parameters and surface quality, as minimizing surface roughness (Ra) was necessary. Therefore, the smaller-is-better S/N ratio is expressed by the following equation [28, 29]:

Smaller-is-better $$S/N = -10 \times log\left(\frac{1}{n}\sum y^2\right)$$ (Eq. 1)

where y is the observed data, n is the number of observations of experiment.

3. RESULTS AND DISCUSSION

In this research, the dry turning process was performed in the L27 orthogonal array, and the results of surface roughness (R_a) for the 27 experiments are shown in Table 4.

Table 4. Experimental results of surface roughness (R_a).

Cutting Parameters				
No.	Spindle Speed N (rpm)	Feed Rate f (mm/rev)	Depth of cut D_c (mm)	Surface Roughness R_a (μm)
1	350	0.039		0.552
2	350	0.060		0.854
3	350	0.108		2.548
4	580	0.039		0.421
5	580	0.060	0.2	0.824
6	580	0.108		2.034
7	920	0.039		0.439
8	920	0.060		2.151
9	920	0.108		2.602
10	350	0.039		0.435
11	350	0.060		0.479
12	350	0.108		1.284
13	580	0.039		0.394
`4	580	0.060	0.35	0.465
15	580	0.108		1.119
16	920	0.039		0.387
17	920	0.060		0.460
18	920	0.108		1.206

19	350	0.039		0.448
20	350	0.060		0.508
21	350	0.108		1.350
22	580	0.039		0.444
23	580	0.060	0.5	0.480
24	580	0.108		1.121
25	920	0.039		0.507
26	920	0.060		0.570
27	920	0.108		2.360

3.1. The Effect of Feed rate Speed on Surface Roughness

As show in Fig 4, Fig 5, Fig 6. shows the relationship between feed rate versus surface roughness (Ra) for different depth of cut at spindle speed 350 rpm, 580 rpm, and 920 rpm respectively during the dry turning process. At the feed rate increases from lower 0.039mm/rev to higher 0.108 mm/rev, the surface roughness also increases. This trend is consistent across all spindle speeds, indicating that higher feed rates result in rougher surfaces. The feed rate (f) has a greater effect than spindle speed (N) on surface roughness (R_a). Khanna et al.[13] also observed that an increase in the feed rate caused the increment of surface roughness due to the feed marks effect. Tefera et al [7] has report that the roughness of the turned surface decreased as the feed was increased. Bouacha et al. [14] reported that the feed rate strongly effects the surface roughness due to its primary function, known as the theoretical geometrical surface roughness. Yousefi et al.[15] reported that when the feed rate increases, the surface roughness average increases. In addition, for this dry turning process perform with tool nose radius r=1.2mm constantly, thus the result can be explained by the equation of surface roughness (R_a): $R_a = f^2/32r$, where f is the feed rate [9]. It clearly reveals that the surface roughness become better with increase of feed rate and with a constant of tool nose radius. The result show that for many levels of depth of cut, spindle speed, constant of tool nose radius of 1.2mm, the best surface roughness is achieved at the lowest level of feed rate.

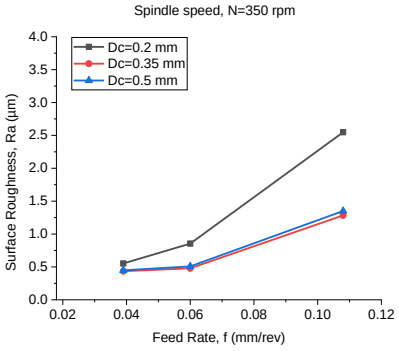


Fig. 4. Effect of feed rate on surface roughness at spindle speed, N=350 rpm

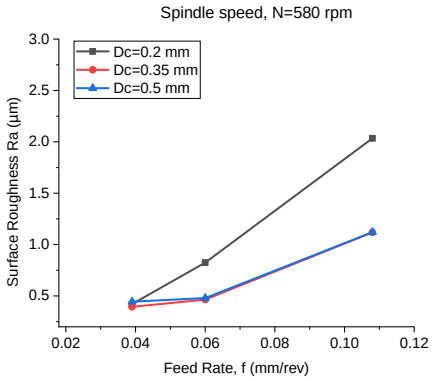


Fig. 5. Effect of feed rate on surface roughness at spindle speed, N=580 rpm

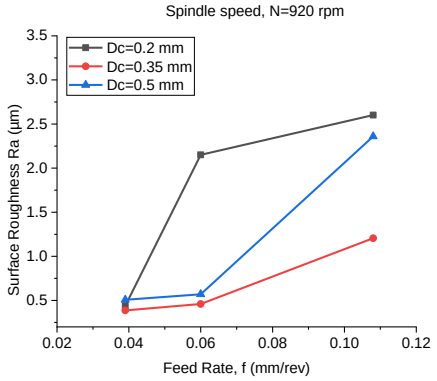


Fig. 6. Effect of feed rate on surface roughness at spindle speed, N=920 rpm.

3.2 The Effect of Spindle Speed on Surface Roughness

As show in Fig. 7, Fig. 8, Fig. 9, the relationship between spindle speed (N) and surface roughness (R_a) for different depths of cut (0.2 mm, 0.35 mm, and 0.5 mm) respectively. The general trend is that the values of surface roughness (R_a) slightly changed with the change of spindle speed (N). For the feed rate of 0.108 mm/rev, the surface roughness (R_a) slightly decreased with the increase in the spindle speed range from 350 to 580 rpm and then increase at a spindle speed of 920rpm.These results have a similar trend compared with the study by Sophal Hai [11].

From Fig 8, the feed rate 0.060mm/rev, surface roughness gradually decreased with an increase in the spindle speed range from 350rpm to 920rpm. However, from the lowest feed rate 0.039mm/rev the surface roughness (R_a) gradually decreased with an increase in the spindle speed (N). Youselfi et al. [15] has reported that, This behavior is related to the decreasing of the cutting forces at higher cutting speeds. Sophal et al [11] has report that the rise of the spindle speed generated the linear decrease in surface roughness due to the cutting force reduced at a higher cutting speed. The thermal softening of the material occurred at a high cutting speed and then led to a decrease in cutting force. Thus, it results in a

good surface roughness. However, in the case of using ceramic and coated carbide inserts, the surface roughness increased with an increase in the cutting speed due to the built-up edge formation (BUE) that occurred during the turning test.

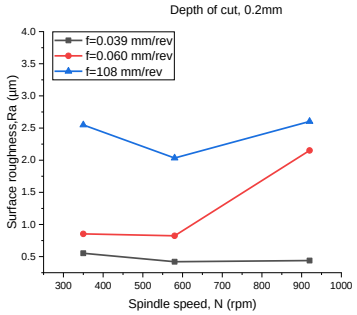


Fig. 7. Effect of spindle speed on surface roughness at depth of cut, Dc= 0.2mm.

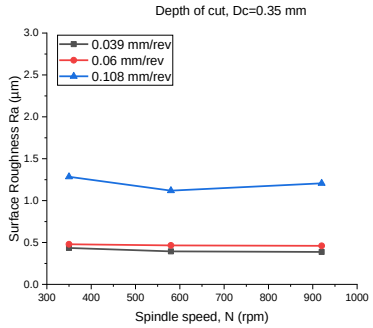


Fig. 8. Effect of spindle speed on surface roughness at depth of cut, Dc= 0.35mm.

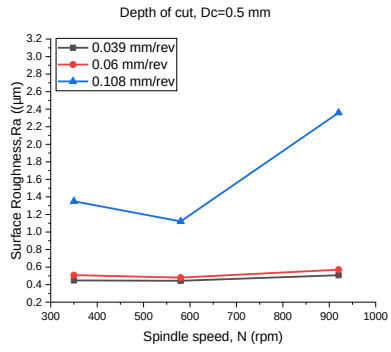


Fig. 9. Effect of spindle speed on surface roughness at depth of cut, Dc=0.5mm

3.3 The Effect of Depth of Cut on Surface Roughness

As show in Fig. 10, Fig. 11, Fig. 12, During dry turning, the relationship between the depth of cut with surface roughness for copper (UNS C10700) at the lower-higher spindle speed. The general trend is that the values of surface roughness (R_a) slightly changed with the change of depth of cut. From the lower feed rate of 0.039mm/rev, the surface roughness decreases with an increase of depth of cut range from 0.2mm to 0.35mm. However, the surface roughness increases with an increase of depth of cut range from 0.35mm to 0.5mm. Terfera et al. [7] reveal that the surface roughness increased as the depth of cut increased because of more material removal with greater cutting forces at a deeper depth of cut. Youselfi et al. [15]. The obtained results show that the surface roughness is decreased about 9% by increasing the cutting depth to 2 mm. Further increasing the cutting depth to about 3mm leads to the surface roughness deterioration. This surface roughness deterioration is mainly due to the chatter phenomenon, which occurs at the high level of the cutting depths.

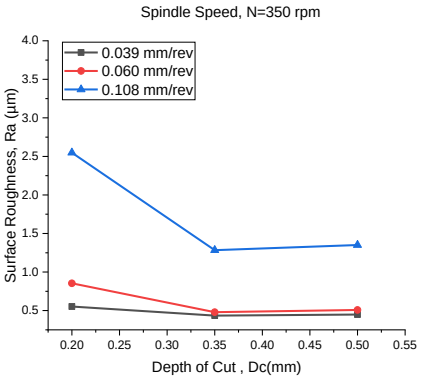


Fig. 10. Effect of depth of cut on surface roughness at spindle speed, N=350rpm.

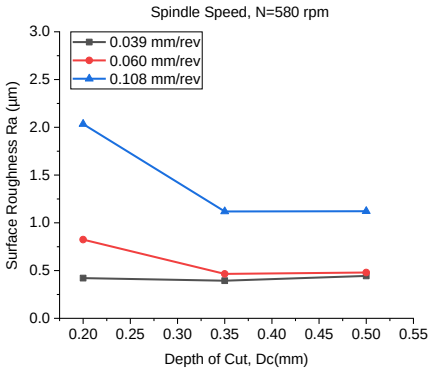


Fig 11. Effect of depth of cut on surface roughness at spindle speed, N=580 rpm.

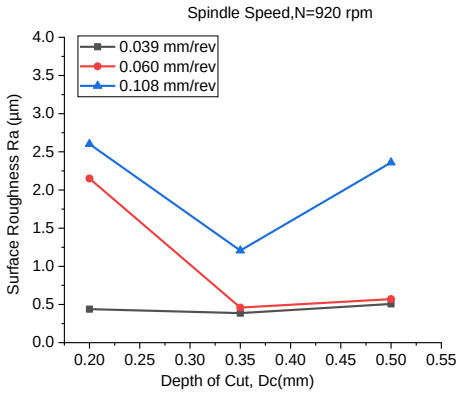


Fig. 12. Effect of depth of cut on surface roughness at spindle speed, N=920 rpm.

3.4 ANOVA and S/N ratio

The data collected were analyzed using analyses of variance (ANOVA) and S/N ratio with surface finish as the response variable and spindle speed N, feed rate f and depth of cut Dc as cutting parameters. The ANOVA model was developed to study the main effects of the independent variables given in Table 5. Magdum et al. [12] The significance level was based on the P-value from ANOVA as:

Insignificant if P > 0.05

Significant if P < 0.05

Table 5 shows the ANOVA of surface roughness (Ra) for the combination of cutting parameters N, f and Dc (CNFD). Based on Table 4.2, the feed rate (f) and depth of cut (Dc) were revealed as significant effect parameters for surface roughness (R_a), while the spindle speed (N) is an insignificant parameter.

Table 5. ANOVA of surface roughness (Ra) for the combination of cutting parameters N, *f*, Dc (CNFD).

Material	Source	DF	Adj SS	Adj MS	F-Value	P-Value	Remarks
Copper (UNS C10700)	*N* (rpm)	2	0.6566	0.3283	2.41	0.115	Insignificant
	f (mm/rev)	2	8.1561	4.0780	29.99	0.000	Significant
	Dc (mm)	2	2.3095	1.1547	8.49	0.002	Significant
	Error	20	2.7193	0.1360	-	-	-
	Total	26	13.8414	-	-	-	-

The response table for this study's S/N ratios is shown in Table 5, for smaller is better. Table 5 presents that the best combination of parameters for smaller is the better characteristic. The Delta shows the scale of effect on surface roughness, which is calculated from the highest and lowest value of each cutting parameter level. The rank is used to indicate which cutting parameters have a highly effect on the surface roughness (R_a). The feed rate (rank1) performs as the highly effected on surface roughness (R_a), the second one is depth of cut (rank 2) and spindle speed have lower effect (rank 3) on surface roughness, which can be observed from the rank table row. The highest S/N ratio value of the feed rate (f), spindle speed (N) and depth of cut (Dc) are illustrated in level 1, level 2 and level 2, respectively. Therefore, the optimal cutting condition during the dry turning process for copper is the following combination of parameters: level 1 of feed rate (f), level 2 of spindle speed (N) and level 2 of depth of cut (Dc). The S/N ratio values revealed in Figure 13 are summarized in Table 6.

Table 6. S/N ratio response table for surface roughness (Ra) with the CNFD

	Cutting Parameters		
Levels	Spindle Speed, *N*(rpm)	Feed Rate f (mm/rev)	Depth of cut, *Dc* (mm)
1	2.2354	7.0385	-0.7090
2	3.2502	3.7033	4.2880
3	0.9860	-4.2703	2.8926
Delta	2.2642	11.3088	4.9970
Rank	3	1	2

Fig 13 shows the graphic of the level values given in Table 6. Therefore, interpretations may be made according to the level values of Spindle Speed, N(rpm), Feed Rate f (mm/rev), and Depth of cut, Dc (mm) factors given in Table 6 and Fig 13 in determining optimum cutting conditions of experiments to be conducted under the same conditions. The average S/N ratio for every level of experiment is calculated based on the recorded value as shown in Table 6. The different values of S/N ratio between maximum and minimum are (main effect) also shown in Table 6. The feed rate and the depth of cut are two factors that have the highest difference between values, 11.3088 and 4.9970 respectively. Based on the Taguchi prediction that the larger different between value of S/N ratio will have a more significant effect on surface roughness (R_a). Thus, it can be concluded that increasing the feed rate will increase the R_a significantly and also the depth of cut. The results of data analysis of S/N ratio for R_a values, which are calculated by Taguchi method, are shown in Table 6 (where df is degree of freedom, F variance ratio, and P significant factor). Thus, it is seen in Fig 13 and Table 6 that, the first level of feed rate factor and second level of depth of cut factor and second level of spindle speed, factor are higher. Consequently, the optimum cutting conditions determined under the same conditions for the experiments to be

conducted will be 0.039 mm/rev for the feed rate, 580rpm for spindle speed, and 0.35mm for the depth of cut.

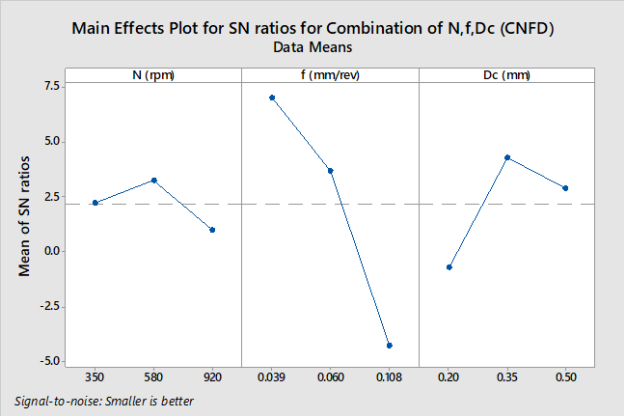


Fig. 13. Main effect plot for average S/N ratios versus cutting parameters for surface roughness in the combination of N, *f*, Dc.

4. CONCLUSIONS

In this research, the impact of the cutting parameters on the surface roughness of high pure copper UNS C10700 with uncoated carbide insert was studied experimentally. The statistical analysis (ANOVA & S/N ratio) were used to investigate and analyze the cutting parameters to improve the machined surface quality. The conclusions of the research were summarized as follows:

The change of feed rate and depth of cut revealed a strong change of surface roughness. However, surface roughness slightly changed with the change of spindle speed. Thus, the feed rate and depth of cut are revealed as cutting parameters that have a highly effect on surface roughness.

From the analysis of variance (ANOVA) of the feed rate (f) and depth of cut (Dc) were revealed as significant effect parameters for surface roughness (Ra), while the spindle speed (N) is an insignificant parameter.

The S/N ratio of analysis with the feed rate (f) and depth of cut (Dc) indicated that the cutting parameters for optimum surface roughness are a combination of a spindle speed of 580 rpm and a feed rate of 0.039 mm/rev for depth of cut 0.35mm.

ACKNOWLEDGMENTS

You may place your acknowledgments here. Write only essential acknowledgments.

REFERENCES

[1] T. Shigemori and K. Sawa, "Characteristics of carbon and copper flat commutator on DC motor for automotive fuel pump," in *Proceedings of the 50th IEEE Holm Conference on Electrical Contacts and the 22nd International Conference on Electrical Contacts Electrical Contacts, 2004.*, 2004: IEEE, pp. 523-527.

[2] S. Płonka, "The influence of selected parameters of finish turning of copper on the surface roughness," *Advances in Manufacturing Science and Technology*, vol. 31, no. 4, pp. 7-19, 2007.

[3] A. Torres, I. Puertas, and C. Luis, "Surface roughness analysis on the dry turning of an Al-Cu alloy," *Procedia engineering*, vol. 132, pp. 537-544, 2015.

[4] V. D. Patel and A. H. Gandhi, "INVESTIGATION OF EFFECT OF CUTTING PARAMETERS AND TOOL NOSE RADIUS ON CUTTING FORCES AND SURFACE ROUGHNESS IN FINISH HARD TURNING OF AISI D2 STEEL WITH CBN TOOL."

[5] M. N. Islam and B. Boswell, "An investigation of surface finish in dry turning," in *Proceedings of the world congress on Engineering*, 2011, vol. 1: WCE 2011, London, UK, pp. 6-8.

[6] V. Sahithi, T. Malayadrib, and N. Srilatha, "Optimization of turning parameters on surface roughness based on Taguchi technique," *Materials Today: Proceedings*, vol. 18, pp. 3657-3666, 2019.

[7] A. G. Tefera, D. K. Sinha, and G. Gupta, "Experimental investigation and optimization of cutting parameters during dry turning process of copper alloy," *Journal of Engineering and Applied Science*, vol. 70, no. 1, p. 145, 2023.

[8] K. Mahajan, M. Sadaiah, and S. Gawande, "Experimental investigations of surface roughness on OFHC copper by diamond turning machine," *International Journal of Engineering Science and Technology*, vol. 2, no. 10, pp. 5215-5220, 2010.

[9] M. P. Gloover, *Fundamentals of modern manufacturing materials, processes, and system* (Printed in the United States of America).

[10] AZoNetwork. "Material selected." https://www.azom.com/article.aspx?ArticleID=6307 (accessed Nov 15 2012.

[11] S. Hai, H.-C. Jung, W.-H. Shim, and H.-G. Shin, "Investigation of surface quality for minor scale diameter of biodegradable magnesium alloys during the turning process using a different tool nose radius," *Metals*, vol. 11, no. 8, p. 1174, 2021.

[12] R. V. B. Magdum and V. R. Naik, "Experimental Study of the Effects of Machining Parameters on the Surface Roughness in the Turning Process," *vol*, vol. 4, pp. 1-8, 2018.

[13] N. Khanna, N. M. Suri, P. Shah, H. Hegab, and M. Mia, "Cryogenic turning of in-house cast magnesium based MMCs: a comprehensive investigation," *Journal of Materials Research and Technology*, vol. 9, no. 4, pp. 7628-7643, 2020.

[14] K. Bouacha, M. A. Yallese, T. Mabrouki, and J.-F. Rigal, "Statistical analysis of surface roughness and cutting forces using response surface methodology in hard turning of AISI 52100 bearing steel with CBN tool," *International Journal of Refractory Metals and Hard Materials*, vol. 28, no. 3, pp. 349-361, 2010.

[15] S. Yousefi and M. Zohoor, "Effect of cutting parameters on the dimensional accuracy and surface finish in the hard turning of MDN250 steel with cubic boron nitride tool, for developing a knowledged base expert system," *International Journal of Mechanical and Materials Engineering*, vol. 14, pp. 1-13, 2019.

[16] R. Singal, M. Singal, and R. Singal, *Fundamentals of machining and machine tools*. IK International Publishing House Pvt. Limited, 2008.

[17] S. Dutta and S. K. R. Narala, "Optimizing turning parameters in the machining of AM alloy using Taguchi methodology," *Measurement*, vol. 169, p. 108340, 2021.

[18] I. Mohsin, K. He, Z. Li, F. Zhang, and R. Du, "Optimization of the polishing efficiency and torque by using Taguchi method and ANOVA in robotic polishing," *Applied Sciences*, vol. 10, no. 3, p. 824, 2020.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Investigating Efficiency of PMSM Using Full State Feedback Control with integral action and Dynamic Compensation

Bunsrung TAING [1*], Sarot SRANG [2], Sokserey SREY [2]

[1] *Faculty of Mechatronics, Information and Communication Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: bunsrung_taing@gsc.itc.edu.kh*

Abstract: Traditional control strategies for Permanent Magnet Synchronous Motors (PMSMs) often struggle to balance dynamic performance and energy efficiency under variable operational demands. This study proposes a Full State Feedback Control (FSFB) system augmented with integral action and dynamic compensation to address these challenges. By deriving a state-space model of the PMSM in dq reference frames and designing a controller to regulate direct-axis current i_d and mechanical speed ω_m, the system achieves rapid transient response, near-zero steady-state error, and robust disturbance rejection. MATLAB/Simulink simulations, validated using real-world motor parameters, demonstrate exceptional performance: at 250 rpm and 500 rpm, i_d converges to zero within seconds, ω_m tracks reference inputs smoothly, and efficiency stabilizes above 73%, peaking at 76% under high-speed conditions. The results highlight the controller's ability to minimize energy losses while maintaining precision, offering a promising solution for electric agricultural machinery and other applications requiring high efficiency. Future work will explore real-time implementation and adaptive enhancements for dynamic environments.

Keywords: Permanent Magnet Synchronous Motor (PMSM), Full State Feedback Control (FSFB), PMSM efficiency, PMSM Control, Power of PMSM.

1. INTRODUCTION

Permanent Magnet Synchronous Motors (PMSMs) have evolved significantly since their early development in the mid-20th century, marked by the introduction of high-energy magnets like Alnico, which enabled more efficient motor designs, mentioned Krishnan et al. [1]. Although early adoption was limited by control and power electronic technologies, PMSMs have become increasingly prevalent in electric vehicles due to their high efficiency, torque capability, and power density. The 2000s marked a turning point, as global efforts to reduce emissions and fossil fuel dependence led to widespread investment in electric propulsion, with PMSMs emerging as a preferred motor type. Notable applications include hybrid vehicles like the Toyota Prius, which leveraged PMSM technology for improved efficiency and lower emissions, said Lu et al. [2].

To maximize PMSM performance in dynamic environments such as agriculture, Full State Feedback (FSFB) control with integral action is implemented to enhance stability, responsiveness, and disturbance rejection. Cheema and Fletcher et al. [3] demonstrated that FSFB enables superior flux and force control with minimal steady-state error across varying speeds.

This thesis aims to model and implement a FSFB control system with integral action for a PMSM, targeting improved dynamic behavior and optimal energy usage under various speed inputs. The study is conducted entirely in MATLAB/Simulink and focuses on regulating the direct-axis current i_d and mechanical speed ω_m. Performance is evaluated based on speed tracking, current response, and efficiency. Motor parameters are adopted from Lu et al. [2] to ensure alignment with real-world validated data.

[1*] Corresponding author: Bunsrung TAING
E-mail: busnrung_taing@gsc.itc.edu.kh; Tel: +855-10234381

2. METHODOLOGY

2.1 Mathematical Modeling of PMSM

The equivalent circuit of a three-phase PMSM in the stationary "abc" reference frame offers a simplified model that captures the motor's essential electrical characteristics. In this circuit, the stator windings are typically represented as three independent phases labeled as a, b, and c. [2]

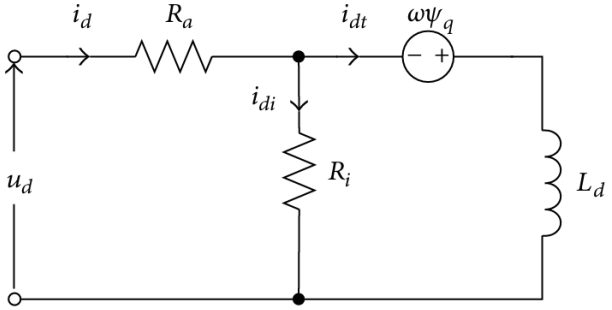


Fig. 1. The d-axis equivalent circuit of synchronous motor.

2.1.1 Electrical equations of PMSM in "abc" stationary Frame

The voltage v, of the stator winding, which is the external supplied voltage, can be defined as the sum of voltage drop across the winding resistance ri, and the induced voltage, or back emf, which is due to the time-varying flux linkage, $\frac{d}{dt}\lambda$.[3]

Therefore, the voltage equation could be written as [4]:

$$v_{abc} = R_s i_{abc} + \frac{d}{dt} L_s i_{abc} + \frac{d}{dt}\lambda_{mabc} \qquad \text{(Eq. 1)}$$

Where:

- λ_{mabc} is the flux established on the stator winding by the permanent magnet associated with the three phases abc.
- $L_s = \begin{bmatrix} l_s & 0 & 0 \\ 0 & l_s & 0 \\ 0 & 0 & l_s \end{bmatrix}$ is the self-inductance of the stator winding matrix.
- $R_s = \begin{bmatrix} r_s & 0 & 0 \\ 0 & r_s & 0 \\ 0 & 0 & r_s \end{bmatrix}$ is the stator resistance matrix.

2.1.2 Mechanical equation of PMSM in "abc" stationary Frames

To complete the dynamic equation of the motor, the mechanical equation based on Newton's second law of motion is integrated into the PMSM model[4].

$$\frac{Jd\omega_m}{dt} = T_e - B_v\omega_m - T_L \qquad \text{(Eq. 2)}$$

$$\frac{d\theta_m}{dt} = \omega_m \qquad \text{(Eq. 3)}$$

Where T_e is the developed electromagnetic torque, T_L is the load torque, B_v is the viscous friction (or damping) coefficient, J is the inertia of the rotor. The relationship between the electrical and mechanical angular speed is:

$$\omega_e = \frac{P}{2}\omega_m \qquad \text{(Eq. 4)}$$

Where P is the number of poles, ω_e is the electrical angular velocity and ω_m is the mechanical angular velocity.

According to Yusivar et al. [4], due to independence with θ_e, the derivative of both the inductance matrix L_s and ω_m are zero. The electromagnetic torque can be obtained as follows:

$$T_e = \frac{P}{2}\lambda_m[\, i_a \;\; i_b \;\; i_c] \begin{bmatrix} -sin\theta_e \\ \frac{1}{2sin\theta_e} + \frac{\sqrt{3}}{2}cos\theta_e \\ \frac{1}{2sin\theta_e} - \frac{\sqrt{3}}{2}cos\theta_e \end{bmatrix} \qquad \text{(Eq. 5)}$$

2.1.3 Reference Frame conversion

Clarke's transformation and the rotational Park transformation are fundamental concepts in the analysis and control of electrical machines, particularly in power electronics and motor drives. Applying these two transformations sequentially will map the "abc" coordinate frame, which is the stationary frame, to "dq" coordinate frame, which is the rotational frame, as shown in **Fig. 2**.

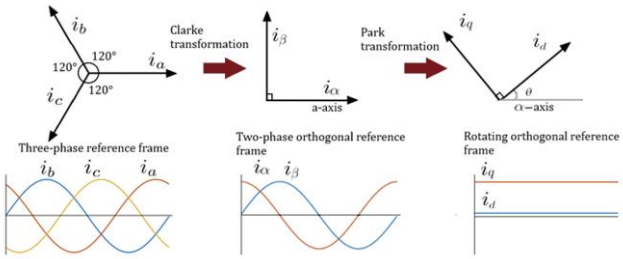


Fig. 2. The three reference frames.[1]

After applying the transformations [4],

$$v_d = r_s i_d + L_d \frac{di_d}{dt} - \frac{L_q \omega_m i_q P}{2} \quad \text{(Eq. 6)}$$

$$v_q = r_s i_q + L_q \frac{di_q}{dt} + \frac{L_d \omega_m i_d P}{2} + \frac{\lambda_m \omega_m P}{2} \quad \text{(Eq. 7)}$$

Performing the linearization by introducing new variables:

$$v_{d0} = -\frac{P}{2}\omega_m L i_q \quad \text{(Eq. 8)}$$

$$v_{q0} = \frac{P}{2}\omega_m L i_d \quad \text{(Eq. 9)}$$

$$V_d = v_d - v_{d0} \quad \text{(Eq. 10)}$$

$$V_q = v_q - v_{q0} \quad \text{(Eq. 11)}$$

So, (Eq. 6) and (Eq. 7) become

$$\dot{\imath_d} = \frac{1}{L}V_d - \frac{r_s}{L}\mathrm{i_d} \quad \text{(Eq. 12)}$$

$$\dot{\imath_q} = \frac{1}{L}V_q - \frac{r_s}{L}i_q - \frac{P\omega_m\lambda_m}{2L} \quad \text{(Eq. 13)}$$

In most cases, manufacturers use surface-mounted permanent magnet synchronous motors(PMSM) [4],

So,

$$L_d = L_q = L \quad \text{(Eq. 14)}$$

After frame conversion into "dq"

$$T_e = \frac{3}{4}P\left[\lambda_m i_q + (L_d - L_q) i_q i_d\right], \quad \text{(Eq. 15)}$$

$$L_d - L_q = 0$$

$$T_e = \frac{3P}{4}\lambda_m i_q \quad \text{(Eq. 16)}$$

$$\rightarrow \quad \frac{d\omega_m}{dt} = \frac{1}{J}\left(\frac{3P}{4}\lambda_m i_q - T_L - B_v\omega_m\right) \quad \text{(Eq. 17)}$$

2.2 Full State Feedback (FSFB) with integral action Controller Design

The general state space differential can be represented:

$$\dot{X} = AX + Bu + Ew \quad \text{(Eq. 18)}$$

$$Y = CX + Du \quad \text{(Eq. 19)}$$

Where, X is the matrix of states, w is the disturbance, u is the control input, Y is the output vector, A is an $n \times n$ constant matrix, B is an $n \times m$ constant matrix, C is a constant matrix, D is a constant(scalar), n is the number of states and m is the number of control inputs.

Using (Eq. 12) (Eq. 13) and (Eq. 17), the system can be reformulated in state-space representation as

$$\begin{bmatrix} \dot{\imath_d} \\ \dot{\imath_q} \\ \dot{\omega_m} \end{bmatrix} = \begin{bmatrix} \frac{-r_s}{L} & 0 & 0 \\ 0 & \frac{-r_s}{L} & \frac{-\lambda_m P}{2L} \\ 0 & \frac{3P\lambda_m}{4J} & -\frac{B}{J} \end{bmatrix} \begin{bmatrix} I_q \\ I_d \\ \omega_m \end{bmatrix} + \begin{bmatrix} \frac{1}{L} & 0 \\ 0 & \frac{1}{L} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_q \\ V_d \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ -\frac{1}{J} \end{bmatrix} T_L \quad \text{(Eq. 20)}$$

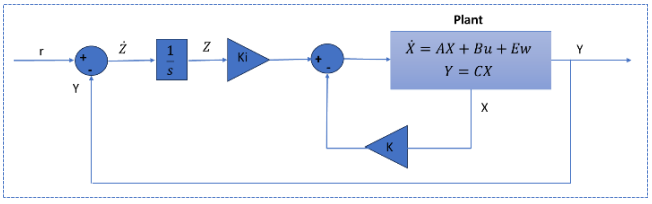


Fig. 3. State feedback with integral control structure.

Let $\dot{Z} = r - Y$, where r is the reference or desired

$$\dot{Z} = r - CX \quad \text{(Eq. 21)}$$

A new state-space equation is obtained.

$$\begin{bmatrix} \dot{X} \\ \dot{Z} \end{bmatrix} = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix} \begin{bmatrix} X \\ Z \end{bmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} u + \begin{bmatrix} 0 \\ I \end{bmatrix} r \quad \text{(Eq. 22)}$$

Where I is the identity matrix.

$$\dot{X}_i = A_i X_i + B_i u + E_i r \quad \text{(Eq. 23)}$$

Let $X_i = \begin{bmatrix} X \\ Z \end{bmatrix}$, $A_i = \begin{bmatrix} A & 0 \\ -C & 0 \end{bmatrix}$, $B_i = \begin{bmatrix} B \\ 0 \end{bmatrix}$ and $E_i = \begin{bmatrix} 0 \\ I \end{bmatrix}$

Assumption: All states are known

Let $K = [K_x \quad -K_i]$, and the control input $u = -KX_i$. Substituting to the equation (Eq. 23)

$$\dot{X}_i = (A_i - B_i K)X_i \qquad \text{(Eq. 24)}$$

The behavior of the closed-loop system is governed by $A_{cl} = (A - BK)$. The eigenvalues λ, of A_{cl} can be positioned arbitrarily by selecting suitable negative values, ensuring that the system converges.[5]

Finally, the desired poles λ can be selected and the values of K can be obtained from the equation below:

$$\det(A_{cl} - \lambda I) = 0 \qquad \text{(Eq. 25)}$$

The solution for K can also be obtained using MATLAB's $place()$ function, which applies the Kautsky–Nichols–Van Dooren (KNVD) algorithm to numerically compute the gain matrix K.

2.3 Performance Evaluation Criteria

Fig. 4 illustrates the power flow diagram of the PMSM during driving mode. As shown, power losses occur at various stages right from the input due to multiple contributing factors.

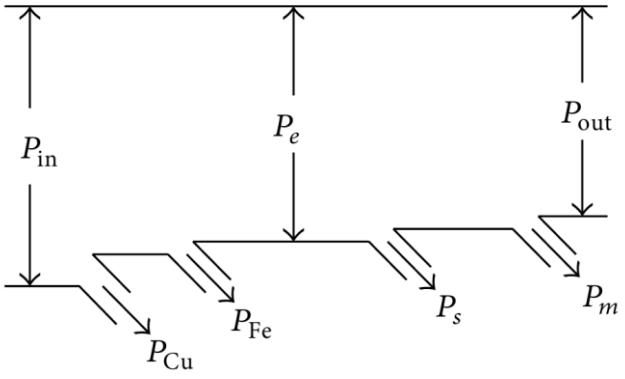


Fig. 4. Power flow diagram of the PMSM.

After implementing current and speed control, it is essential to evaluate the motor's energy performance by determining its power and efficiency. As stated in Lu et al. [2], the power can be calculated using the following formula:

$$P_{in} = 1.5(v_d i_d + v_q i_q) \qquad \text{(Eq. 26)}$$

$$P_{out} = P_e - P_m - P_s \qquad \text{(Eq. 27)}$$

Where P_{in} is the power input for the motor, P_e is the electromagnetic power, P_m is the mechanical loss, P_s is the stray loss in this case P_s is neglectable and P_{out} is the output power or mechanical power.

Finally, the efficiency can be calculated using formula:

$$efficiency = \frac{P_{out}}{P_{in}} \times 100 \qquad \text{(Eq. 28)}$$

2.4 MATLAB/Simulink Implementation

To evaluate and validate the theoretical models and control strategies discussed, MATLAB/Simulink is employed as the primary simulation environment. This section outlines the implementation of the permanent magnet synchronous motor (PMSM) model, the control algorithms, and the associated power and efficiency calculations. By simulating motor behavior under various operating conditions, the system's performance including torque response, power losses, and efficiency can be quantitatively assessed. The simulation also enables comparative analysis between different desired step speed inputs, such as 250 rpm and the maximum rated speed at 500 rpm.

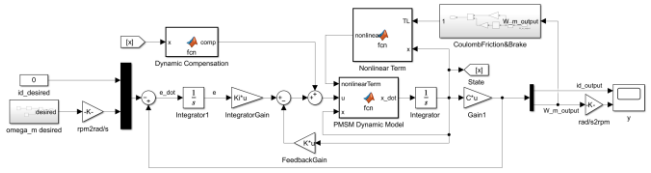


Fig. 5. i_d and ω_m Full State Feedback with integral action control and Dynamic Compensation.

The **Fig. 5** illustrates the implementation of full state feedback control targeting two specific desired variables: i_d and ω_m. In this configuration, the nonlinearities and disturbance terms that were initially eliminated during the linearization process are reintroduced into the PMSM dynamic model. The feedback gain plays a critical role in ensuring system convergence, while the integral gain further enhances the system's ability to accurately track and stabilize at any given reference input.

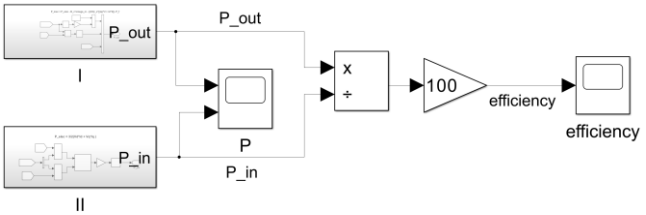


Fig. 6. Power and efficiency calculation.

In this setup, the motor parameters are adopted from the experimental study conducted by Lu et al. [2]. The specific values are presented in **Table 1**.

Table 1. Specifications of the surface permanent magnet synchronous motor.

Motor Parameters	Values
Rated phase voltage(V)	21.6
Rated phase current(A)	30
Rated speed(rpm)	500
Rated torque(N·m)	30
Pole pairs	23
Armature resistance(Ω)	23
d-axis inductance(H)	7.6e-5
q-axis inductance(H)	7.6e-5
Magnet flux-linkage (Wb)	0.0204
Iron consumption resistance(Ω)	$0.006\omega + 1.5$

Where ω is the electrical angular velocity.

3. RESULTS AND DISCUSSION

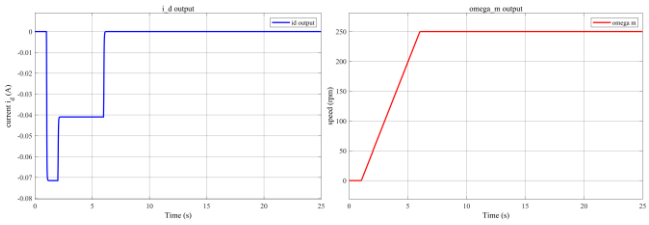


Fig. 7. The output of current i_d and ω_m for the desired speed at 250 rpm.

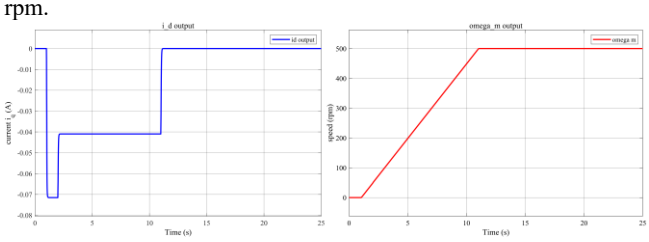


Fig. 8. The output of current i_d and ω_m for the desired max speed 500 rpm.

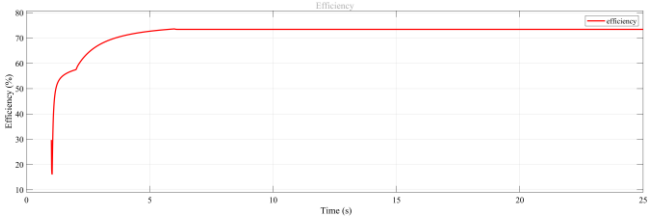


Fig. 9. Efficiency of PMSM for 250 rpm desired speed.

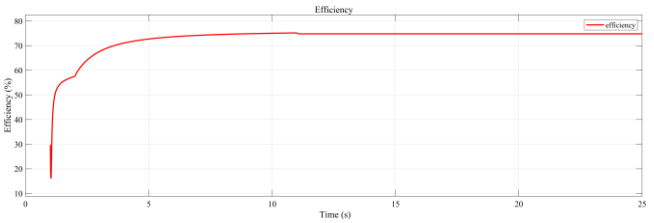


Fig. 10. Efficiency of PMSM for max speed (500 rpm).

- Case 1: Desired speed 250 rpm
  In **Fig. 7**, the motor operates at a speed of 250 rpm. The current i_d quickly stabilizes to near zero after a minor transient, indicating effective decoupling and control. The motor speed rises smoothly to the target value within the first six seconds and then maintains a stable value. Correspondingly, the efficiency in **Fig. 9** starts from around 29%, peaks near 74%, and remains steady around 73.5%, demonstrating that the control system ensures high efficiency at rated conditions.
- Case 2: Desired speed 500 rpm
  In **Fig. 8**, the motor reaches the target speed of 500 rpm smoothly within 12 seconds. The current i_d shows an initial transient spike close to 0.07 A, followed by rapid settling to zero, indicating an effective transient response and minimal steady-state error.
  The corresponding efficiency plot in **Fig. 10** demonstrates outstanding system performance. The efficiency rapidly rises to 76% and remains steady at 75% throughout the operation, reflecting very low losses and excellent power conversion. This implies that at 500 rpm, the PMSM operates near its optimal performance range under the given control strategy.

3. CONCLUSIONS

In conclusion, this study successfully implemented a Full State Feedback (FSFB) control strategy with integral action and dynamic compensation for a Permanent Magnet Synchronous Motor (PMSM), aiming to evaluate its dynamic performance and efficiency under varying speed conditions. Through MATLAB/Simulink simulation, the system demonstrated excellent transient and steady-state responses across different desired speeds, including 250 rpm and 500 rpm setpoints.

The simulation results revealed that the FSFB control approach, complemented by integral compensation, ensured rapid convergence of the current i_d and speed ω_m to their desired values with minimal overshoot and very small steady-state error. Furthermore, the efficiency analysis showed that

the system maintained high operational efficiency reaching up to 73-75% shortly after startup and remaining stable thereafter. These findings confirm the controller's robustness and its ability to maintain energy-efficient operation across different load conditions.

Overall, the proposed control scheme demonstrates a viable solution for improving the energy efficiency of PMSM-based electric drivetrains, particularly in applications requiring precision, such as electric agricultural machinery. Future work may explore real-time implementation, incorporate load disturbances or noise, and investigate adaptive or nonlinear control strategies to enhance system performance in more dynamic environments.

ACKNOWLEDGMENTS

I would like to express sincere gratitude to Dr. Sarot SRANG, research advisor, for his invaluable guidance, encouragement, and technical insight throughout the development of this study. Special thanks also go to Mr. Sokserey SREY, co-advisor, for his continuous support, constructive feedback, and mentorship during the research process. I am gratefully acknowledging the financial support provided by The University of Sydney, which made this research possible.

REFERENCES

[1] R. Krishnan, *Permanent Magnet Synchronous and Brushless DC Motor Drives*, 1st ed. CRC Press, 2017. doi: 10.1201/9781420014235.

[2] D. Lu, M. Ouyang, J. Gu, and J. Li, "Optimal Velocity Control for a Battery Electric Vehicle Driven by Permanent Magnet Synchronous Motors," *Mathematical Problems in Engineering*, vol. 2014, no. 1, Jan. 2014, doi: 10.1155/2014/193960.

[3] M. A. M. Cheema and J. E. Fletcher, *Advanced Direct Thrust Force Control of Linear Permanent Magnet Synchronous Motor*. in Power Systems. Cham: Springer International Publishing, 2020. doi: 10.1007/978-3-030-40325-6.

[4] F. Yusivar, N. Hidayat, R. Gunawan, and A. Halim, "Implementation of field oriented control for permanent magnet synchronous motor," in *2014 International Conference on Electrical Engineering and Computer Science (ICEECS)*, Kuta, Bali, Indonesia: IEEE, Nov. 2014, pp. 359–362. doi: 10.1109/ICEECS.2014.7045278.

[5] K. Ogata, *Modern control engineering*, 5th ed. Prentice Hall, 2022.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Predictive Modeling of Garment Trade Flows in Cambodia: A Data-Driven Approach to Forecasting Export Trends

Dara Em [1], Vesal KHEAN [2], Sokkhey PHAUK [2*]

[1] *Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: phauk.sokkhey@itc.edu.kh*

Abstract: *The Garment sector has been the backbone of the Cambodian economy for more than two decades. This sector is the largest export and constitutes the bulk of its manufacturing sector. More people are employed in the garment sector than any other industrial sector in Cambodia and, as a result, Cambodia's rapid growth in attracting garment factories have led to rapid progress in alleviating poverty. The purpose of this study is to create a predictive modeling framework that will use a data-driven method to estimate Cambodia's garment trade flows, and giving factory and government useful information. Optimizing trade tactics and identifying important indicator in economic that impacting export patterns are the main goals. Data engineering is the first step in this organized workflow technique, where raw datasets from trade databases, economic indicators, and the globe market trends are preprocessed and cleaned for analysis. Use exploratory data analysis (EDA) and visualization know patterns, correlations, and trends within the dataset, and forming the foundation for predictive model development. To forecast export patterns, machine learning algorithms such as ensemble approaches, Support Vector Regression (SVR), XGBoost, and regression models are implemented. Historical trade data is used to train and validate these models, and accuracy is ensured by performance indicators like mean absolute error (MAE), Mean Absolute Erro (MAE), and R-squared. Finding practical insights and suggestions is the last stage. The study identifies the main economic factors affecting Cambodia's garment exports by examining model results. The results provide useful information for factory, trade analysts, and government organizations to enhance policymaking, maximize supply chain tactics, and reduce risks related to fluctuations in the global market.*

Keywords: Cambodia Garment Exports; Forecasting; Linear Regression (LR); Support Vector Regression (SVR); XGBoost; Economic Indicators; Machine Learning; Global Supply Chain; Inflation; Export Performance.

1. INTRODUCTION

The garment industry stands as a critical driver of Cambodia's economic development, serving as the country's largest export sector and primary source of industrial employment. Accounting for approximately 16% of GDP and employing over 800,000 workers, the sector has been instrumental in Cambodia's economic transformation since the 1990s [1]. The United States and European Union (EU) are Cambodia's two largest export markets for garments, together accounting for over 70% of total garment exports in 2023 [2].

This study seeks to address these gaps by pursuing three main objectives: First, to identify and analyze the most significant economic indicators affecting Cambodia's garment exports, including trade policies, labor market conditions, and global demand trends [2]. Second, to evaluate the predictive accuracy of various modeling approaches using historical trade data from 2010 to 2023 [3]. Third, to develop evidence-based recommendations for industry stakeholders and policymakers to enhance the sector's resilience and competitiveness [4].

The scope of this research encompasses quantitative analysis of Cambodia's garment export performance, with particular focus on major destination markets (EU and U.S.)

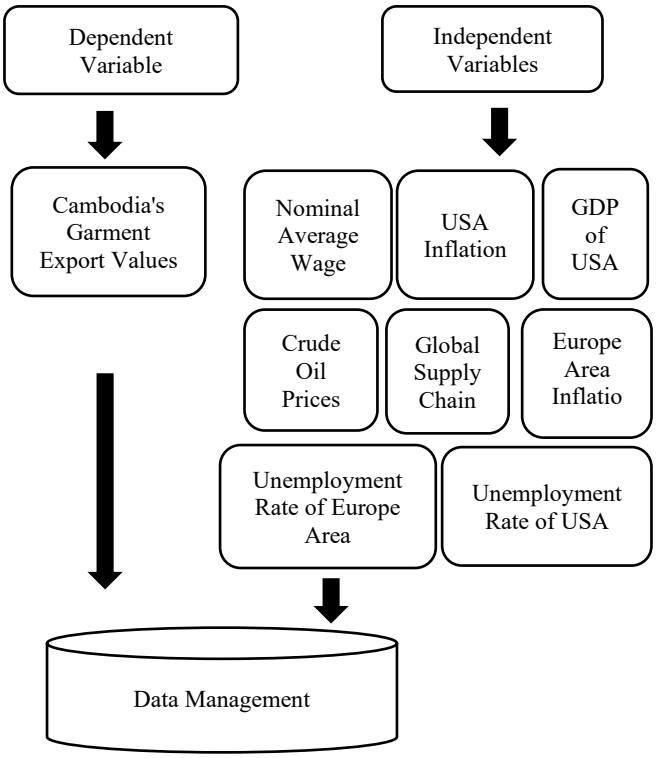
[1*] Corresponding author: Sounean Heng
E-mail: heng_sounean@itc.edu.kh; Tel: +855-12 311 110

and key competitive factors [2]. The study will employ both traditional statistical methods and machine learning techniques to provide comprehensive insights into trade flow dynamics [3][4]. Findings from this research are expected to contribute to more informed decision-making among manufacturers, exporters, and government agencies involved in Cambodia's vital garment sector [1][2].

2. METHODOLOGY

2.1 Data Sources and Variable Selection

Dataset Collation



Selecting an appropriate dataset for forecasting Cambodia's garment exports requires careful consideration of data availability, temporal coverage, and relevance to global economic drivers. Given the complexity of the garment trade, which is influenced by both domestic and international factors, it is essential to integrate a diverse range of economic variables. However, public datasets on Cambodian trade and labor markets are often fragmented, and comprehensive datasets on garment exports are typically limited. As a result, we compiled a robust, time-series dataset from authoritative sources to enhance the reliability and precision of our analysis. The dataset spans from 2010 to 2024 and includes monthly data with no missing values, ensuring the accuracy of the forecast model. The dataset includes the following variables:

Dependent Variable:

Cambodia's Garment Export Values (USD): This is the total monthly export value of garments from Cambodia, sourced from the International Trade Centre (ITC) Trade Map. This variable captures the primary outcome of interest—garment export performance—and is critical for understanding the export dynamics of Cambodia's key sector [5]

Independent Variables:

Nominal Average Wage (USD): Sourced from Cambodia's Ministry of Labour and Vocational Training (MLVT), this variable tracks the nominal wages within Cambodia's garment sector. Wages are an important determinant of production costs and competitiveness, as higher wages can increase the cost of labor and influence the price of Cambodian garments in global markets [6]. Global Supply Chain: While specific global supply chain metrics can vary, this variable captures broader global trends in trade flows, logistics, and the availability of raw materials. source from UNCTAD (United Nations Conference on Trade and Development) [7]. USA Inflation: Sourced from the U.S. Bureau of Labor Statistics, this variable reflects inflation in the U.S. economy, which impacts consumer purchasing power and demand for imported goods, including garments [8]. Europe Area Inflation: Sourced from Eurostat, this variable measures the inflation rate within the Eurozone. Like U.S. inflation, European inflation impacts demand for Cambodian exports, as higher inflation can reduce consumers' disposable income and alter buying patterns, especially in discretionary sectors such as apparel [9]. GDP of USA: The Gross Domestic Product (GDP) of the United States, sourced from the World Bank, is a critical indicator of the economic performance of the U.S., which is one of Cambodia's largest export markets. U.S. GDP growth reflects the health of the U.S. economy, directly influencing the purchasing power and demand for imported garments from Cambodia [10]. Unemployment Rate of USA: Sourced from the U.S. Bureau of Labor Statistics, this variable tracks the unemployment rate in the U.S., reflecting the labor market's health and indirectly signaling trends in consumer purchasing behavior. Higher unemployment can reduce consumer spending, affecting demand for Cambodian exports [11]. Unemployment Rate of Europe Area: Sourced from Eurostat, this variable track unemployment trends within the European Union. Unemployment rates are closely linked to economic conditions in Europe, impacting the purchasing power of consumers and, consequently, the demand for Cambodian garments in European markets [12]. Crude Oil Prices (Average): Sourced from the World Bank, this variable represents the monthly average price of crude oil, which affects production, shipping, and transportation costs. Fluctuations in crude oil prices can influence the cost of

manufacturing garments and their shipment to international markets, thereby affecting export performance [13].

2.2 Data Per-processing

We used the Interquartile Range (IQR) method to identify and remove outliers in dataset. For each feature (column) in dataset, the IQR method first calculates the first quartile (Q1) and the third quartile (Q3) [14].

$$IQR = Q3 - Q1 \text{ (Eq. 1)}$$

Determining Outlier Boundaries:

Lower Bound = Q1 - 1.5 * IQR
Upper Bound = Q3 + 1.5 * IQR

Scaling, in data science and machine learning, refers to transforming the values of features in dataset to a specific range. This is often a crucial step in preparing data for modeling, especially when using algorithms that are sensitive to the scale of features, such as those you are using (Linear Regression, SVR, and XG boost). When features have different ranges (e.g., Garment Export might be in the Millions while 'Unemployment Rate of USA' is a percentage), models may give undue importance to features with larger values, even if they are not inherently more significant. Scaling ensures that all features are treated equally during model training [15].

$$z = (x - \mu) / \sigma \text{ (Eq. 2)}$$

2.3 Train-test split

The accuracy of every forecast can be determined by the model performance on new data which has not been used during the fitting of the model. Therefore, one of the important steps in machine learning techniques is to split the data into train and test sets. In this study, a common split is used which indicates the proportion of 75% of the data to the training set, and 25% to the test set. A train-test split was employed to assess the model's generalization ability and prevent overfitting.

2.4 Forecasting model

This section presents the details of the models we use to forecast Garment sector. These models include the Linear Regression, SVR, XGboot models.

2.4.1 Build linear Regression

Linear Regression assumes a linear relationship between the predictor variables (features like Global Supply Chain, USA Inflation, etc.) and the target variable (Cambodia Export Garment). It tries to find the best-fitting line that minimizes the difference between the predicted and actual values [16].

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_n x_n + \varepsilon \text{ (Eq. 3)}$$

2.4.2 SVR

SVR model aims to predict the "Cambodia Export Garment" value based on several predictor variables like "Global Supply Chain," "USA Inflation," "GDP of USA," etc. For predict the garment export value for a new data point, the model projects the data point into the same higher-dimensional space using the kernel function for time series data are assumed as follow [17].

$$f_{(x)} = W^T \varphi(x) + b \text{ (Eq. 4).}$$

For W is the weight vector l dimension, $\varphi(x)$ is a function that maps x to the feature space with l dimensions, b is biased. The function of $\varphi(x)$ shows a point in the higher dimensional feature space, the result of the mapping of the input vector in the lower dimensional input space.

2.4.3 XG Boost

XG Boost (Extreme Gradient Boosting) is another ensemble method that builds trees sequentially. Each new tree tries to correct the errors made by the previous trees, leading to a highly accurate model. It uses gradient boosting to optimize the model's performance [18].

$$Y = Y_0 + \alpha_1 f_1(x) + \alpha_2 f_2(x) + \cdots + \alpha_n f_n(x) \text{ (Eq. 7).}$$

2.6 Measure of forecast error

This study will be using the MAE for calculates the average of the absolute differences between the predicted values and the actual values. It treats all errors equally [19].

$$MAE = \frac{1}{n}\sum_{t=1}^{n} Y_{actual,t} - Y_{forecast,t} \text{(Eq. 8).}$$

This study will be using MAPE for calculates the average of the percentage errors between the predicted values and the actual values. It provides a relative measure of error, making it easier to compare across different datasets [19].

$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \frac{(Y_{actual,t} - Y_{forecast,t})}{Y_{forecast,t}} \text{(Eq. 9).}$$

This study will be using R-squared for measure that represents the proportion of the variance in the dependent variable 'Cambodia Export Garment') that is explained by the independent variables [20].

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \text{ (Eq. 10).}$$

Where:

SS_{res} : The sum of squares of residuals, also called the residual sum of squares.

SS_{tot} : The total sum of squares (proportional to the variance of the data.

3. RESULTS AND DISCUSSION

Figure. 1 illustrates the forecasted values of garment exports using a Linear Regression (LR) model, along with the actual observed values. The forecasting was performed using independent input variables, including department flow data and lost time due to production delays. In the graph, the blue line with circular markers represents the actual garment export values, while the red dashed line with 'x' markers represents the predicted values from the LR model. As shown, the LR model closely follows the actual data trend.

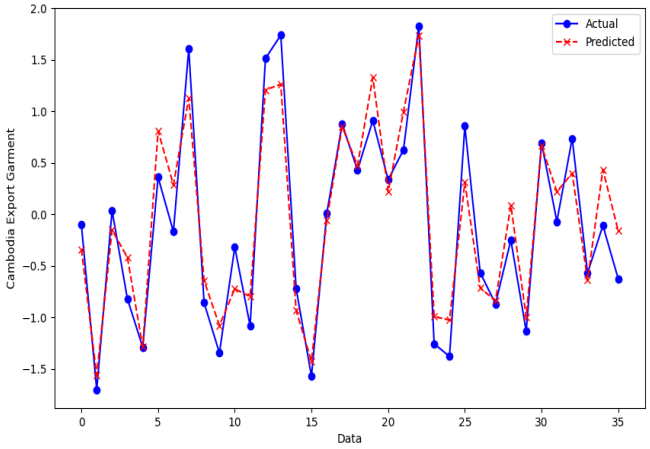


Fig. 1. Graph of forecasted and actual values using linear Regression (LR).

Figure. 2 displays the comparison between actual and predicted values of garment exports using the Support Vector Regression (SVR) model. The forecasting process utilized independent variables, specifically department flow data and lost time due to production delays. In the figure, the blue line with circular markers represents the actual garment export values, while the red dashed line with 'x' markers indicates the predicted values generated by the SVR model. The graph shows that the SVR model is able to capture the general trend and fluctuations in the actual data with a relatively close fit. Although minor deviations are present, the predicted values largely follow the pattern of the actual exports, demonstrating the model's effectiveness in forecasting within this context.

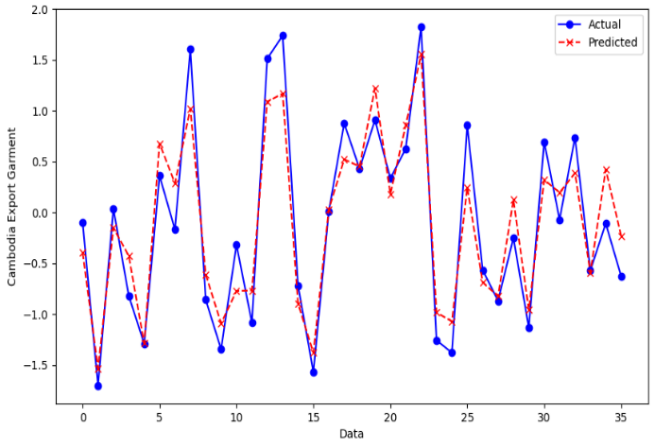


Fig. 2. Graph of forecasted and actual values using SVR.

Figure. 3 Graph of forecasted and actual value using XGBoost. Figure 3 presents the comparison between the actual and predicted values of Cambodia's garment exports using the XGBoost regression model. The forecasting is based on independent variables, including department flow data and time lost due to production delays. In this graph, the blue line with circular markers denotes the actual export values, while the red dashed line with 'x' markers represents the predicted values generated by the XGBoost model. The figure demonstrates that XGBoost is highly effective in capturing the overall trends and variations in the actual data. The close alignment between the two lines, especially around peak and trough points, indicates that XGBoost performs well in modeling and forecasting garment exports with high accuracy.

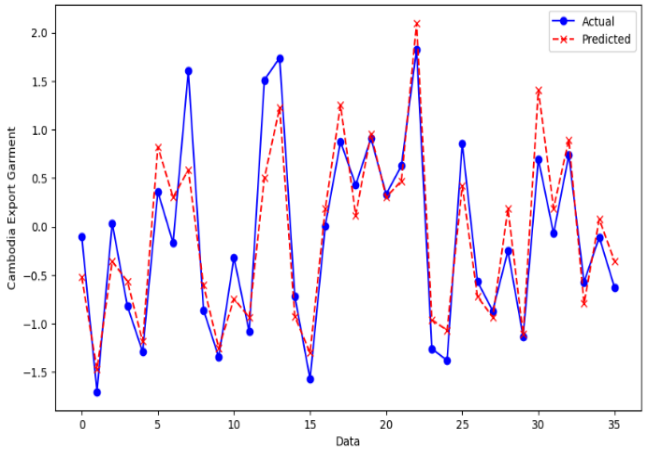


Figure. 3. Graph of forecasted and actual value using XGBoost.

Table 1. Model Performance Comparison

Metric	MAE	MAPE	R-squared
Linear Regression	0.26	1.02	0.90
SVR	0.33	1.32	0.83
XG Boost	0.31	1.47	0.84

Table 1 summarizes the performance metrics—Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and R-squared—of the three forecasting models used to predict Cambodia's garment exports: Linear Regression (LR), Support Vector Regression (SVR), and XGBoost. Among the three, Linear Regression demonstrated the best overall performance, achieving the lowest MAE (0.26), and lowest MAPE (1.02), along with the highest R-squared value (0.90). This indicates that LR was most effective in minimizing prediction errors and explaining the variance in the dataset. The SVR model also performed reasonably well, with an R-squared of 0.83, but showed slightly higher error values (MAE = 0.33, MAPE = 1.32), suggesting less accuracy and a weaker fit compared to LR. On the other hand, while XGBoost had a slightly better R-squared value (0.84) than SVR, it recorded the highest MAPE (1.47) among the three models. This suggests that XGBoost was less effective in minimizing forecasting errors despite moderately capturing the variance. Overall, Linear Regression outperformed both SVR and XGBoost in terms of both error minimization and model fit, making it the most suitable model for this specific forecasting task.

4. CONCLUSIONS

This study explored the forecasting of Cambodia's garment exports by integrating a robust set of economic indicators and applying comparative modeling techniques. A carefully curated time-series dataset from 2010 to 2024 was developed, incorporating both domestic and international variables known to influence garment trade performance. These included labor market indicators (such as nominal average wages), macroeconomic factors (such as inflation and GDP in the U.S. and Europe), global supply chain trends, and crude oil prices. The inclusion of such variables ensured a comprehensive view of the multifaceted drivers behind Cambodia's export performance. Three forecasting models—Linear Regression (LR), Support Vector Regression (SVR), and XGBoost—were evaluated for predictive accuracy using key performance metrics. The results clearly indicated that the Linear Regression model achieved the highest accuracy, with the lowest Mean Absolute Error (0.26), and Mean Absolute Percentage Error (1.02), along with the highest R-squared value (0.90). While SVR and XGBoost captured general trends, they exhibited higher error rates, with XGBoost performing slightly better in variance explanation ($R^2 = 0.84$) but falling short in error minimization.

In summary, the combination of thoughtful feature selection and rigorous model evaluation led to the conclusion that Linear Regression is the most reliable model for forecasting Cambodia's garment exports in this context. These findings provide a valuable foundation for evidence-based policy recommendations and strategic planning for industry stakeholders aiming to enhance the competitiveness and resilience of Cambodia's key export sector.

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to Dr. PHAUK Sokkey and Mr. KHEAN Vesal for their invaluable support, encouragement, and insightful guidance throughout the course of this research. Their expertise and mentorship have been instrumental in shaping the direction and quality of this study. I would also like to extend my sincere appreciation to the Department of Applied Mathematics and Statistics (AMS) for providing the academic foundation, resources, and environment necessary to conduct this research. The support and encouragement from the department have been essential to the completion of this work. Finally, I am thankful to everyone who contributed directly or indirectly to the success of this study.

REFERENCES

[1] World Bank. (2023). Cambodia Economic Update: From Recovery to Reform. World Bank Group. https://openknowledge.worldbank.org/handle/10986/40869

[2] International Labour Organization (ILO). (2021). Cambodia Garment and Footwear Sector Bulletin. https://www.ilo.org/asia/publications

[3] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: The state of the art. International Journal of Forecasting, 14(1), 35–62. https://doi.org/10.1016/S0169-2070(97)00044-7

[4] Ala'raj, M., Abbod, M. F., & Taylor, J. (2021). Time series forecasting of economic indicators using machine learning techniques. Journal of Big Data, 8, 67. https://doi.org/10.1186/s40537-021-00454-2

[5] International Trade Centre (ITC). (2023). Trade Map - International Trade Statistics. https://www.trademap.org/
[6] Ministry of Labour and Vocational Training (MLVT). (2023). Labor Market Information Cambodia. http: //www.mlvt.gov.kh/
[7] United Nations Conference on Trade and Development (UNCTAD). (2023). Global Supply Chain and Trade Trends. https://unctad.org/
[8] U.S. Bureau of Labor Statistics. (2023). Consumer Price Index (CPI) – Inflation. https://www.bls.gov/cpi/
[9] Eurostat. (2023). Euro Area Inflation Rate. https: // ec. europa.eu/eurostat/
[10] World Bank. (2023). Gross Domestic Product of the UnitedStates.https://data.worldbank.org/indicator/NY.GDP.MKTP.CD?locations=US
[11] U.S. Bureau of Labor Statistics. (2023). Unemployment Rate – United States. https://www.bls.gov/cps/
[12] Eurostat. (2023). Unemployment Rates - Euro Area. https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Unemployment_statistics
[13] World Bank. (2023). Crude Oil Prices Monthly Averages .https://www.worldbank.org/en/research/commodity markets
[14] Alabrah, A. (2023). An improved CCF detector to handle the problem of class imbalance with outlier normalization using IQR method. Sensors, 23(9), 4406.
[15] Ahsan, M. M., Mahmud, M. A. P., Saha, P. K., Gupta, K. D., & Siddique, Z. (2021). Effect of data scaling methods on machine learning algorithms and model performance. Technologies, 9(3), 52.
[16] L. H. Hasibuan, D. M. Putri, and M. Jannah. (2022). Simple Linear Regression Method to Predict Cooking Oil Prices in the Time of Covid-19, 81–94.
[17] Hanifah Muthiah, Umu Sa'adah, Achmad Efendi. (2021). Support Vector Regression (SVR) Model for Seasonal Time Series Data.
[18] M. Amjad, I. Ahmad, M. Ahmad, P. Wróblewski, P. Kamiński, and U. Amjad, "Prediction of pile bearing capacity using XGBoost algorithm: modeling and performance evaluation," Appl. Sci., vol. 12, no. 4, p. 2126, 2022.
[19] de Myttenaere, A., Golden, B., Le Grand, B., & Rossi, F. (2016). Mean Absolute Percentage Error for regression models. Neurocomputing, 192, 38–48. https://doi. Org / 10.1016/j.neucom.2015.12.114
[20] A.Y. J Akossou, R.Palm. (2013). Impact of Data structure on the Estimators R-squares And Adjusted R-square in Linear Regression, 2-3. https: //www .researchgate.net/profile/Arcadius-Akossou/publication. 2-3

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Reconstructing Higher-Order Interactions in Algal Communities

Danut CHHORN [1*], Sokkey PHAUK [1], Timoteo CARLITTI [2], Malia LIN [3]

Department of Applied Mathematics and Statistics, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Department of Science, University of Namur, 5000 Namur Belgium*

[*] *Corresponding author: danutchhorn@gmail.com*

Abstract: *In algal systems, where species interactions go beyond paired relationships, higher-order interactions (HOIs) are especially important in determining the dynamics and structure of ecological communities. Although pairwise interactions are the focus of classic ecological models, new research indicates that HOIs, like three-way interactions, can have a big impact on ecosystem functioning, biodiversity, and community stability. To understand how these intricate relationships influence population dynamics and community formation, we examine the function of HOIs in algal communities in this work. We model the dynamics of algal communities using both synthetic and empirical data. Records of species abundance from carefully monitored lab tests and field observations of algae communities in freshwater environments are examples of empirical data. A variety of interaction scenarios, such as paired and higher-order interactions, are simulated using synthetic datasets in a range of environmental settings. The first step we used dataset that uniform distribution includes: 7 of Species abundance over time, Growth rates and carrying capacities (1,100) for each species, the initial conditions in the interval(10,100), the growth rate is randomly uniform distribution between (0,1),Pairwise interaction matrices (A) and higher-order interaction tensors (H), we embarked on an exploration using Lotka-Volterra equations as our baseline model. And the second step will be experimenting with empirical data. To characterize the dynamics of algal communities, we create a mathematical model that considers both pairwise and higher-order interactions. The ordinary differential equations (ODEs) that form the basis of the model show that the pace at which species densities vary is impacted by Logistic growth, pairwise interactions, Higher-order interactions. When higher-order interactions are considered, it can be difficult to determine the kind of contacts (such as pairwise or three-body) present in coupled oscillator systems of algal communities since various interactions might result in comparable dynamical states and bifurcations. After we experimented with synthetic data with optimization methods like: Ordinary Least Square Regression, Ridge Regression, and LASSO and Adaptive LASSO to refer (reconstructed of pairwise and three body interactions) between microbial from time series data. We can find the best performance method for classifying the types of interactions and infer the probabilities of the existence for pairwise and three-body couplings. The best performance is Adaptive lasso method is better than others method, it achieved the lowest Mean Squared Error (MSE) of* ***0.001603****, an R-squared value of* ***0.9985****, indicating an excellent fit, and minimal prediction errors as shown by its Root Mean Square Error (RMSE) of* ***0.00959****, Mean Absolute Error (MAE) of* ***0.00158****, and SMAPE of* ***0.002514****. and we will apply all those methods with algal communities' data (real data), we will show how the method can help understanding the complexity of algae systems.*

Keywords: Higher-Order Interactions; Lotka-Volterra; Pairwise Interactions; Ordinary Differential Equations (ODE); OLS, Ridge, LASSO

1. INTRODUCTION

Algal communities, found in oceans, lakes, and rivers, play an important role in the health of aquatic ecosystems. These communities consist of many different species of algae that interact with each other in complex ways. Often, we think of these interactions as just pairwise, where two species might compete for resources or cooperation. However, in real-life ecosystems, interactions often involve more than two species at once, which are known as higher-order interactions.

In the case of algal communities, these higher-order interactions might involve multiple species working together or competing in groups. These interactions can have a significant impact on things like algal blooms (which can affect water quality) or the stability of the ecosystem.

Understanding the complex interactions within algal communities, particularly higher-order interactions, is crucial for assessing their growth, stability, and response to environmental changes. Currently models focus on pairwise interactions, Traditional pairwise interaction models may fail to capture complex dynamics. However, in this case, we will study and develop a methodology that can accurately infer both pairwise and higher-order interactions from time series data and evaluate their impact on ecosystem stability, specifically algal populations.

The objective of this study is to understand how algal populations grow and respond to ongoing environmental changes. Identify the type of interaction (pairwise and three-body) and infer the topology and strength of the couplings between algal species using time series data. Examine the relationship between higher-order interaction and stability and develop a new method for finding out not only pairwise interactions but also higher-order interactions in algal communities.

2. METHODOLOGY

2.1 Model Lotka-Volterra dynamics Community

The Lotka-Volterra model is a foundational model in population dynamics by incorporating, commonly used to describe predator-prey or competitive species interactions.

$$\dot{x}_i = f_i(x_i) + \sum_{j=1}^{N} A_{ij}{}^{(1)} x_i x_j + \sum_{j<k, k\neq i, j\neq i} A_{ijk}{}^{(2)} x_i x_j x_k$$

$f_i(x_i)$ This is a nonlinear function that depends only on the variable X_i.

$\sum_{j=1}^{N} A_{ij}{}^{(1)} x_i x_j$ This term captures pairwise interactions.

$\sum_{j<k, k\neq i, j\neq i} A_{ijk}{}^{(2)} x_i x_j x_k$ This is a third-order interaction term, representing triplet interactions.

2.2 Optimization Methods

Ordinary Least Square: is a method used in linear regression to find the best-fitting line (or hyperplane) through a set of data points by minimizing the sum of squared residuals.

$$\hat{\beta} = arg_{\beta} min \|y - X\beta\|_2^2$$

Ridge Regression: is a type of linear regression that adds a penalty on the size of coefficients to reduce overfitting and handle multicollinearity (when predictors are correlated).

$$\hat{\beta}^{ridge} = arg_{\beta} min \{\|y - X\beta\|_2^2 + \alpha \|\beta\|_2^2\}$$

Lasso Regression is a regularized linear regression method that encourages sparsity (i.e., setting some coefficients to exactly zero), which is useful for feature selection.

$$\hat{\beta}^{lasso} = arg_{\beta} min (\|y - X\beta\|_2^2 + \alpha \|\beta\|_1)$$

Adaptive Lasso is an extension of Lasso regression that improves variable selection consistency by applying data-dependent weights to the L1 penalty term.

$$\hat{\beta}^{ALasso} = arg_{\beta} min \left(\|y - X\beta\|_2^2 + \alpha \sum_{j=1}^{p} w_j |\beta_j|_1 \right)$$

[1]* Corresponding author: DANUT CHHORN
E-mail: danutchhorn@gmail.com; Tel: +855-87 90 26 24

3. RESULTS AND DISCUSSION

The study case, we have addressed synthetic data of microbial interaction (7 species) and real data with algal communities (8 species), for synthetic data we have been set parameters growth rate R for all species have been selected from uniform distribution in the interval (0,1), and the carrying capacity K from uniform distribution (1,100) and condition interval (10,100) with time evolution of the variable xi(t), with i=1…..7. *(Parameters that are used in synthetic data).*

Number Species	1	2	3	4	5	6	7
Carrying capacity	88.7647	3.8387	49.5002	17.6248	97.8894	71.5568	50.5467
Growth Rate	0.6099	0.6177	0.8594	0.8055	0.5767	0.1829	0.2399
Initial condition	30	45	32	50	55	30	40

For real dataset provides a type of algal strain called Diatoms that was isolated from a single phytoplankton simply collected near the **Thomtonback Wind Farm** at (Southern Bight of the North Sea) spring bloom in Match 2013. And the raw it has 1962 rows and 16 features. And describe the estimated population parameters for the 8 species used in biodiversity experiments, I0 is the cell volume (Initial condition) that had value around (3940-752767) R is the intrinsic growth rate that had around (0.19-0.51), K is the carrying capacity that had around (406 957 000-10583 784 289), and Species are focused on different algal strains.

Code	Name	Species	I0	R	K
1	s8	*Coscinodiscus sp.*	367008	0.33	948 839 324
2	s10	*Ditylum sp.*	24757	0.27	862 505 671
3	s17	*Bacillaria sp.*	6448	0.51	406 957 000
4	s18	*Odontella sp.*	752767	0.24	10 583 784 289
5	sE18	*Thalassiosira sp.1*	25727	0.28	108 096 248
6	s23	*Gyrosigma sp.*	3940	0.19	651 492 127
7	s31	*Guinardia sp.*	14001	0.19	1 176 672 117
8	s32	*Thalassiosira sp.2*	12556	0.43	629 079 021

The communities for diversity level 2,4 and 6 or communities for Pairwise interactions and Higher-Order interactions (three body interactions). Numbers refer to the species code.

Pairwise Interactions	Higher-Order Interaction	Higher-Order Interaction
Diversity=2	Diversity=4	Diversity=6
1+3	1+2+4+8	1+2+3+4+5+7
2+4	2+4+5+6	1+2+4+5+6+7
7+8	3+4+5+8	1+3+4+5+6+7
1+7	2+3+4+5	1+2+3+5+6+7
3+5	1+2+3+6	1+2+3+5+6+8
5+6	3+5+7+8	2+3+5+6+7+8
4+6	2+3+4+7	1+3+4+5+6+8
3+8	1+5+6+8	2+3+4+5+6+7
2+5	1+3+5+6	2+4+5+6+7+8
1+3	1+2+4+7	3+4+5+6+7+8

We need to reconstruct pairwise interaction and three-body, it based on optimization via the methods Ordinary Least Square regression, Ridge regression, Lasso and Adaptive lasso. And the result of reconstruction of synthetic data:

Among the four regression models evaluated Adaptive Lasso, OLS (Ordinary Least Squares), LASSO, and Ridge Regression the Adaptive Lasso model on **Table 1** demonstrated superior predictive performance across all statistical metrics. It achieved the lowest *Mean Squared Error (MSE) of 0.001603,* an *R-squared value of 0.9985*, indicating an excellent fit, and minimal prediction errors as shown by its RMSE of *0.00959*, *MAE of 0.00158*, and *SMAPE of 0.002514.*

In contrast, the OLS model, while still strong, had a higher *MSE of 0.017603* and a slightly lower *R-squared of 0.9795*, suggesting slightly less accuracy and a greater average error compared to Adaptive Lasso.

Ridge Regression also underperformed relative to Adaptive Lasso and OLS, with an *MSE (1.209106)*, *R-squared (0.8843)*, RMSE (5.420943), *MAE (2.667501)*, and the highest *SMAPE (28.541884)*

Similarly, The LASSO Regression model performed poorly, with a dramatically higher *MSE of 90.586046* and a much lower *R-squared of 0.6107*, indicating that it failed to capture the underlying patterns in the data effectively. Its high *RMSE (9.51767)*, *MAE (5.034199)*, and *SMAPE (20.718666)* further confirm its weak performance.

So, the Adaptive Lasso clearly outperforms the other models in both goodness-of-fit and prediction accuracy, making it the most effective approach for this dataset.

Fig 1: it demonstrates: the first plot graph original of an ecological mode defines pairwise interaction and higher order interactions by overtime (day) between each species. And last one after we reconstructed using better optimization method like Adaptive Lasso that is the dot line showing best fit with original of Lotka-Volterra Dynamic model.

Producing solid and trustworthy results requires an understanding of these methodological constraints, which include overfitting, linearity assumptions, multicollinearity, autocorrelation, and problems with data quality. To increase the validity of their findings, we will use various modeling approaches, validation procedures, and suitable diagnostic checks. The legitimacy and reproducibility of study findings are further improved by openly disclosing these limitations.

We will apply real data using the same Model and optimization methods for reconstructing and limitation of apply model just identify pairwise term three body tensors in algal communities of each algal strain and how to reconstruct them.

Table 1. Showing the statistical metric performance of optimization methods on synthetic data there are seven species)

Statistical Metric	MSE	R Square	RMSE	MAE	SMAPE
Adaptive lasso	0.001603	0.9985	0.00959	0.00158	0.002514
OLS	0.017603	0.9795	0.132677	0.059993	0.259419
LASSO Regression	90.586046	0.6107	9.51767	5.034199	20.718666
Ridge Regression	1.209106	0.8843	5.420943	2.667501	28.541884

Fig. 1. Showing Species populations over time that using Lotka-Volterra Dynamic model with time series (day).

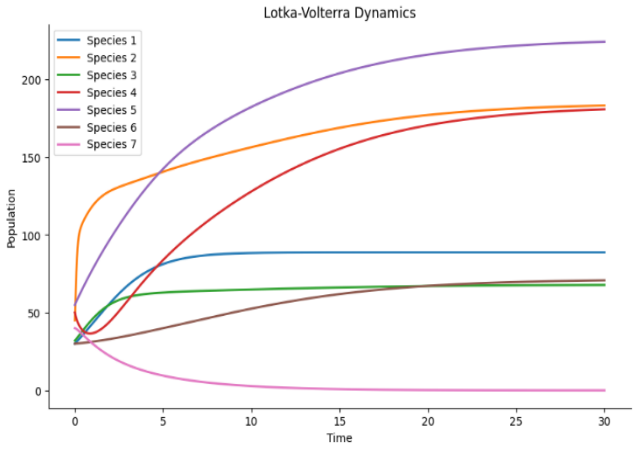


Fig. 2. Showing Species populations over time that using Lotka-Volterra Dynamic model with time series (day). And comparing original and after reconstructing successfully.

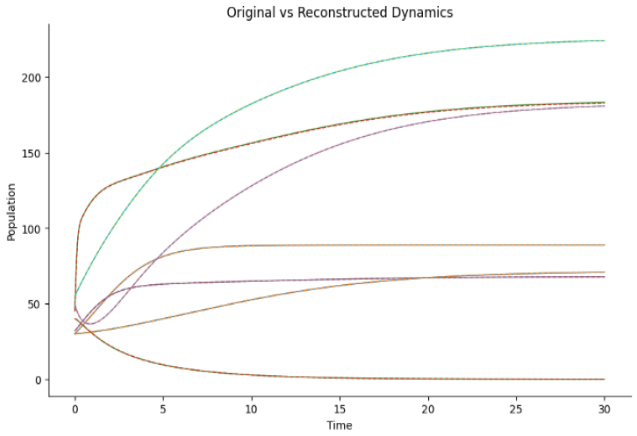


Fig. 3. Comparing network graphs of pairwise interaction, higher-order interaction and mixtures. That check with True with reconstruction optimization methods.

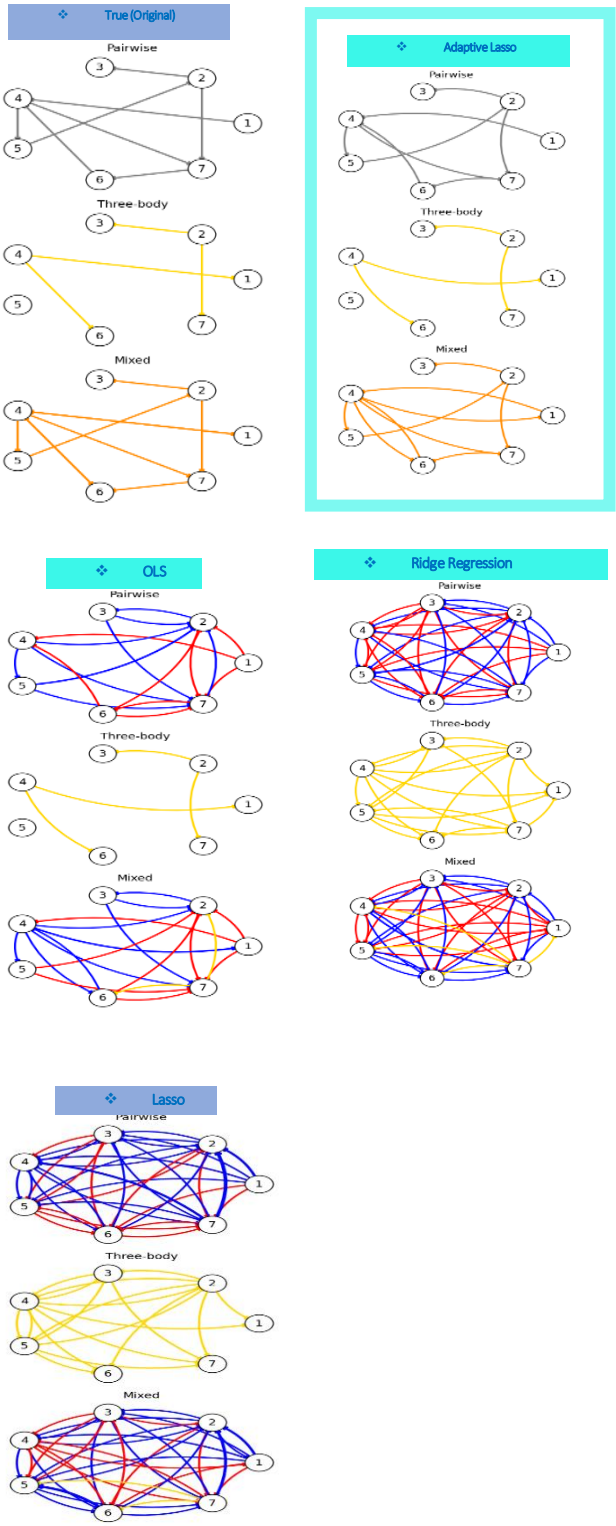


Fig. 4. Checking Error with radio of model complexity (M/H) that comparing three methods (OLS, Ridge and Lasso)

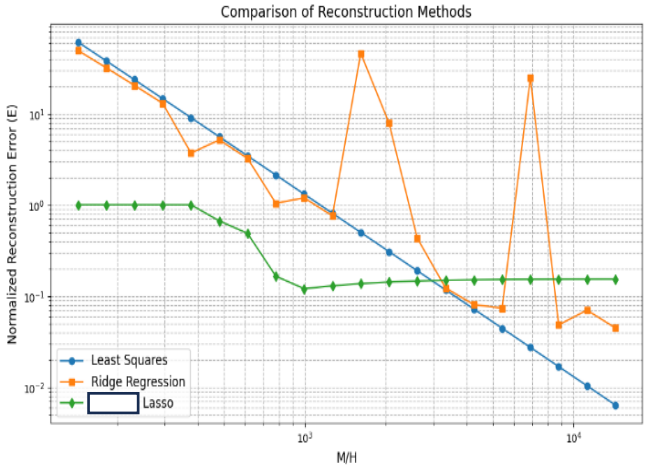
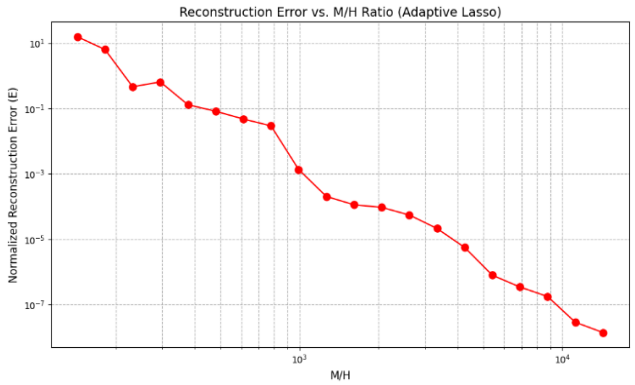


Fig. 5. Checking Error with radio of model complexity (M/H) of Adaptive Lasso: is the best performance.



4. CONCLUSIONS

This study evaluated and compared four regression-based optimization techniques Ordinary Least Squares (OLS), Ridge Regression, LASSO, and Adaptive LASSO for reconstructing pairwise and three-body interactions in a synthetic ecological dataset derived from a Lotka-Volterra dynamic model. Adaptive LASSO consistently outperformed the other methods across all statistical evaluation metrics. It achieved the lowest Mean Squared Error (MSE) of 0.001603, an R-squared value of 0.9985, indicating an excellent fit, and minimal prediction errors as shown by its RMSE of 0.00959, MAE of 0.00158, and SMAPE of 0.002514 and errors of reconstruction decreases while M/H increase that is better than other methods.

And showing a strong alignment between the original ecological dynamics and the Adaptive LASSO-reconstructed model, especially in modeling higher-order species interactions over time. And to sum up, demonstrate that: Adaptive LASSO is the most robust and accurate method among those evaluated for reconstructing both pairwise and higher-order ecological interactions. Its ability to combine variable selection and coefficient shrinkage with adaptive weights provides superior generalization and interpretability.

Moving forward, this modeling framework will be extended to real-world ecological datasets to evaluate its performance in empirical.

Environmental agencies can better combat algal blooms by incorporating scientific results into management and policy frameworks. These tactics prioritize integrated watershed methods, health protection, climate adaptation, monitoring, and nutrient control. To protect aquatic ecosystems and general welfare, effective implementation depends on interdisciplinary cooperation, ongoing research, and stakeholder participation.

ACKNOWLEDGMENTS

I would like to thank my supervisor and co-supervisor for guidance, support and valuable feedback during my research, and University of Namur, Department of Science, in Belgium for providing computational resources and research facilities. This work would not have been possible without their assistance.

REFERENCES

[1] T. Gibbs, S. A. Levin, and J. M. Levine, "Coexistence in diverse communities with higher-order interactions," *bioRxiv*, 2022. https://doi.org/10.1101/2022.03.04.483022

[2] A. D. Letten and D. B. Stouffer, "The mechanistic basis for higher-order interactions and non-additivity in competitive communities," *Ecology Letters*, vol. 22, no. 3, pp. 423–436, 2019. https://doi.org/10.1111/ele.13211

[3] F. Malizia, A. Corso, L. V. Gambuzza, G. Russo, V. Latora, and M. Frasca, "Reconstructing higher-order interactions in coupled dynamical systems," *Accepted Manuscript*, 2024. https://github.com/LValentinaGambuzza/Code-for-Reconstructing-higher-order-interactions-in-coupled-dynamical-systems

[4] R. Mulas, D. Horak, and J. Jost, "Graphs, Simplicial Complexes and Hypergraphs: Spectral Theory and

Topology," in *Higher-Order Systems*, F. Battiston and G. Petri, Eds. Springer, 2022, pp. 1–58. https://doi.org/10.1007/978-3-030-91374-8_1

[5] L. Peel, T. P. Peixoto, and M. De Domenico, "Statistical inference links data and theory in network science," *Nature Communications*, vol. 13, p. 1331, 2022. https://doi.org/10.1038/s41467-022-28941-9

[6] W. Su, S. Hata, H. Kori, H. Nakao, and R. Kobayashi, "Pairwise vs Higher-order interactions: Can we identify the interaction type in coupled oscillators from time series," *arXiv*, 2025. https://arxiv.org/abs/2503.13244

[7] M. Timme and J. Casadiego, "Revealing networks from dynamics: an introduction," *Journal of Physics A: Mathematical and Theoretical*, vol. 47, no. 34, p. 343001, 2014. https://doi.org/10.1088/1751-8113/47/34/343001

[8] J. M. Levine, J. Bascompte, P. B. Adler, and S. Allesina, "Beyond pairwise mechanisms of species coexistence in complex communities," *Nature*, vol. 546, no. 7656, pp. 56–64, 2017. https://doi.org/10.1038/nature22898

[9] A. D. Letten and D. B. Stouffer, "The mechanistic basis for higher-order interactions and non-additivity in competitive communities," *Ecology Letters*, vol. 22, no. 3, pp. 423–436, 2019. https://doi.org/10.1111/ele.13211

[10] H. Mickalide and S. Kuehn, "Higher-order interaction between species inhibits bacterial invasion of a phototroph-predator microbial community," *Cell Systems*, vol. 9, no. 6, pp. 521–533, 2019. https://doi.org/10.1016/j.cels.2019.11.004

[11] R. Mulas, D. Horak, and J. Jost, "Graphs, simplicial complexes, and hypergraphs: Spectral theory and topology," in *Higher-Order Systems*, F. Battiston and G. Petri, Eds. Springer, 2022, pp. 1–58. https://doi.org/10.1007/978-3-030-91374-8_1

[12] C. Shen, K. Lemmen, J. Alexander, and F. Pennekamp, "Connecting higher-order interactions with ecological stability in experimental aquatic food webs," *Ecology and Evolution*, vol. 13, p. e10502, 2023. https://doi.org/10.1002/ece3.10502

[13] J. T. Wootton, "Predicting direct and indirect effects: An integrated approach using experiments and path analysis," *Ecology*, vol. 75, no. 1, pp. 151–165, 1994. https://doi.org/10.2307/1939391

Materials Science and Structure (MSS)

2 Sub-parallel Sessions

"MSS-1: Materials Science and Engineering"

Chair: Dr. LIM Sovanvichet, Co-chair: Dr. OENG Thaileng

No.	Topic
1	Analysis of Iron Ores and Slags from Archeological Sites by Using Energy Dispersive X-ray Spectroscopy and X-ray Diffraction Technique *Authored by: Sreynoch DUM, Chhor Yi LY, Marany SAT, Navin CHHEAT, Nhornhoem PICH, Bunsong SEY, Sirisokha SEANG, Kakda KERET, Chanthourn THUY, Tharith SRIV*
2	Investigation on Mechanical Properties and Antimicrobial Activity of Natural Rubber Latex Examination Gloves Incorporated with Pomelo Peel Powder *Authored by: Likea HENG, Soma Tey NUON, Seka KONG, Laymey SRENG, Sovathana PHUONG, Chakrya THEAP, Chanthol PENG, Phanny YOS*
3	Degradation Study of Few-Layer SnP2X6 (X=S, Se) in Ambient Conditions Using OM, AFM, and SEM *Authored by: Sok Heng HOR, Tharith SRIVE*
4	Lithology, Geochemistry, Hydrothermal Alteration and Mineralization of Cu-Mo Porphyry Deposits at Canada Wall Prospect, Andong Meas District, Ratanakiri Province, Cambodia *Authored by: Mengly IEM, Yana CHHOEUN, Bunlim HEAN, Bunchoeun PICH, Sirisokha SEANG, Kimhouy OY*
5	Design and Modelling of Cost-effective Manufacturing of Glass Tempering Furnace Used in Industrial Applications *Authored by: Rayuth KEAT, Raveth HIN, Chansopheak SEANG*
6	Investigation of Utilization of Natural Oils as Processing Aids and Activators Alternative to Naphthenic Oils in Rubber Composites *Authored by: Sreynich KONG, Yukleav NAT, Phanny YOS, Laymey SRENG*
7	Mineralogical and Physicochemical Characterizations of Silica Sand by Using Microbeam, XRD, and XRF *Authored by: Muylin MEUN, Chhor Yi LY, Navin CHEAT, Marany SAT, Nhornhoem PICH, Vichet THI, Sok Heng HOR, Tharith SRIV*
8	Determination of Crystallite Size of Transition Metal Oxides (NiO, Mn_3O_4, and α-Fe_2O_3) Using the Scherrer Formula and Williamson–Hall Plot *Authored by: Sokly SREANG, Chhor Yi LY, Navin CHEAT, Piseth LIM, Tharith SRIV*
9	Utilization of Fly Ash as an Eco-friendly Filler in Styrene Butadiene Rubber Compound for Tire Tread Application *Authored by: Kakda YIEM, Yukleav NAT, Laymey SRENG, Phanny YOS*
10	Development of Low-carbon Concrete using Biomass Burning Ash for Port and Harbor Facilities *Authored by: Sovannarith CHEA, Yamaguchi TOSHINOBU, Akira YOSHIKAZU, Kimura YOKINOBU, Kawakami TAKASHI*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Analysis of Iron Ores and Slags from Archaeological Sites by Using Energy Dispersive X-ray Spectroscopy and X-ray Diffraction Technique

Sreynoch Dum[1], Chhor Yi Ly[2], Marany Sat[1], Navin Chheat[1], Nhornhoem Pich[1], Bunsong Sey[2], Sirisokha Seang[4], Kakda Kret[4], Chanthourn Thuy[3, *], Tharith Sriv[1, 2, *]

[1]*Department of Physics, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*

[2] *Graduate Program in Physics, Graduate School of Science, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*

[3] *Royal Academy of Cambodia, Russian Federation Blvd., Sen Sok, Phnom Penh, Cambodia*

[4] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding authors: tharith@rupp.edu.kh, chanthourn@yahoo.com*

Abstract: *Iron (Fe) is essential in numerous applications, including machinery, tools, and structural components for buildings, bridges, aircraft, and ships. Archaeological evidence in Cambodia such as agricultural tools, weapons, statues, and iron bars used in temple construction demonstrates that iron has played a crucial role in the country's historical development. Recognizing this, several scientific studies have been conducted at archaeological sites, particularly in the provinces of Preah Vihear, Siem Reap, Kampong Thom, and Oddar Meanchey. In-situ investigations like Preah Khan Kampong Svay, areas surrounding Phnom Daek, and within Angkor monuments have analyzed iron ore and slag samples using scanning electron microscopy (SEM) equipped with energy-dispersive X-ray spectroscopy (EDS). However, iron ores and slags from Preah Vihear have remain under-investigated, resulting in limited scientific evidence for archaeological and metallurgical research in this region. Motivated by this gap, the present study focuses on analyzing the elemental compositions and mineralogical phases of iron ore and slag samples collected from various archaeological sites in Preah Vihear province, including Russey Treb, Phnom Daek, and Phumi Boeng, using advanced techniques such as EDS and X-ray diffraction (XRD). EDS results show that Sample #1 from Phnom Daek contains the highest iron content (37.5%), followed by Sample #4 from Russey Treb (Fe = 27.7%). Other samples from Russey Treb exhibit lower iron concentrations: Sample #2 (7.7%), Sample #3 (24.8%), Sample #5 (0.5%), Sample #6 (2.1%), and Sample #7 (0.9%). Sample #8 from Phumi Boeng shows the iron content at just 0.8%. Additional elements such as oxygen (O), silicon (Si), and aluminum (Al) were observed in all samples. XRD analysis further reveals the presence of hematite (Fe_2O_3) and magnetite (Fe_3O_4) in Sample #1, while quartz (SiO_2) is identified in Samples #2 and #8 alongside hematite. Samples #3 through #6, which are preliminarily identified as iron slags, contain fayalite (Fe_2SiO_4) and various other compounds: magnetite in Sample #3; hematite, magnetite, and quartz in Sample #4; quartz and ferrosilite ($FeSiO_3$) in Sample #5; and pyroxmangite ($MnSiO_3$) and magnetite in Sample #6. Sample #7 contains pyroxmangite and quartz. The variations in elemental and mineralogical composition across sites suggest differences in ore quality and smelting processes. Among all studied samples, Sample #1 from Phnom Daek is determined to have the highest quality iron ore. These findings provide valuable insights into the chemical and mineralogical characteristics of Cambodian iron ores and contribute to a deeper understanding of ancient iron production in the region. Moreover, the study supports future research and potential development in Cambodia's mining and metallurgical sectors.*

Keywords: Iron ores; archaeological; XRD; EDS

1. INTRODUCTION

Iron has historically served many applications in tools, construction, and weaponry, which has been supported the rise of civilizations across the world. In Cambodia, the most famous and significant monument is Angkor Temple, which is primarily located in the Angkor complex, Siem Reap, Cambodia. Since the temple and others in the complex were mainly constructed from stone, iron tools such as chisels were essential for quarrying the stone blocks [1]. Therefore, iron products played a vital role in the construction of Khmer monuments, indicating that iron production actively played a crucial role during the Khmer Empire (802-1431). In addition to tools used for stone quarrying and processing, it is also likely that iron was employed for agricultural activities and military equipment [2]. Iron fragments, which are believed to

[*] Corresponding authors: Tharith Sriv, Chanthourn Thuy
E-mails: tharith@rupp.edu.kh; Tel: +855-85289791, chanthourn@yahoo.com; Tel: +855-11 424767

be components of armatures, have been discovered at the ancient Royal Palace within the Angkor complex [3]. Evidence of iron production is further supported by the discovery of iron slags by products of the smelting process found in numerous slag dumps at Preah Khan of Kompong Svay [4, 5].

The Iron Industries of Angkor Project (INDAP) covers Cambodia's northern province of Preah Vihear, focusing on three locations within the enclosure walls of the Angkorian Preah Khan complex and two sites approximately 30 km to the east, near Phnom Daek also known as Iron Mountain. Using a scanning electron microscope equipped with an energy-dispersive spectroscope (SEM-EDS), the project has provided valuable insights into iron production, which played a significant role in the historical development of Cambodia [6]. Moreover, iron slags and ores have been analyzed from sites including Preah Khan of Kompong Svay, the area surrounding Phnom Daek, and the Angkor monuments. The chemical compositions of fayalite and wüstite present in the iron slags, as well as magnetite found in the iron ores, were identified by SEM-EDS [7]. Several locations surrounding Phnom Daek such as Phnom Daek West (Trapeang Sanlong , O Rumchek , Russei Prei , Svay Damnak Thmei , and Sre Po), Phnom Daek South (Prei Russei , Tonle Bak , and Chhay), Phnom Daek East (Boeng, Ta Tong, Rung Reung, and Phum Sanlong) have been selected for a study about iron-production sites by using inductively coupled plasma atomic emission spectroscopy (ICP-AES) for the quantification of the major elements and mass spectrometry (ICP-MS) for minor and trace elements [2]. However, iron ores and slags in Preah Vihear such as Russey Treb, Phumi Boeng have not widely been investigated. In this study, advanced instruments such as EDS and x-ray diffraction (XRD) technique were employed to investigate elemental and mineralogical compositions of iron ores and slags from different archaeological sites in Preah Vihear province such as Russey Treb, Phnom Daek, and Phumi Boeng.

2. METHODOLOGY

2.1 Sample Collection

Eight samples of iron ores and slags were collected from three archaeological sites in Preah Vihear province: Russey Treb, Phnom Daek, and Phumi Boeng. The samples were labeled as Sample#1 (S1) to Sample#8 (S8) based on their collection order and origin. Specifically, S1 was collected from Phnom Daek, while S2 to S7 were obtained from Russey Treb, and S8 was taken from Phumi Boeng.

2.2 Characterization Techniques

In this study, the collected iron ores and slags samples were polished before being chemically and structurally analyzed by using EDS (Thermo Scientific Axia ChemiSEM) and XRD (Malvern Panalytical Aeris) system with CuK^{α} radiation of the wavelength (λ) of 0.15406 nm. The techniques provide insights into the crystalline structure and elemental compositions. It should be noted that XRD patterns were analyzed by using MDI Jade 6 software, which allowed identification of crystal phases as well as structures present in the samples. The following section provides key results of the analyses as well as interprets and explains significance of the findings.

3. RESULTS AND DISCUSSION

3.1 EDS Results

Table 1 presents the EDS results, revealing a wide range of elemental concentrations across iron samples from S1 to S8. Notably, S1 exhibits the highest iron concentration (37.5%), followed by S4 (27.7%), S3 (24.8%), and S2 (7.7%). These samples also show prominent percentages for oxygen (O), indicating the presence of oxide compounds in each sample. In contrast, S5 through S8 display lower iron concentrations of 0.5%, 2.1%, 0.9%, and 0.8%, respectively. Additionally, the EDS spectra and elemental mapping, shown in Figure 1, clearly illustrate both elemental distributions and intensity variations across the samples. The EDS spectra revealed peaks of iron in all samples aside from silicon (Si), and aluminum (Al) across most samples, especially in S1 to S4. Minor elements, including nickel (Ni), phosphorus (P), sulfur (S), calcium (Ca), copper (Cu), potassium (K), magnesium (Mg), Chlorine (Cl), and titanium (Ti), were also detected with varying intensities. The elemental maps represent the distribution of elements within each sample. S1, S3 and S4 show dense, iron-rich zones with red and orange mapping hues indicating high Fe concentrations. S2 and S5 to S8 appear mostly green, indicating lower metal content and a dominance of lighter elements such as O, Si, and Al.

Table 1. Elemental composition of iron ores and slags from different sites by using EDS

Element	Atomic Percentage (%)							
	S1	S2	S3	S4	S5	S6	S7	S8
Fe	37.5	7.7	24.8	27.7	0.5	2.1	0.9	0.8
O	61.2	89.6	58.5	57.2	78.1	83.2	86.8	75.3
Si	0.7	1.6	12.2	10.7	13.6	7.2	6.6	13.9
Al	0.1	0.5	1.6	1.6	4.6	3.7	3.3	9.5
Ni	0.1	0.4	0.1	0.1	0.1	1.6	0.7	-
P	-	-	-	-	-	-	-	-
S	-	0.1	-	-	-	-	0.2	-
Mn	0.3	-	1.9	1.7	0.6	0.1	0.2	-
Ca	-	0.0	0.6	0.7	0.3	0.1	0.1	0.1
Cu	-	-	0.2	1.0	-	-	-	-
K	-	-	0.1	0.1	0.1	-	-	0.
Mg	-	-	-	-	1.9	1.9	1.2	0.2
Ti	-	-	-	-	-	-	-	0.1
Cl	-	-	-	-	0.1	-	-	-

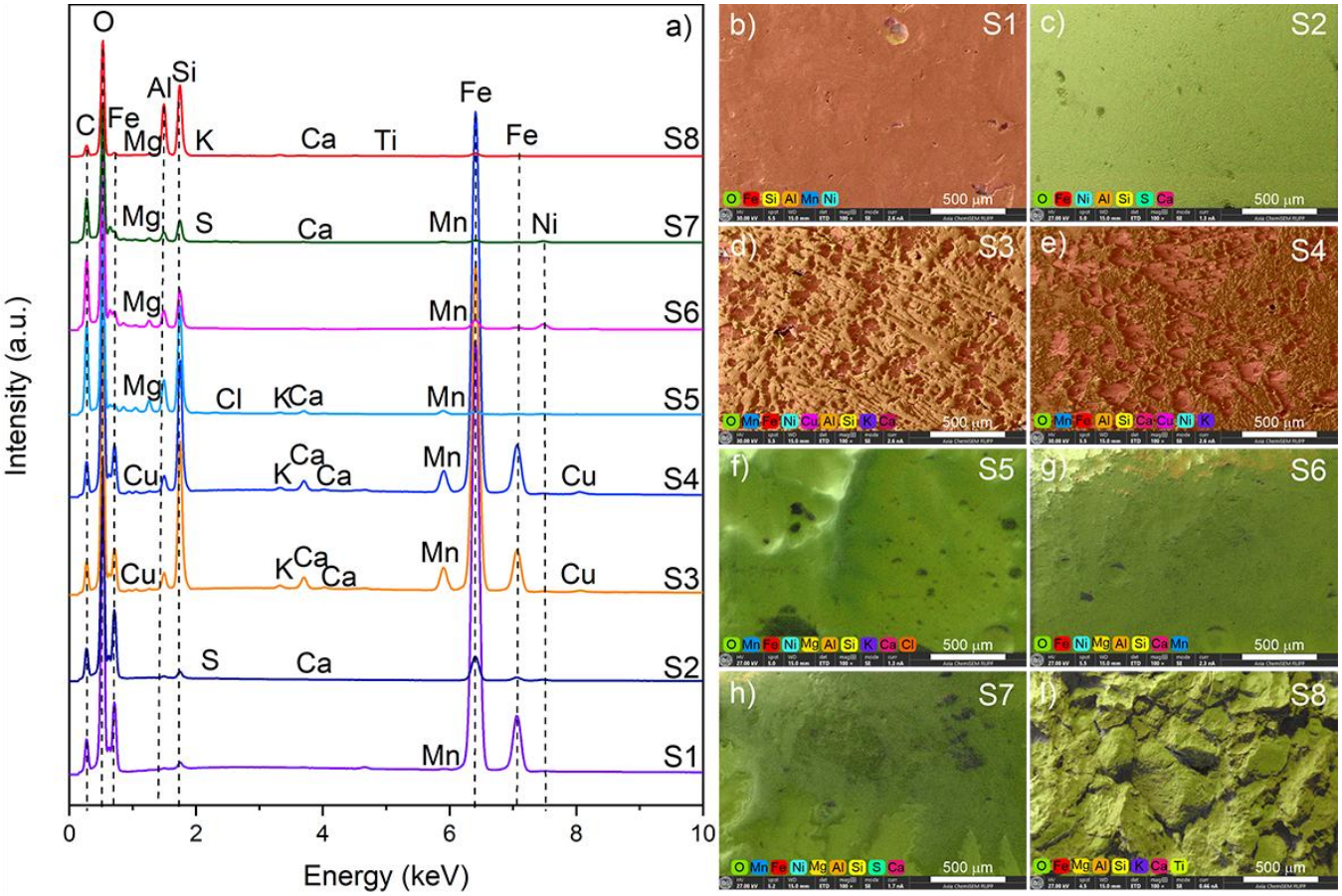


Fig. 1. a) EDS spectra of samples collected from Preah Vihear and EDS mapping of the samples labeled as b) S1, c) S2, d) S3, e) S4, f) S5, g) S6, h) S7, l) S8

S8, in particular, exhibits a highly porous and irregular texture, which may contribute to its low elemental concentration. Collectively, the EDS data and mapping provide valuable insights into the chemical composition, and elemental distribution present in the samples.

3.2 XRD Results

As mentioned earlier, XRD technique was utilized to obtained XRD patterns of iron ores and slags as shown in Figure 2. The XRD analysis revealed the presence of hematite (Fe_2O_3) and magnetite (Fe_3O_4) in S1, as shown in Figure 2a, while quartz (SiO_2) is observed in S2 and S8 in addition to hematite. Moreover, S3 composes of fayalite (Fe_2SiO_4) and magnetite, while fayalite, magnetite, hematite, and quartz were observed in S4. Furthermore, S7 composes of pyroxmangite ($MnSiO_4$) and quartz, while pyroxmangite is observed in S6 in addition to fayalite and magnetite. In addition, fayalite, quartz, and ferrosilite ($FeSiO_3$) were detected in S5. Analysis of S8 revealed a low Fe content but a high Si content, along with other elements such as Al, Ni Ca, K, Mg, and Ti. XRD data confirmed the presence of SiO_2 and Fe_2O_3, which matches the high Si concentration. This suggests that S8 could potentially be blast furnace slag (BFS), as BFS typically contains: SiO_2 (30–36%), CaO (32–45%), Al_2O_3 (10–16%), MgO (6–9%), Fe_2O_3 (0.2–1.5%) [8]. The sample from Phnom Daek (S1) was revealed to have a high percentage of iron according to EDS data, while the XRD pattern showed the presence of hematite and magnetite. These findings are consistent with previous findings [7], which suggest that S1 is iron ore. The Russey Treb samples (S2 to S7) showed inconsistencies between the samples in both EDS and XRD data. S2 had a high level of Fe concentration after S3 and S4, while the XRD pattern revealed the presence of Fe_2O_3 and SiO_2. This implies that most of the Fe in the sample was mainly Fe_2O_3, which suggests that S2 was iron ore. The percentage of Fe in S3 and S4 is relatively higher than anything else in the sample, except oxygen. However, the XRD pattern confirmed the presence of fayalite with high peak intensities. It should be noted that fayalite is normally known to be a byproduct of iron smelting, as quartz sands were historically used as fluxing agents in early iron production [9]. This additionally confirms that S3 and S4

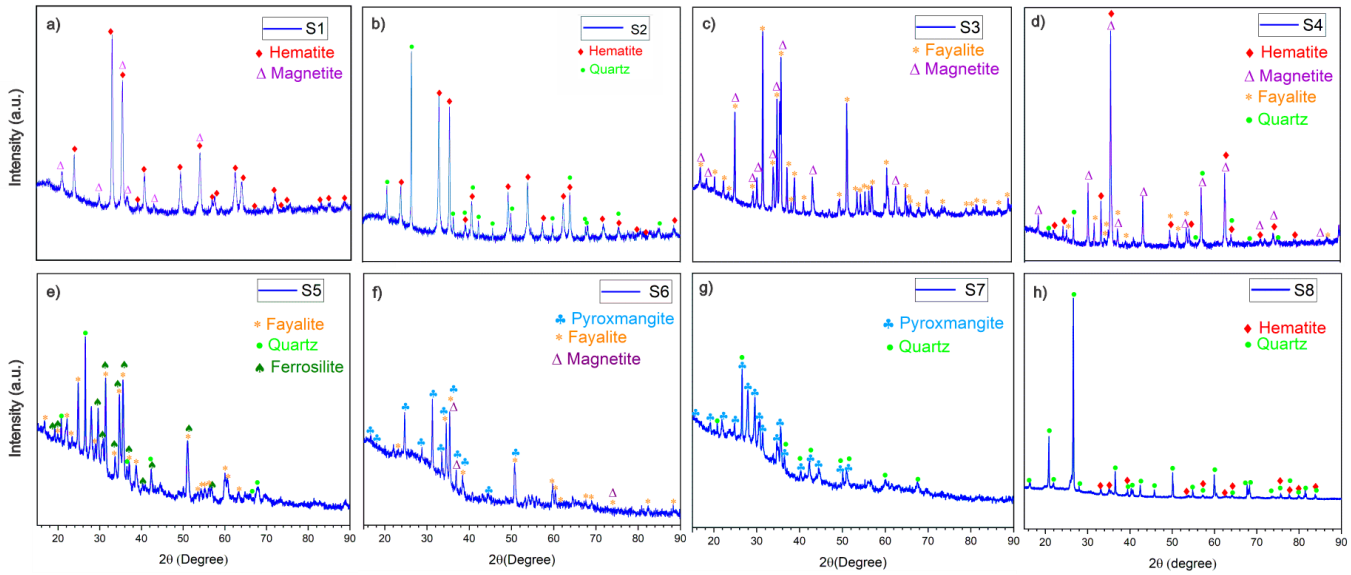


Fig. 2. XRD patterns of iron ore and slag samples collected from Preah Vihear, illustrating the elemental composition present in each material.

were slag from the smelting process. On the other hand, S5 and S6 both contained fayalite, while ferrosilite was additionally found in S5. Pyroxmangite was, on the other hand, found in S6. The XRD analysis of S7, on the other hand, revealed only the presence of pyroxmangite and quartz.

4. CONCLUSIONS

This study presents the elemental compositions and mineralogical proportions of iron ore and slag samples collected from archaeological sites in Preah Vihear province, Cambodia, specifically Phnom Daek, Russey Treb, and Phumi Boeng. The analyses were conducted using EDS and XRD techniques. EDS results indicate that the sample collected from Phnom Daek (S1) has a high iron content (37.5%), suggesting a high-quality iron ore. In contrast, samples from Russey Treb and Phumi Boeng exhibit lower iron concentrations, several of which were identified as iron slags due to their elevated silicon and manganese contents, as well as the presence of fayalite, quartz, and pyroxmangite as confirmed by XRD analyses. These findings highlight the variability in ore quality across the sites and provide evidence of localized smelting activities in the region. The identification of hematite, magnetite, and fayalite in various samples supports the interpretation of both raw iron ores and by-products of ancient iron production. Overall, this study contributes to the ongoing archaeological and metallurgical research in Cambodia.

ACKNOWLEDGMENTS

This study was supported by the Royal Government of Cambodia through the Higher Education Improvement Project (IDA Credit No. 6221-KH) and Swedish International Development Cooperation Agency (SIDA) through Sweden and Royal University of Phnom Penh's Pilot Research Cooperation Programme (Sida Contribution No. 11599).

REFERENCES

[1] E. Uchida and I. Shimoda, "Quarries and transportation routes of Angkor monument sandstone blocks," *J. Archaeol. Sci.*, vol. 40, no. 2, pp. 1158–1164, Feb. 2013, doi: 10.1016/j.jas.2012.09.036.

[2] M. Hendrickson, S. Leroy, Q. Hua, K. Phon, and V. Voeun, "Smelting in the Shadow of the Iron Mountain: Preliminary Field Investigation of the Industrial Landscape around Phnom Dek, Cambodia (Ninth to Twentieth Centuries a.d.)," *Asian Perspect.*, vol. 56, no. 1, pp. 55–91, 2017.

[3] M. Polkinghorne, B. Vincent, N. Thomas, and D. Bourgarit, "Casting for the King: The Royal Palace bronze workshop of Angkor Thom," 2014, doi: 10.3406/befeo.2014.6179.

[4] C. Jacques, P. Lafond, and T. White, *The Khmer empire : cities and sanctuaries, fifth to the thirteenth centuries*. River Books, 2007.

[5] M. Hendrickson, Q. Hua, and T. O. Pryce, "Using In-Slag Charcoal as an Indicator of 'Terminal' Iron Production within the Angkorian Period (10th–13th Centuries AD) Center of Preah Khan of Kompong Svay, Cambodia," *Radiocarbon*, vol. 55, no. 1, pp. 31–47, Jan. 2013, doi: 10.2458/azu_js_rc.v55i1.16152.

[6] T. O. Pryce *et al.*, "The Iron Kuay of Cambodia: tracing the role of peripheral populations in Angkorian to colonial Cambodia via a 1200 year old industrial landscape," *J.*

Archaeol. Sci., vol. 47, pp. 142–163, Jul. 2014, doi: 10.1016/j.jas.2014.04.009.

[7] E. Uchida, M. Murasugi, A. Kuroda, and Y. Lu, "Chemical Compositional Signatures of ConstituentMinerals of Iron Slags and Ores from the Khmer Monuments," *Heritage*, vol. 2, no. 2, Art. no. 2, Jun. 2019, doi: 10.3390/heritage2020105.

[8] S. Teir, S. Eloneva, C.-J. Fogelholm, and R. Zevenhoven, "Dissolution of steelmaking slags in acetic acid for precipitated calcium carbonate production," *Energy*, vol. 32, no. 4, pp. 528–539, Apr. 2007, doi: 10.1016/j.energy.2006.06.023.

[9] T. Rehren, M. Charlton, S. Chirikure, J. Humphris, A. Ige, and H. A. Veldhuijzen, "Decisions set in slag: the human factor in African iron smelting," in *In: LaNiece, S and Hook, DR and Craddock, PT, (eds.) Metals and Mines - Studies in Archaeometallurgy. (pp. 211-218). Archetype/British Museum: London, United Kindom. (2007)*, S. LaNiece, D. R. Hook, and P. T. Craddock, Eds., London, United Kindom: Archetype/British Museum, 2007,pp.211218. https://discovery.ucl.ac.uk/id/eprint/93663/

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Investigation on Mechanical Properties and Antimicrobial Activity of Natural Rubber Latex Examination Gloves Incorporated with Pomelo Peel Powder

Likea Heng [1], Soma Tey Nuon [1], Seka Kong [1], Laymey Sreng [2], Sovathana Phuong [3], Chakrya Theap [3], Chanthol Peng [3] Phanny Yos [2,3*]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3]*Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: phannyyos@itc.edu.kh*

Abstract: *The growing issue of agricultural waste has led to increased interest in repurposing byproduct for renewable applications. One such waste product is pomelo (Citrus maxima) peel, which is often discarded despite its rich bioactive compounds and potential functional applications. This study explores the incorporation of pomelo peel powder into natural rubber latex (NRL) examination gloves to enhance their mechanical properties and antimicrobial activity. By repurposing agricultural waste as a natural addictive, this study corresponding with the growing demand for eco-friendly innovations in the rubber and healthcare industries. Different concentration (2 and 6 parts per hundred parts of rubber by weight, phr) of pomelo peel powder blended into NRL formulations and processed into gloves. The modified gloves were then tested for mechanical properties and antimicrobial effectiveness. Result showed that an optimal concentration of pomelo peel powder improved the gloves' mechanical properties while introducing antimicrobial functionality. Higher concentrations slightly reduced elasticity but enhanced bacterial inhibitory. These findings suggest that pomelo peel powder is a promising natural antimicrobial agent that can be integrated into NRL gloves without significantly compromising their mechanical performance. This approach not only improve glove functionality but also provides a sustainable solution for agricultural waste management, supporting circular economic principles. The study highlights the potential for turning waste into value-added products, contributing to both environmental sustainability and improved hygiene standards in medical and food-handling industries.*

Keywords: Natural rubber latex gloves; Pomelo peel powder; Antimicrobial agent

1. INTRODUCTION

Natural rubber is harvested as latex from rubber trees and processed into products of varying quality. One basic method allows latex to solidify naturally, but adding acid speeds up coagulation, producing lower-quality 'cup lumps' commonly collected by small-scale farmers. These are later processed into raw rubber products like sheets and blocks.

Cambodia's rubber industry expanded rapidly after the 1990s with the introduction of Economic Land Concessions (ELCs), which utilized large forest areas for plantations. By 2021, plantation areas grew from 51,000 hectares in 1985 to over 300,000 hectares. Initially state-run and poorly managed, the industry improved following privatization in 1993, leading to increased productivity and a stronger contribution to Cambodia's economy. By 2022, rubber exports generated around $470 million, becoming the country's second-largest agricultural export after rice [1].

In healthcare and food industries, hands are a major source of pathogen transmission, and the problem is worsened by multidrug-resistant (MDR) microbes. Although guidelines for hand hygiene and glove use exist, adherence is inconsistent. Gloves can help reduce direct pathogen transfer but may act as a "second skin," increasing contamination risks

[1*] Corresponding author: Phanny Yos
E-mail: phannyyos@itc.edu.kh*; Tel: +855-12 870 256*

when improperly used. Healthcare workers' gloves often become contaminated through contact with patients or their surroundings, while in the food industry, bacteria like *E. coli* causes significant foodborne illness, leading to high healthcare and economic costs. Though gloves are commonly used to prevent cross-contamination, improper use can lower hygiene practices, making antimicrobial gloves a promising alternative [2]. This study aims to enhance the safety and performance of natural rubber latex examination gloves by incorporating pomelo peel powder (PPP) at 2, 3, and 4 phr. It evaluates the impact of PPP on the gloves' mechanical properties and investigates its antimicrobial activity against *Staphylococcus aureus* and *Escherichia coli* to assess its potential to inhibit bacterial growth.

2. METHODOLOGY

2.1 Materials

High ammonia natural rubber latex (60%), and chemical agents such as sulfur (S), potassium hydroxide (KOH), potassium oleate (PO), zinc oxide (ZnO), zinc diethyldithiocarbamate (ZDEC), Phenolic Antioxidant (PA), triton x-100 and Calcium Carbonate ($CaCO_3$) were purchased from Thailand. Calcium nitrate ($Ca(NO_3)_2$) were obtained from Elago Enterprise Pty Ltd, 5 The Cloisters, Cherrybrook, N.S.W. 2126, Australia. Chloroform was purchased from Ho Chi Minh city, Vietnam. Polyethylene glycol (PEG) was purchased from Wuxi Chemical Equipment Co., Ltd, China. Distilled water was supplied by laboratory of Faculty of Chemical and Food Engineering, ITC. The raw material, pomelo peel, was a commercially available local product obtained from the Chbar Ampov market. Pomelo peel powder (PPP) was prepared by drying in oven for 24h at 50℃ then grinded by using grinder machine and sieved to obtain particles with a diameter of less than 75 microns. To achieve an even finer particle size, the PPP was further ground using an agate mortar before subjected in ball mill with PEG and distilled water for 24h to further reduce its particle size and improve dispersion. The bacterial strains were provided by Platform for Aquatic Ecosystem Research (PAER Lab).

2.2 Latex solution compounding

Latex solution compounding formulation in part per hundred rubber (phr) is shown in Table 1. The process begins with a 30-minute stirring of the latex at a speed of 230 rpm to achieve a uniform mixture. Following this, ingredients are slowly added using a dropper to avoid chemical shock in the latex solution, at two-minute intervals, while stirring is maintained for another hour to promote even distribution. Next, the PPP filler solution is gradually introduced, and the mixture is stirred for an additional hour to ensure thorough blending. This is followed by the pre-vulcanization step, which initiates the formation of cross-links in the rubber matrix, enhancing its strength and flexibility. During this step, the latex compound is kept in a water bath while being stirred at 55 ℃, with progress checked every 10 minutes by mixing chloroform ($CHCl_3$) and latex in a 1:1 ratio until level 2 or 3 is reached. The final stage involves allowing the mixture to mature at room temperature for 24 hours, enabling its physical and chemical properties to fully develop before further use [3].

Table 1. Formulation of latex solution preparation

Ingredients	Amount (phr)
NR latex	100
KOH	0.5
Potassium laurate	0.25
ZnO	1
ZDEC	1.5
Sulfur	1.5
Antioxidant	1
PPP	2, 3, 4

2.3 Coagulation solution preparation

The coagulation formulation in grams is shown in Table 2. Distilled water was heated in a water bath to 55 °C, then calcium nitrate ($Ca(NO_3)_2$), calcium carbonate ($CaCO_3$), and Triton X-100 were added to the distilled water while stirring. The solution was kept in the water bath for a minimum of 10 hours before use for homogeneous mixture [3].

Table 2. Formulation of coagulant solution preparation

Ingredients	Amount (g)
Distilled water	100
$Ca(NO_3)_2$	10
$CaCO_3$	0.8
Triton x-100	0.06

2.4 Plate dipping process

The process of preparing a latex film begins with the filtration of the latex solution, during which dirt and impurities are removed using a filtration sieve. Next, the plate is prepared by washing and drying a porcelain plate at 100 ℃, it is then taken out and maintained at a temperature of 65-70 ℃. In the coagulation dipping step, the plate is dipped into a coagulant solution for 10 seconds (in and out), dried at 75 ℃

for 10 minutes, removed from the oven, and kept at 55-58 ℃. The latex dipping step follows, where the plate is dipped into latex for 10 seconds (in and out), then allowed to dry in the oven for 18 minutes at 100 ℃. After that, it was taken out and kept in room temperature until it reached 55-58 ℃. A second dipping is performed using the same procedure as the first dipping. Afterward, water leaching is carried out by immersing the latex-coated plate in heated distilled water at 70 ℃ for 3 minutes. The plate is then subjected to oven drying at 100 ℃ for 20 minutes. Finally, in the film removal step, the plate is cooled to room temperature, dusted with baby powder, and the latex film is stripped off the plate [3].

2.5 Thickness measurement

Thickness measurement was performed according to ASTM D412-16. A film from each concentration of PPP solution (2,3, and 4) was cut into five pieces using a dumbbell cutter. The thickness of each piece was measured using a thickness gauge. The thickness of each specimen is measured at three points, the centre and both ends of the narrow section of the film. The median of these three measurements was used to calculate the cross-sectional area. While the difference between the thickest and thinnest points is more than 0.08 mm, the specimen is discarded. The width is determined by the distance between the die's cutting edges in the narrow section.

2.6 Tensile properties

Tensile properties, including tensile strength and elongation at break, were evaluated according to ASTM D412 using a universal testing machine (LFM-TOP-50). The test specimens were cut in alignment with the dipping direction of the film using a standard cutter, Die D.

2.7 Aging test

Aging test was conducted in accordance with ASTM D 573. Samples were placed in an air oven at 100 ℃ for 22 hours [4]. After aging, they were kept at room temperature for one day before being subjected to property testing, including thickness, tensile strength, and elongation at break. The resulting data were then compared with those of unaged samples.

2.8 Antimicrobial activities of PPP films

The *E. coli* and *S. aureus* strain were prepared by using pure culture from stock bacteria suspension *E. coli* K12 and *S. aureus* A1CC at -80°C. One loop of pure bacteria was taken and streaked on the prepared Muller-Hinton agar, then was incubated in incubator at 37°C for 24 hours. After 24h, bacteria inoculum of 10^8CFU/ml were prepared by using phosphate buffer solution (PBS) and were adjusted to 0.5McFarland turbidity standard. The cultured was then used for antimicrobial activity test.

The antimicrobial test was conducted according to the modified standard method CLSI. First, MHA agar was prepared by suspended 38g of MHA in one litter of distilled water. The solution of MHA was then heated with agitation until boiled for 1 minute, in order to completely dissolved the medium. The MHA was kept to cool to room temperature, then it was poured into plastic petri dish 99mm dimension, at least 4mm in dept. The plates were kept until MHA solidified. Second, the prepared bacteria strain of concentration 10^8CFU/ml was inoculated into the MHA plates by streaking using swab. Third, 4phr PPP incorporated and control film which has no filler (PPP), were cut into small piece with dimension 30mm×30mm. Three test pieces were cut from 3 different films of PPP incorporate and control film for triplicating the test. Fourth, the test pieces were placed on the MHA with inoculated bacteria, one for each plate. Lastly, the plates were placed in 37°C incubator for 24h. The zone of inhibition was measured using caliper and ZI data were recorded after 24h of incubation.

The minimum inhibitory concentration (MIC) was done by spot test method. The spot test was conducted by preparing 0.5g the paste (ball milled PPP solution), then it was diluted 5 times with sterile de-ionized water, and agitated using voltage to fully dissolved the solution in to sterile DI water. After that, 5µl and 10µl of diluted solution was added onto two different MHA plates for three spots. Lastly, the plates were incubated at 37°C for 24h. After 24h, the zone of inhibition was measured and recorded.

3. RESULTS AND DISCUSSION

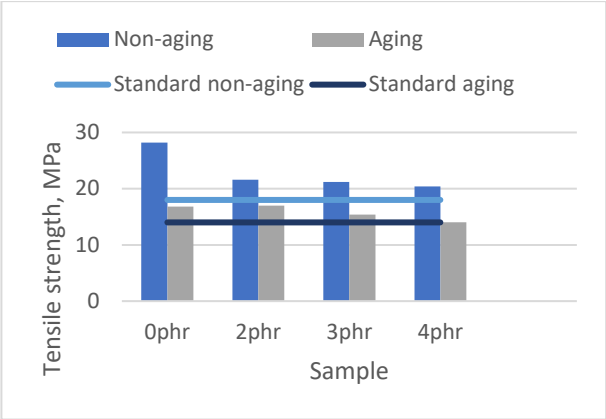


Fig. 1 Tensile strength of non-aging and aging of NRL films containing varying contents of PPP. The light and dark blue solid lines represent minimum requirements for non-aging and aging medical examination latex gloves, respectively, according to ASTM D3578-01.

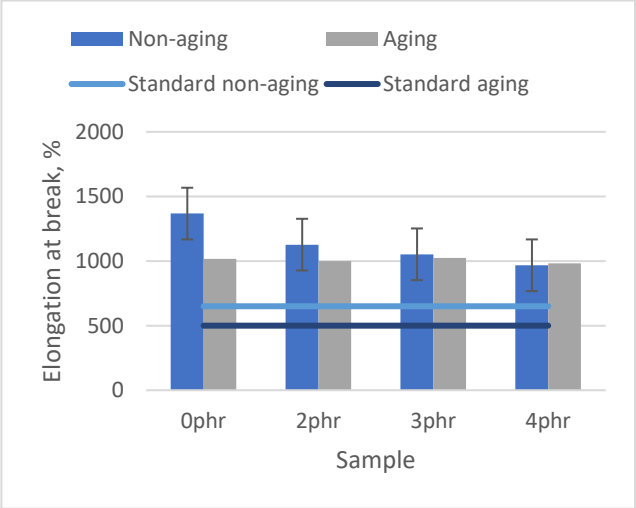


Fig. 2 Elongation at break of non-aging and aging of NRL films containing varying contents of PPP. The light and dark blue solid lines represent minimum requirements for non-aging and aging medical examination latex gloves, respectively, according to ASTM D3578-01.

The tensile properties, including tensile strength and elongation at break, of NRL films incorporating varying concentrations of PPP at 0, 2, 3, and 4 phr are presented in Fig. 1. and 2. The results demonstrate a general decline in tensile strength with increasing PPP content. This reduction may be attributed to unfavorable filler-filler interactions and the presence of air bubbles, which can lead to void formation within the film matrix. Despite the observed decrease, all PPP-incorporated films fulfilled the minimum tensile strength requirements specified for medical examination gloves. Furthermore, the findings suggest that higher PPP loading is feasible in NRL glove production to improve antimicrobial efficacy. Elongation at break also showed a decreasing trend with increasing PPP concentration, likely due to the restricted mobility of natural rubber chains caused by the incorporation of PPP particles, which interfere the mobilization of NRL chains along the direction of external force [3].

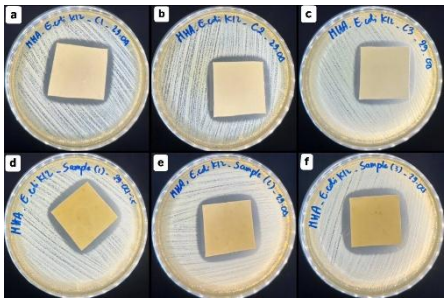


Fig. 3 Images of zone of inhibition against *E. coli*: (a), (b), and (c) are control films, (d), (e), and (f) are PPP incorporated films.

Fig 3. presents the bacterial growth on Mueller Hinton Agar (MHA), with control samples and PPP-incorporated films positioned at the center of the plates. The control films (a, b, and c) exhibited smaller zones of inhibition against *E. coli* compared to the PPP-containing films (d, e, and f), which displayed significantly larger and clearer zones, indicating enhanced antimicrobial activity. These results suggest that the incorporation of PPP imparts antibacterial properties to the films, likely due to the presence of bioactive compounds capable of suppressing bacterial proliferation. Furthermore, the formation of clear inhibition zones implies that the antimicrobial agents are diffusible, migrating from the NRL film matrix into the surrounding agar to inhibit bacterial growth [3].

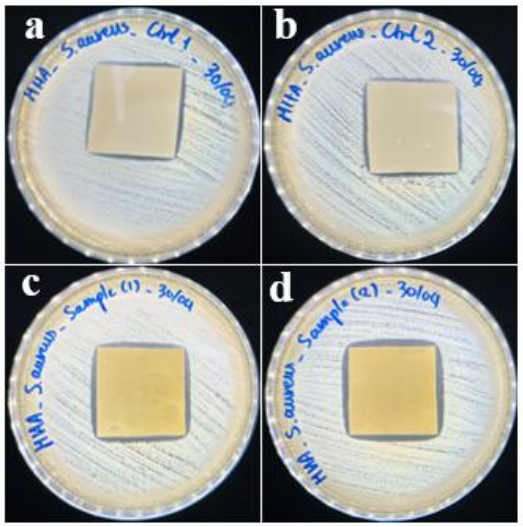


Fig. 4 Images of zone of inhibition against *S. aureus*: (a) and (b) are control films, (c), and (c) are PPP incorporated films.

Fig. 4 illustrates the growth pattern of *Staphylococcus aureus* on Mueller Hinton Agar (MHA), with both control samples

and PPP-incorporated films positioned centrally on the plates. The control films (a, and b) produced relatively smaller inhibition zones possibly due to ZnO used in latex compounding which has antimicrobial properties [3], compared to the PPP films (c and d), which inhibited broader and more pronounced clear zones-evidence of enhanced antibacterial efficacy.

Overall, the results against both *E. coli* and *S. aureus* demonstrate that films incorporated with PPP are effective in inhibiting the growth of both Gram-negative and Gram-positive bacteria.

4. CONCLUSION

Adding pomelo peel powder (PPP) to natural rubber latex films improved their ability to fight bacteria. Although the films became slightly weaker with more PPP, they still met the standard for medical gloves. The PPP films showed clear zones of inhibition against both *E. coli* and *S. aureus*, proving they can stop the growth of both Gram-negative and Gram-positive bacteria. This suggests that PPP can be used to make antimicrobial gloves.

ACKNOWLEDGMENTS

The work was funded by Takahashi Project No. 14-003-215. All laboratory experiments were conducted in Laboratory of Rubber Processing of ITC, and antimicrobial activity testing was carried out by PAER Lab.

REFERENCES

[1] 'UPSTREAM AND MIDSTREAM ANALYSIS OF NATURAL RUBBER MARKET SYSTEM IN EASTERN CAMBODIA'.

[2] R. A. Reitzel, T. L. Dvorak, R. Y. Hachem, X. Fang, Y. Jiang, and I. Raad, 'Efficacy of novel antimicrobial gloves impregnated with antiseptic dyes in preventing the adherence of multidrug-resistant nosocomial pathogens', *Am. J. Infect. Control*, vol. 37, no. 4, pp. 294–295, May 2009, doi: 10.1016/j.ajic.2008.07.003.

[3] Kevin Gwee, 'The preparation of antimicrobial natural rubber (NR) latex films for stethoscope diaphragm cover application', PhD thesis, Universiti Sains Malaysia, Penang, Malaysia, 2021.

[4] A. Thumwong, J. Darachai, N. Thamrongsiripak, S. Tokonami, T. Ishikawa, and K. Saenboonruang, 'Fruit Peel Powder as Natural Antioxidant and Reinforcing Bio-Filler in Natural Rubber Latex Gloves: Cases of Mangosteen, Pomelo and Durian', *Antioxidants*, vol. 12, no. 5, p. 1119, May 2023, doi: 10.3390/antiox12051119.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Degradation Study of Few-Layer SnP_2X_6 (X=S, Se) in Ambient Conditions Using OM, AFM, and SEM

Sok Heng Hor[1] and Tharith Sriv[1, 2, *]

[1]*Graduate Program in Physics, Graduate School of Science, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*
[2]*Department of Physics, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*

[*]*Corresponding author: tharith@rupp.edu.kh*

Abstract: *Two-dimensional (2D) materials are of an ultrathin nature and have unique properties that are extremely useful for advanced applications such as electronics, optoelectronics, photonics, energy storage, sensors, and biomedicine. However, some of the 2D materials encounter limited applications due to instability in ambient conditions. For this reason, before further investigation of their physicochemical properties, scientists usually study the stability of 2D materials. Tin chalcogenophosphates [SnP_2X_6, (X=S, Se)] are indirect bandgap 2D semiconductors that have attracted much interest for advanced applications in optoelectronic devices due to their high carrier mobility. Additionally, these 2D semiconducting materials exhibit high dielectric constants, nonlinear optical properties, and other physical properties for electronics, photocatalysis, photodetectors, photovoltaics, and other potential applications. However, the electronic and optical properties of the mechanically exfoliated few-layer SnP_2X_6 might change in ambient conditions. Furthermore, the presence of the metal Sn atom and chalcogen vacancies X of 2D SnP_2X_6 can lead to degradation under ambient conditions. Therefore, in this work, we investigated the degradation behavior of few-layer SnP_2X_6 (X=S, Se) utilizing a combination of advanced non-destructive microscopies, namely optical microscopy, atomic force microscopy (AFM), and scanning electron microscopy (SEM). A comparison of OM, AFM, and SEM images of exfoliated samples after being left for one day and several more days in ambient conditions revealed that few-layer SnP_2X_6 samples suffer degradation in ambient conditions. Furthermore, we notice that few-layer SnP_2Se_6 samples are significantly more susceptible to degradation compared to those of SnP_2S_6 samples. This gives evidence of fast degradation in few-layer SnP_2X_6 in ambient conditions and suggests that further study of mechanically exfoliated few-layer SnP_2X_6 samples must be conducted in vacuum conditions or encapsulated. Moreover, the device designs by 2D SnP_2X_6 require a consideration of their degradation, especially when the devices are to be used in ambient conditions.*

Keywords: Degradation; SnP_2S_6; SnP_2Se_6; AFM; SEM

1. INTRODUCTION

Graphene, black phosphorus (BP), silicene, germanene, hexagonal boron nitride (h-BN), two-dimensional (2D) transition metal dichalcogenides (TMDCs), and other 2D materials are of an ultrathin nature and have unique properties that are extremely useful for advanced applications such as electronics, optoelectronics, photonics, energy storage, sensors, catalysis, and biomedicine [1]. However, some of the 2D materials encounter limited applications due to instability in ambient conditions [2]. For this reason, before further investigation of their physicochemical properties, the stability study of 2D materials is systematically performed [3, 4]. As a novel class of the 2D materials, tin chalcogenophosphates [SnP_2X_6, (X=S, Se)] family with space group of R3 (No. 146) and rhombohedral crystal structure are indirect bandgap 2D semiconductors that have attracted much interest for advanced applications in electronic and optoelectronic devices due to their high carrier mobility [5, 6]. Additionally, these 2D semiconducting materials exhibit high dielectric constants, nonlinear optical properties, and other physical properties for photocatalysis, photodetectors, photovoltaics, and other potential applications [7, 8, 9, 10, 11, 12, 13, 14, 15]. The carrier mobilities and bandgaps, for instance, of tin thiophosphate (SnP_2S_6) and tin selenophosphate (SnP_2Se_6) in monolayer form are known to be 890 $cm^2 v^{-1} s^{-1}$, 819 $cm^2 v^{-1} s^{-1}$, and 2.244 eV, 1.568 eV, respectively [5, 6]. However, the electronic and optical properties of the mechanically

[1]* Corresponding author: Tharith Sriv
E-mail: tharith@rupp.edu.kh; Tel: +855-85 289 791

exfoliated few-layer SnP_2X_6 might change in ambient conditions [2, 16]. Furthermore, the presence of the metal Sn atom and chalcogen vacancies (X) of 2D SnP_2X_6 can lead to degradation under ambient conditions [6, 17, 18].

Given these challenges, it is essential to investigate the degradation behavior of few-layer SnP_2S_6 and SnP_2Se_6 to ascertain their suitability for long-term applications in ambient environments. Ambient factors such as oxygen, humidity, and light can induce various degradation mechanisms, impacting the structural, chemical, and functional integrity of these materials [19].

In this work, we investigate the degradation of few-layer SnP_2S_6 and SnP_2Se_6 in ambient conditions utilizing a combination of advanced microscopies, namely optical microscopy, atomic force microscopy (AFM), and scanning electron microscopy (SEM). A comparison of OM, AFM, and SEM images revealed clear signs of degradation of few-layer SnP_2X_6 in ambient conditions. The rests of the paper are organized as follow. Section 2 describes how mechanically exfoliated samples are prepared and measured by using OM, AFM and SEM. Section 3 provides scientific results of the degradation study by the microscopic techniques, and further discusses the results, while the last section gives conclusion remarks.

2. METHODOLOGY

2.1 Preparation of Samples

Fresh few-layer SnP_2X_6 (X=S, Se) samples were mechanically exfoliated from high-quality single crystals (HQ Graphene) onto clean silicon substrates with a 280-nm silicon dioxide layer. Several exfoliated SnP_2S_6 and SnP_2Se_6 samples were then further studied to find out the degradation behavior, as mentioned earlier.

2.2 Degradation Investigation

To investigate the degradation behavior of the exfoliated few-layer SnP_2S_6 and SnP_2Se_6 samples, we utilized a combination of advanced microscopic techniques, namely OM for monitoring color, contrast, and visible morphological changes; AFM (Nanosurf FlexAFM) for measuring surface topography and roughness, revealing nanoscale degradation features; and SEM (ThermoFisher Axia ChemiSEM) for providing high-resolution imaging, for further revealing degradation signs.

3. RESULTS AND DISCUSSION

Figure 1 illustrates the OM and AFM images of few-layer SnP_2S_6. Figure 1*a* is the optical image obtained right after the mechanical exfoliation (within the first hour as time was spent to find the sample on the SiO_2 substrate). This optical image of the sample is assumed to be at 0 hour. Figure 1*d* is the corresponding AFM image observed about 2 hours and 42 minutes after being exposed to ambient air. Figures 1*b-c* are optical images checked after being exposed for about 1 day and 56 minutes and 7 days 1 hour and 59 minutes, respectively, whereas Figures 1*e-f* are AFM images of the sample obtained after being exposed to ambient air. According to these optical and AFM results, few-layer samples are significantly degraded as observed in the white spots on the surface of the sample.

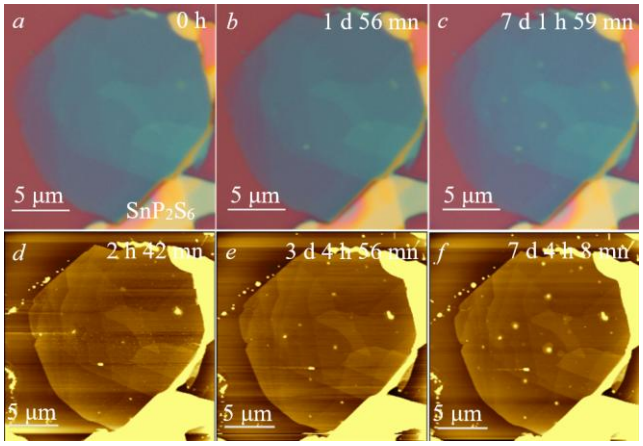


Fig. 1. Optical (*a-c*) and AFM (*d-f*) images of few-layer SnP_2S_6 (Sample 1) on SiO_2/Si substrate after being exposed to ambient air for: *a*) 0 hour, *b*) 1 day and 56 minutes, *c*) 7 days 1 hour and 59 minutes, *d*) 2 hours and 42 minutes, *e*) 3 days 4 hours and 56 minutes, *f*) 7 days 4 hours and 8 minutes

To clearly understand the degradation of few-layer SnP_2S_6 samples, we studied several more few-layer samples. Figures 2 and 3 further illustrate signs of degradation. As shown in Figure 2, which consists of optical and AFM images of Sample 2, significant change of the optical contrasts (almost unseen at thinner part as indicated by the white arrow in Figure 2.c) clearly indicates the degradation. The sign of degradation can even be more clearly observed through white spots in AFM images (Figures 2*d-f*). This confirms the degradation results from Figure 1. Similarity is observed in Figure 3, where a good SnP_2S_6 sample (Sample 3) appeared significantly different after being exposed to ambient air for about 38 days.

As part of a more systematic study, the scanning electron microscopy was further utilized as mentioned earlier. The results of the SEM obtained for Samples 1 and 2 (Figures 4) reveal clear black spots on the surface of the samples. This confirms degradation of mechanically exfoliated few-layer SnP_2S_6 samples on SiO_2/Si substrates. The results suggest a careful consideration before the atomically thin SnP_2S_6 materials are used for the designs of microelectronic devices such as field-effect transistors, photodetectors, or sensors. Additionally, vacuum condition or encapsulation of the samples by more stable materials such as graphene, MoS_2, WS_2 samples are to be considered before atomically SnP_2S_6 samples are further characterized.

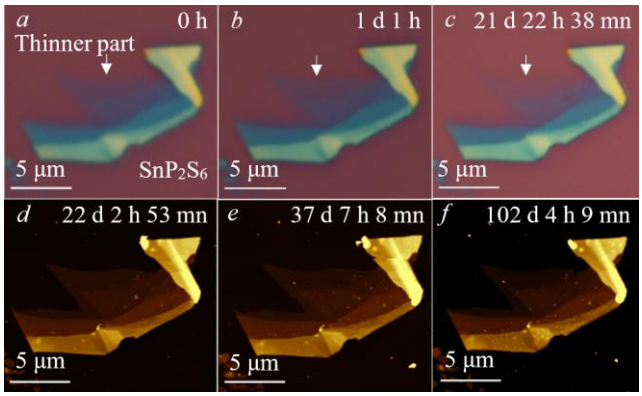


Fig. 2. Optical (*a-c*) and AFM (*d-f*) images of few-layer SnP_2S_6 (Sample 2) on SiO_2/Si substrate after being exposed to ambient air for: *a*) 0 hour, *b*) 1 day and 1 hour, *c*) 21 days 22 hours and 38 minutes, *d*) 22 days 2 hours and 53 minutes, *e*) 37 days 7 hours and 8 minutes, and *f*) 102 days 4 hours and 9 minutes

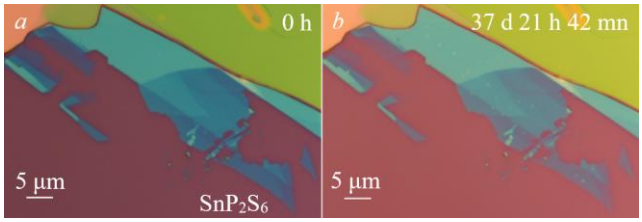


Fig. 3. Optical images of few-layer SnP_2S_6 (Sample 3) on SiO_2/Si substrate at different times: a) 0 hour, b) 37 days 21 hours and 42 minutes

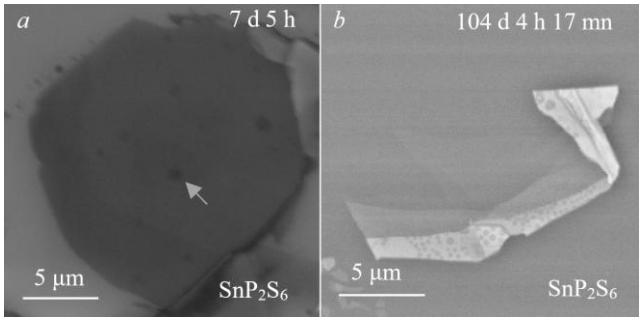


Fig. 4. SEM images of few-layer SnP_2S_6 after being exposed to ambient air for: *a*) 7 days and 5 hours (Sample 1) and *b*) 104 days 4 hours and 17 minutes (Sample 2). White arrow is placed in Figure 4.a to guide the eye to one degradation spot. Similar spots on the surface of the sample are also the sign of degradation

As mentioned earlier, degradation investigation was similarly done for atomically thin SnP_2Se_6 samples. Figure 5 consists of optical and AFM results of the measurements done for a SnP_2Se_6 sample. In this figure, although the optical images do not clearly show the degradation due to small size of the sample, the AFM results clearly indicate the degradation (Figure 5*d-f*). This is observed from the dramatic alteration in surface roughness of the sample observed in Figures 5*d-f*.

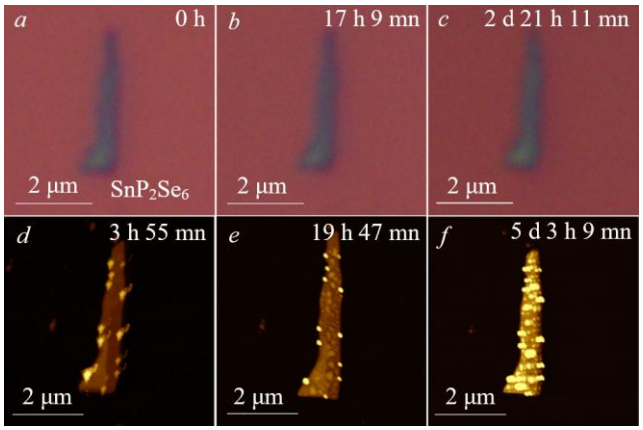


Fig. 5. Optical and AFM images of few-layer SnP_2Se_6 (Sample 4) on SiO_2/Si substrate after being exposed to ambient air for: a) 0 hour, *b*) 17 hours and 9 minutes, *c*) 2 days 21 hours and 11 minutes, *d*) 3 hours and 55 minutes, *e*) 19 hours and 47 minutes, and *f*) 5 days 3 hours and 9 minutes

The following figure additionally illustrates degradation of atomically-thin SnP_2Se_6 samples. In this figure, optical contrast change is clearly observed (Figure 6.*b*) as guided by the dashed line, indicating significant degradation of the sample after about 11 days in air.

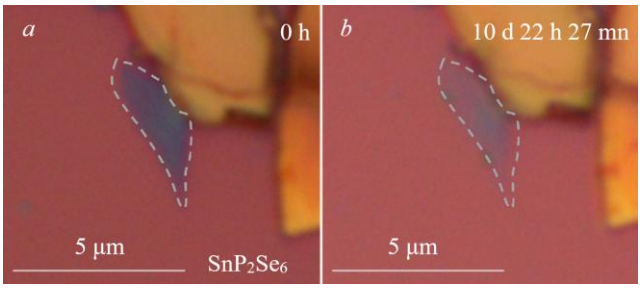


Fig. 6. Optical images of few-layer SnP_2Se_6 (Sample 5) on SiO_2/Si substrate after being exposed to ambient air for: a) 0 hour, and *b*) 10 days 22 hours and 27 minutes

Figure 7 shows the optical images of few-layer SnP_2Se_6 (Sample 6) on SiO_2/Si substrate after being exposed to ambient air. According to this figure, the degradation happened on the surface, which additionally confirms the previous results.

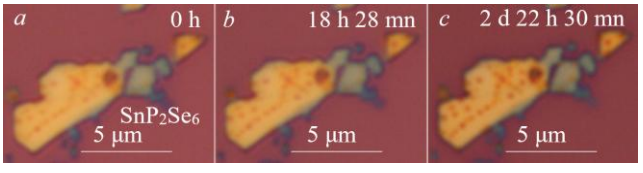


Fig. 7. Optical images of few-layer SnP_2Se_6 (Sample 6) on SiO_2/Si substrate after being exposed to ambient air for: *a*) 0 hour, *b*) 18 hours and 28 minutes, and *c*) 2 days 22 hours and 30 minutes

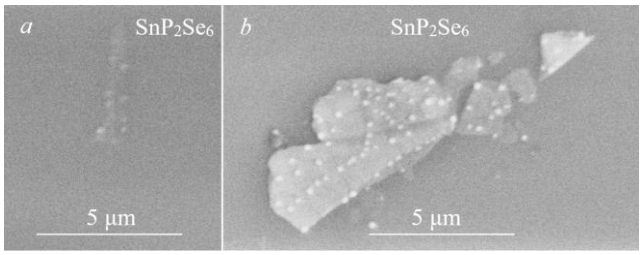


Fig. 8. SEM images of few-layer SnP_2Se_6 (Samples 4 and 6) after being exposed to ambient air for 5 days 22 hours and 35 minutes

Figures 8*a* and *b* are the SEM images of Samples 4 and 6, respectively, measured after about 6 days. According to the results, mechanically exfoliated SnP_2Se_6 samples significantly degrade. Similar to the case of sulfide samples shown in Figures 1-4, this requires careful consideration in the device designs and further scientific studies of the materials.

4. CONCLUSIONS

We investigated the ambient degradation of few-layer SnP_2X_6 (X = S, Se) using optical microscopy, atomic force microscopy, and scanning electron microscopy. Our results reveal that few-layer SnP_2X_6 undergoes significant degradation under ambient conditions, as evidenced by pronounced morphological changes and increased surface roughness. Notably, SnP_2Se_6 exhibits a higher susceptibility to environmental degradation than SnP_2S_6. These findings indicate that few-layer SnP_2X_6 compounds are inherently unstable in ambient environments. To improve their stability and performance, SnP_2X_6 materials should be stored in vacuum desiccators, and future scientific studies on the materials should be conducted under vacuum conditions to minimize surface degradation. Furthermore, degradation effects must be considered in the design and application of SnP_2X_6-based devices, particularly for use in ambient environments.

ACKNOWLEDGMENTS

This work was supported by the Royal Government of Cambodia through the Higher Education Improvement Project (IDA Credit No. 6221-KH) and Swedish International Development Cooperation Agency (SIDA) through Sweden and Royal University of Phnom Penh's Pilot Research Cooperation Programme (Sida Contribution No. 11599).

REFERENCES

[1] S. Singh, K. Verma, and C. Prakash, "Materials Horizons: From Nature to Nanomaterials Advanced Applications of 2D Nanostructures Emerging Research and Opportunities." Springer Nature Singapore Pte Ltd, Singaore (2021)

[2] Q. Li, Q. Zhou, L. Shi, Q. Chen, and J. Wang, "Recent advances in oxidation and degradation mechanisms of ultrathin 2D materials under ambient conditions and their passivation strategies," 2019, *Royal Society of Chemistry*. doi: 10.1039/c8ta10306b.

[3] C. Dong, L. S. Lu, Y. C. Lin, and J. A. Robinson, "Air-Stable, Large-Area 2D Metals and Semiconductors," Apr. 17, 2024, *American Chemical Society*. doi: 10.1021/acsnanoscienceau.3c00047.

[4] X. Wang, Y. Sun, and K. Liu, "Chemical and structural stability of 2D layered materials," *2d Mater*, vol. 6, no. 4, Jul. 2019, doi: 10.1088/2053-1583/ab20d6.

[5] M. Lin *et al.*, "Two-dimensional nanoporous metal chalcogenophosphates MP_2X_6 with high electron mobilities," *Appl Surf Sci*, vol. 493, pp. 1334–1339, Nov. 2019, doi: 10.1016/j.apsusc.2019.07.115.

[6] F. Li, X. Bo, C. Wu, Y. Zhang, S. Li, and H. Wu, "High carrier mobility and strong anisotropy in the family of monolayer SnP_2X_6 (X=S, Se, Te)," *Comput Mater Sci*, vol. 244, Sep. 2024, doi: 10.1016/j.commatsci.2024.113245.

[7] J. Hu *et al.*, "2D semiconductor SnP_2S_6 as a new dielectric material for 2D electronics," *J Mater Chem C Mater*, vol. 10, no. 37, pp. 13753–13761, Jul. 2022, doi: 10.1039/d2tc01340a.

[8] C. Zhu *et al.*, "Strain-Reduced Inversion Symmetry in Ultrathin SnP_2Se_6 Crystals for Giant Bulk Piezophotovoltaic Generation," *ACS Nano*, Jan. 2025, doi: 10.1021/acsnano.4c12605.

[9] V. K. Sangwan *et al.*, APPLIED SCIENCES AND ENGINEERING "Bulk photovoltaic effect and high mobility in the polar 2D semiconductor SnP_2Se_6" 2024, doi: 10.1126/sciadv.ado8272

[10] S. Wang *et al.*, "A high-performance self-powered photodetector based on SnP_2S_6 in the visible light region," *J Mater Chem C Mater*, vol. 11, no. 48, pp. 16941–16949, Nov. 2023, doi: 10.1039/d3tc03312k.

[11] B. X. Zhu *et al.*, "Two-dimensional SnP_2Se_6 with gate-tunable Seebeck coefficient for telecommunication band photothermoelectric detection," *InfoMat*, Oct. 2024, doi: 10.1002/inf2.12600.

[12] Y. Jing, Z. Zhou, J. Zhang, C. Huang, Y. Li, and F. Wang, "SnP_2S_6 monolayer: A promising 2D semiconductor for photocatalytic water splitting," *Physical Chemistry Chemical Physics*, vol. 21, no. 37, pp. 21064–21069, 2019, doi: 10.1039/c9cp04143e.

[13] C. Y. Zhu *et al.*, "Two-dimensional semiconducting SnP_2Se_6 with giant second-harmonic-generation for monolithic on-chip electronic-photonic integration,"

Nat Commun, vol. 14, no. 1, Dec. 2023, doi: 10.1038/s41467-023-38131-2.

[14] D. Chica *et al.*, "SnP_2Se_6: A Chiral 2D Semiconductor for High-Performance Electronics and Optoelectronics," Oct. 17, 2022. doi: 10.21203/rs.3.rs-2143024/v1.

[15] Y. Zhang *et al.*, "2D Ruddlesden-Popper perovskite sensitized SnP_2S_6 ultraviolet photodetector enabling high responsivity and fast speed," *Nanoscale Horiz*, vol. 8, no. 1, pp. 108–117, Nov. 2022, doi: 10.1039/d2nh00466f.

[16] G. D'olimpio *et al.*, "Tin Diselenide ($SnSe_2$) Van der Waals Semiconductor: Surface Chemical Reactivity, Ambient Stability, Chemical and Optical Sensors," Feb. 01, 2022, *MDPI*. doi: 10.3390/ma15031154.

[17] Y. Zhang *et al.*, "Inversion symmetry broken 2D SnP_2S_6 with strong nonlinear optical response," *Nano Res*, vol. 15, no. 3, pp. 2391–2398, Mar. 2022, doi: 10.1007/s12274-021-3806-0.

[18] L. Ma, W. P. Xu, J. X. Cao, and X. L. Wei, "Surface Oxidation and S Atom Vacancy in Tuning the Photoelectric Properties of Monolayer SnP_2S_6," *Journal of Physical Chemistry C*, vol. 127, no. 41, pp. 20523–20529, Oct. 2023, doi: 10.1021/acs.jpcc.3c04947.

[19] S. Vranic, R. Kurapati, K. Kostarelos, and A. Bianco, "Biological and environmental degradation of two-dimensional materials," Mar. 01, 2025, *Nature Research*. doi: 10.1038/s41570-024-00680-5.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Lithology, Geochemistry, Hydrothermal Alteration and Mineralization of Cu-Mo porphyry deposits at Canada Wall prospect, Andong Meas District, Ratanakiri Province, Cambodia

Mengly IEM [1*], Yana CHHOEUN [1], Bunlim HEAN [1], Bunchoeun PICH [1], Sirisokha SEANG [1,2], Kimhouy OY [1,3]

[1] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Economic Geology Laboratory, Department of Earth Resources Engineering, Faculty of Engineering, Kyushu University*

[*] *Corresponding author: iemmengly142@gmail.com*

Abstract: *Canada Wall porphyry Cu-Mo project is a mineral exploration concession of Angkor Resource Crop, which is located in the Andong Meas district, Ratanakiri Province, Cambodia. The main purpose of this research is to identify the copper and molybdenite deposit, focusing on the lithology, geochemistry, hydrothermal and ore mineralization in the Canada Wall project according to the sample in the drill hole CW15-005D. The analysis methods consist of hand specimen description and x-ray fluorescence were used to interpret the lithology that consists of granite, granodiorite, basaltic trachyandesite and basaltic andesite. According to the result of petrography (thin section), Ore X-press analysis and X-ray diffraction (XRD), the alteration mineral patterns can be classified as phyllic alteration consisting of sericite +quartz +pyrite +montmorillonite and propylitic alteration consisting of chlorite +goethite +albite +epidote +calcite +quartz. Phyllic alteration (plagioclase +sericite and quartz) suggests that host rocks were altered by hydrothermal fluid at the temperature of approximately 200 ℃ to 350℃ under neutral conditions. The results from the Polish section, the ore mineral consists of pyrite chalcopyrite sphalerite molybdenite magnetite chalcocite and pyrrhotite which occur in veins, veinlets, and dissemination. Then, Geochemical analyses reveal that the granitoid is classified as medium to high-k calc-alkaline, I-type granite, peraluminous, and part of the magnesian series. They were formed in a subduction-related tectonic setting, likely a volcanic arc. Trace elements spider diagrams normalized to primitive mantle display strong enrichment in large-ion lithophile elements such as Rb, Ba and K and depletion in some high-field strength elements like Ti, suggesting that magmas were generated in a subduction related tectonic setting. Overall, these findings suggest the study area has potentially copper and molybdenite ore bearing, it is suitable for further exploration endeavors.*

Keywords: Geochemistry; Hydrothermal alteration; Porphyry deposit; Copper-molybdenite, Ratanakiri

1. INTRODUCTION

Cambodia is located in the southern part of Indochina Terrane, mainland of SE Asia. The mainland SE Asia comprises several Gondwana-derived terranes including Indochina, South China, Sibumasu and West Myanmar Terranes, which were assembled and amalgamated by subduction-collision processes during Late Palaeozoic-Mesozoic time. The Indochina Terrane is made up of several tectonic units, which host mineralization belts such as Truong Son Fold Belt (TSFB), Loei Fold Belt (LFB), Dalat-Kratie Belt (DKB) [1]. The Canada Wall area in NE Cambodia lying on DKB consists of Triassic to Cretaceous sedimentary rocks intruded by Cretaceous (125-75Ma) volcano-plutonic rocks, which are overlain by Quaternary intraplate basalts. The DKB extends across the region from Cambodia to southern Vietnam and overlies the southern continuation of the TSFB and LFB. The Dalat–Kratie and Southeast China magmatic belts may have been closely associated during the Jurassic-Cretaceous, and were later separated by a Cenozoic SE Asia extrusion as a result of the Himalaya-India collision [2] (Fig. 1). The lithology consists in the Canada Wall are granodiorite, diorite, and andesite, the alteration zone in drill hole CW21-010D are phyllic and potassic alteration zone and geochemically, the granodiorite plots within the medium-K calc-alkaline range and shows I-type affinity [3]. According to previous studies, the Angkor Gold Corp has discovered a

[*] Corresponding author: Mengly IEM
Email: iemmengly142@gmail.com; Tel: +855-12 712 747

mineral discovery at the Andong Meas south tenement that is economically viable for investigation.

The main objective is to define the located of molybdenite and copper resources. This study is focused on the lithology, geochemical characteristic, hydrothermal alteration style, and ore mineralization form in intrusive and extrusive rock in the research area in drill hole CW15-005D.

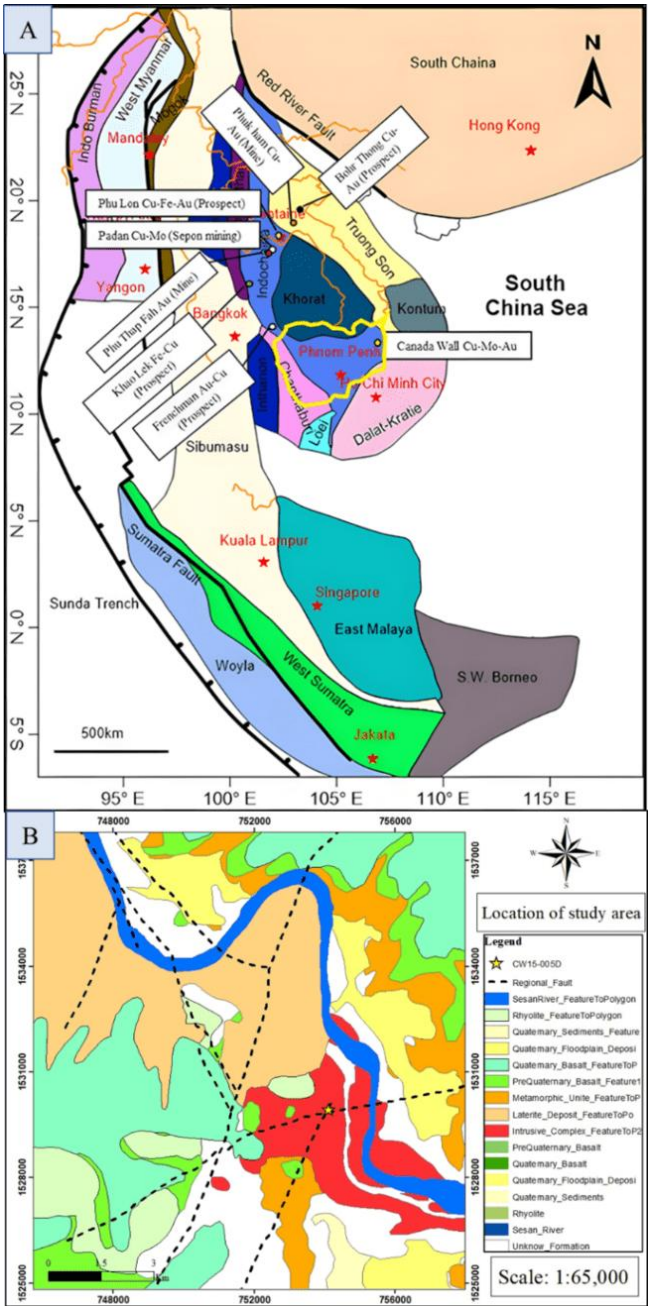


Fig. 1. (A) The regional map shows tectonic divisions of the Indochina Terrane including the mineralization belt of Cambodia, Vietnam, Laos, and Thailand [2], [4]. (B) Location of Canada Wall Porphyry Cu-Mo prospect and drill hole.

2. METHODOLOGY

2.1 Research stage

This research is start from literature review such as regional studies, geological map, previous studies and satellite image interpreted by remote sensing. The core logging work is conducted for a week to record the geological information and sample selection. It has been divided into the type of experiments like fresh rock and less altered for lithology identification and geochemical characterize, altered rock for hydrothermal alteration and ore associated with rock for ore mineralization. In the final stage, the results were interpreted.

2.2 Sampling and Analytical Method

The total depth in drill hole CW15-005D is 251m. The samples were conduct ninety-eight (98) samples from this hole. The lithology was identified by hand spacemen. The fresh rock and less altered was selected ten (10) samples for X-ray fluorescence (XRF) to identified the major and trace element to characterize the geochemistry. The altered rock was taking thirty (30) samples for thin section, X-ray diffraction (XRD) and ore x-press to identified the altered mineral which are importance to characterize type of hydrothermal alteration. Last but not least, the ore that associated with rock was selected third teen (13) sample for polish section to identified the ore that present with rock along fracture, vein and disseminated on rock. The method of petrography and polish section conduced in microscopy lab (116B), XRF was conducted in 301H and ore x-press was prepared in laboratory of nanostructure and chemical analysis were conducted, at Institute of Technology of Cambodia (ITC) and XRD method was conducted economic geology laboratory at Kyushu University in Japan.

3. RESULTS AND DISCUSSION

3.1. Lithology Identification

The granodiorite is medium to coarse-grained, medium gray and consist of quartz, plagioclase, k-feldspar, muscovite and biotite. Andesite is fine to medium-grained, dark grey and consist of quartz, plagioclase and biotite. Granite is a type of igneous rock that is formed from the slow crystallization of molten magma deep within the Earth's crust. It is composed mainly of quartz, feldspar, and mica, and is known for its durability and strength. Other minerals that may be present in smaller quantities include amphibole, pyroxene, and biotite.

3.2. Whole rock geochemistry

The major elements of the igneous rocks in study area have been analyze by X-ray Fluorescence (XRF). The result from the analysis is all summary in the **Table 1**. The percentage of major chemical composition are plotted in TAS diagram for identification and classification of plutonic rock and volcanic rock by using the percentage of total alkalis and silica in each igneous rock.

Table 1. Major elements analyses of intrusive and volcanic rocks from drill hole CW15-005D, Abbreviation: GRN: Granite, GRD: Granodiorite, BTA: Basaltic trachyandesite, BA: Basaltic andesite.

ID	D5-76	D5-92	D5-105	D5-156	D5-240	D5-45	D5-165
Rock Type	GRN	GRD		GRN		BTA	BA
SiO_2	70.2823	69.8936	71.7350	69.0259	70.3943	52.5687	53.6109
Al_2O_3	14.3761	13.9887	15.0689	15.1547	15.1152	17.5870	16.5600
K_2O	6.1107	4.0701	4.5062	4.8001	5.5660	4.1121	2.5199
CaO	1.7308	1.7522	1.6885	1.8097	1.6515	4.8750	6.5255
TiO_2	0.1669	0.1757	0.1811	0.1906	0.2225	1.1089	0.9746
MnO	0.0234	0.0191	0.0202	0.0251	0.0293	0.1488	0.1353
Fe_2O_3	1.1762	2.7383	1.3050	1.7944	1.1625	7.8264	7.4456
Na_2O	2.8360	3.2695	3.8891	3.8165	3.6071	1.7691	1.9869
MgO	0.5434	0.5081	0.4036	0.5546	0.6137	6.8371	5.4651
Rb_2O	0.0151	0.0120	0.0101	0.0116	0.0142	0.0262	0.0131
Y_2O_3	0.0013	0.0018	0.0022	0.0072	0.0024	0.0033	0.0018
ZnO	0.0044	0.0025	0.0026	0.0039	0.0037	0.0242	0.0118
Nb_2O_5	n.d	n.d	0.0016	n.d	n.d	n.d	n.d
LOI	1.4818	1.5800	0.6979	1.3101	0.9013	2.0800	2.4800
Rb (ppm)	12.7685	10.1471	9.2356	10.6072	12.9847	23.9576	11.9788
Y	n.d	n.d	1.7324	5.6696	1.8899	2.5986	1.4174
Cu	n.d	n.d	9.6663	28.3597	8.7875	21.9688	14.2198
Rb	12.7685	10.1471	9.2356	10.6072	12.9847	23.9576	11.9788
Nb	n.d	n.d	1.1185	n.d	n.d	n.d	n.d
Na2O+K2O	8.9467	7.3396	8.3953	8.6166	9.1731	5.8812	4.5068
FeOT (wt%)	1.0586	2.4645	1.1745	1.6150	1.0463	7.0438	6.7010
FeOT/(MgO +FeOT) (wt%)	0.6608	0.8291	0.7442	0.7444	0.6303	0.5074	0.5508

The major elements of these rocks are high SiO2 (69.02 to 71.73wt%, mean 70.35wt%), high Na2O+K2O (8.39 to 9.17wt%, mean 8.78wt%). According to the results from (Fig. 2) sow the rock unite such as granite, granodiorite [7], basaltic trachyandesite and basaltic andesite [8]. In addition, all the sample plutonic rock plot of alumina saturation with alkalinity of the granitoid showing majority of the samples are peraluminous [9] and within I-type granite affinity (Fig. 3) .

Additionally, the I-type granite in the study area (Fig.4A) show the concentrate of magnesian that compare to ferroan rocks [11]. After that, the granodiorite that in the location of study are including tholeiitic and calc-alkaline magmatic [12] differentiation trends (Fig. 4B) was used to estimate the source of magmatic series of granitoids in the study area. With increasing oxygen fugacity during magmatic differentiation, magnetite will crystallize and separate from the melt, resulting in the enrichment of SiO2 and a change of the magma composition towards the calc-alkaline series. Trace element analyses were used in this study to investigate the tectonic origins of intrusive rocks in the study area. On the tectonic discrimination Y+Nb vs Rb and Y vs. Nb diagram [13], the intrusive rocks from the research area were plotted on the volcanic arc setting (Fig. 5). The location of study area form in volcanic arc granitoid (Fig. 6) [4], [5], [6].

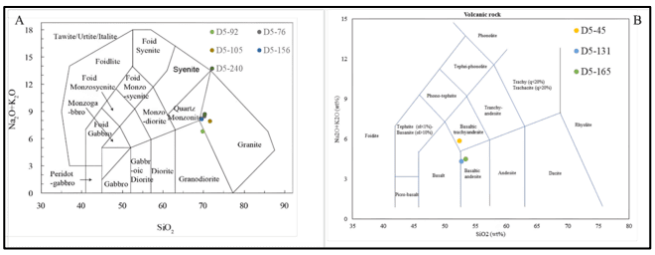


Fig. 2. Classification diagram based on total alkali versus SiO2. A). TAS plutonic rock classification diagram [7]. B). TAS volcanic rock classification diagram [8].

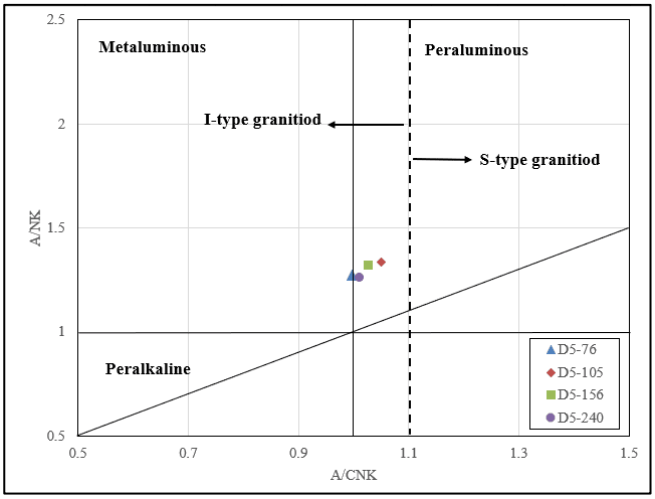


Fig. 3. A/CNK [= molar Al2O3/(CaO+K2O+Na2O)] versus A/NK [=molar Al2O3/(K2O+Na2O)] diagram [9], [10]

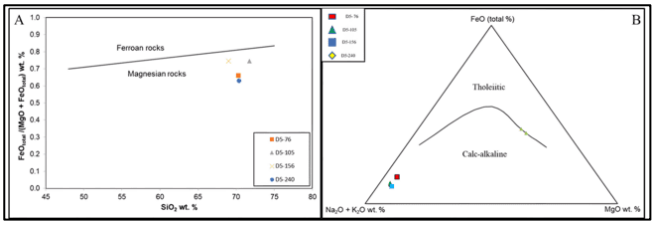


Fig. 4. A). FeOtotal/ (MgO + FeOtotal) versus SiO2 wt% diagram for geochemical classification of granitoid proposed [11]. B). AFM graphic showing calc-alkaline and tholeiitic magmatic differentiation tendencies [12]

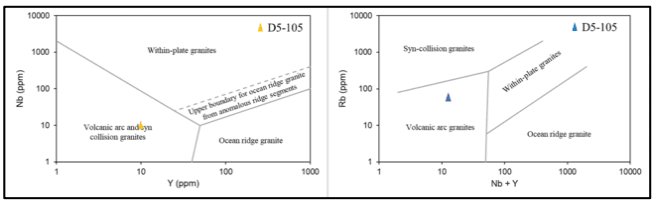


Fig. 5. Tectonic discrimination diagrams for plutonic rocks [13]

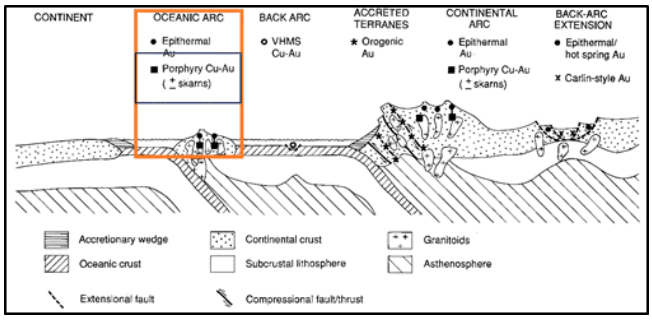


Fig. 6. Porphyry copper deposits in the context of plate tectonics [5]

3.3. Hydrothermal alteration

The altered rock at depth from 0m to 191m contain of altered minerals such as sericite which replaces the primary mineral plagioclase feldspar, quartz, pyrite and chlorite mineral altered from biotite. And the depth from 191m to 251m contain of altered mineral included quartz, sericite, chlorite, epidote and calcite. The secondary mineral such as chlorite and epidote are altered from biotite. The calcite mineral from as vein in granitoid and andesite (Fig. 7).

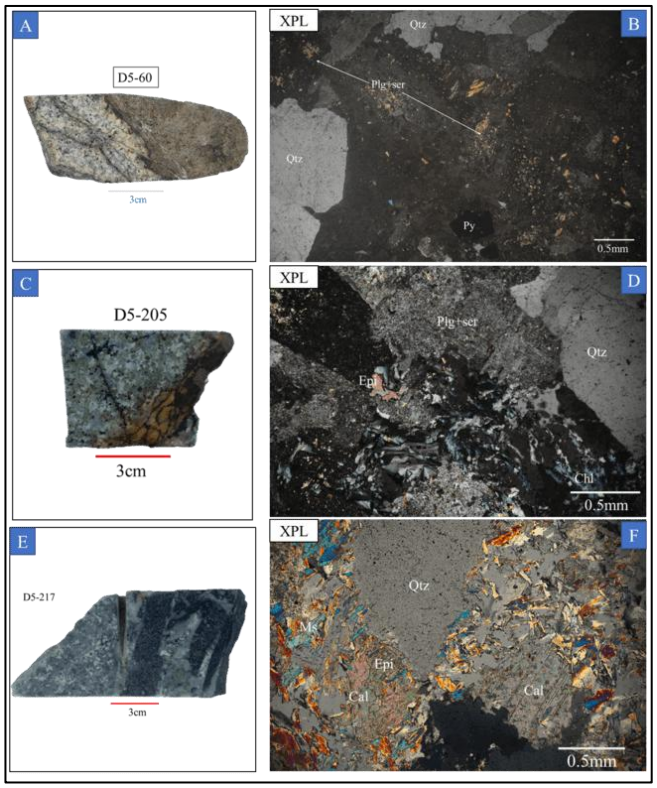


Fig. 7. Representative photographs and photomicrographs of porphyry-style alteration in the study area. (A),(C)&(E). Hand spacemen sample. (B). showing the mineral quartz and sericite replace plagioclase. (D). showing the mineral quartz, sericite replace plagioclase, chlorite and epidote. (F). showing the mineral quartz, muscovite, calcite and epidote. Abbreviations: Qtz=Quartz; Plg=Plagioclase; Chl=Chlorite; Ser=Sericite; Epi=Epidote, Cal=Calcite and Mus=Muscovite.

Base on the results from ore x-press spectrometer (Fig. 8), D5-32.25 and D5-60 consist of muscovite and quartz. Muscovite is characterized by absorption features in the 1400nm and 2200nm regions. the wavelength absorption bands of quartz in sample D5-144 and D5-157.2 is about 900nm regions. The sample D5-205 and D5-240 consist of montmorillonite and illite. Montmorillonite was presence of broad and deep-water absorptions at 1400 nm and 1900 nm and an AI-OH absorption around 2200 nm Montmorillonite is commonly found in soils as a weathering product of alumino-silicates. Illite has similar properties to muscovite and is differentiated by the lesser abundance of interlayer K. The main difference with illite is found only at the 1900nm absorption feature which is much shallower in the muscovite, Al-OH absorption feature in the 2200nm region. And the sample D5-205, D5-206 consist of chlorite and goethite. Chlorite is characterized by absorption of Mg-OH features in the 2250 – 2300 nm region due to the hydroxyl molecule. Goethite is the mineral displays trans opaque behavior. Both the Fe3+ ion and Fe2+ ion occur, creating features at 550 nm and 900 nm [14].

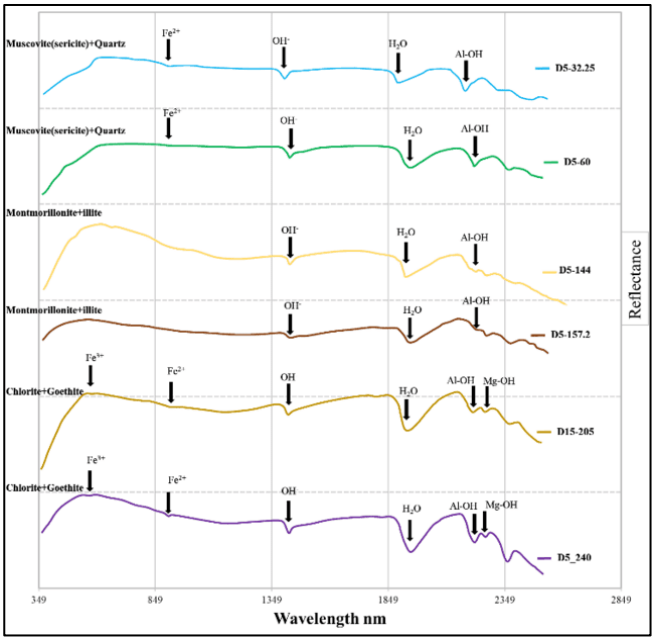


Fig. 8. All the wavelength absorption bands of the mixed mineral in drill hole CW15-005D.

According to the result from X-ray diffraction (XRD) showing the minerals include quartz, plagioclase and K-feldspar and the secondary mineral like illite-mica(sericite), chlorite, albite and montmorillonite. After that, the associated ore like pyrite also was shown in these results (Fig. 9).

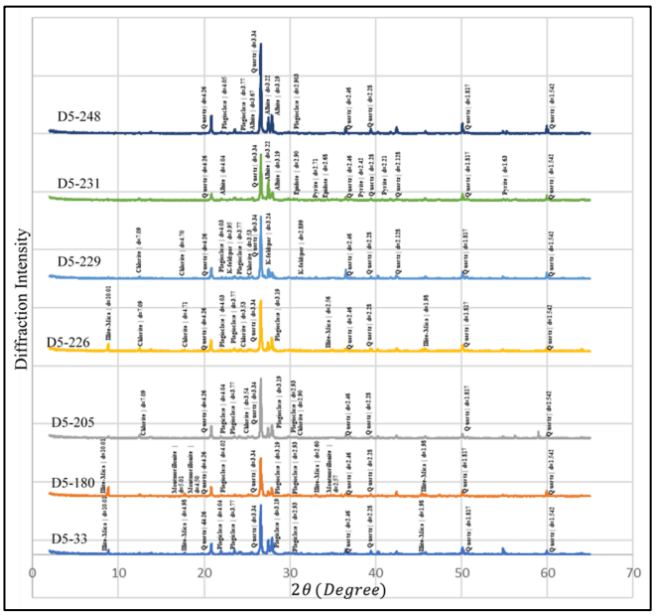


Fig. 9. The intensity of reflections responds to minerals contain in all samples from XRD analysis

Base on the results from petrography, ore x-press and X-ray diffraction, the depth from 0m to 205m are present the mineral include quartz, plagioclase, pyrite, sericite, montmorillonite and some chlorite. So, these primary mineral and secondary minerals can classify it is in phyllic alteration. And the depth rang from 205m to 251m is propylithic consisting of quartz, plagioclase, sericite, chlorite, calcite, goethite, epidote and some k-feldspar [15]. Phyllic alteration consists of sericite, quartz, and chlorite, suggesting that the host rocks beneath the surface were altered by hydrothermal fluids at temperatures ranging between 200°C and 350°C and under neutral pH conditions (Fig. 10) [16].

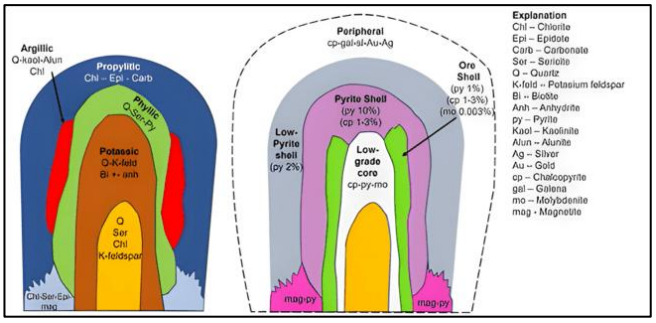


Fig. 10. Hydrothermal alteration zones associated with porphyry copper deposit [15]

3.4. Ore mineralization

The mineralogy includes pyrite, chalcopyrite, pyrrhotite, sphalerite, magnetite, molybdenite and chalcocite. The ore minerals are disseminated in the phyllic alteration and propylitic alteration, as well as in quartz vein and altered granodiorite and andesite. The gangue includes quartz and calcite.

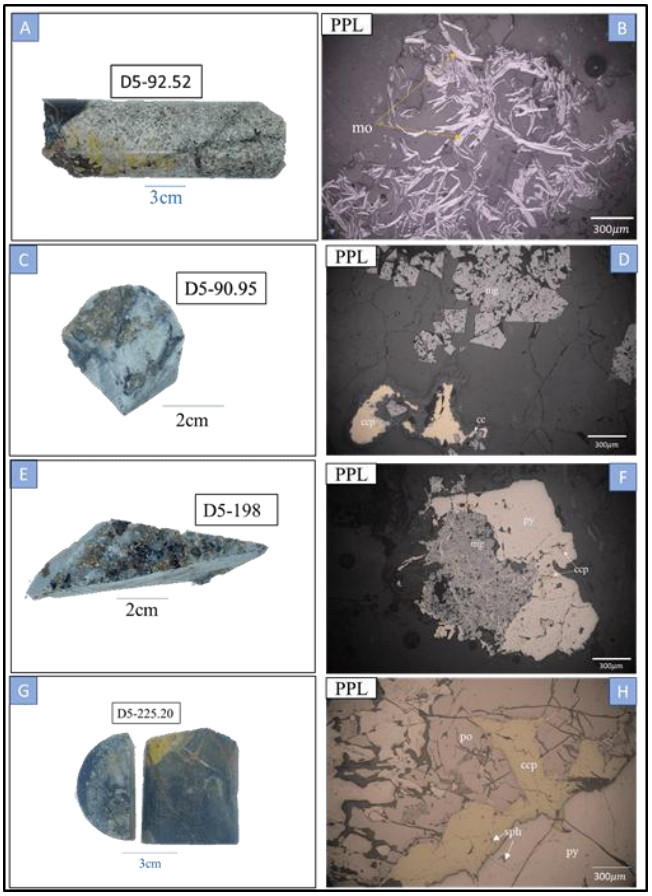


Fig. 11. (A), (C), (E) & (G). photograph of hand spacemen, (B). Photomicrograph showing the massive molybdenite, (D). Photomicrograph showing the magnetite and chalcocite replace chalcopyrite. (F). Photomicrograph of magnetite, pyrite and chalcopyrite occurred as a vein within magnetite. (H). Photomicrograph of chalcopyrite, pyrite, pyrrhotite and sphalerite.

Chalcopyrite is a form of gold-yellow mineral with irregular shapes, anhedral shape, relief is medium, and have no pleochroism. In plan-polarized light, the interference color is brown and weakly non-isotropic. It is regularly observed in exsolution intergrowths with other copper sulfides and pyrrhotite and occurs in sphalerite as chalcopyrite disease (Fig. 11D&H)

Pyrite appears with physical properties of pale yellow distinctly yellowish in color, and isotropic. Pyrite in plan-polarized light under the microscopic are the appearance of pale yellow, has no pleochroism, and crystal form as euhedral (Fig. 11F&H).

Pyrrhotite is a creamy pink to brown color in PPL and may show weakly to strongly pleochroism, which means it can appear in different colors depending on the direction of

light, and exhibits strong anisotropy, meaning it changes color when the stage is rotated under polarized light XPL. This anisotropy can range from yellow-grey to greenish- grey or greyish-blue. Sphalerite appears blackish grey under a microscope, indicating isotopism, irregular shape, and no bireflectance, occurred within pyrite, sphalerite appears as an exsolution texture Sphalerite is isotropic, appearing dark under polarized light (Fig. 11H)

Molybdenite typically appears as thin, platy crystals with a metallic luster. These crystals can range in color from a silvery-gray to a blackish color. In reflected light, molybdenite has a brassy yellow to grayish-white color, with a strong metallic luster. Molybdenite exhibits weak pleochroism, meaning its color can vary slightly depending on the direction of light. Molybdenite has perfect cleavage in one direction, resulting in thin (Fig. 11B).

Magnetite showed as grey with slight brownish and no bireflectance in plane-polarized light. It has no change in brashness, is dark brown, and it's isotropic in cross-polarized light. Magnetite has a high reflectivity when compared to other minerals using reflected light microscopy (Fig. 11F).

Chalcocite show as bluish-grey and no bireflectance in plane-polarized light. It weak to distinct anisotropy, between emerald green and light pink, difficult to observe with exactly crossed nicols. It replaces the chalcopyrite (Fig. 11D).

4. CONCLUSIONS

According to the explanation of drill hole CW15-005D, it is reasonable to assume that the lithology, hydrothermal alteration responsible for ore mineralization of the Cu-Mo porphyry deposit are as follows:

- The examination of the area's lithology consisted of many varieties, including granite, granodiorite, basaltic trachyandesite and basaltic andesite. The rock unit is mostly crossed by quartz veins and veinlets formed by sulfide minerals.
- Based on geochemistry characterization, the granitoid in the study area is owing to I-type granite, magnesian, peraluminous, and calc-alkaline series which originated form closely link with volcanic arc related subduction zone magmas.
- This research has two types of alteration zones: phyllic, which consists of sericite, quartz, montmorillonite and pyrite alteration minerals, and propylitic, which consists of chlorite, goethite, epidote, albite, k-feldspar, quartz, and calcite. Furthermore, sericite and chlorite are found in altered rock and can be identified using thin sections.
- Mineralization consists of pyrite, sphalerite, molybdenite, pyrrhotite, chalcopyrite, chalcocite and magnetite.

This research provides a comprehensive understanding of the lithology, geochemistry, hydrothermal alteration, and ore mineralization within the Canada Wall Porphyry Cu-Mo-Au prospect, aiding in the development of geological models for the exploration company to define the formation of copper, molybdenite and gold deposit.

ACKNOWLEDGMENTS

I gratefully acknowledge the support of the Laboratory-Based Education (LBE) promoted by the Project for Strengthening Engineering Education and Research for Industrial Development in Cambodia, Institute of Technology of Cambodia (ITC). I would like to express my gratitude to all lecturers of the faculty of Geo-resources and Geo-technical Engineering. We extend our heartfelt gratitude Angkor Gold Corp, for graciously permitting and supporting our field research in the Canada Wall Porphyry Cu-Mo-Au deposit, Ratanakiri, Cambodia.

RECOMMENDATION

- To comprehend tectonic, lithogeochemistry of intrusive and volcanic rocks, and whole-rock major elements oxides, trace elements, and rare earth elements, intrusive rocks should undergo Inductively Coupled Mass Spectrometry (ICP-MS).
- Selection more drill hole to do experiment to confirm more detail both of hydrothermal alteration and geochemistry.
- Fluid inclusion and sulfur isotope studies related to alteration and mineralization should be conducted at each step of vein and alteration in order to track the source of hydrothermal fluids and their evolution, as well as create a genetic model of the deposit.

REFERENCES

[1] I. A. N. Metcalfe, "Gondwana dispersion and Asian accretion: an overview.," *Gondwana Dispers. Asian accretion*, 1999.

[2] K. Zaw *et al.*, "Tectonics and metallogeny of mainland Southeast Asia - A review and contribution," *Gondwana Res.*, vol. 26, no. 1, pp. 5–30, 2014, doi: 10.1016/j.gr.2013.10.010.

[3] P. Vannak, U. G. Mada, S. Sirisokha, and K. Yonezu, "Lithology, Hydrothermal Alteration, and Ore Mineralization of Canada WallPorphyry Cu-Mo-Au prospect, Ratanakiri, Cambodia," no. September, 2024.

[4] Seang Sirisokha, "Lithogeochemistry, Hydrothermal Alteration, Mineralization, Fluid Inclusion and Sulfur Isotope Study of the Halo Porphyry Copper-Molybdenum Prospect, Northeast Cambodia," 2019, [Online]. Available: https://doi.org/10.15017/2534418

[5] D. I. Groves, R. J. Goldfarb, M. Gebre-Mariam, S. G.

Hagemann, and F. Robert, “Orogenic gold deposits: a proposed classification in the context of their crustal distribution and relationship to other gold deposit types,” *Ore Geol. Rev.*, vol. 13, no. 1–5, pp. 7–27, 1998, doi: 10.1016/S0169-1368(97)00012-7.

[6] R. D. Raju, “I-, M-, a- and S-Type Granitoids: Their Attributes and Mineralization, With Indian Examples,” *Jour. Econ. Geol. Georesource Manag.*, vol. 5, no. 1 & 2, pp. 1–23, 2003.

[7] K. G. Cox, J. D. Bell, and R. J. Pankhurst, *The Interpretation of Igneous Rocks*. 1979. doi: 10.1007/978-94-017-3373-1.

[8] R. . Van Bemmelen, “A guide to the chemical classification of the common volcanic rocks.,” *Can. J. Earth Sci.*, vol. 8, no. 5, pp. 523–548, 1999.

[9] B. W. Chappell and A. J. R. White, “I- and S-type granites in the Lachlan Fold Belt,” *Trans. R. Soc. Edinb. Earth Sci.*, vol. 83, no. 1–2, pp. 1–26, 1992, doi: 10.1017/S0263593300007720.

[10] P. D. Maniar and P. M. Piccoli, “Geological Society of America Bulletin Tectonic discrimination of granitoids Subscribe Tectonic discrimination of granitoids,” *Geol. Soc. Am. Bull.*, vol. 101, no. 5, pp. 635–643, 1989, doi: 10.1130/0016-7606(1989)101<0635.

[11] C. D. Frost, “A geochemical classification for feldspathic igneous rocks,” *J. Petrol.*, vol. 49, no. 11, pp. 1955–1969, 2008, doi: 10.1093/petrology/egn054.

[12] D. Marshall, “Ternplot: An excel spreadsheet for ternary diagrams,” *Comput. Geosci.*, vol. 22, no. 6, pp. 697–699, 1996, doi: 10.1016/0098-3004(96)00012-X.

[13] J. A. Pearce, N. B. W. Harris, and A. G. Tindle, “Trace element discrimination diagrams for the tectonic interpretation of granitic rocks,” *J. Petrol.*, vol. 25, no. 4, pp. 956–983, 1984, doi: 10.1093/petrology/25.4.956.

[14] Hunt et al, “Chapter 4 Remote Sensing Background Electromagnetic Radiation Theory and the Interaction with Materials Background on Electromagnetic Radiation,” no. 1, pp. 49–109, 1975.

[15] J. M. Gumbert, “Lateral and vertical alteration-mineralization zoning in porphyry ore deposits,” *Econ. Geol.*, vol. 65, no. 4, 1970.

[16] G. J. Corbett and T. M. Leach, “Southwest Pacific rim gold–copper systems: structure, alteration and mineralization.,” *Soc. Econ. Geol.*, vol. Special Pu, no. May 1997, p. 236, 1998.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Design and modelling of cost-effective manufacturing of glass tempering furnace used in industrial applications.

Rayuth KEAT [1*], Raveth HIN [1], Chansopheak SEANG [1]

[1] *Research Unit Materials Science and Structure, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: rayuth.keat@gsc.itc.edu.kh*

Abstract: *This study presents the design and modelling of a cost-effective glass tempering furnace. The furnace design is based on critical parameters, including the casing, insulation system, heating mechanism, electrotechnical components, and safety control systems. The key considerations in the design process include material selection, heating control, and precise calculations to ensure optimal performance in industrial settings. The furnace was developed to achieve a uniform heating profile, ensuring efficient preheating and chemical glass tempering by salt bath KNO_3. The system's temperature readings show 300°C during the preheating phase and 600°C within the heating chamber, providing the necessary conditions for effective ion exchange in chemical tempering using a KNO_3 salt bath. Ion exchange significantly enhances the mechanical strength and durability of glass, improving resistance to scratches, cracks, and impacts by several times. This furnace is designed to operate systematically and efficiently, maintaining a constant temperature at any preset value. The integration of a controlled heating system allows precise thermal management, ensuring consistency and quality in the tempering process. Additionally, the furnace's heat retention capability, uniform heating rate, and controlled atmosphere contribute to improved energy efficiency. The cost-effective design of the glass tempering furnace reduces manufacturing and operational costs, making it a viable solution for industrial applications. Its long lifespan, enhanced safety features, and advanced control systems ensure reliability and sustainability. The findings of this study demonstrate that an optimized furnace design can improve energy efficiency and production quality while maintaining affordability. Future developments may focus on further automation and real-time monitoring systems to enhance operational efficiency and reduce human intervention in the process.*
.

Keywords: Glass Tempering, Furnace Design, Ion exchange, Cost-Effect Manufacturing, and Industrial Application.

1. INTRODUCTION

The glass tempering furnace heat treatment technology involves various aspects of glass processing and treatment methods [1]. Mathematical modelling and simulation techniques have been applied to optimized glass furnace design and operation [2]. developed a segregated modeling approach for continuous heat treatment furnaces glass, three combination modes convection, conduction, and radiation models to accurately predict product temperatures [3]. Industrial furnaces have become vitally important equipment since there involved in the production of many consumer products, such as foods, beverages, containers, machining tools, infrastructure materials among others applications [4]. In such furnaces, heat generation technology is mainly based on gas heating as well as electricity. Operation costs related to energy consumption are significant in most high-temperature furnaces; therefore, there is potential for energy saving. Heat treatment has been defined as controlled heating and cooling operation applied to metals and alloys in solid state so as to obtain the desired properties [5].

The furnace is the most important equipment used in the heat treatment processes [6]. Furnace should uniformly perform their heat transfer process. Non-uniform temperature distribution inside the heating chamber, resulting in low-quality glasses, such as making them easy to break. The traditional furnaces often struggle with high energy consumption and inconsistent temperature distribution, Leading to potential quality issues and increased costs. This paper addresses these challenges by introducing a new

[1*] Corresponding author: Rayuth Keat
E-mail: rayuth.keat@gsc.itc.edu.kh; Tel: +855-92 770 335

practical design and modelling approach for a cost-effective glass tempering furnace. The motivation behind this work is based on two main points. The primary objectives are, first, to develop a furnace glass that maintains a consistent temperature profile throughout the cycle; and second, to ensure that the design is economically feasible for large-scale industrial implementation. By focusing on design parameters such as outer casing, insulation properties, heating elements, electrotechnical systems, and safety controls, the proposed furnace promises improved performance and reliability. A controlled environment is crucial for achieving the chemical ion exchange using a KNO_3 salt bath which not only improves the tempering process but also enhances the glass's mechanical properties significantly. The construction utilized various materials, including stainless streel, mild steel plates, angle steel bars, refractory bricks, fiberglass, refractory cement, hinges, and heating elements.

2. METHODOLOGY

2.1 Design of furnace glass

The conceptual design of the furnace glass, which resembles a box, or a microwave oven used for cooking. However, the focus is on evaluating material properties parameters, furnace features, functional systems, and performance testing of the furnace glass as shown **Figures 1**. The furnace glass was divided into two sections: the preheating chamber and the main heating chamber. The internal dimensions of preheating chamber are 450 mm x 450 mm x 500 mm (length x width x height), while the main heating chamber measures 450 mm x 450 mm x 600 mm. Different materials were used for thermal insulation, including an outer casing made of stainless steel, a middle layer of glass-wool, and an inner lining of firebrick.

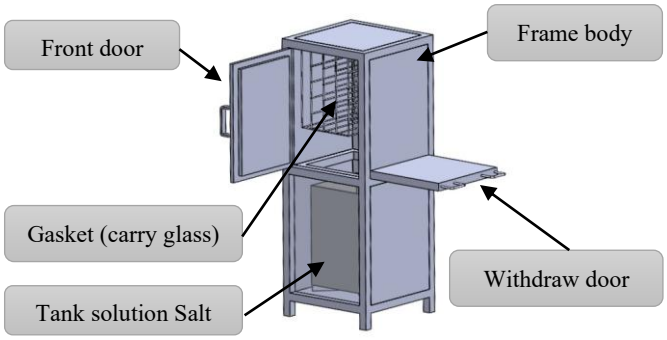


Fig.1. 3D model of furnace glass with detailed specifications of its subcomponent. (Internal size: 450 x 450 x 500, Heating chamber size: 450 x 450 x 600, Thickness (t): 3) (Unit: mm)

2.2 Material properties

The properties of engineering materials, particularly metals and their alloys, are fundamentally determined by their microstructure characteristics [6]. The main body, or outer casing, of the chamber was constructed from SS304 stainless steel plate with the thickness 3mm to 4mm. Chemical composition of SS304 includes 0.08%C, 18-20%Cr, 66.345-74%Fe, 0.2%Mn, 8-10.5%Ni, 0.045%P, 1%Si, and 0.030% S. Furthermore, the physical properties of SS304 stainless steel are presented in **Table 1**.

Table 1. The stainless steel SS304

Physical Properties	Specification
Density, $\boldsymbol{g/cm^3}$	7.93
Melting point, ℃	1398-1454
Magnetic in annealed condition	No
Magnetic permeability	1.02 (approximate)
Specific heat capacity, $\boldsymbol{J/(kg.K)}$	500 at 0-100℃
Electrical resistivity, $\boldsymbol{\mu\Omega.m}$	0.73 at 20℃
Modulus of elasticity (Elastic modulus), GPa	193
Thermal diffusivity, $\boldsymbol{mm^2/s}$	3.84 at 20-100℃
Thermal conductivity ($\boldsymbol{W/m.k}$)	16.3 at 100
	21.5 at 500℃
Mean coefficient of thermal expansion, ($\boldsymbol{10^{-6}/K}$)	17.2 at 0-100℃
	17.8 at 0-315℃
	18.4 at 0-538℃

Table 2 presents the insulation properties of the ceramic fiber blanket, code LYTX-1260T. The material is characterized by high tensile strength, low shrinkage, excellent hot strength, high resiliency, low thermal conductivity, and low heat storage, as detailed in the table below [7].

Table 2. The insulation of ceramic fiber blanket code LYTX-1260T.

Items	Units	Index
Temperature	℃	1260
Shrinkage on heating (%)	℃ × h	1000x24h ≤ −3
Density	kg/m^3	96
Reshrinking temperature	℃	500
Thermal conductivity($\boldsymbol{128kg/m^3}$)	W/m.K	0.07
Tensile strength (25mm thick)	MPa	0.06

Table 3. showed the refractory firebrick insulated power with light warm air duct and steam pipeline, and inside of furnace used the firebrick insulation refractory selected excellent quality high with about 65-75 percent alumina of Mullite brick.

Table 3. Specification of Mullite Brick MU65/MU75

Items		MU 65	MU 70	MU 75
Chemcial Composition	Al_2O_3%	≥65	≥70	≥75
	SiO_3%	≤33	≤26	≤24
	Fe_2O_3%	≤1.0	≤0.6	≤0.4
Bulk Density $\boldsymbol{g/cm^3}$		≥2.55	≥2.55	≥2.55
Reheating Linear Change (%) 1500℃ × $\boldsymbol{2h}$		0~+0.4	0~+0.4	0~+0.4
Apparent porosity (%)		≤17	≤17	≤18
Thermal Conducvity (W/m.K) 1000℃		1.84	1.95	1.95
Cold crushing strength, MPa		≥60	≥80	≥80
Thermal Expansion Coefficient (x10-**6**℃-1)		0.6	0.6	0.55
0.2MPa Refractoriness Under Load T0.6℃		≥1600	≥1600	≥1650
Thermal Shock Resistance 1000℃ water cycles		≥18	≥18	≥18

2.3 Mathematical

In a solid, the flow of heat by conduction is the result of the transfer of vibrational energy from one molecule to the next, and in fluids it occurs in addition because of the transfer of kinetic energy. Stead-State conduction is a form of conduction **Fourier's Law** in one dimension [8].

$$q = \frac{k.A.\Delta T}{d} \qquad \text{(Eq. 1)}$$

Where,

q is the heat transfer rate (W)
k is thermal conductivity of material (W/m. K)
A is the cross − sectional area through which heat is beinig transferred (m^2)
ΔT is the temperature difference across the material (K)
d is the thickness of the material (m)

Condition 1: Determine the thickness of each layer by specifying the desire temperature of each layer, assuming a heat transfer rate $q = 1000W$ and cross-sectional area through the wall of $A = 1m^2$.

1-D Heat conduction through composite wall (Steady state heat transfer through composite wall using the mathematical). An induction furnace wall is made up of layers, inside, middle and outer layers with thermal conductivity K1, K2, and K3 respectively as shown in **Table 4**, **Table 5**, and **Table 6**. Determine the thickness of material.

Case 1:

Table 4. Material data analysis

Materials	Thermal conductivity (W/m.K)	T_i (℃)	T_o(℃)	ΔT(℃)
Ceramic brick	0.0941	300	250	50
Glass wool	0.07	250	30	220
Air	0.06	30	25	5

Case 2:

Table 5. Material data analysis

Materials	Thermal conductivity (W/m.K)	T_i (℃)	T_o(℃)	ΔT(℃)
Ceramic brick	0.0941	1500	1200	300
Glass wool	0.07	1200	700	500
Air	0.06	700	600	100

Case 3:

Table 6. Material data analysis

Materials	Thermal conductivity (W/m.K)	T_i (℃)	T_o(℃)	ΔT(℃)
Ceramic brick	0.0941	1500	1200	300
Glass wool	0.07	1200	100	1100
Air	0.06	100	25	75

Condition 2: Calculate the thickness of one layer as function another layer.

$$\frac{q}{A} = \frac{T_{inner} - T_{outer}}{\frac{L1}{K1} + \frac{L2}{K2} + \frac{L3}{K3}} \qquad \text{(Eq. 2)}$$

3. RESULTS AND DISCUSSION

The results of the determines, based on mathematical equation 1 and the applied temperature differences obtained from result 1 in **Figure 2**, indicate the required layer thicknesses during the preheating stage in the first section of the furnace. With a maximum temperature of 300℃ and minimum 25℃ (ambient), the calculated thickness for the refractory, insulation, and air layer are 4.705 mm, 20.105 mm, and 20.405 mm, respectively. During the main heating phase, as shown in Result 2 in **Figure 3,** the internal temperature ranges from 1500℃ to 600℃, resulting in layer thicknesses of 18.23 mm (refractory) 63.23 mm (insulation), and 69.23 mm (air). For the heat source area, illustrated in **Figure 4,** with a temperature gradient from 1500℃ to 25℃, the same thicknesses are observed: 18.23 mm for refractory, 63.23 for insulation, and 69.23 mm for air. The combined results from all sections are presented in **Figure 5**.

For the purposes of this discussion, the new practical design and modelling of furnace glass will be constructed using a 2mm thick flat sheet of mild steel, 50.8 mm angle iron, 25.4 mm rectangular pipe, refractory cement, silica sand, sodium silicate as a binder, fiberglass, insulating tape, and other supplementary materials [9]. The furnace was tested at a set temperature of 1000 °C with a tolerance of ±5 °C. In comparison, another heat treatment furnace was designed using a 7 mm thick flat sheet of mild steel, 5 mm thick angle bars, fiberglass, and bricks within the heating chamber, and it was capable of reaching a maximum temperature of 1200 °C [6].

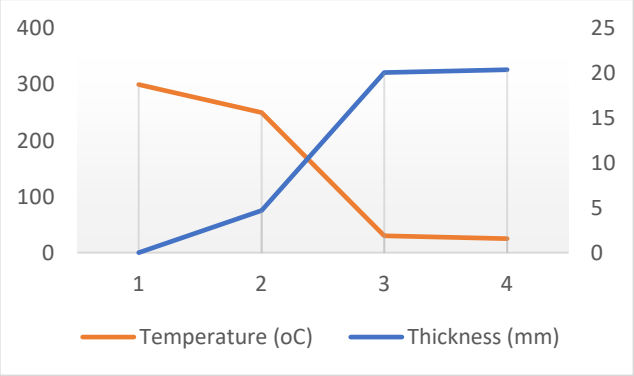


Fig.2. Preheat temperature 300℃ of vertical, and thick horizontal axis.

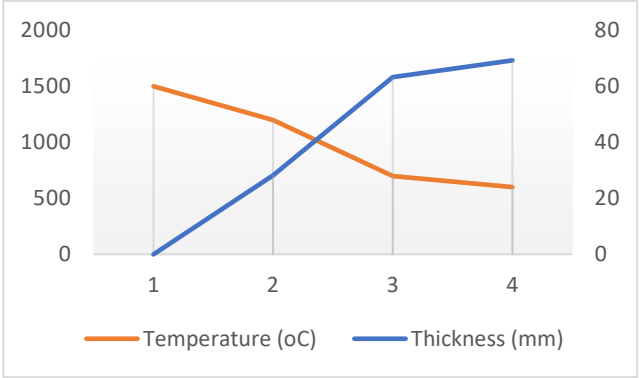


Fig.3. During heat temperature 600℃ of vertical, and thick horizontal axis.

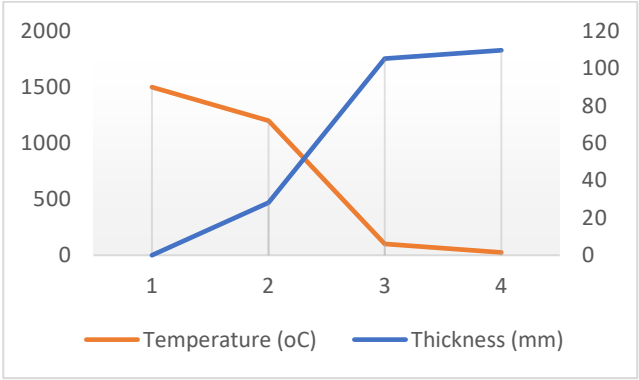


Fig.4. Heat sources temperature 1500℃ of vertical, and thick horizontal axis.

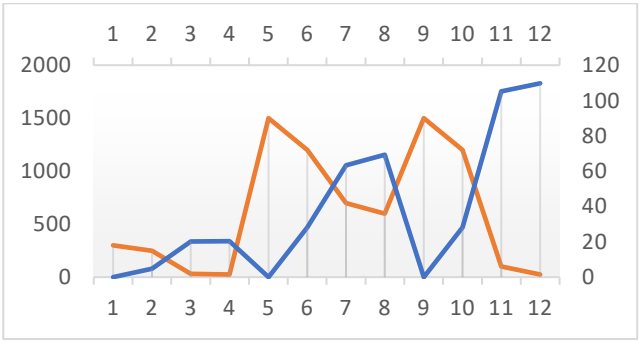


Fig.5. Combination of heating differences preheat, during heat, and heat source temperature **Figure 2**, **Figure 3**, and **Figure 4**.

4. CONCLUSIONS

In this study on the Desing and modelling of cost-effective manufacturing of glass tempering furnace used in industrial applications. The conclusion of this paper can be concluded as follows:

(1) Though it was designed for controlled drying and heating materials and metals within the temperature range of 0℃ to 1500℃ but can equally be adapted for used in other drying and heating operations of the same temperature range, such as production of materials.

(2) The result obtained the temperature difference based on the diversity of thickness of material such as insulation, refractory, and ceramic brick.

(3) The lower cost of design of the furnace glass coupled with its good heating retaining capacity, uniform heating rate, controlled atmosphere, safety, ease of operation and maintenance justifies the usage.

(4) Furthermore, in order to reduce the dead weight of the furnace, the future models should use light sheet metal for the casting and substitute more fibre glass for bricks and insulation.

ACKNOWLEDGMENTS

Thanks to my lovely family, they are always encouraging and supported. And great thanks to my advisor and co-advisor for reviewing and giving good feedback on the important points in this paper. Especially, I would like to honor thank for financial supporting from leaders of National Polytechnic Institute of Cambodia.

REFERENCES

[1] Q. Ma, Z. Liu, Y. Zhong, Y. Lang, X. Liu, and H. Fang, "Thermal optimization of induction-heated pulling-down furnace for quartz glass rod fabrication," *Appl. Therm. Eng.*, vol. 141, pp. 1–9, Aug. 2018.

[2] M. P. M. Ersin Yıldız ∗, Altu ˘g Melik Ba ¸s ol, "Segregated modeling of continuous heat treatment furnaces," *ELSEVIER*, p. 106993, 2020.

[3] R. H. Rayuth Keata,*, Chansopheak Seangb, "2D Structural modelling and analysis of heat transfer through three walls for furnace glass using ANSYS program," in *The ESSD 2022 Proceeding*, 2022, pp. 171–177.

[4] I. D. Palacio-Caro, P. N. Alvarado-Torres, and L. F. Cardona-Sepúlveda, "Numerical simulation of the flow and heat transfer in an electric steel tempering furnace," *Energies*, vol. 13, no. 14, 2020.

[5] R. Purushothaman, "Evaluation and Improvement of Heat Treat Furnace Model," no. June, p. 120, 2008.

[6] S. C. Ikpeseni and N. Okereke, "Development and Performance Evaluation of a Heat Treatment Furnace," vol. 6, no. 05, pp. 496–503, 2017.

[7] L. Shangdong Luyang share co., "CERAMIC LUYANG.pdf." 2000.

[8] Yunus_A_Cengel_Yunu_Cengel, "-3 Heat and Other Forms of Energy 6 1-4 The First Law of Thermodynamics 11 Energy Balance for Closed Systems (Fixed Mass) 12 Energy Balance for Steady-Flow Systems 12 Surface Energy Balance 13 STEADY HEAT CONDUCTION 127 3-1 Steady Heat Conduction in Plane Walls 128."

[9] O. G. Babatunde *et al.*, "Development of a Novel Oven-Furnace," *J. Miner. Mater. Charact. Eng.*, vol. 05, no. 02, pp. 62–73, 2017.

[10] https://theworldmaterial.com/304-stainless-steel/.

[11] https://huameiworld.com/glass-wool/high-temperature-glass-wool-board.html.

[12] https://rsrefractories.com/refractories_bricks/Muyllite-Brricks.html.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Investigation of Utilization of Natural Oils as Processing Aids and Activators Alternative to Naphthenic Oils in Rubber Composites

Kong Sreynich [1], Nat Yukleav [1], Phanny Yos [2,3*], Laymey Sreng [2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3*] *Material Science and Structure Research Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: phannyyos@itc.edu.kh*

Abstract: *The Cambodian economy, until recently, was highly dependent on the agricultural sector. Rubber has been a significant source of foreign cash, government revenue, employment opportunities, and agro-industrial development. Natural rubber is a polymeric material composed of hydrocarbon chains possessing high flexibility and green strength making them suitable for various applications including automotive tires, construction materials and a lot of other materials. In rubber compounding, to facilitate handling and manufacturing it is important to use rubber processing aids. Rubber processing aid is a substance improves processability of rubber compounding by improving the rubber compounds' flow, and dispersion of fillers. Petroleum-based oils are commonly used as a processing aid and plasticizer in rubber composites. This research aims to use natural processing oils instead of hydrocarbon oils due to limited supplies and the toxicity of hydrocarbon oils. This study examined natural rubber of grade TSR10 from Cambodia with loadings of rice bran oil (RBO) and crude palm oil (CPO) that varied from 5 to 20 phr (part per hundred rubber) and were analyzed on the physical, and mechanical properties. The results showed that the loading of processing oils improved on rubber composite in terms of hardness, tensile strength, elongation at break, abrasion resistance index, and thermal ageing. As the results, hardness of rubber is reduced 8.78% when using 10phr of crude palm oil and 13.47% while using 10phr of rice bran oil which is in the range of the standard requirement. The elongation at break with 10phr of CPO increases 23.85% and 39.29% when using 10phr of RBO. The abrasion resistance index of rubber composite with 10 phr increased 14.54% and 20.71% while using 10phr of RBO and it has higher thermal resistance in comparing to the rubber composite without processing aid. Based on the results, CPO and RBO in a rubber composite presents feasibility as a renewable raw material that can serve as an alternative to petrochemical oils in various applications.*

Keywords: Mechanical properties; processing aid; natural rubber; natural oils; vulcanization

1. INTRODUCTION

Rubber has historically been a significant source of foreign cash, government revenue, employment opportunities, poverty alleviation for smallholders, agro-industrial development, environmental conservation, and regional development. Rubber plantations, like natural forests, serve an important purpose. Extensive research has been conducted on their functions, including rainfall, runoff, evaporation, transpiration, and soil moisture [1]. Rubber has become a major crop in Cambodia and is expected to be the second-largest source of income, following rice paddies. Cambodia has 404,578 hectares of rubber plantations as of 2022 and the annual production was approximately 382,000 tons [2]. Rubber processing aids are additives that enhance the flow, dispersion, and overall processability of rubber

3* Corresponding author: Dr. Yos Phanny
E-mail: phannyyos@itc.edu.kh*; Tel: +855-16 870 256*

compound, making it easier to handle and manufacture. Rubber processing oil improves processability without adverse affecting physical and mechanical properties of rubber composite. Petroleum-based oils are commonly used as a plasticizer in rubber composites. Petroleum-based oils, including paraffinic, naphthenic, and aromatic oils, are commonly used as rubber processing aids due to their compatibility with rubbers. However, their high content of carcinogenic polycyclic aromatic hydrocarbons (PAHs) has been linked to harm to humans and the environment [3]. The European Parliament and Council Regulation Concerning the reach regulation (EC 18/09/2006) banned highly aromatic oils with polycyclic aromatic hydrocarbons on January 1, 2010 [4]. Several attempts have been made to use vegetable oils or derivatives in rubber compounding. However, there are only a few studies on crude palm oil and rice bran oil [5,6, 7]. Rice bran oil can substitute aromatic oil, antioxidants, and fatty acids in natural rubber's sulfur vulcanization recipe. Natural rubber(NR) has minimal impact on the curing characteristics and physical qualities of the vulcanizate[7]. Palm oil and rice bran oil are very heat resistant, making it suitable for high-temperature applications[7]. Palm oil is often more affordable than petroleum-based oils, making it a cost-effective alternative for rubber compounding. Palm oil improves the processability of rubber compounds, lowering energy consumption and production costs [3]. Palm oil-based processing aids are less harmful and safer for the environment and human health than petroleum-based oils [3]. This study will determine characteristics of the natural rubber composites with crude palm oil and rice bran oil on properties such as cure characteristic, hardness, swelling index, tensile strength, elongation at break, tear strength, abrasion resistance index, and aging. The findings of the rubber composites properties will be then compared with the rubber composite with naphthenic oil.

2. METHODOLOGY

2.1 Material preparations

Natural rubber (TSR10) was obtained from Chob, Tboung Khmum province. Carbon Black, and N-Isopropyl-N'-phenyl-p-phenylenediamine (IPPD) were supported by Rubber Processing Lab of Institute of Technology of Cambodia, Zinc oxide (ZnO) and sulfur were obtained from Xilong Scientific Co., Ltd. Stearic acid was supplied by Merck KGaA, Germany. Crude palm oil (CPO) was obtained from MRT-TCC Joint Venture factory in Preah Sihanouk Province, and rice bran oil (RBO) was supplied by Sokhak rice bran oil factory in Battambang Province.

2.2 Rubber compounding process

Composites of carbon black and elastomers produced by two roll mill mixing were prepared composites with diverse processing oils such as crude palm oil, rice bran oil and naphthenic oil (NO). The amount of each type of oil was varied from 5phr, 10phr, 15phr, and 20phr. Different oil/NR composites were prepared by following the standard procedure ASTM D 2084. The cure characteristics (ASTM D 2084) were carried out using Montech Moving Die Rheometer (MDR 3000) at 150 °C for scorch time (ts_2) and cure time (tc_{90}). The compounds were placed in a hot press molding and cured at 150° C and under pressure 1000psi at the cure time (tc_{90}) (vulcanization optimum) that were obtained by MDR. The formulations of the diverse oil/NR composites were summarized in Table2.1.

Table 2.1. Rubber compounding formulation

Ingredients	NPO	CPO	RBO	NO
NR	100	100	100	100
ZnO	5	5	5	5
Stearic acid	2	2	2	2
Carbon black	50	50	50	50
CPO	0	5,10,15,20	0	0
RBO	0		5,10,15,20	0
NO	0	0	0	5,10,15,20
CBS	1.5	1.5	1.5	1.5
IPPD	1.5	1.5	1.5	1.5
Sulfur	2.5	2.5	2.5	2.5

* Amount as parts per hundred rubbers (phr)

2.3 Hardness

Rubber composite's hardness was measured in aligned with the standard test method for rubber property- Durometer Hardness, ASTM D-2240, using the durometer Shore A. The test specimen interchangeably, shall be at least 6.0 mm in thickness. Indentor extension and shape shall be in accordance with minimum shock and the shore A unit value of each sample shown on the durometer screen was recorded.

2.4 Abrasion Resistance index

The mass loss due to abrasion was evaluated with a Din Abrader according to ASTM Standards D 5963 Test Methods for Rubber Products. This test method provides procedures for preparing cylindrical test pieces of specified dimensions from vulcanized thermoset rubbers or thermoplastic elastomers and for evaluating their abrasion resistance by sliding a test piece across the surface of an abrasive sheet attached to a rotating drum. The result is expressed as abrasion resistance index as (Eq. 1):

$$ARI = (Vs / Vt) \times 100 \quad \textbf{(Eq. 1)}$$

$$V = m / d \quad \textbf{(Eq2)}$$

Where: Vs is the volume loss of standard sample

Vt is the volume loss of tested sample

m is mass loss sample or standard rubber,

d is density of sample or standard rubber

2.5 Tensile strength and elongation at break

The determination of tensile strength and ultimate elongation start with the test pieces taken from the sample sheet in dumbbell shape was followed by the standard ASTM D 412-06. The time interval between vulcanization and testing shall be in accordance with IS0 1826.

2.6 Aging properties

Aging test was carried out based on the ASTM D 573. Rubber composites were placed in the air oven at 105 ºC for 24 h after which they were subjected to properties tests including hardness, tensile properties, and elongation at break. The obtained data was then compared with the unaged samples.

2.7 Statistical analysis

Each analysis was conducted triplicate, and the results were then given as averages with standard deviations (mean ± STD) attached. One-way ANOVA, and Duncan were used for statistical comparisons, and p-values less than 0.05 were considered significant difference. IBM SPSS Statistics for Windows (Version 25) was used for analyzing the data and graphs were plot using Microsoft Excel 2024.

3. RESULTS AND DISCUSSION

3.1 Curing characteristics

The processability of rubber compounds can be determined by evaluating the cure characteristics such as cure time (tc_{90}) as shows in Fig 1. Based on the Fig.1. shown that increasing amount of processing aids significant effected (p-value = 8.78E-18 < 0.05) on the cure time of the rubber compound. In addition, when using different type of processing aids also significant impacted on the cure time of rubber compound (p-value = 8.7833E-18 < 0.05). When the processing aids are not used the cure time was 5.1minutes. However, the increasing the number of natural oils resulting in decreasing the cure time. For crude palm oil cure time decreased which range from 4.58 to 4.1 minutes and RBO the cure time declined in the range of 4.42 to 4.22 minutes. In contrary, increasing amount of NO caused the cure time increase from 5.28 to 5.5 minutes. The results indicated that rubber compound with CPO and RBO had slightly faster cure time compared to NO. Rubber compound utilizing crude palm oil (CPO) and rice bran oil cures faster than naphthenic oil due to these natural oils contain unsaturated fatty acids and stearic acid which it is a co-activator in rubber composite. These fatty acids actively contribute to the vulcanization process, increasing filler-rubber interaction and boosting cross-linking [8].

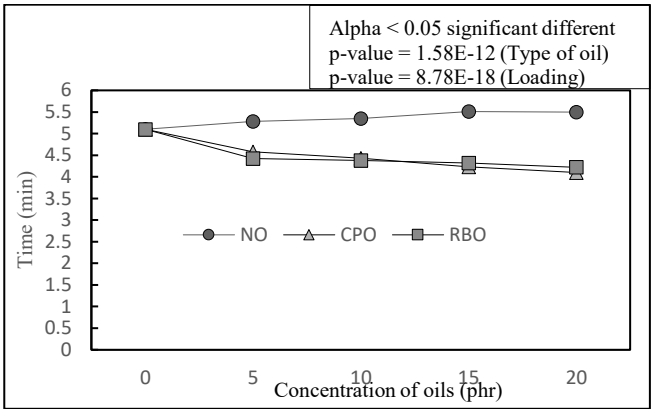


Fig 1. Cure time of rubber composite with different processing aids and different ratios

3.2 Hardness

The results of hardness of rubber composites with CPO, RBO and NO filled vulcanizates tend to decrease while the increasing the concentration of processing aids as shown in Fig 2. Based on the ANOVA analysis demonstrated that the hardness property of vulcanizate rubber was significantly different when rubber compound incorporated with different processing aids (p-value = 6.49E-7 < 0.05) and different concentration (p-value = 2.39E-13) in 95% confident interval. The hardness before aging without the processing aid was 69.78 Shore A. However, hardness of rubber composite with CPO decreased from 67.3 to 57.25 Shore A, and the hardness of vulcanizate rubber with RBO raged from 60.65 to 53.43 Shore A, and hardness of rubber composite with NO declined from 64.57 to 57.8 Shore A when the concentrations of oils were increased from 5phr to 20 phr. In the Fig 3. illustrated that the hardness of rubber composite after aging at 100°C for 24 hours raised from 69.78 to 73.25 Shore A while processing aids were not used, and similarly for rubber composite with processing aids were grown after exposing to high temperature. The increase in hardness during aging in rubber materials is mostly caused by chemical changes such as cross-linking and oxidation. Heat, oxygen, and other environmental conditions cause polymer chains to produce more cross-links over time [9]. This approach lowers the material's elasticity and makes it tougher.

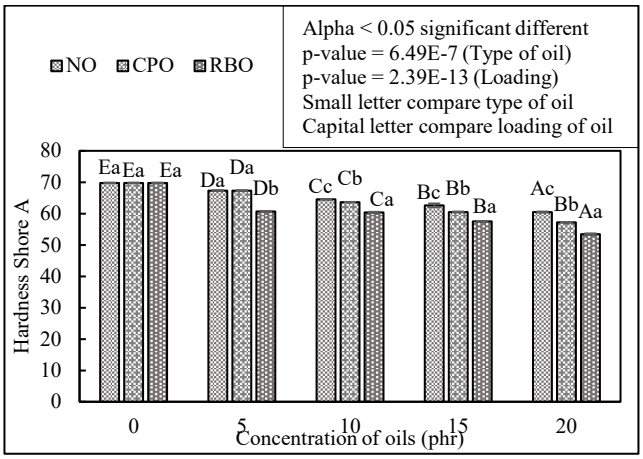


Fig2. Hardness of rubber composite with different processing aid and different ratio before aging

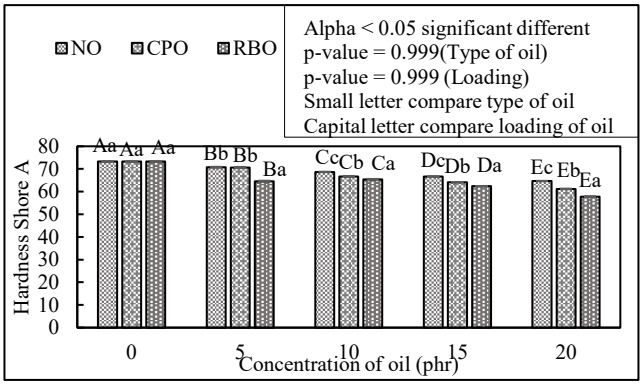


Fig 3. Hardness of rubber composite with different processing aid and different ratio after aging

3.3 Abrasion resistance index

The abrasion resistance index of rubber composites is illustrated in Fig 4**.** The results of ANOVA analysis shown that abrasion resistance index was not significant changed by the use of different processing oils (p-value = 0.37 > 0.05), however, it was significant different in term of using different concentrations of oils (p-value = 0.0002 < 0.05) in the 95% confident interval. The abrasion index of rubber composite with CPO was increased with the raising number of oils which was in the ranged from 100.33% to 133.75%. Similar to CPO, the abrasion resistance index of rubber composite with RBO was greater when larger amount of RBO were used which its value was grown from 115.97% to 142.27%. Like CPO, the value of abrasion resistance index of rubber composite with NO was also raised from 103.54% to 125.15% when oils are increased from 5phr to 20phr.According to the observation on the results, the rubber composite with rice bran oil was greater resistance than CPO, and rubber composite with CPO was better resistance than NO. This results also shown that the abrasion index tends to increase while the increasing of oils. Processing oils, act as plasticizers, lowering the viscosity of the rubber compound during mixing, which enables fillers like carbon black to disperse more uniformly and improves filler-rubber interaction [10]. This process was often responsible for the increase in abrasion index that occurs when oil was added to rubber compounds.

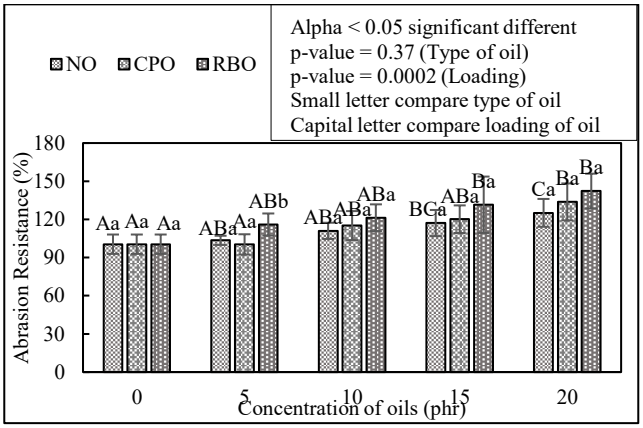


Fig 4. Abrasion resistance of rubber composite with different processing aid and different ratio

3.4 Tensile strength and elongation at break

The results for tensile strength are shown in Fig 5. As the results demonstrated that tensile strength of rubber composite was not significantly impacted by using different processing aids (p-value = 0.12 > 0.05), however the variety of concentration of oils caused the tensile strength significantly changed (p-value = 0.02 < 0.05). Tensile strength of rubber composite with CPO ranges from 22.85 MPa to 21.35 MPa, and range from 23.99 MPa to 22.59 when using RBO. For rubber composite with NO, its values were between 22.89 MPa and 21.4 MPa. Based on the observation, the tensile property tended to a little lower when increasing concentration of all three types of the processing aids. Moreover, in the Fig 6**.** has shown that tensile strength decreases after conducting to 24 hours in high temperature. However, the decreasing was improved with the larger used of oils. Tensile strength decreases predominantly as a result of chemical degradation processes such oxidation, chain scission, and cross-linking. These actions destruction the rubber's molecular structure, lowering its tensile strength. However, increasing processing aids can help to minimize this reduction by enhancing filler dispersion and rubber matrix flexibility. Oils operate as plasticizers, lowering the viscosity of the rubber compound and improving the interaction between the rubber and fillers. This better contact

can improve the rubber's mechanical qualities, particularly tensile strength, even after age [11].

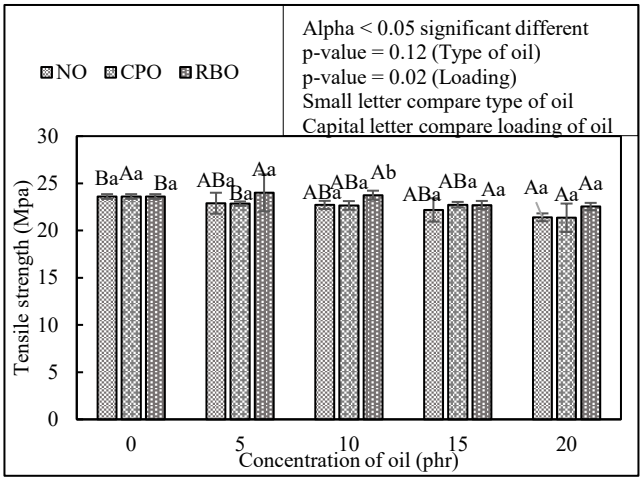


Fig 5. Tensile strength with different processing aid and different ratio before aging

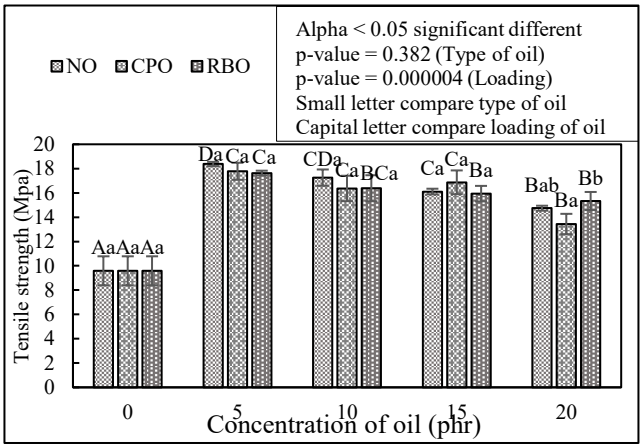


Fig 6. Tensile strength with different processing aid and different ratio after aging

In Fig 7**.** Illustrated the results for elongation at break of rubber composite which incorporated with CPO, RBO and NO. The elongation at break of rubber composite was significant effected with using different types of oils (p-value = 0.0013 < 0.05) and different ratios (p-value = 2.9578E-7 < 0.05). Elongation was raised simultaneously, the increasing of processing aids. As the results, elongation of rubber composite with CPO increased from 481.21 % to 610.93%, and the elongation raised from 539.08% to 686.40% while using RBO. For rubber composite with NO, the tensile strength grown within 466.01% to 529.02%. When greater amounts of oils used in rubber compound, the plasticizing impact of the oil causes an increase in elongation at break. As processing aids, oils lower the rubber compound's viscosity, facilitating easier movement of the polymer chains. Due to its improved mobility, the rubber matrix is more flexible and can withstand greater tensile stresses [12]. On the other hand, the Fig 8. revealed that after aging the elongation was decreased. However, the decreasing was mitigated through the raising amounts of oils. Oils can mitigate this decline by increasing the flexibility of the rubber matrix. Oils reduce the viscosity of the rubber compound while enhancing the dispersion of fillers. This enhanced dispersion minimizes stress concentration spots while increasing the elasticity of the rubber, allowing it to retain higher elongation qualities even after age [13]. According to the results, the tensile strength of rubber composite with plant oils both crude palm oil and rice bran oil is improved better in comparing to the rubber composite using naphthenic oil.

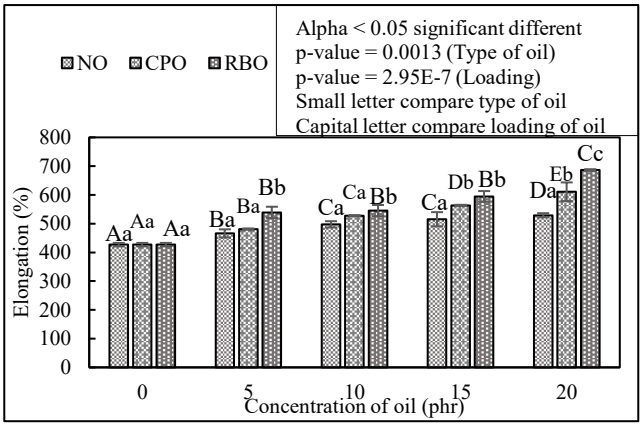


Fig 7. The elongation at break with different processing aid and different ratio before aging

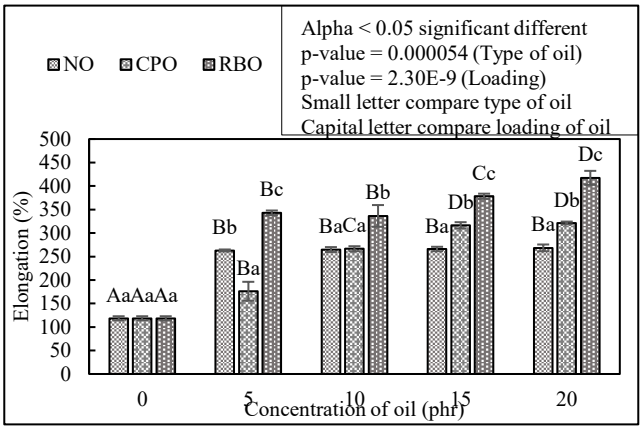


Fig 8. The elongation at break with different processing aid and different ratio after aging

4. CONCLUSIONS

In conclusion, this study demonstrated that the utilizing crude palm oil and rice bran oil were obtained the improvement of the processing of rubber compounding, physical, and mechanical properties similar or better than that using naphthenic oil. At 10phr of processing oil was considered as the optimum condition. Crude palm oil and rice bran oil were better than naphthenic oil since it reduced the cure time of rubber composite. When using 10phr of oils, the tensile strength of crude palm oil was 22.66MPa almost the same naphthenic oil which was 22.71 MPa and rice bran oil was 23.74 MPa. Moreover, Elongation of crude palm oil and rice bran oil was better than naphthenic oil which were 527.62%, 594.64 %, and 497.53% respectively. In addition, crude palm oil and rice bran oil were better than naphthenic oil in term of abrasion resistance. The abrasion resistance index of rubber composite with crude palm oil was 115.12%, rice bran oil was 121.32% while naphthenic oil only 111.01%. It was suggested that the both natural oils investigated can be employed as processing aids with no adverse effect on the qualities of the vulcanizates produced. The results show that rice bran oil and crude palm oil were superior alternative processing aids than petroleum-based naphthenic oils, which have been linked to cancer. In addition, palm oil has the finest heat resistance. Furthermore, the cure characteristics and mechanical qualities of the crude palm oil-based vulcanizate indicate that it could be employed as a co-activator in rubber compounding.

ACKNOWLEDGMENTS

The work was supported by Rubber Processing Lab, and MRT-TCC Join Venture was gratefully acknowledged.

REFERENCES

[1] Y. Dararath, N. Top, and V. Lic, "Rubber Plantation Development in Cambodia : At What Cost ?," *Econ. Environ. Progr. Southeast Asia*, no. September, 2011, [Online]. Available: http://www.eepsea.org/

[2] Im Hol and Serey Mardy, "A Review on the Status of Rubber Plantations in Cambodia," *Int. J. Sustain. Appl. Sci.*, vol. 2, no. 2, pp. 123–130, 2024, doi: 10.59890/ijsas.v2i2.1356.

[3] H. M. Hanif, K. C. Yong, and S. Y. Lee, "Effects of a newly developed palm-Based processing aid and curing systems on the physical properties of nitrile rubber composites," *Sains Malaysiana*, vol. 49, no. 9, pp. 2187–2196, 2020, doi: 10.17576/jsm-2020-4909-16.

[4] D. Kusuma Arti and W. Donwang, "Influence of Processing Oil Content on Rubber-Filler Interactions in Silica/Carbon Black-Filled Natural Rubber Compounding," *E3S Web Conf.*, vol. 481, pp. 1–10, 2024, doi: 10.1051/e3sconf/202448106007.

[5] D. Lee and S. H. Song, "Investigation of epoxidized palm oils as green processing aids and activators in rubber composites," *Int. J. Polym. Sci.*, vol. 2019, no. 1, p. 2152408, 2019.

[6] T. Ushikusa *et al.*, "Investigation of Crude Palm Oil as an Alternative Processing Oils in Natural Rubber : Effect of the Unsaturated Fatty Acid Investigation of Crude Palm Oil as an Alternative Processing Oils in Natural Rubber : Effect of the Unsaturated Fatty Acid," *Mater. Sci. Eng.*, 2019, doi: 10.1088/1757-899X/548/1/012009.

[7] W. G. D. Jayewardhana, G. M. Perera, D. G. Edirisinghe, and L. Karunanayake, "Study on natural oils as alternative processing aids and activators in carbon black filled natural rubber," *J. Natl. Sci. Found. Sri Lanka*, vol. 37, no. 3, pp. 187–193, 2009, doi: 10.4038/jnsfsr.v37i3.1212.

[8] H. Mohd-Hanif, K. C. Yong, and S. Y. Lee, "Palm-based process aid as an alternative to naphthenic oil for deproteinised natural rubber compounding," *J. Rubber Res.*, vol. 25, no. 4, pp. 299–311, 2022.

[9] M. A. G. Sánchez, D. H. Giraldo-Vásquez, and R. M. Sánchez, "The effect of thermal ageing on the mechanical properties of natural rubber-based compounds used for rubber bearings," *J. Eng. Technol. Sci.*, vol. 53, no. 3, 2021, doi: 10.5614/j.eng.technol.sci.2021.53.3.10.

[10] W. Arayapranee, "Rubber Abrasion Resistance," *Abrasion Resist. Mater.*, 2012, doi: 10.5772/30556.

[11] R. Boillat-Newport, S. P. Isanaka, and F. Liou, "Impact of Delayed Artificial Aging on Tensile Properties and Microstructural Evolution of Directed Energy Deposited Scalmalloy®," *Appl. Sci.*, vol. 15, no. 7, p. 3674, 2025.

[12] X. Li *et al.*, "Study on Increasing the Binding Amount of Rubber and Reinforcing Filler by Adding Aromatic Solvent Oil," *Polymers (Basel).*, vol. 14, no. 13, 2022, doi: 10.3390/polym14132745.

[13] H. Harahap, M. Lubis, N. Hayat, and Sukardi, "Effect of aging process on elongation at break and morphology of natural rubber latex film filled with nanocrystalline cellulose and alkanolamide," *Int. J. ChemTech Res.*, vol. 9, no. 11, pp. 201–205, 2016.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Mineralogical and Physicochemical Characterizations of Silica Sand by Using Microbeam, XRD, and XRF

Muylin Meun[1], Chhor Yi Ly[2], Navin Cheat[1], Marany Sat[1], Nhornhoem Pich[1], Vichet Thi[2], Sok Heng Hor[2], Tharith Sriv[1,2,*]

[1] *Department of Physics, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*
[2] *Graduate Program in Physics, Graduate School of Science, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*

[*]*Corresponding author: tharith@rupp.edu.kh*

Abstract: *Silica sand, also known as white sand or industrial sand, mainly consists of silicon and oxygen with low impurities. Different percentages of impurities, grain sizes, and surface morphologies lead to different industrial applications such as optical fibers, electronics, photovoltaic cells, aside from typical applications in glassmaking, ceramics, construction materials, water filtration, sports fields, paints and coatings, and abrasives. In previous studies, elemental compositions and grain sizes from silica sand in Kandal and Sihanoukville have been investigated by using x-ray fluorescence (XRF) spectroscopy and x-ray diffraction (XRD) technique. However, sand from Kampot, Koh Kong, and Kep was not considered. This study investigates the mineralogical and physicochemical properties with special focus on surface morphology, elemental compositions, and grain sizes of sand collected from Kep, Kampot, Kandal, Koh Kong, and Sihanoukville by using handheld XRF, XRD, and microbeam techniques such as energy-dispersive x-ray spectroscopy (EDS) and scanning electron microscopy (SEM). The XRF results revealed the percentages of Si of 94.79% (Kep), 97.89% (Kampot), 86.12% (Kandal), 98.28% (Koh Kong), and 98.14% (Sihanoukville). Additional results from EDS with the suppression of carbon element during the measurements reveal the percentages of 98.80% (Kep), 99.39% (Kampot), 99.8% (Sihanoukville), 94.83% (Kandal), and 99.18% (Koh Kong) of quartz (SiO_2). The XRD results confirmed the presence of quartz of crystal structure known as hexagonal for all samples collected from different sites, with an additional presence of albite ($NaAlSi_3O_8$) of triclinic structure in the sample from Kandal. The SEM results illustrate that the average grain sizes of the sand from Kep, Kampot, Kandal, Koh Kong, and Sihanoukville are of 1.04* mm, *0.49* mm, *0.37* mm, *0.26* mm, *and 0.23* mm, *respectively. These scientific findings provide valuable information on high qualities of silica sand, which lead to several significant applications such as glassmaking and foundry industry for every sand except Kandal. Despite having a relatively low Si concentration, the Kandal samples can still be utilized in the construction industry, water filtration, and sport fields.*

Keywords: Silica sand; microbeam; XRD; XRF

1. INTRODUCTION

Silica sand is known as white sand found on beaches, coastal areas, and riversides, and is commonly found in nature as quartz (SiO_2). Silica sands are valued for both chemical and physical properties. These include high-grade silica sands, essential raw materials in the world's industrial development, particularly for photovoltaic devices in solar panel production [1] and glassmaking [2]. The high-quality silica sand consists of quartz (SiO_2) content between 95% and 99.9%, and a small percentage of impurities such as Fe_2O_3, even down to 0.01% [3]. Due to its worldwide availability and low cost, quartz has been used in the foundry industry (containing approx. 95% SiO_2) [4]. Besides its chemical properties, sand plays an important role in water filtration due to its physical properties, which depend on the grain sizes used for slow-sand filtration applications [5]. In addition, silica sand has been used in the construction industry as a fine aggregate in concrete [6, 7]. In Egypt, silica sand, located in Sinai and along the Red Sea coast (Wadi El Dakhal), with SiO_2 ranging from 98.60% to 99.00%, has been contributed to and increased by adding value to the raw silica sand in the chemical leaching process to produce a high SiO_2 content. The leaching process with HCl, (HCl & H_2SO_4), (HCl, HF, HNO_3), [HCl, HF and $(COOH)_2$] acids results indicate upgrading the SiO_2 content from 98.60 to 99.99% by x-ray fluorescence (XRF) spectroscopy, and x-ray diffraction (XRD) technique and other techniques [8]. Moreover, the raw silica sand from the

[1]* Corresponding author: Tharith Sriv
E-mail: tharith@rupp.edu.kh

Ivorian sedimentary basin possibly provides solutions for producing soda-lime glass from local silica [9]. Furthermore, there was a study about silica sand properties and its particle size by using XRF, XRD analysis, and Frantz isodynamic magnetic separator in Kandal and Sihanoukville, Cambodia [10, 11]. However, silica sand from Kep, Kampot, and Koh Kong provinces has not been investigated.

In this study, advanced instruments such as handheld XRF, energy dispersive x-ray spectroscopy (EDS), XRD technique, and scanning electron microscope (SEM) were utilized to investigate the elemental compositions, crystal structures, surface morphology, and grain sizes of silica sand collected from Kep, Kampot, Kandal, Koh Kong, and Sihanoukville provinces.

2. METHODOLOGY

2.1 Geological Map of Sample Sites

The sand samples were collected by using non-probability sampling technique, specifically known as convenience sampling, from several locations such as Kep, Kampot, Kandal, Koh Kong, and Sihanoukville provinces, Cambodia, and labeled as Kep, KP, KD, KK, and SN, respectively. The specific locations of the samples were detailed in Figure 1 (a) to (e).

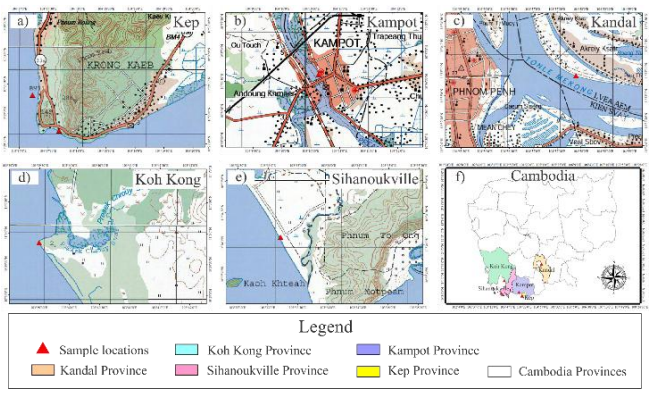


Fig. 1. Sampling sites in (a) Kep, (b) Kampot, (c) Kandal, (d) Koh Kong, and (e) Sihanoukville provinces, Cambodia. Specific locations are marked, with a separate legend provided for symbol interpretation

2.2 Sample Measurements

The intact samples were characterized using handheld XRF (X-MET8000) to determine the elemental proportions after being cleaned in an ultrasonic cleaner to remove fine contaminants, including dirt, organic matter, and residuals from the sand surface, and being fully dried. For further characterizations, SEM equipped with EDS (Thermo Scientific Axis ChemiSEM SEM/EDS) was utilized to understand its surface morphology and elemental distribution within the samples. The grain sizes of each sample as revealed by SEM were further analyzed by using ImageJ software. Additionally, all samples were ground into powder to increase surface area and level of homogeneity before XRD measurements using Malvern Panalytical Aeris systems ($\lambda = 0.15406$ nm) to determine the crystal phases and structures. The XRD results were analyzed by utilizing MDI Jade 6 software.

3. RESULTS AND DISCUSSION

3.1 XRF Analysis of Silica Sand Samples

Table 1 mainly shows the results from XRF characterization of the studied silica sand. It illustrates the average elemental compositions, highlighting a high percentage of silicon (Si) in all the samples, with values of 94.79%, 97.89%, 86.81%, 98.28%, and 98.14% for Kep, Kampot, Kandal, Koh Kong, and Sihanoukville, respectively. The high concentration of Si indicates that all samples are of high quality. Trace amounts of impurity elements, including iron (Fe), lead (Pb), and aluminum (Al), and other low-concentration elements were also detected.

Table 1. The average percentages of elemental composition of Silica Sand in each province from XRF measurement

Element	Kep%	KP%	KD%	KK%	SN%
Si	94.79	97.89	86.81	98.28	98.14
Fe	3.51	0.64	6.89	0.89	0.84
Pd	0.24	0.24	0.46	0.23	0.27
Al	0.97	0.80	0.20	0.09	0.33
P	0.14	-	-	-	0.17
Zr	0.06	0.06	-	-	-
Nb	0.02	0.02	-	-	0.02
Co	0.02	0.02	-	-	0.02
W	0.01	0.01	-	-	0.02
Re	0.02	0.01	0.01	0.01	0.03
Ni	0.02	0.01	-	0.01	0.01
Mo	0.02	0.03	-	0.09	0.04

3.2 EDS Analysis of Silica Sand Samples

For more systematic studies, as mentioned in Section 2, EDS analysis was carried out to determine the elemental

composition of the samples. Table 2 presents the average atomic percentages of elemental compositions for all silica sand samples. The results show that oxygen (O) has the highest concentration across all samples, with values of 65.87% (Kep), 66.26% (KP), 63.22% (KD), 66.12% (KK), and 66.55% (SN), indicating the dominance of oxide compounds. It should be noted that silicon (Si) also exhibited high proportions across all samples, with values of 32.93%, 33.13%, 31.61%, 33.06%, and 33.25%, respectively. Small amounts of impurities, including aluminum (Al), iron (Fe), titanium (Ti), sodium (Na), and other minor elements, were also detected, most likely present as their respective oxides. These impurities are consistent with the natural mineralogical variations typically observed in silica sand, as long been understood. It is noteworthy that during the SEM/EDS measurements, oxygen and carbon (C) EDS signals were suppressed due to the presence of carbon tape being used in the measurement process to mount the samples on the stubs. By calculating the average atomic percentage, the results show that the SiO_2 content in SN samples has the highest value of 99.80%, followed by KK (99.18%), KP (99.39%), Kep (98.80%), and KD (94.83%). Iron (Fe) was detected in the KD (0.80%), Kep (0.50%), KK (0.26%), and KP (0.37%), except SN samples.

Table 2 The average atomic percentage of elemental compositions of Silica Sand in each province

Element	Kep%	KP%	KD%	KK%	SN%
Si	32.93	33.13	31.61	33.06	33.25
O	65.87	66.26	63.22	66.12	66.55
Al	0.42	0.37	2.13	0.49	0.1
Fe	0.50	0.16	0.8	0.26	-
Mg	0.11	-	0.1	0.07	-
Ti	0.01	-	0.03	-	-
K	0.07	-	0.81	-	-
Ca	0.02	-	0.23	-	-
Zn	0.02	-	-	-	-
In	0.02	-	-	-	-
Ni	0.03	0.08	-	-	-
Na	-	-	0.6	-	-

The iron (Fe) content in sand influences its suitability for glass production, as metallic oxides in silica sand typically produce colored glass. Therefore, iron concentration is the key factor in determining whether a sand source is appropriate for manufacturing clear or colored glass [12]. Window glass sand typically contains 0.1% to 0.5% iron oxide (Fe_2O_3), while darker glass, such as that used for green bottles, may contain 2% to 3% iron oxide [13].

Figure 2 presents the EDS elemental mapping of silica sand samples collected from various provinces in Cambodia. The elemental mapping images illustrate the spatial distribution of elements within the samples. Silicon (Si), aluminum (Al), and iron (Fe) were primarily present in the composite matrix and were uniformly distributed throughout the samples.

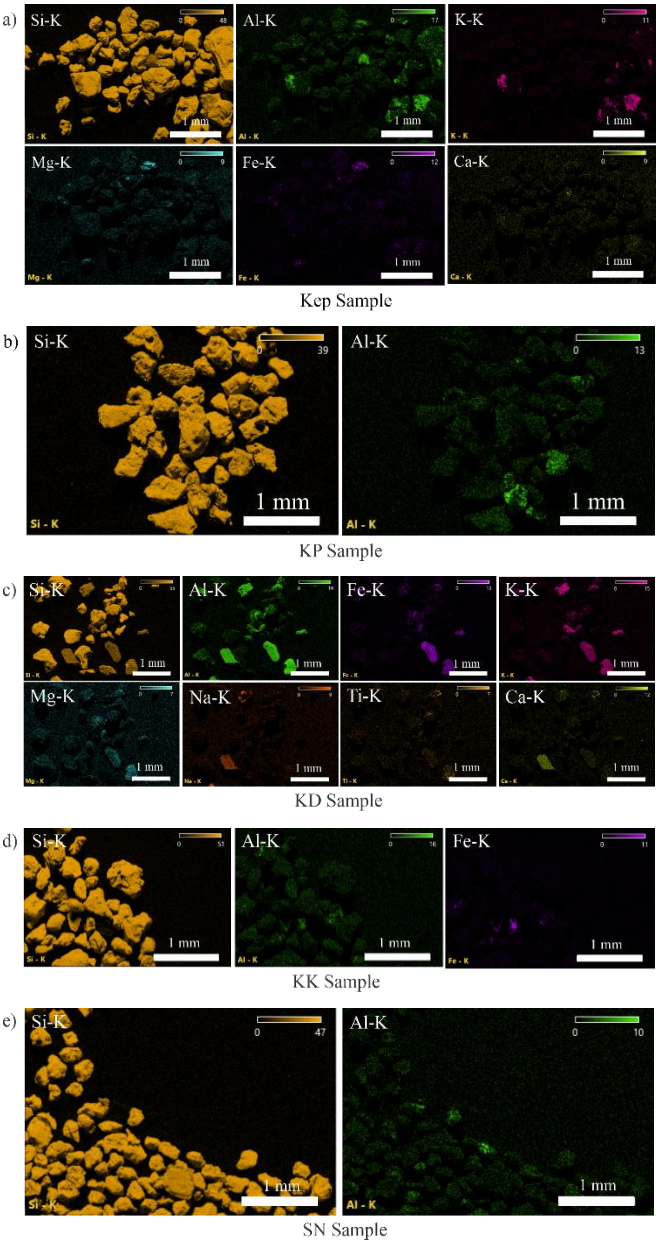


Fig. 2. EDS mapping of silica sand samples labeled as (a) Kep, (b) KP, (c) KD, (d) KK, and (e) SN, respectively

Figure 3 shows the characteristic energy peaks corresponding to each detected element. These spectra confirm the presence of Si, O, Al, Fe, and other minor elements, with each peak appearing at specific energy levels unique to the respective element. These findings are consistent with the quantitative EDS results presented in Table 2, further supporting the high silica content and the presence of oxide-based impurities, which influence the overall purity and potential industrial applications of the samples.

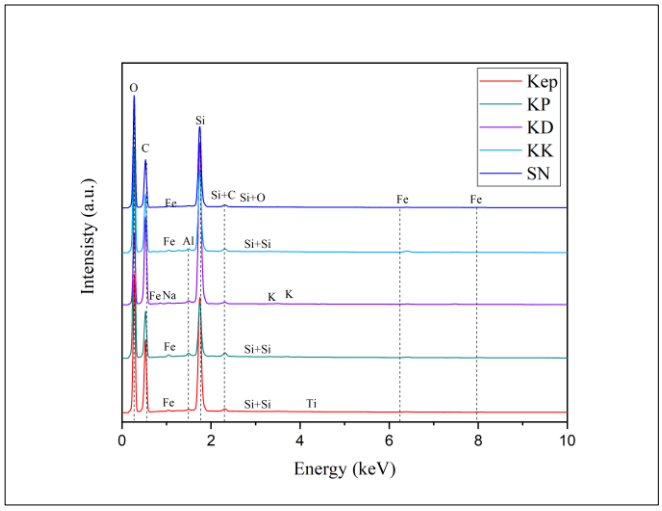


Fig. 3. EDS spectra of silica sand collected from five provinces, including Kep, Kampot, Kandal, Koh Kong, and Sihanoukville

3.3 XRD Analysis of Silica Sand Samples

XRD analysis was conducted to identify the crystal phases, and crystal structures present in all the samples. Figure 4 illustrates the XRD patterns of silica sand samples collected from Kep, Kampot, Kandal, Koh Kong, and Sihanoukville provinces. The diffraction peaks observed in all patterns are attributed to quartz (SiO_2), indicating the dominance of high-purity silica sand across all samples.

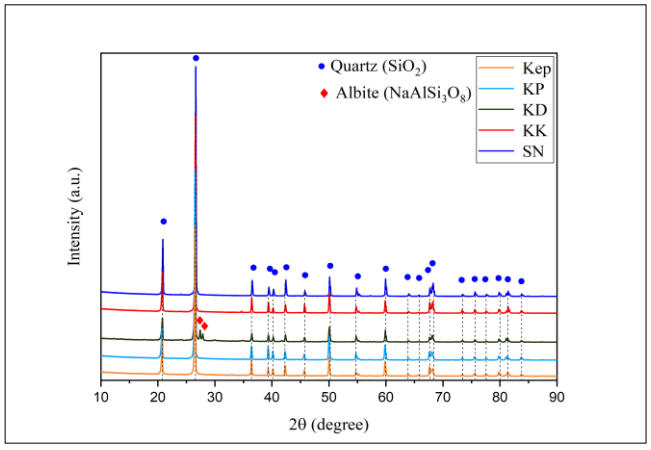
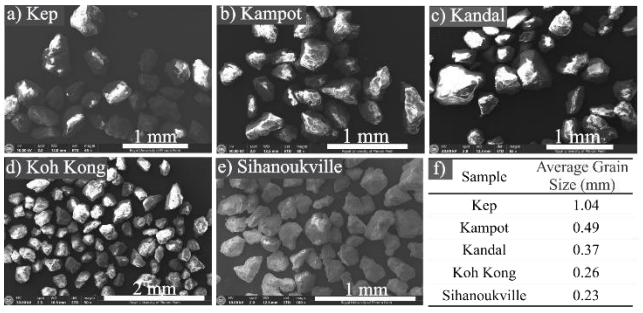


Fig. 4. XRD patterns of silica sand from different provinces in Cambodia

Notably, the samples from Kandal Province also exhibit additional peaks corresponding to albite ($NaAlSi_3O_8$), suggesting the presence of a secondary mineral phase. These findings are consistent with the EDS results in Table 2, which revealed a relatively lower silicon content in the Kandal samples compared to others. Furthermore, among all the samples, the silica sand samples from Koh Kong exhibit the highest silicon concentration, followed by those from Sihanoukville, Kampot, Kep, and Kandal. These significant results are more suitable for high-tech applications [14].

3.4 Grain Sizes of Silica Sand Samples

SEM images of the Kep, KP, KD, KK, and SN samples are presented in Figure 5 (a), (b), (c), (d), and (e), respectively. Figure 5 (f) illustrates the average grain sizes of these samples. The results show that the Kep sample has the largest average grain size of approximately 1.04 mm, followed by the KP, KD, KK, and SN samples of approximately 0.39 mm, 0.37 mm, 0.26 mm, and 0.23 mm, respectively. It has been reported that grain sizes ranging from 150 to 400 µm are suitable for foundry industries [15], and larger grain sizes are widely known to be advantageous for the construction industry. In addition, the KK and SN grain sizes are good for water filtration, as reported that the optimal sand filtration size ranges from 0.35 mm to 0.15 mm, with a uniformity coefficient ideally below 2 [5]. Based on the SEM results, the sand samples from these provinces show potential for industrial applications, including in foundry, construction, and high-tech sectors.



f)

Sample	Average Grain Size (mm)
Kep	1.04
Kampot	0.49
Kandal	0.37
Koh Kong	0.26
Sihanoukville	0.23

Fig. 5. SEM images of silica sand. (a) Kep, (b) Kampot, (c) Kandal, (d) Koh Kong, and (e) Sihanoukville

4. CONCLUSIONS

The mineralogical and physicochemical properties of silica sand samples collected from various provinces, including Kep, Kampot, Kandal, Koh Kong, and Sihanoukville, were revealed through systematic studies, in which handheld XRF, EDS, SEM, and XRD were utilized. The XRF, EDS, and XRD results indicated different

proportions of Si across all samples. Moreover, the XRD pattern of the KD sample exhibited new diffracted peaks that correspond to albite ($NaAlSi_3O_8$). Furthermore, SEM results mainly and accurately revealed the average grain sizes of approximately 1.04 mm, 0.49 mm, 0.37 mm, 0.26 mm, 0.23 mm for Kep, KP, KD, KK, and SN, respectively. These results provide significant scientific evidence for silica sands in the studied sites and reveal their potential applications such as high-tech applications, photovoltaic devices, water filtration, glass making, and the foundry industry. In other words, future research should investigate the potential of Cambodian silica sand as a raw material for high-purity silica production, supporting the development of a local semiconductor materials industry and exporting to other countries.

ACKNOWLEDGMENTS

This work was supported by the Royal Government of Cambodia through the Higher Education Improvement Project (IDA Credit No. 6221-KH) and Swedish International Development Cooperation Agency (SIDA) through Sweden and Royal University of Phnom Penh's Pilot Research Cooperation Programme (Sida Contribution No. 11599).

REFERENCES

[1] T. Buonassisi *et al.*, "Chemical natures and distributions of metal impurities in multicrystalline silicon materials," *Prog. Photovolt. Res. Appl.*, vol. 14, no. 6, pp. 513–531, 2006, doi: 10.1002/pip.690.

[2] P. R. Thio, K. B. Koffi, K. D. Konan, and K. A. Yao, "Production of High-Purity Silica Sand from Ivorian Sedimentary Basin by Attrition without Acid Leaching Process for Windows Glass Making," *J. Miner. Mater. Charact. Eng.*, vol. 9, no. 4, Art. no. 4, Jul. 2021, doi: 10.4236/jmmce.2021.94024.

[3] A. Burkowicz, K. Galos, and K. Guzik, "The Resource Base of Silica Glass Sand versus Glass Industry Development: The Case of Poland," *Resources*, vol. 9, no. 11, Art. no. 11, Nov. 2020, doi: 10.3390/resources9110134.

[4] J. Svidró, J. Tamás Svidró, and A. Diószegi, "The role of purity level in foundry silica sand on its thermal properties," *J. Phys. Conf. Ser.*, vol. 1527, no. 1, p. 012039, Apr. 2020, doi: 10.1088/1742-6596/1527/1/012039.

[5] "New_Horizons_for_Slow_Sand_Filtration.pdf." Accessed: Apr. 19, 2025. [Online]. Available: https://library.uniteddiversity.coop/Water_and_Sanitation/BioSand_Filters/New_Horizons_for_Slow_Sand_Filtration.pdf

[6] K. Turk and M. L. Nehdi, "Flexural toughness of sustainable ECC with high-volume substitution of cement and silica sand," *Constr. Build. Mater.*, vol. 270, p. 121438, Feb. 2021, doi: 10.1016/j.conbuildmat.2020.121438.

[7] R. Malathy *et al.*, "Use of Industrial Silica Sand as a Fine Aggregate in Concrete—An Explorative Study," *Buildings*, vol. 12, no. 8, Art. no. 8, Aug. 2022, doi: 10.3390/buildings12081273.

[8] G. Mohamed, T. Hamed, and M. Fathy, "Advanced Chemical leaching process of Wadi El Dakhal silica sand to produce a high concentration of SiO 2 for electronic industry," Mar. 20, 2023, *Research Square*. doi: 10.21203/rs.3.rs-2685990/v1.

[9] R. P. Thio, D. K. Konan, B. K. Koffi, and A. K. Yao, "Characterization of raw silica sand from the Ivorian sedimentary basin for silica glass making".

[10] "(PDF) Comparison of Silica Sand Properties from Kandal Province, Cambodia and Tapah, Perak, Malaysia and Characterization of Soda Lime Silicate Glass Produced From Cambodian Silica Sand," *ResearchGate*, Mar. 2025, doi: 10.4028/www.scientific.net/AMR.858.248.

[11] S. Ros, "Pre-feasibility study of silica sand processing plant in Sihanoukville province, Cambodia," Bangkok, Thailand, 2018. doi: 10.58837/CHULA.THE.2018.265.

[12] "Silica-Sand-March-2010-Report-Final.pdf." Accessed: Apr. 28, 2025. [Online]. Available: https://gwp.uk.com/wp-content/uploads/2016/04/Silica-Sand-March-2010-Report-Final.pdf

[13] "Uses and Specifications of Silica Sand," Alberta Geological Survey. Accessed: Apr. 28, 2025. [Online]. Available: https://ags.aer.ca/publications/all-publications/esr-1971-04

[14] M. H. Tania, A. S. W. Kurny, and F. Gulshan, "The Prospect of Bipinganj Sand in High Tech Applications," *Procedia Eng.*, vol. 90, pp. 172–175, 2014, doi: 10.1016/j.proeng.2014.11.832.

[15] "Foseco Ferrous Foundryman's Handbook - Google Books." Accessed: Apr. 28, 2025. [Online]. Available: https://books.google.com.kh/books?hl=en&lr=&id=1BoUNFn6S6MC&oi=fnd&pg=PR11&dq=Foseco+Ferrous++Foundryman%E2%80%99s+Handbook&ots=-QZK1PCFi-&sig=Xk8_Fk6J5YNNxf9hUhYEbySOZr0&redir_esc=y#v=onepage&q=Foseco%20Ferrous%20%20Foundryman%E2%80%99s%20Handbook&f=false

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Determination of Crystallite Size of Transition Metal Oxides (NiO, Mn_3O_4, and α-Fe_2O_3) Using the Scherrer Formula and Williamson–Hall Plot

Sokly Sreang[1], Chhor Yi Ly[2], Navin Chheat[1], Piseth Lim[2], Tharith Sriv[1,2,*]

[1]*Department of Physics, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*
[2]*Graduate Program in Physics, Graduate School of Science, Royal University of Phnom Penh, Russian Federation Blvd., Tuol Kork, Phnom Penh, Cambodia*

[*]*Corresponding author: tharith@rupp.edu.kh*

Abstract: *Transition metal oxides (TMOs) are compounds of transition metals and oxygen that have unique physical, chemical, and electrical properties. Due to the wide range of applications and their interesting properties, TMOs have been studied, where crystal structure, crystallite size, particle size and shape, space group were considered, by using several techniques such as scanning electron microscopy (SEM), energy dispersive x-ray spectroscopy (EDS), atomic force microscopy (AFM), x-ray diffraction (XRD) technique, transmission electron microscopy (TEM) . However, Scherrer equation and Williamson-Hall plot, which are widely used for estimating the crystallite size of nanomaterials, have not been used for Nickel Oxide (NiO), Manganese Tetroxide (Mn_3O_4), and Alpha-phase Ferric Oxide (α-Fe_2O_3). It should be noted that several factors such as synthesis temperature, reaction time, precursors and solvents, annealing or heat treatment can affect the crystallite size. The current study is motivated by an investigation of the crystallize sizes of NiO, Mn_3O_4, and α- Fe_2O_3 nanoparticles (NPs) by using Scherrer equation and Williamson-Hall plot. The materials were synthesized by using co-precipitation method. The synthesized NiO, Mn_3O_4 and α-Fe_2O_3 NPs were then confirmed by using SEM, EDS, and XRD technique. The crystallite sizes of these three TMOs were finally determined by extracting the XRD data using Scherrer equation and Williamson-Hall plot. The SEM results show that the particle sizes of NiO, Mn_3O_4 and α-Fe_2O_3 NPs range from 55 nm to 169 nm, 79 nm to 141 nm, and 68 nm to 163 nm, respectively. In addition, according to the EDS and XRD results, the desired TMO nanoparticles were perfectly obtained. Finally, the results from Scherrer equation indicate that the crystallite sizes of NiO, Mn_3O_4, and α-Fe_2O_3 are 4.7709* nm*, 15.2309* nm*, and 33.8360* nm, respectively, whereas *those obtained from Williamson-Hall plot are 4.1769* nm*, 15.0711* nm*, and 43.3293* nm*, respectively. These crystallite sizes differ slightly. This is due to consideration of strain within Williamson-Hall plot in the samples, which leads to better results as compared with those from Scherrer equation. As crystallize size of a nanomaterial can affect the surface area, reactivity, mechanical strength, and quantum effects aside from catalytic activities, optical absorption, magnetic behavior as well as electrical conductivity and absorption of greenhouse gas, knowing the exact crystallize sizes of these TMOs allows one to effectively enhance the aforementioned properties, which consequently enhances device designs and other applications including potential use in the absorption of CO_2.*

Keywords: Crystallite size; X-ray; Scherrer; Williamson-Hall

1. INTRODUCTION

Transition metal oxides (TMOs) are the oxides of d-block elements in the periodic table with partially filled *d*-sub-shell [1]. They display a variety of technically significant electronic features, including charge-density, insulator-metal transitions, magnetic, dielectric and ferroelectric, catalytic, superconductive, redox, thermoelectric, and multiferroic characteristics [2]. Different TMOs including NiO, Mn_3O_4 and α-Fe_2O_3, which are in nanoscales, have been used in different applications and observed to be of nanoparticle structure. Nanomaterials are defined by the International Organization for Standardization, or ISO, as "material with any external dimension in the nanoscale or having an internal structure or surface structure in the nanoscale," since the nanoscale is only 1 to 100 nm (roughly) [3]. Nanoparticles are of great scientific interest, physical and chemical properties of which have been studied such as surface area, optical properties, and uniformity. In the nanoscale their properties depend on discrete molecular or atomic phenomena that come from the different sample growth and synthesis.

[1]* Corresponding author: Tharith Sriv
E-mail: tharith@rupp.edu.kh; Mobile: +855-85289791

Crystallite size is one of the dimensional impacts on the properties of nanoparticles, which has been focused on by scientists. To estimate the crystallite size of nanomaterials, the techniques typically involve Scherrer formula and Williamson-Hall plot through XRD data. It should be noted that XRD technique provides information on structures, phases, preferred crystal orientations (texture), and other structural parameters, such as average grain size, crystallite size, strain, and crystal defects [4]. Scherrer formula and Williamson-Hall plot have been utilized to determine crystallize sizes and strain of materials. The formula was found by Paul Scherrer, which can lead to the result of crystallite sizes from the broadening of the main x-ray diffraction peak. The equation is appropriately applied particles of average size up to 100-200 nm [5]. On the other hands, Williamson-Hall plot is recognized as a method introduced by G.K. Williamson and W.H Hall in 1935 to determine XRD line broadening in materials [6]. The method involves plotting the full width at half maximum (FWHM) of the x-ray peaks against the diffraction angle. This plot yields a straight line, where the slope and intercept provide information about the lattice strain and crystallite size, respectively. The crystallite size of several nanoparticles has been determined using these two techniques, however determination of the crystallize size of NiO, Mn_3O_4, and α-Fe_2O_3 by using these techniques remains unobserved, at least to the best of our knowledge.

In this study, nanoparticles such as NiO, Mn_3O_4 and α-Fe_2O_3 samples were first synthesized by using co-precipitation method and characterized with SEM, EDS and XRD. Then Scherrer formula and Williamson-Hall plot were utilized to determine the crystallite sizes and micro-strain of these three TMOs based on data from XRD measurements. The rests of the paper are organized as follow. Section 2 gives details about synthesis, and characterizations of the studied TMO nanoparticles, while Section 3 provides results from the characterizations along with a discussion. The last section gives concluding remarks.

2. METHODOLOGY

2.1 Synthesis of Nikel Oxide (NiO)

NiO nanoparticles was synthesized through co-precipitation method by mixing 9.303 g of Ni $(NO_3)_2$ in 60 ml of deionized water (DI) and stirred with a magnetic stirrer for 30 minutes at 50 ℃. Then, 3 g of NaOH was mixed with 150 ml of DI water and stirred on the same magnetic stirrer at 50 ℃ for 10 minutes. After this, the stirred NaOH solution was slowly added into Ni $(NO_3)_2$ solution and stirred for another 1 hour. DI water was then used to wash the salt precipitate using filter paper for filtration, and then dried in an oven for 11 hours at 150 ℃. Finally, the dried material was ground for 30 minutes before being fired in a tunnel furnace at 300 ℃ for 2 hours to obtain NiO nanoparticles [7].

$$Ni(NO_3)_2 \cdot 6H_2O + 2\,NaOH \rightarrow Ni(OH)_2 + 2\,NaNO_3 \quad \text{(Eq.1)}$$

$$Ni(OH)_2 \rightarrow NiO + H_2O \quad \text{(Eq.2)}$$

2.2 Synthesis of Manganese Tetroxide (Mn_3O_4)

In the case of Mn_3O_4 nanoparticles, the co-precipitation method was similarly utilized. First, 3.56 g of $MnCl_2 \cdot 4H_2O$ was mixed with 80 ml of DI water and stirred for 1 h on a magnetic stirrer. During this time, 1.49 g of NaOH was added to 20 ml of DI water. After stirring for 1 hour with $MnCl_2$, a micropipette was utilized to pump 0.14 ml of H_2O_2 into the solution. The NaOH solution was added dropwise and stirred for 6 hours. After that, DI water was added to the resulting mixture of Mn_3O_4 and NaCl several times to dilute NaCl. The solution was next filtered with filter paper for NaCl to be totally removed before being dried at 60 ℃ for 12 hours. At the end of the process, Mn_3O_4 of 1.235 g was obtained [8].

$$3MnCl_2 + 2NaOH + O_2 \rightarrow Mn_3O_4 + 2NaCl + H_2O \quad \text{(Eq.3)}$$

2.3. Synthesis of Iron Ferric Oxide (α- Fe2O3)

In the synthesis of α-Fe_2O_3 nanoparticles, a co-precipitation method that involves mixing chemical elements together to obtain a new element of interest. First, 7.1410 g of $FeCl_3$ was dissolved in 130 ml of DI water and stirred using a magnetic stirrer for 1 hour. Then, 3.168 g of NaOH was dissolved in 20 ml of DI water and stirred using a stirrer. The NaOH solution was slowly added to the $FeCl_3$ solution. During the addition, the speed of the magnetic stirrer was increased to ensure that the reaction between the solutions took place well. After mixing the solution, the temperature was slowly increased to 80 ℃ within 3 hours until salt precipitates appear on the surface of the chemical elements. The resulting NaCl salt was washed from $Fe(OH)_3$ with DI water three to four times and filtered through filtered paper. The resulting solution was dried in an oven for 12 hours and then ground for 5 minutes. Finally, $Fe(OH)_3$ was burned in a tunnel furnace at 500 ℃ for 4 hours for the final Fe_2O_3 nanoparticles to be obtained. Below summarizes the chemical reactions [9].

$$2FeCl_3(aq) + 6NaOH(aq) \rightarrow 2Fe(OH)_3(s) + 6NaCl(aq) \quad \text{(Eq.4)}$$

$$2Fe(OH)_3(s) \rightarrow Fe_2O_3(s) + 3H_2O(g) \quad \text{(Eq.5)}$$

2.4. Characterization Techniques

The nanoparticle samples were characterized by utilizing a scanning electron microscopy/ energy dispersive x-ray spectroscopy (Thermo Scientific Axia ChemiSEM SEM/EDS) system. Additionally, XRD system (Malvern Panalutical Aeris) was employed to confirm the success of the syntheses of NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles. ImageJ was further utilized for nanoparticle size analysis.

2.5. The Scherrer Equation

Scherrer equation was used to estimate the crystallite sizes of NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles. It should be noted that the equation is typically utilized to determine the crystallite sizes of homogeneous nanoparticles with diameters below 200 nm. The equation can be written as,

$$D = \frac{K\lambda}{\beta \cos\theta}, \qquad \text{(Eq. 6)}$$

where K is Scherrer constant = 0.9 (referring to shape of particles), λ is the x-ray wavelength, θ is peak position (radians), β is full width at haft maximum (broadening of peak), and D is the mean size of crystal.
Taking natural log(ln) on both sides of the equation, applying product rule of log, and rearrange it results in,

$$\ln\beta = \ln\frac{1}{\cos\theta} + \ln\frac{K\lambda}{D}. \qquad \text{(Eq. 7)}$$

Set $y = \ln\beta$, $a = 1$, $x = \ln\frac{1}{\cos\theta}$, and $b = \frac{K\lambda}{D}$, Eq. 7 then becomes linear of the form $y = ax + b$.

2.6. Williamson-Hall Plot

Williamson-Hall (W-H) plot is another method that is used to obtain the crystallite size of nanoparticles. Additionally, the method can also be used to determine the strain, which is formed in the nanostructure due to the different kinds of defect energy density model. The W-H technique deconvolutes the broadening peaks into the lattice strain and crystallite size, as demonstrated in [10]. The representing expression can be written as,

$$\beta_T = \beta_D + \beta_\varepsilon, \qquad \text{(Eq. 8)}$$

where β_T is total broadening, β_D is broadening due to crystallite size, and β_ε is broadening due to strain. Since, $\beta_D = \frac{K\lambda}{D\cos\theta}$, and $\beta_\varepsilon = 4\varepsilon\tan\theta$, the expression can be rewritten as,

$$\beta_T = \frac{K\lambda}{D\cos\theta} + 4\varepsilon\tan\theta. \qquad \text{(Eq. 9)}$$

Multiplying it by $\cos\theta$ gives,

$$\beta_T\cos\theta = \varepsilon(4\sin\theta) + \frac{K\lambda}{D}. \qquad \text{(Eq. 10)}$$

This similarly becomes a linear equation ($y = ax + b$) with $y = \beta_T\cos\theta$, $a = \varepsilon$, $x = 4\sin\theta$, and $b = \frac{K\lambda}{D}$.

3. RESULTS AND DISCUSSION

3.1. SEM/EDS Data Analyses

As mentioned earlier, upon the synthesis of the NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles, the samples were characterized using SEM and the images were further processed by using ImageJ. The results (Figure 1.a) shows that NiO nanoparticle sizes range from 55 nm to 169 nm with an average size of 122 nm. Similar particle average size of 118 nm within the range of 79 nm- 141 nm were obtained for the case of Mn_3O_4 nanoparticles, and 116 nm between 68 nm and 163 nm for that of α-Fe_2O_3 nanoparticles (Figure 1.b-c).

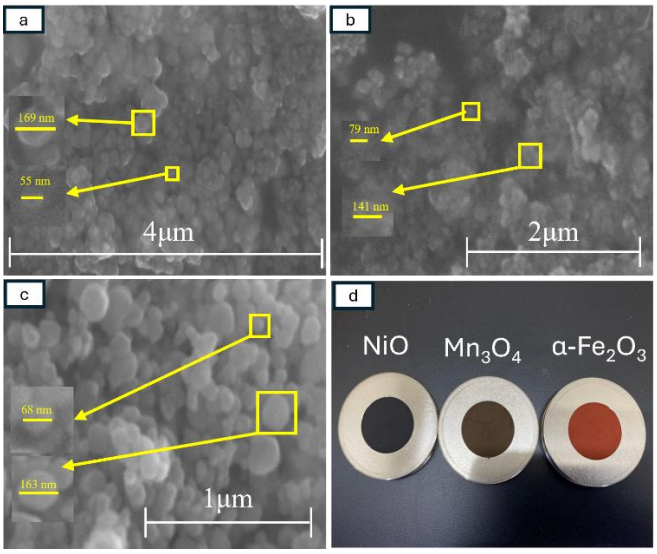


Fig. 1. SEM images of a. NiO, b. Mn_3O_4, and c. α-Fe_2O_3 nanoparticles. d. shows the image of TMO powders, which appear black (NiO), brown (Mn_3O_4), and orange (α-Fe_2O_3) to the naked eye

Table 1. Atomic and weight percentages of NiO, Mn_3O_4, and α-Fe_2O_3 nanoparticles obtained from EDS measurements

Materials	Element	Atomic %	Weight %	Net Counts
NiO	Ni	38.4	69.6	1 885 837
	O	61.6	30.4	688 023
Mn_3O_4	Mn	40.2	69.8	3 016 075
	O	59.8	30.2	1 194 291
$\alpha-Fe_2O_3$	Fe	37.1	67.3	4 019 098
	O	62.9	32.7	1 638 563

Table 1 illustrates the results from EDS characterization of NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles. According to the data in the table, the atomic percentages of Ni and O in NiO

nanoparticles are 38.4 and 61.6. Those of Mn and O in Mn_3O_4 nanoparticles are 40.2 and 59.8, whereas 37.1 percent of Fe and 62.9 percent of O make α-Fe_2O_3 nanoparticles. These results confirm the stoichiometry of the resulting nanoparticles synthesized using the co-precipitation method.

Figure 2 illustrates the EDS mapping results from the energy dispersive x-ray spectroscopy done on powders consisting of nanoparticles being studied. Figure 2.a shows the spectra of the synthesized nanoparticles (NiO, Mn_3O_4, and α-Fe_2O_3). The carbon peak, which appears in each spectrum, was attributed to the carbon tape used during the EDS measurements. Figures 2.b, c and d illustrate the presence of each element composed in NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles, respectively. This additionally confirm the successful syntheses of the target materials with a high level of purity for the study.

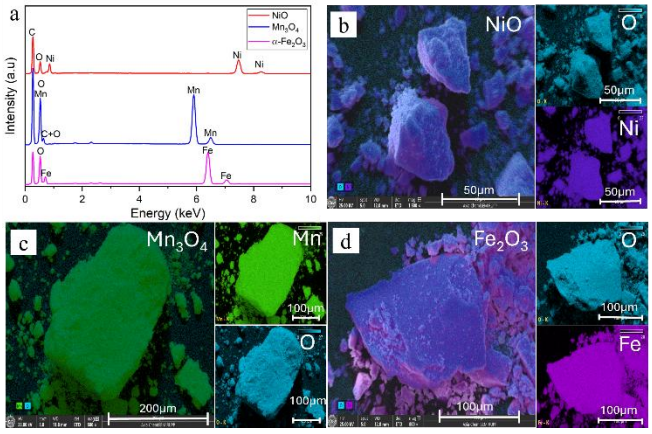


Fig. 2. a. EDS spectra of NiO, Mn_3O_4, α-Fe_2O_3 nanoparticles. b. EDS mapping of NiO nanoparticles. c. EDS mapping of Mn_3O_4 nanoparticles. d. EDS mapping of α-Fe_2O_3 nanoparticles.

3.2. XRD Data Analysis

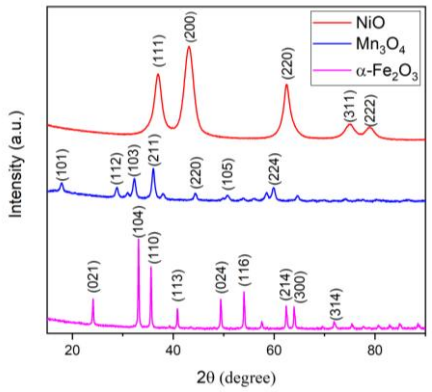


Fig. 3. XRD patterns of NiO, Mn_3O_4, α-Fe_2O_3 nanoparticles

The samples underwent XRD with the 2θ range between 20° and 90° at the laboratory temperature. The XRD data was analyzed by using MDI Jet 6 software. The results of the analyses further confirm the presence of the synthesized NiO, Mn_3O_4, and α-Fe_2O_3 nanoparticles as all peaks, which are presented in Figure 3 perfectly match with those from the database of MDI Jet 6 software. In addition, the analyses indicate that the crystal structure of NiO is cubic (space group: $Fm\bar{3}m$) with lattice constants of 4.179 Å. That of Mn_3O_4 is tetragonal (space group: $I4_1/amd$) with the lattice constants of a=b=5.765 Å and c=9.442 Å. Finally, the α-Fe_2O_3 nanoparticles have a rhombohedral crystal structure (space group: $R\bar{3}c$) with the lattice constants of a=b=c=5.424 Å.

3.3. Determination of Crystallite Sizes of the Nanoparticles Using Scherrer Formula and Williamson-Hall Plot

In the Scherrer equation, by plotting and fitting (Figures 4 a,b and c) the data extracted from the XRD results, the crystallite size can be determined from the intercept of the linear equation. It should be noted that the x-ray wavelength is 0.15406 nm. The results of the crystallite sizes are shown in Table 2.

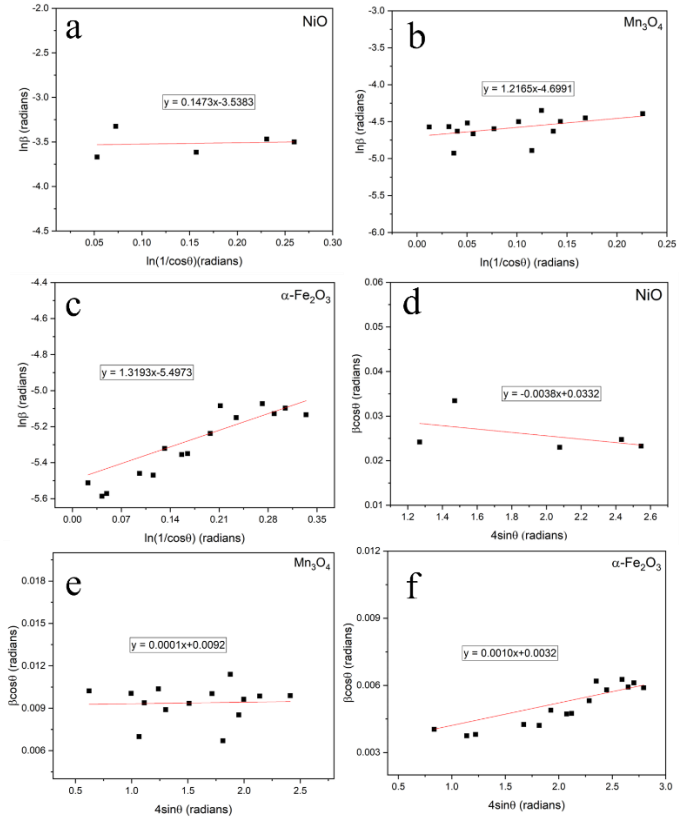


Fig. 4. The Scherrer formula plot by Origin software for a. NiO, b. Mn_3O_4 c. α-Fe_2O_3 nanoparticles and Williamson-Hall plot for d. NiO, e. Mn_3O_4 and f. α-Fe_2O_3

As expressed earlier, Eqs (7, 8, 9) are Williamson-Hall equation. To do further W-H analysis, a plot is drawn with

$4\sin\theta$ along the x-axis, and $\beta\cos\theta$ along the y-axis for all the coverage peaks. From Figures 4 d, e, f. the crystalline size was estimated from the y-intercept and micro-strain (ε) from the slope of the linear fit to the data [11].

Table 2. Crystallite size and micro-strains determined by Williamson-Hall plot and Scherrer formula

Sample	Scherrer D (nm)	Williamson-Hall (nm)	
		D (nm)	Strain
NiO	4.7709	4.1769	- 0.0038
Mn_3O_4	15.2309	15.0711	- 0.0001
α-Fe_2O_3	33.8360	43.3293	0.0010

Table 2 summarizes the crystallite sizes and micro-strains from the Scherrer and Williamson-Hall plot in nanometers. The results of crystallite sizes obtained from Scherrer and Williamson-Hall plot are slightly different due to the micro-strain in nanoparticles. Micro-strains of NiO and Mn_3O_4 take negative value representing a compressive strain, whereas that of α-Fe_2O_3 is positive, which represents tensile strains in the samples [12].

4. CONCLUSION

NiO, Mn_3O_4 and α-Fe_2O_3 nanoparticles were successfully synthesized by using co-precipitation method at laboratory temperature. The presence of pure NiO, Mn_3O_4, and α-Fe_2O_3 nanoparticles, which are ascribed to cubic, rhombohedral, and tetragonal crystal systems, respectively, is confirmed by XRD and EDS analyses. The XRD data allowed determination of crystallite sizes to be 4.7709 nm, 15.2309 nm, and 33.8360 nm from Scherrer formula, 4.1769 nm, 15.0711 nm, 43.3293 nm from W-H plot for NiO, Mn_3O_4 and α-Fe_2O_3, respectively. The micro-strains values obtained by Williamson-Hall plot are -0.38×10^{-2}, -0.01×10^{-2}, and 0.10×10^{-2}, respectively. The positive value of the latter is attributed to the value of crystallite size D and crystalline micro-strain ε. When the strain decreases, it leads to small difference between the results from Scherrer and those from Williamson-Hall plot and vice versa. To conclude, Scherrer and Williamson-Hall plot are appropriate and give accurate results of crystallite sizes of nanoparticles synthesized by using co-precipitation method and other methods that lead to the particle sizes below 200 nm with controlled synthesized parameters, as confirmed by the current results.

ACKNOWLEDGMENTS

This study was supported by the Royal Government of Cambodia through the Higher Education Improvement Project (IDA Credit No. 6221-KH) and Swedish International Development Cooperation Agency (SIDA) through Sweden and Royal University of Phnom Penh's Pilot Research Cooperation Programme (Sida Contribution No. 11599).

References

[1] A. Tyagi, S. Banerjee, J. Cherusseri and K. Kar, "Characteristics of transition metal oxides.," Handbook of Nanocomposite Supercapacitor Materials, no. 1, pp. 91-123, 2020.

[2] J. Goodenough, "Perspective on engineering transition-metal oxides," Chemistry of Materials 26, no. 2, pp. 820-829, 2014.

[3] I. O. f. Standardization, "Nanotechnologies — Vocabulary — Part 1: Core Terms," Geneva, 2023.

[4] A. Bunaciu, E. UdriŞTioiu and H. Aboul-Enein, "X-ray diffraction: instrumentation and applications," Critical reviews in analytical chemistry, no. 4, pp. 289-299, 2015.

[5] U. a. G. N. Holzwarth, "The Scherrer equation versus the'Debye-Scherrer equation," Nature nanotechnology, no. 5, pp. 534-534, 2011.

[6] G. a. H. W. Williamson, "X-Ray Line Broadening from Filed Aluminium and Wolfram," Acta Metallurgica, no. 6, pp. 22-31, 1953.

[7] A. K. Atul, S. K. Srivastava, A. K. Gupta and N. Srivastava,, "Synthesis and characterization of NiO nanoparticles by chemical co-precipitation method: an easy and cost-effective approach," Brazilian Journal of Physics, no. 7, p. 2, 2022.

[8] B. Raj, A. Asiri, J. Wu and S. Anandan, "Synthesis of Mn3O4 nanoparticles via chemical precipitation approach for supercapacitor application," Journal of Alloys and Compounds, no. 8, pp. 234-240, 2015.

[9] M. Farahmandjou, and F. Soflaee, "Synthesis and characterization of α-Fe2O3 nanoparticles by simple co-precipitation method," Physical Chemistry Research 3.3, no. 9, pp. 191-196, 2015.

[10] H. R. K. M. a. A. S. Irfan, "X-ray peak profile analysis of CoAl2O4 nanoparticles by Williamson-Hall and size-strain plot methods," Modern Electronic Materials, no. 10, pp. 31-40, 2018.

[11] B. a. H. B. Rajesh Kumar, "X-ray peak profile analysis of solid-state sintered alumina doped zinc oxide ceramics by Williamson–Hall and size-strain plot methods," Journal of Asian Ceramic Societies, no. 11, pp. 94-103, 2017.

[12] S. I. M. S. J. a. A. M. H. Yusuf, "Scherrer andWilliamson-Hall estimated particle size using XRD analysis for cast aluminum alloys," Journal of Theoretical and Applied Physics,, no. 12, 2024.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Title: Utilization of Fly Ash as an Eco-friendly filler in Styrene Butadiene Rubber compound for tire tread application

Kakda Yiem [1], Yukleav Nat [2], Laymey Sreng [3], Phanny Yos [4*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: phannyyos@itc.edu.kh*

Abstract: *This study explores the use of Fly Ash, a byproduct of coal combustion, as an eco-friendly and cost-effective filler in Styrene-Butadiene Rubber (SBR) compounds for tire tread applications. Tire treads play a critical role in a tire's performance, safety, and longevity. Traditionally, Carbon Black is used as the primary filler, but replacing part of it with Fly Ash offers both environmental and economic advantages. Fly Ash mainly contains oxides such as 52% SiO_2 and 26% Al_2O_3. Before use, Fly Ash was dried and ground to reduce its particle size to under 75 µm. Its moisture content was found to be between 1.4–1.6%. Rubber compounds were prepared with 100 parts per hundred rubber (pphr) SBR and 50 pphr of total filler in various ratios of Carbon Black to Fly Ash (CB/FA): CB0FA50, CB10FA40, CB20FA30, CB30FA20, CB40FA10, and CB50FA0, including a control with no filler (CB0FA0). These were mixed using a two-roll mill and vulcanized following ASTM standards using a conventional vulcanization system. Cure characteristics were determined using a Moving Die Rheometer (MDR), and samples were cured in a hydraulic hot press at 1000 psi. The mechanical properties of the vulcanized compounds were evaluated. Hardness ranged from 55 to 62 Shore A, and tensile strength ranged from 15 to 25 MPa. Compounds with CB20FA30, CB30FA20, CB40FA10, and CB50FA0 met the required tire tread standards (55–65 Shore A hardness and ≥15 MPa tensile strength). All formulations, except the control without fillers, met the minimum elongation standard of 420%. Additionally, the compression set of the CB40FA10 compound was comparable to CB50FA0, both around 12%, which is within the acceptable range of under 20%. Overall, the CB40FA10 formulation emerged as the most optimized blend for tire tread applications, balancing performance and sustainability. These findings indicate that Fly Ash can serve as a viable partial replacement for Carbon Black in SBR compounds, reducing environmental impact and production costs while maintaining or improving the mechanical properties of tire treads. Further research is recommended to refine the formulation and assess long-term durability in real-world conditions.*

Keywords: *ASTM, SBR, MDR, Tire Tread, Hardness, Tensile Strength, Elongation, Compression set, Cure Times.*

1. INTRODUCTION

1.1. Background

The demand for pneumatic tires is rising as a result of the global increase in the desire to possess a car. The output of vehicles worldwide rose by 25% between 2010 and 2017 [1,2]. However, subsequent years saw a decline, with production dropping to 92.2 million units in 2019 and further to 77.7 million units in 2020, largely due to the impacts of the COVID-19 pandemic. A gradual recovery followed, with 80.2 million vehicles produced in 2021 and 85 million in 2022. Projections for 2025 estimate production levels around 88.7 million units, reflecting ongoing challenges such as supply chain disruptions and geopolitical factors. These fluctuations in vehicle production directly influence the demand for pneumatic tires, underscoring the interconnected nature of the automotive and tire industries [3]. Pneumatic tires are made up of an outside surface known as tread, which is supported by a steel cord reinforcing belt. The tread comes into direct touch with the road surface, it must be sturdy enough to prevent rips or splits. Additionally, lowering rolling resistance is crucial for tread since it is linked to fuel consumption, with higher rolling resistance translating into greater energy waste. Tread is made up of a rubber compound. There are two main types of rubber, synthetic rubber and natural rubber. Natural rubber is collected from natural rubber tree which must be planted in tropical areas. As a consequence, natural rubber production is constrained and unable to meet market demand. World War II saw the development of synthetic rubber. The synthetic rubber known as Styrene Butadiene Rubber (SBR) is created by polymerizing the monomers butadiene and styrene [4]. SBR is produced in such large quantities, it has emerged as one of the most significant synthetic rubbers globally (Ciesielski, 1999). In tire applications, SBR is typically utilized in tire tread because of its strong abrasion resistance and stability [5]. In rubber compound, SBR was blended with chemical agents (Ingredients), Carbon black (CB) has long been a common reinforcing filler in SBR compound. However, the search for substitute fillers that can lessen the environmental impact has been prompted by the growing demand for eco-friendly and sustainable products. Many studies have found that Fly ash from power plant can be used in SBR compounds to reduce depend on carbon black while providing cost and environmental improvements. In addition to this, it improves on mechanical properties which is the reason why Fly ash is investigated compared with Carbon black in SBR compounds for tire tread application [6].

1.1. Objective

In the study aime to evaluate on SBR compound blend with Carbon black and Fly ash as filler. The following goals are to be met in conjunction with this study :

The main goal of this studies is to determine the effect of using Fly ash from power plant as a filler blend with Butadiene Rubber (SBR) compound to improve mechanical and physical properties for Tread tire application.

Secondary goals are to reduce fly ash waste to Environment and reduce cost in rubber compounds.

2. METHODOLOGY

2.1. Raw material, Chemical agents, and Equipment

Raw material fly ash was collected from Sailun cart tire's power plant located in Svay Rieng Province, Cambodia. Caborn black N330 was purchased from YChem Solutions Co., Ltd in Thailand. Zinc oxide (ZnO) and Sulfur were purchased from Xilong Scientific Co., Ltd. In addition to this, Thermo Fisher Scientific (Heysham) imported IPPD, whereas CBS was initially manufactured in China and supplied by Unibond (Danyang) Advanced, Rubber Co., Ltd. Merck KGaA of Germany provided the stearic acid.

In rubber compound, two roll mill was used to mix between raw rubber, fly ash, and other chemical reagents. Studies on cure characteristic, Moving Dia Rheometer was utilized and vulcanization or forming was carried out with a hydraulic hot press. Mechanical properties were conducted using durometer for hardness testing, oven for compression set and universal testing machine for testing on tensile strength and elongation at break.

2.2. Frame work

In the experiment divided 4 main stages: stage-1: sample preparation/characterization, stage-2: compounding press, stage-3: curing (cure characteristic and vulcanization) and stage-4: mechanical properties were studies for evaluation on hardness, tensile strength, elongation, and compression set. Flow chart of experiment was shown in **figure 2.2.1**

[1]* Corresponding author: Phanny Yos
Email: phannyyos@itc.edu.kh *; Tel: +855-12 870 256*

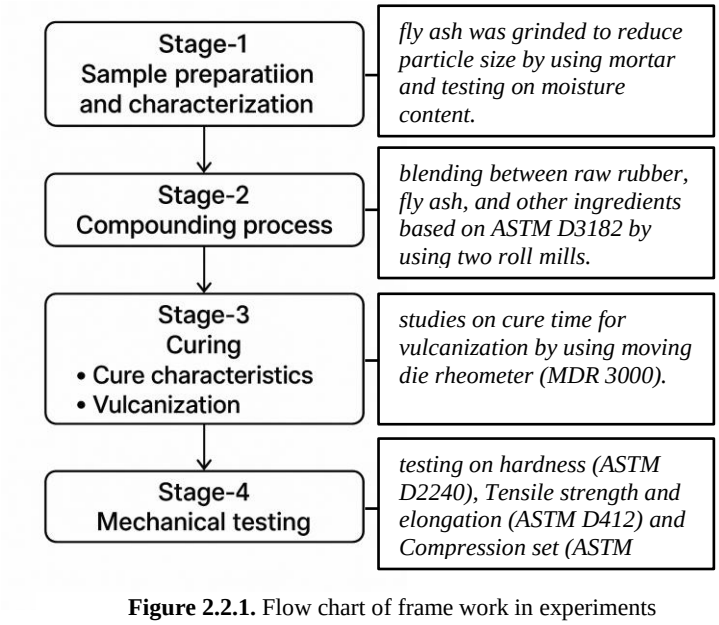


Figure 2.2.1. Flow chart of frame work in experiments

2.3. Experimental

2.3.1. Sample preparation and characterization

Fly ash was dried in oven at temperature 105 ± 5°C for 24 hours. After fly ash was grinded to reduce particle size by using mortar 75µm. In addition to this testing on characterization, moisture content was evaluated by using oven at temperature 105 ± 5°C for 2 hours. Then, place it in a desiccator for 15 minutes to prevent moisture absorption and weigh it to determine the mass after dehydration.

2.3.2. Compounding process

A two-roll mill consists of two horizontally placed hollow metal rolls rotating towards each other. The distance between the mill rolls can be varied and this gap is known as nip. The speed difference between the rolls is called friction ratio and it allows the shearing action. The back roll moves in friction ratio is 2:1:6. The first step in the rubber processing process is to set the mill to a 2 mm opening and run the rubber through the rolls twice without banding. The rubber is then banded while three-quarter cuts are made from either side of the mill, which has been adjusted to a 1 mm opening. Zinc oxide (ZnO) is next added, and then there are more three-quarter incisions. The same cutting technique is then used to add stearic acid. The mill pan is swept to include all of the pigment in the batch after the accelerator is properly applied to ensure little loss. Three-quarter cuts are again made. A subsequent round of three-quarter cuts is made after adding sulfur and any debris that has fallen into the mill pan. Following its removal from the mill, the rolled stock is run through the mill six times in an endwise direction. In the end Following its removal from the mill, the rolled stock is run through the mill six times in an endwise direction. In the end, the mill is set up to permit a minimum stock thickness of 6 mm. The stock is then cycled through the rolls four times, folding every time.

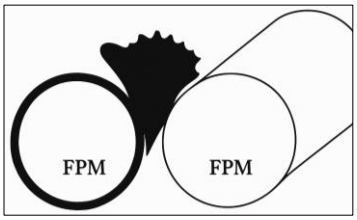


Fig 2.3.2.1: Two roll mills

Table 2.3.2.1: Rubber formulation for tire treads application

Ingredients	(CB/FA) in ratio (pphr)	Times (mn)
SBR	100	7
ZnO	3	1
Stearic Acid	2	1
Carbon Black	0/0/10/20/30/40/50	5
Fly Ash	0/50/40/30/20/10/0	5
IPPD	1	3
CBS	1.6	2
Sulfur	1.4	6

2.3.3. Curing

After compounding process, rubber samples are typically stored for 24 hours in fridge with temperature 4 ± 2°C to stabilize material. Then rubber was kept at room temperature 23 ± 2°C for at least 1 to 2 hours before testing on cure characteristics. Cure characteristic is defining cure time for vulcanization by using Moving Die Rheometer (MDR 3000). The rubber was catted around 5 g put it into MDR based on ASTM D5289 [7]. When the cure time study was performed. Vulcanization was started with pressure 1000 psi by using hydraulic hot press to sharp other specimens for mechanical testing.

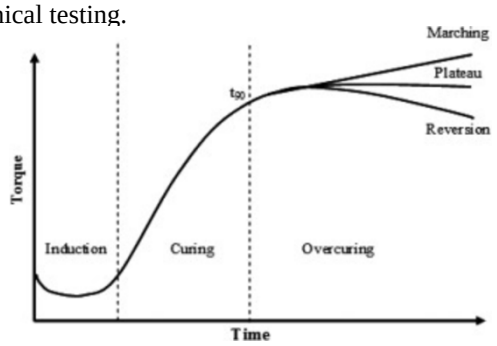


Figure 2.3.3.1: Cure time analysis in moving die rheometer

2.3.4. Hardness

Rubber hardness is the measure of a rubber material's resistance to surface deformation or indentation under load. Rubber hardness can be measured via durometer or Shore Hardness, which assesses the rubber's resistance to indentation, which is crucial for thin or small rubber samples. A durometer, also known as the "Shore hardness," determines the hardness value. This device applies a specific force to the material's surface and measures the depth of the indentation. The Shore scale is divided into categories, such as Shore A and Shore D, to suit different material types. Shore A is used for softer rubbers, like elastomers. while Shore D is used for harder, more rigid materials (2025). Range of Shore hardness is between 0-100. Rubber's hardness was tested based on ASTM D2240 [8].

- Scope: durometer is used to teste on indentation
- Procedure: specimen's dimension, cylindrical disc with thickness ≥ 6 mm and diameter ≥ 12 mm. Condition humidity is 50 ± 5% and temperature is 23 ± 2°C. Indentation depth of durometer is 0-2.5 mm

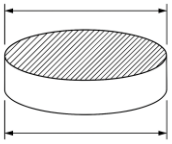


Fig. 2.3.4.1: Cylindrical specimen

2.3.5. Tensile strength and elongation

Rubber tensile strength refers to the maximum tensile load a rubber material can withstand before breaking, making it a crucial sign of toughness, flexibility, and mechanical qualities. In industrial applications, it is essential for maintaining the durability and deformation resistance of rubber components under stress. Elongation at break, tensile modulus, polymer strength, and stress-strain behavior analysis using machine. Dumbbell specimen is a common method. Tensile strength and elongation were tested based on ASTM D412 [15].

- Scope: universal testing machine with speed 500 ± 50 mm/min
- Procedure: specimen's dimension, dumbbell with overall length 155 mm, narrow section 33 mm, gauge length 25 mm, gauge width 6 ± 0.3 mm, and thickness 2 ± 0.2 mm. Condition humidity is 50 ± 5% and temperature is 23 ± 2°C.

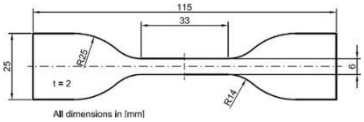


Fig. 2.3.5.1. Dumbbell specimen

2.3.6. Compression set

Compression set is defined as the permanent deformation remaining in a rubber specimen after it has been subjected to a compressive force for a specified time and temperature. Rubber was tested on compression set based on ASTM D395 [9].

Formulation to calculation compression set

$$\text{Compression set } \% = \frac{t_0 - t_f}{t_0 - t_s} \times 100 \quad \text{(Eq 1)}$$

Where:

- t_0 original thickness of specimen
- t_f: final thickness after recovery
- t_s: thickness of the spacer used (compressed thickness)

3. RESULTS AND DISCUSSION

3.1. Cure Characteristic

Studies on cure characteristics are necessary to define cure time for vulcanization process. In the experiment, using differences filler it can affect on cure time due to mechanical properties. In the rubber compound, Fly ash and filler were added (CB/FA) in ratio. The result in **table 3.1.1**, was shown that increasing of amount carbon black decrease the cure time. The compound with 0 pphr Carbon Black and 50 pphr Fly Ash (CB0FA50) has the highest cure time of 11.47 minutes compared to the other compounds. For compound the containing only 50 pphr carbon black without fly ash (CB50FA0) has the lowest cure time of 7.04 minutes. At this point it shown that when using carbon black more than fly ash in rubber compound leads to faster vulcanization because surface activities of carbon black, combined with its thermal conductivity, enhances heat transfer during compounding due to the friction between the two roll-mills and the rubber's rubber.

Table 3.1.1: Cure time for vulcanization

Compounds ratio (CB/FA)	Cure time (mn)
CB0FA50	11.47
CB10FA40	9.67
CB20FA30	8.83
CB30FA20	8.16
CB40FA10	7.48
CB50FA0	7.04

3.2. Hardness

In the experiment, hardness was conducted. In **Fg. 3.2.1**, was shown that using carbon black and fly ash in ratio for all compounds increase hardness. In rubber compounds, increasing of amount carbon black and decreasing of amount fly ash in each compound leads to higher hardness due to reinforcing nature of carbon black particle (below than 0.1µm). it is well known that the lower particle size, the higher surface area. High surface area makes more interconnected filler network that restricts the movement of the polymer chain. In the study, fly ash with a particle size of approximately 75 µm was used as the filler. However, fly ash continued to be used in treads tire application. In commercial tire tread, hardness is 60 ± 5 shore-A. As a results, the compounds containing 20, 30, 40, and 50 pphr of carbon black and 30, 20, 10, and 0 pphr of fly ash (55.65 vs 58.4 vs 61.3 vs 62.68 shore-A, $p \leq 0.05$), respectively, were within the standard range.

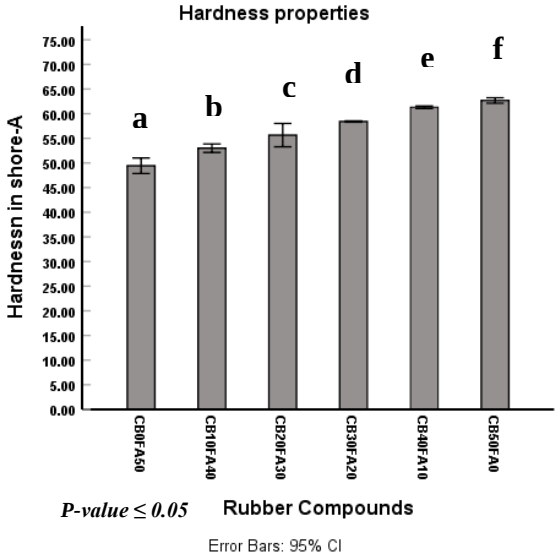


Fig. 3.2.1: Hardness testing (CB/FA) in ratio

3.3. Tensile strength

Tensile strength was tested to evaluate tire treads, because it is an important parameter for assessing durability and performance of the tire under stress, tensile strength is related to hardness when increasing hardness, tensile strength also increases. In commercial tire tread, tensile strength is 15 Min. Pa. As a results in **Fg.3.3.1**, the compounds containing 20, 30, 40, and 50 pphr for carbon black and 30, 20, 10, and 0 pphr of fly ash (15.42 vs 21.39 vs 23.18 vs 25.54 Mpa, $p \leq 0.05$), respectively, were within the standard range. For the compound contains less than 20 pphr of carbon black, the tensile strength falls outside the standard range application. At this point, it was shown that carbon black is an indispensable ingredient in tire industry.

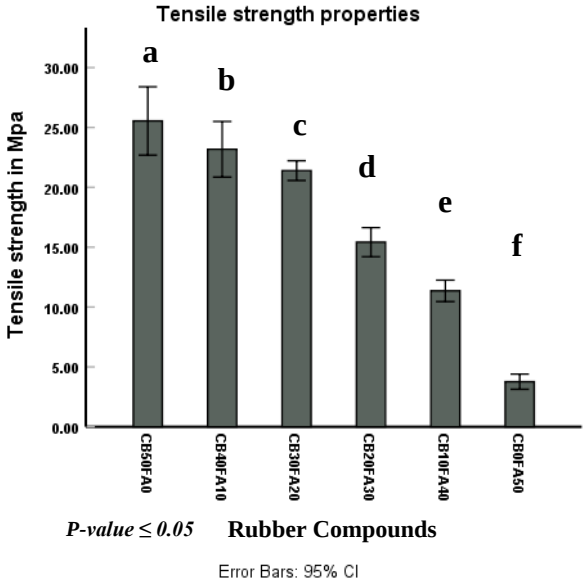


Fig. 3.3.1: Tensile strength testing (CB/FA) in ratio

3.4. Elongation

In the **Fig.3.4.1**, elongation result was shown that the compound CB10FA40 had the highest elongation is 835.89%, indicating excellent flexibility. As carbon black content increased and stearic acid decreased, elongation values generally dropped. The compound CB50FA0 had a much lower elongation 565.05%, showing reduced flexibility due to higher stiffness. The lowest value 594.63% was seen in CB0FA50, which lacked both additives. These results suggest that a balanced ratio of carbon black and fly ash enhances elasticity, while too much carbon black reduces it.

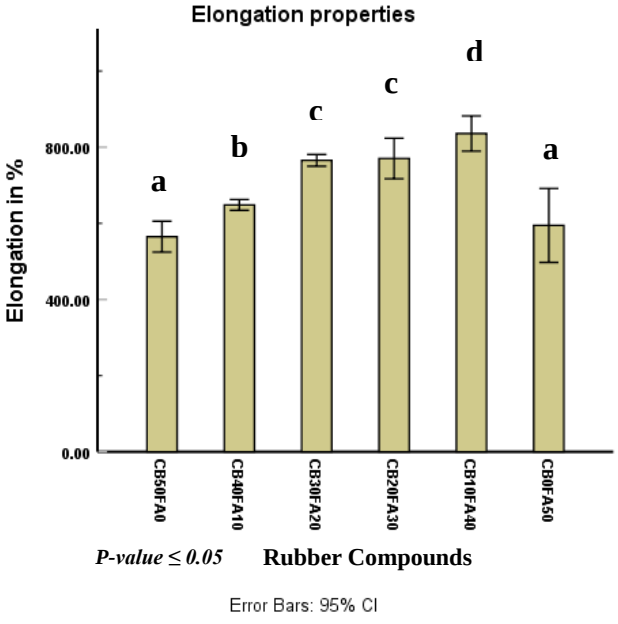


Fig. 3.4.1: Elongation testing (CB/FA) in ratio

3.5. Compression set

In the **Fig.3.5.1**, compression set was observed that compounds CB10FA40 to CB50FA0 meet the standard requirement (≤ 20%). Indicating good resistance to permanent deformation crucial for tire durability. CB0FA50 exceed the limit with values of 21.94%. this tread confirms that high carbon black reinforces the rubber and improve the compound's ability to resist permanent deformation under

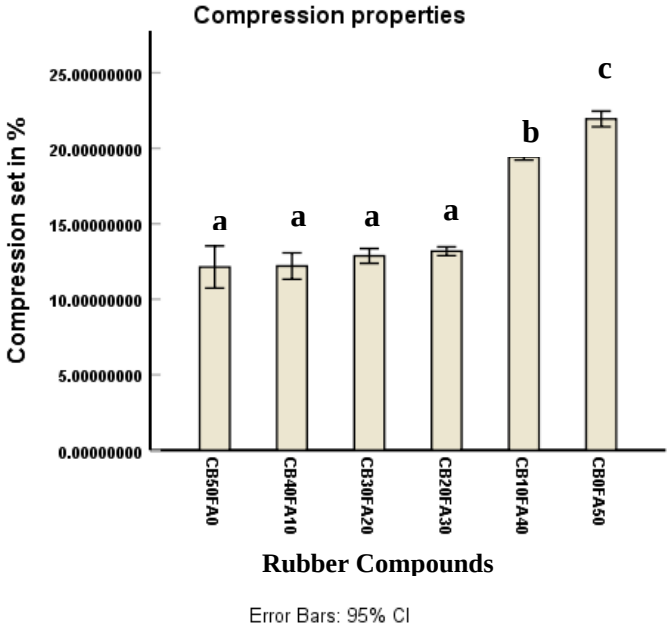


Fig. 3.5.1: Compression set testing (CB/FA) in ratio

4. CONCLUSION

Through this case study was shown that the compounds containing 20, 30, 40, and 50 pphr for carbon black and 30, 20, 10, and 0 pphr of fly ash, respectively were within the standard range for tire treads application. But the compound CB20FA30 is considered the optimal formulation with 55.65 shore-A hardness, 15.41 Mpa tensile strength, and 13.18% compression set. Theses parameters meet the standard requirement (60 ± 5 shore-A for hardness, Min. 15 Mpa for tensile strength, and below than 20% for compression sets). Therefore, compound containing 20 pphr carbon black and 30 pphr fly ash (CB20FA30) is the best optimization for tire tread application, because it reduces material costs and helps decrease fly ash waste in the environment. However, to ensure tire performance, safety, and quality, further research in needed to optimize the formulation, long term durability, and comprehensive evaluations including abrasion, tear resistance, rebound resilience, are essential to fully validate this compound for commercial use in tire.

ACKNOWLEDGMENTS

The work was supported by HEIP Project #12 Cambodia Natural Rubber. In addition, the research is partially supported by the Faculty of Food and Chemical Engineering and Faculty of Geo-resources and Geotechnical Engineering at the Institute of Technology of Cambodia.

REFERENCES

[1] A. Benazzouk, O. Douzane, K. Mezreb, and M. Queneudec, "Physico-mechanical properties of aerated cement composites containing shredded rubber waste," *Cement and Concrete Composites*, vol. 24, no. 4, pp. 345–354, 2002.

[2] Baranwal, K. C., Benarey, H. A., and Ohm, R. G., *Rubber Compounding: Chemistry and Applications*, 1st ed. New York, NY, USA: Marcel Dekker, 2001.

[3] European Automobile Manufacturers Association, *The Automobile Industry Pocket Guide 2023*. Brussels, Belgium: ACEA, 2023. https://www.acea.auto/publication/automobile-industry-pocket-guide-2023

[4] International Organization of Motor Vehicle Manufacturers, *Global Vehicle Sales Statistics 2005–2018*. Paris, France: OICA, 2018b. http://www.oica.net/category/sales-statistics

[5] ISO 23529:2016, *Rubber — General procedures for preparing and conditioning test pieces for physical test methods*, International Organization for Standardization, Geneva, Switzerland, 2016.

[6] K. C. Baranwal, H. A. Benarey, and R. G. Ohm, *Rubber Compounding: Chemistry and Applications*, 1st ed. New York, NY, USA: Marcel Dekker, 2001.

[7] ASTM D5289-21, *Standard Test Method for Rubber Property—Vulcanization Using Moving Die Rheometer (MDR)*, ASTM International, West Conshohocken, PA, USA, 2021.

[8] ASTM International, *ASTM D2240-15: Standard Test Method for Rubber Property—Durometer Hardness*, West Conshohocken, PA, USA: ASTM International, 2023. https://www.astm.org/d2240-15.html

[9] ASTM International, *ASTM D395-18: Standard Test Methods for Rubber Property—Compression Set*, West Conshohocken, PA, USA: ASTM International, 2023. https://www.astm.org/d0395-18.html

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Development of low-carbon concrete using biomass burning ash for port and harbor facilities

Sovannarith Chea [1*], Yamaguchi Toshinobu [2], Akira Yoshikazu [2], Kimura Yokinobu [2], Kawakami Takashi[2]

[1] *Graduate School of Science and Engineering, Engineering Major, Ocean Civil Engineering Program, Kagoshima University, Korimoto 1-21-24, Kagoshima City, Kagoshima Prefecture, Japan*
[2] *Department of Ocean Civil Engineering, Kagoshima University, Korimoto 1-21-24, Kagoshima City, Kagoshima Prefecture, Japan*

[*]*Corresponding author: sovan-narith@yoneg-net.co.jp, sovannarith.edu@gmail.com*

Abstract: *The Japanese government aims for carbon neutrality by 2050 and a 46% reduction in greenhouse gas emissions by 2030 (compared to 2013 levels). With thermal power plants contributing ~40% of annual CO_2 emissions, efforts are underway to adopt renewable energy and improve thermal power technologies. This research focuses on creating low-carbon concrete by reducing natural fine aggregate and cement use, lowering CO_2 emissions, and reusing biomass power plant waste. The study examines concrete with bottom ash and fly ash admixtures compared to traditional concrete, assessing strength, wave-dissipating block performance, CO_2 reduction effect, and economic feasibility. Goals include achieving compressive strengths of over 18 N/mm² at 28 days and 3.7 N/mm² at 3 days, suitable for wave-dissipating blocks for workability and durability. Develop admixture concrete using bottom ash and fly ash, compare its strength with industry-standard concrete, and evaluate the performance and durability of wave-dissipation blocks made from both materials through ocean exposure and drop test. In this study, We investigated the production and quality of wave-dissipating blocks using environmentally friendly concrete that effectively utilizes biomass combustion ash. As a result, it is possible to apply it to wave-dissipating blocks of concrete using concrete using biomass ash. The concrete using biomass ash used in this study can reduce CO_2 emissions by up to about 32% compared to concrete using ordinary Portland cement.*

Keywords: biomass ash, wave-dissipating block, low-carbon concrete, CO_2 reduction

1. INTRODUCTION

The Japanese government aims for carbon neutrality by 2050 and a 46% reduction in greenhouse gas emissions by 2030 (compared to 2013 levels). With thermal power plants contributing ~40% of annual CO2 emissions, efforts are underway to adopt renewable energy and improve thermal power technologies [1,2]. This research focuses on creating low-carbon concrete by reducing natural fine aggregate and cement use, lowering CO2 emissions, and reusing biomass power plant waste [3]. The study examines concrete with bottom ash and fly ash admixtures compared to traditional concrete, assessing strength, wave-dissipating block performance, CO2 reduction effect, and economic feasibility. Goals include achieving compressive strengths of over 18 N/mm² at 28 days and 3.7 N/mm² at 3 days, suitable for wave-dissipating blocks for workability and durability. Develop admixture concrete using bottom ash and fly ash, compare its strength with industry-standard concrete, and evaluate the performance and durability of wave-dissipation blocks made from both materials through ocean exposure and drop test.

2. Material properties

2.1 Woody biomass combustion ash and bottom ash psychical properties

The physical properties of woody biomass fly ash and bottom ash were measured using Japanese standards. For fly ash (FA), JIS A 6201 and JIS R 5201 were used, focusing on density measurement with a Le Chatelier flask and kerosene or light oil as per JIS standards. The density was found to be 2.26 g/cm³. For bottom ash (BA), JIS A 1109 and JIS A 1102 were applied, assessing density, water absorption, and fine modulus. The saturated surface density (Ds) was 2.55 g/cm³,

[1*] Corresponding author: Sovannarith CHEA
E-mail: sovan-narith@yoneg-net.co.jp; Tel: +81-90 2715 9475

dry density (Dd) 2.45 g/cm³, and water absorption rate 2.00%. The fine modulus from the sieve analysis was 2.84, highlighting proper particle gradation for improved concrete properties. [4]

2.2 Evaluation of the effect of admixture of woody biomass combustion ash on mortar quality

To confirm the effect of biomass (BM) ash mixing on mortar quality, flow value tests and compressive strength tests were carried out in accordance with the mix design value of concrete on a total of 7 levels of mortar, including blast-furnace slag cement type B (C) as a cement, 3 types of mortar with mass replacement to BM-FA in 0%, 30%, and 50% of the cement, and three types of water-binder ratios (W/B) of 50%, 60% and 70%. Fine aggregate was used BM-BA. Table 1 shows material properties of mortar. In addition, the mortar specimens were cured in 20˚C water for 3, 7 and 28 days. The slump flow test of mortar is conducted to assess the workability and flowability of the mortar. For this case, the mixture proportions of the mortar were transformed from the mix design value of the concrete successfully made for this research. It was found that it was necessary to increase the admixture to obtain the same flow value in mortar using BM-FA. The compressive strength is presented in Fig.1. However, when comparing the compressive strength of W/B60% BM-FA0% and W/B50% BM-FA30% and also between W/B70% BM-FA0% and W/B60% BM-FA30%, which had roughly the same compressive strength at 28 days of age, in both cases, it can be seen that the amount of unit cement in concrete mixed with BM-FA decreased by about 100 kg/m³. Therefore, the strength of BM-FA is increased by the pozzolanic reaction, and it is considered that it is effective as a binder.[5]

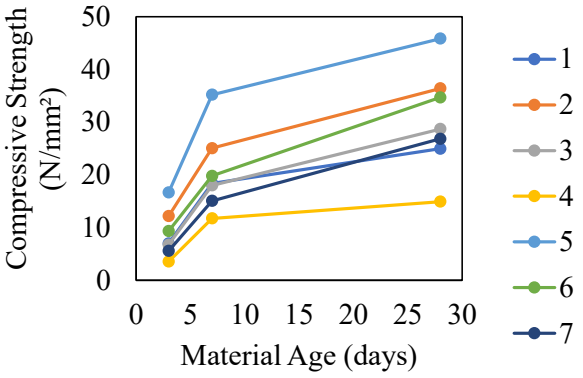


Fig. 1. Mortar compressive strength

2.3 Evaluation of the effect of admixture of woody biomass combustion ash on concrete quality

To evaluate the effects of BM ash mixing on concrete quality, various tests (slump test, air volume measurement, chemical admixture quantification, concrete temperature and compressive strength) were conducted. Starting with a water-to-binder (W/B) ratio of 50%, 60% and 70%, BM-BA was used to replace fine aggregates, while BM-FA was replaced with the amount of 0%, 30% and 50% in cement for binder. Concrete workability was considered crucial for the upcoming wave-dissipating block construction, it was set with target values of 12cm ±2.5cm for slump and 4.5% ±1.5% for air volume. Table 3 shows materials of concrete, and table 4 shows mixture proportions of concrete. The water content, sand-to-aggregate ratio (s/a), and chemical admixture content were all adjusted during concrete production to achieve the desired workability.

Table 1. Mortar material properties

Material	Types	Symbol	Density
Cement	Blast-furnace cement	C	3.04
	Biomass Fly ash	FA	2.26
Fine aggregate	Biomass Bottom ash	BA	2.45
Chemical admixture	AE water reducing	ad	1.11
Water	Industrial water	W	1.00

Table 2. Mortar mixture proportions

№	W/B (%)	FA (%) Sub.rate	Unit (kg/m³)				ad
			C	FA	W	BA	
1	70	0	501	0	350	1188	C×1.5%
2	60	0	559	0	335	1180	C×1.0%
3	60	30	390	168	335	1133	C×2.0%
4	60	50	279	279	335	1103	C×2.0%
5	50	0	655	0	328	1119	C×1.0%
6	50	30	459	197	328	1064	C×1.0%
7	50	50	328	328	328	1028	C×2.0%

Table 3. Concrete material properties

Material	Types	Symbol	Density
Cement	Blast-furnace cement	C	3.04
	Biomass Fly ash	FA	2.26
Fine aggregate	Biomass Bottom ash	BA	2.45
Coarse aggregate	Crushed stone 2005	G	2.71
Chemical admixture	AE water reducing	ad	1.11
Water	Industrial water	W	1.00

Table 4. Concrete mixture proportions

№	W/B (%)	S/a (%)	FA (%) Sub.rate	Unit(kg/m³)					ad
				C	FA	W	BA	G	
1	70	41.9	0	286	-	200	678	1041	C×1.5%
2	60	40.4	0	312	-	187	659	1075	C×1.0%
3	60	39.5	30	218	94	187	633	1075	C×2.0%
4	60	38.8	50	156	156	187	616	1075	C×2.0%
5	50	39.2	0	366	-	183	625	1075	C×1.0%
6	50	37.3	30	256	110	183	594	1075	C×1.0%
7	50	37.1	50	183	183	183	574	1075	C×2.0%

Fig.2 presents the measured slump values and air volume results that are in between the required range. Even in the case of concrete, the amount of admixture tended to increase when BM-FA was used. The compressive strength results are presented in Fig.3. The compressive strength required for making the wave-dissipating block is based on achieving a 3-day compressive strength of over 3.7 N/mm² for workability and 28-day compressive strength of over 18 N/mm² for durability. The strength of concrete mixed with BM-FA, as well as mortar, decreased. However, since all the mixture proportions meet the requirements, W/B60% BM-FA50% in No. 4 was selected due to its higher quantities of fly ash and bottom ash compared to the others.

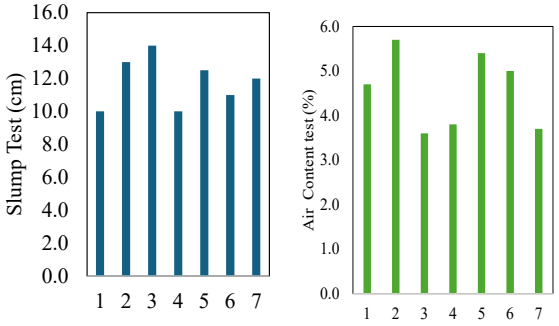


Fig. 2. Concrete Slump value and Air content results

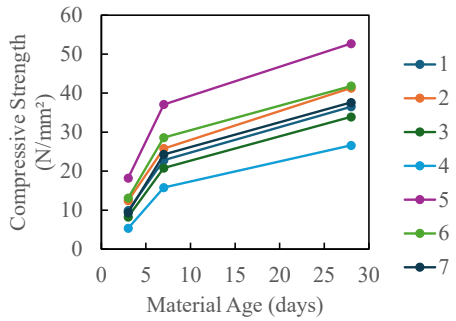


Fig. 3. Concrete compressive strength

3. Wave-dissipating block

3.1 On-site construction test

Two types of concrete, ordinary concrete (OC) and biomass-cement concrete (BmA), were produced at a plant to create wave-dissipating blocks for comparison. Both were designed to achieve a compressive strength of 18 N/mm² at 28 days and over 3.7 N/mm² for demolding after 3 days. The BmA had a W/B of 60% and 50% BM-FA replacement, while the OC had a nominal strength of 18 N/mm² and compound strength of 25.5 N/mm². Test manufacturing evaluated workability, strength development, and other properties, identifying quality issues and proposing solutions. Four wave-dissipating blocks were produced on-site. Two were made with OC and two with BmA. Table 5 shows materials of concrete and Table 6 shows mixture proportions of concrete. W/C of OC is 65.6%, and it of BmA is 60%. The mixture proportion of BmA is the same as No.4 in Table 4.

Table 5. Wave-dissipating block concrete mixture proportions

No	W/C (%)	S/a (%)	FA (%) Sub.rate	Unit(kg/m³)								
				C	FA	W	BA	SS	CS	GCS	G	ad
1) OC	65.6	44.1	0	261	-	171	-	393	164	247	1057	2.61
2) BmA	60	38.8	50	156	156	187	616	-	-	-	1075	6.24

Table 6. Wave-dissipating block concrete material properties

Material	Types	Symbol	Density
Cement	Blast-furnace cem	C	3.04
	Biomass Fly ash	FA	2.26
Fine aggregate	Biomass Bottom ash	BA	2.45
	Sea Sand	SS	2.46
	Crushed Sand	CS	2.85
	Galliba Sand	GCS	2.65
Coarse aggrgate	Crushed stone 2005	G	2.71
Chemical admixture	AE High water red.	ad	1.11
Water	Industrial water	W	1.00

Fig.5 presents the temporally changes slump values and air volume measured. Initial slump value was 18cm in BmA and 6.5cm in OC, and initial air volume was 2.5% in BmA and 4.9% in OC. Although it is not possible to make a clear consideration because the initial slump values are different, the amount of slump loss during the same slump is about the same, and BmA is expected to decrease in slump to the same extent as OC. Fig.4 shows the assembly of the formwork and the concrete pouring process. In the experiment, a 2-ton wave-dissipating block was manufactured. Regarding workability, no significant difference was observed between OC and BM-FA.



Fig. 4. Wave-dissipating block preparation and concrete pouring

The appearance after demolding is also shown in the photo. Minor air bubbles and water stains were observed on the legs and the top of the abdomen of both blocks; however, these imperfections were minimal, and both types of concrete exhibited a similar finished shape. Fig.6 shows the change in

compressive strength over time. Although the initial strength of BmA is smaller than that of OC, it was confirmed that the compressive strength of the material at 1 year of age is almost the same as that of OC, and the pozzolanic reaction of BM-FA is expected to gradually progress, resulting in high-quality concrete in the long term.

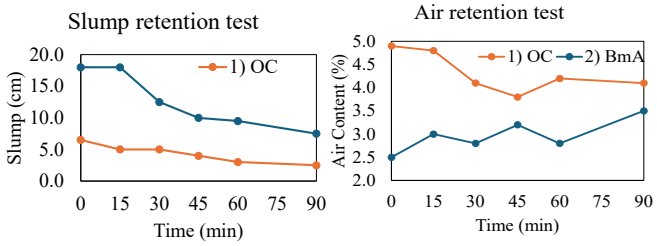


Fig. 5. Wave-dissipating block slump and air retention results

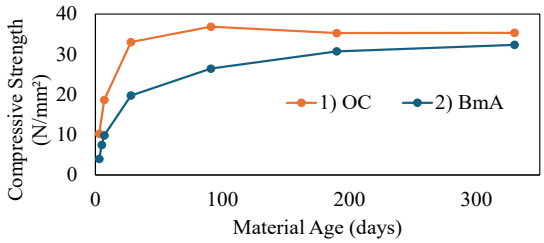


Fig. 6. Wave-dissipating block concrete compressive strength results

3.2 Exposure test under seawater

Wave-dissipating blocks not only protect coastal infrastructure but also serve as artificial habitats, with ordinary concrete offering long-term colonization potential and fly ash concrete promoting early-stage marine biodiversity due to its lower alkalinity, denser microstructure, and sustainability, making it ideal for eco-conscious coastal projects. Composite breakwaters with wave-dissipating blocks are increasingly utilized in deep-sea regions of Japan to withstand severe wave conditions. Fig. 7 shows appearance of wave-dissipating blocks after 1 year exposure. Sessile organisms or biofouling are attached to both wave-dissipating blocks at similar levels. The organisms that can be confirmed at this time are barnacles, Padina arborescens Holmes and Colpomenia sinuosa, and although it is a one-year result, no damage was confirmed to the blocks, and they remained in a healthy state.

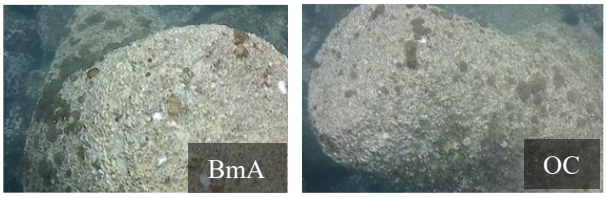


Fig. 7. Wave-dissipating block concrete condition inside the ocean

3.3 Drop test

Concrete blocks are subjected to both static and dynamic loads. Static loads arise from the weight of overlying blocks and the structure itself, while dynamic loads are caused by pulsating wave forces, collisions between blocks, and other external impacts. Among these, impact loads such as collisions involve extremely large impact energy, making it important to evaluate the impact resistance of wave-dissipating block [6]. Therefore, in this study, an impact test was conducted by dropping a block onto one of its legs. The experimental setup is shown in the figure 9. A load cell was installed at the collision point to measure the maximum impact load during the impact. Additionally, a strain gauge was attached to the wave-dissipating block at the position shown in fig.9 to measure deformation during the impact. Both measurements were taken using a dynamic strain meter at 1 ms intervals.



Fig. 8. Wave-dissipating blocks for the drop test



Fig. 9. Drop test experiment and preparation

Furthermore, to observe the failure at the collision point, the impact was recorded using a high-speed camera. For the test method, the wave-dissipating block was lifted to a predetermined height and then subjected to an impact load through free fall. The time-dependent changes in the load applied during the collision are shown in fig.10 and fig.11. It was confirmed that the applied load increased as the drop height increased. OC failed at 378 kN (drop height 40 cm), and BM-FA failed at 329kN (drop height 30cm). The failure mode was compressive failure at the loading point. Since no bending failure occurred in the abdomen area for either specimen, it is inferred that both blocks did not suffer significant damage. The strain tended to increase with the rise

in impact load. Even at a maximum impact load at 300kN, the strain was approximately 66μ, and 60μ for BM-FA. The relationship between tensile stress and tensile strain for both concretes is presented. Considering that the tensile strain limit is estimated to be around 150μ, it was found that both concretes maintain a safety margin of more than twice the limit for bending failure. Based on these results, it is concluded that BM-FA is also suitable for regular use in block production. Fig.12 shows the relationship between the impact load and the tensile strain at the maximum bending position. The load-strain relationship was generally the same regardless of the type of concrete. The maximum strain at fracture was 96μ for OC and 75μ for BmA. Here, Fig. 13 shows the relationship between tensile strength and strain. The stress at the time of strain generated at the time of fracture is 3.32 N/mm^2 for OC and 2.86 N/mm^2 for BmA, and the safety factor until tensile fracture was 1.5 for OC and 2 for BmA.

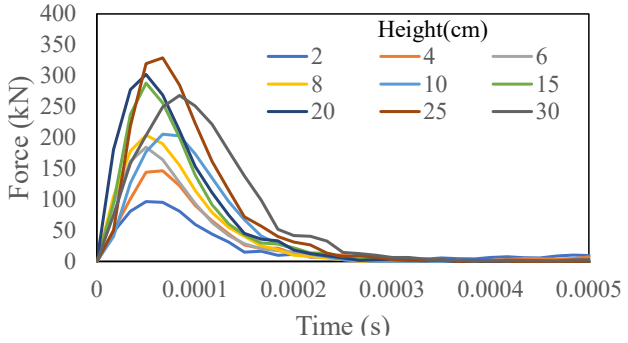


Fig. 10. Impulse depends on high of BmA

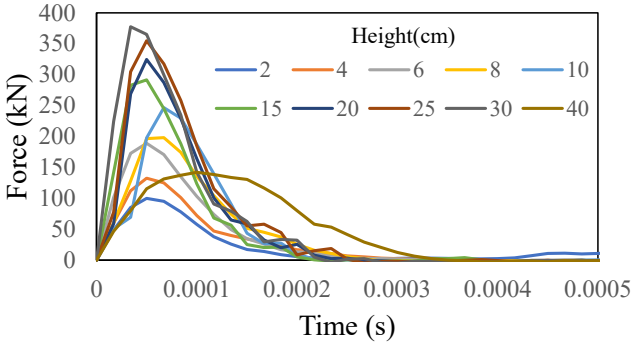


Fig. 11. Impulse depends on high of OC

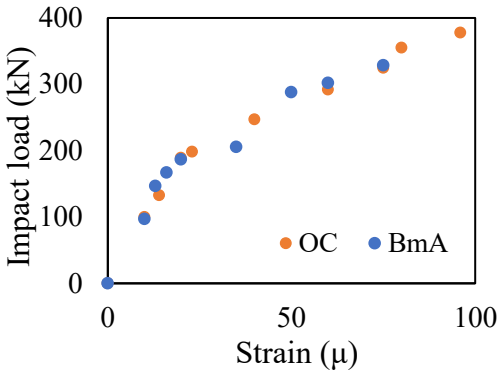


Fig. 12. Impact load of tensile strain

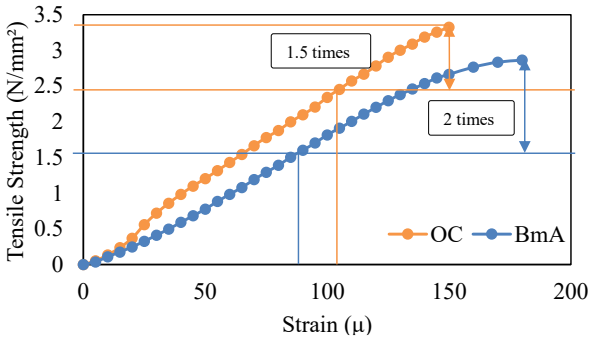


Fig. 13. Young modulus of tensile strength

4. CONCLUSIONS

In this study, the investigated the production and quality of wave-dissipating blocks using environmentally friendly concrete that effectively utilizes biomass combustion ash. The results obtained are shown below.

1. It is possible to apply it to wave-dissipating blocks of concrete using concrete using biomass ash.

2.The concrete using biomass ash used in this study can reduce CO_2 emissions by up to about 32% compared to concrete using ordinary Portland cement. [7]

ACKNOWLEDGMENTS

This thesis represents the culmination of two years of hard work, learning, and growth, made possible by the support of many individuals. I would like to expresses deep gratitude to Professor Akira for his invaluable guidance and mentorship, and to Professor Yamaguchi for his critical help with admission and scholarship support. Special thanks are extended to Prof. Kawakami and Mr. Shiragawa for their

assistance with logistics and experiments, and to fellow classmates for their friendship and support.

Especially thankful to Mr. Naoki Yonemori, whose provides the Yonemori scholarship program, mentorship, and personal support had a transformative impact. Mr. Maniwa is also acknowledged for his crucial role in facilitating research coordination and communication. Appreciation is given to the employees of Nikken Material for their support and collaboration during part-time work and research activities.

Finally, I would like to thank to my family and friends for unwavering emotional support, and Kagoshima University for providing an excellent academic environment.

To everyone who has contributed to this journey in any way—thank you from the bottom of my heart

REFERENCES

[1] Purton, Michael. Cement is a big problem for the environment. Here's how to make it more sustainable. World Economic Forum. [Online] September 13, 2024. https://www.weforum.org/stories/2024/09/cement-production-sustainable-concrete-co2-emissions/#:~:text=It's%20been%20said%20that%20if,the%20world's%20total%20CO2%20emissions.

[2] 国土交通省総合政策局政策課. 国土交通白書. 令和 6 年版国土交通白書. [Online] 国土交通省, 2024. https://www.mlit.go.jp/hakusyo/mlit/r03/hakusho/r04/html/n1111000.html.

[3] CO2 メディア. 日本の二酸化炭素排出の原因とは？排出量を部門別にランキング. RVSTA. [Online] RVSTA, September 04, 2023. https://www.tansomiru.jp/media/basic/mag_1031/

[4] 土木学会コンクリート委員会コンクリート標準示方書改訂小委員会, コンクリート標準示方書　施工編, Tokyo: 公益社団法人 土木学会, September 2023.

[5] A. Ahmed, "Chemical Reactions in Pozzolanic Concrete," Mod. Approaches Mater. Sci., vol. 1, no. 4, pp. 128–133, 2019, doi: 10.32474/mams.2019.01.000120.

[6] Mitsui, Jun, Masato Yamamoto, Satoshi Noboru, and Ichiro Nishiwaki. 2011. "Impact Force Analysis on Wave Dissipating Concrete Blocks During Rocking Motion." Coastal Engineering Proceedings (32): 44.

[7] "低炭素型コンクリートの普及促進に向けて.pdf," 東京, 2020. [Online]. Available: https://www.nikkenren.com/sougou/10thaniv/pdf/05-06-20.pdf.

Materials Science and Structure (MSS)

2 Sub-parallel Sessions

"MSS-2: Structure and Infrastructure Engineering"

Chair: Dr. DOUNG Piseth, Co-chair: Dr. CHHUN Kean Thai

No.	Topic
1	Resilient Modulus of Lime-treated Materials for Pavement Design in the Cambodian Context *Authored by: Chan BUOY, Kuchvichea KAN, Bertrand FRANÇOIS*
2	Deploying ANPR Technology to Detect Traffic Violation at Intersections in Phnom Penh *Authored by: Sambath KEO, Vengkheang PHUN, Yat YEN*
3	3D Road Surface Reconstruction: A Case Study of Indoor Corridor Using RPLiDAR 360° A2M8 *Authored by: Phearun MAO, Veng Kheang PHUN 1, Narith SAUM*
4	Assessing the Impact of Motorcycle Emissions on Urban Air Quality: A Case Study at the Institute of Technology of Cambodia *Authored by: Lihong PHINY, Veng Kheang PHUN, Narith SAUM, Sokhit PLACK*
5	Correlation Between Dynamic Cone Penetrometer (DCP) and California Bearing Ratio (CBR) for Subgrade Soil Materials in Cambodia *Authored by: Bopheaktra PHO, Narith SAUM, Tetsya SOK, Lyhour CHHAY*
6	Geotechnical Properties of Clay Soil with River Sand for Sustainable Road Embankment Applications *Authored by: Samnang SOK, Mankong YIN, Sarong RI, Kean Thai CHHUN*
7	Machine Learning-Based Prediction of Public Bus Ridership: A Case Study in Phnom Penh City *Authored by: Piseth SRORN, Veng Kheang PHUN, Narith SAUM*
8	Evaluation of Modified Asphalt Binders Incorporating Recycled Plastic Waste for Enhanced Pavement Performance *Authored by: Chy TENG, Kuchvichea KAN, Phanny YOS*
9	Improving the Durability and Strength of Expansive Soils Stabilized with Emulsion Asphalt and Cement *Authored by: Vanntheng YOEUNTHY, Kuchvichea KAN, Bertrand FRANÇOIS, Kean Thai CHHUN*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Resilient modulus of lime-treated materials for pavement design in the Cambodian context

BUOY Chanbo[1*]; KAN Kuchvichea[1]; Bertrand François[2]

[1] *Faculty of Materials and Structural Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *ArGEnCo Département, Université de Liège, Liège, Belgium*

[*] *Corresponding author: buoychanbo@gmail.com*

Abstract: *Cambodia faces challenges in road infrastructure due to weak subgrade conditions and seasonal fluctuations. Lime stabilization is widely used to improve soil strength, particularly in expansive subgrades, by reducing shrinkage and swelling. However, the impact of moisture intrusion and the behavior of lime-treated soils under cyclic loading remains unclear. The subgrade is generally characterized by the Resilient Modulus (MR), determined under laboratory conditions.*

This study evaluates the MR of untreated and lime-treated soil samples using cyclic triaxial testing, repeated load Unconfined Compressive Strength (UCS) testing, and cyclic loading with a California Bearing Ratio (CBR) machine. The triaxial cyclic test, considered the standard method for MR determination due to its ability to simulate realistic in-situ stress conditions, serves as a benchmark for evaluating the accuracy and practicality of the simpler cyclic UCS and CBR-based methods.

Mechanical performance was assessed by comparing MR and permanent strain under each testing method. Analysis of lime-treated soils under simulated in-situ loading provided insights into their long-term behavior and durability. Results indicate that a 3% lime treatment with a 7-day curing period significantly increases MR and reduces permanent strain across all testing methods. A strong correlation was found between MR values from the cyclic CBR and UCS tests with those obtained from the triaxial cyclic test, supporting the feasibility of using more accessible methods for MR estimation in pavement design.

These findings provide a cost-effective and practical approach to characterizing subgrade performance in Cambodia, contributing to more durable and resilient road infrastructure. Further research is recommended to validate field performance and assess the long-term effects of lime stabilization under actual traffic and environmental conditions.

Keywords: Resilient Modulus; lime-treated; cycle loading; CBR test; Unconfined Compressive Strength (UCS) Test; Triaxial test.

1. INTRODUCTION[1]

In Cambodia, subgrades are often composed of expansive clayey soils that are highly sensitive to seasonal moisture variations. The country experiences distinct wet and dry seasons. The rainy season leads to excessive moisture absorption, causing subgrade swelling and a decrease in bearing capacity, while the dry summer season results in moisture loss, shrinkage, and cracking. These moisture fluctuations degrade pavement performance over time, leading to increased maintenance costs and reduced service life.

Expansive soils undergo volume change due to changing moisture conditions. If not treated, these soils cause severe distress to lightly loaded structures like pavements, compound walls, single-story buildings, etc [1]. Lime treatment is often used to treat these soils for pavement applications. Lime basically alters the chemical nature of expansive soils, increases the shear strength, and reduces the plasticity, thereby effectively controlling the swelling behavior [2]. The lime stabilization process essentially is a two-stage process: an immediate exchange reaction and long-term pozzolanic reaction [3]. When lime ($Ca(OH)_2$) is added to clayey soil, it raises the soil pH (above 12), causing cation exchange between calcium ions (Ca^{2+}) from lime and the cations (like Na^+, K^+, H^+) adsorbed on the clay surface, and flocculation and agglomeration of clay particles. This immediate reaction reduces soil plasticity and increases workability [4]. For long-term reaction, lime reacts with siliceous and aluminous materials naturally present in soils to form cementitious compounds like calcium silicate hydrate (CSH) and calcium aluminate hydrate (CAH). These pozzolanic reactions lead to long-term strength improvement in the treated soil [4].

Construction of roads, especially low-volume roads has a low budget, and sophisticated tests of materials are excluded

[1*] Corresponding author: Buoy Chanbo
E-mail: buoychanbo@gmail.com; Tel: +855-87 432 487

from the design process. This statement also concerns asphalt and soil testing. Quality control in laboratories depends on fast and uncomplicated tests that give application to the design process results. Between them most popular soil testing technique is the CBR test, which in many variations has been successfully introduced in many countries over the last decades[6]. Such success results in many different proposition of interpretation of the CBR test results[7]. Traditionally, the CBR test is performed under static conditions, providing an index value that is used empirically in pavement thickness design, such as in the AASHTO 1993 method or British Road Note 31. However, the static CBR test does not account for the effects of repeated traffic loading, which can lead to progressive plastic deformation and failure in subgrade materials over time [8]. To address this limitation, the Repeated Load CBR (RL-CBR) has been developed. This method applies cyclic loading to the specimen to better simulate field conditions and evaluate the permanent and resilient deformation characteristics of subgrade soils. RL-CBR provides insights into accumulated plastic strain, resilient modulus, and fatigue behavior under traffic-like conditions [9], [10].

The Unconfined Compressive Strength (UCS) test is one of the most fundamental laboratory methods used to determine the strength of cohesive soils, particularly in stabilization studies. The test measures the maximum axial stress that a cylindrical soil specimen can withstand under unconfined conditions, meaning without any lateral confinement, until failure occurs [11]. Cyclic loading UCS tests are less standardized than static UCS, but they have been implemented in experimental studies to evaluate how lime-treated soils perform under long-term traffic simulation. These tests often use waveform-controlled loading (e.g., sinusoidal or haversine) at low frequencies and allow researchers to track stiffness degradation and failure over thousands of load cycles [12].

In addition, the Consolidated Undrained (CU) triaxial test is used to provide more comprehensive strength parameters under controlled confining pressures and undrained conditions. This test allows the measurement of both total and effective stresses and is particularly useful in evaluating shear strength parameters (cohesion and friction angle) before and after lime treatment. Moreover, CU tests help assess excess pore pressure development, stress path evolution, and overall improvement in effective strength behavior of lime-stabilized soils, which is critical for long-term pavement stability.

This study investigates the effectiveness of lime stabilization for subgrade improvement in Cambodia by evaluating resilient modulus (MR), strength behavior, and deformation characteristics using cyclic CBR, cyclic UCS, and CU triaxial tests. These tests replicate field stress conditions and provide insight into the mechanical performance of both untreated and treated soils under repeated loading.

2. METHODOLOGY

2.1 Properties of Soil and Compaction Characteristics

The natural Cambodia soil used in this study was characterized following ASTM standards to determine its physical properties. Grain size distribution was determined using sieve analysis in accordance with ASTM D422-63[13], while the finer fraction was analyzed using hydrometer testing. The specific gravity of soil particles was established following ASTM D854-14[14] providing essential data for evaluating compaction and other engineering parameters. To assess the soil's plasticity characteristics, Atterberg limits were measured according to ASTM D4318-17[15], which allowed for classification and evaluation of the soil's consistency behavior.

To evaluate compaction behavior, the Modified Proctor Compaction Test was conducted in accordance with ASTM D1557 [16]. This test provided the optimum moisture content (OMC) and maximum dry density (MDD) of the natural soil, which are critical for preparing test specimens and simulating field compaction conditions. A second Standard Proctor test was then conducted on the lime-treated 3% mixture to determine its modified OMC and MDD. These changes reflect the soil’s altered compaction behavior and increased flocculation due to lime treatment, which can significantly influence the performance of subgrade layers in pavement structures.

Table 1. Properties of the soil

Soil properties	Values
Atterberg limit	
Liquid limit (%)	26.57
Plastic limit (%)	12.07
Plastic index (%)	14.50
Specific gravity	2.650
Grain distribution	
Clay (%)	12.47
Silt (%)	29.09
Sand (%)	58.44
Compaction test characteristics (natural soil)	
Maximum dry unit weight (kN/m3)	2.10
Optimum moisture content (%)	7.87
Compaction test characteristics (natural soil with lime 3%)	
Maximum dry unit weight (kN/m3)	2.02
Optimum moisture content (%)	9.50

2.2 lime material

In this study, 3% hydrated lime by dry weight of soil was selected for stabilization, based on prior studies and preliminary tests. Lime was manually mixed with the soil to ensure uniform distribution. The lime-treated samples were then sealed and cured for 7 days at room temperature to allow for pozzolanic reactions, enhancing soil strength and stiffness for further mechanical testing.

2.3 The Resilient Modulus (MR)

The Resilient Modulus (MR) is defined as the ratio of the applied cyclic deviator stress to the recoverable axial strain during repeated loading. It is mathematically expressed as[5] :

$$M_r = \frac{\sigma_d}{\epsilon_n}$$

Where σ_d is the axial deviatoric stress
ϵ_n is the resilient axial strain
M_r is Resilient Modulus

This relationship is derived from the principles of linear elastic theory applied under cyclic loading, where only the recoverable (elastic) portion of deformation is considered [5]. The resilient modulus thus provides a more realistic measure of a soil's stiffness under conditions simulating repeated traffic loading, unlike static modulus or strength values that do not account for cyclic effects.

The resilient modulus was calculated using the empirical equation developed by Haghighi et al. [17], which estimates the elastic modulus under cyclic loading based on data from the confined repeated load test:

$$E_{equ} = 1.513.(1 - v^{1.104}).\frac{\Delta\sigma_p.r}{\Delta u^{1.012}}$$

Where E_{equ} is the equivalent elastic modulus in MPa, v is the Poisson's ratio, $\Delta\sigma_p$ is the change in plunger stress between the maximum and minimum load in a loading cycle (MPa), r is the plunger radius (typically 24.8 mm), $\Delta u^{1.012}$ is the corresponding change in recoverable displacement during the cycle (mm)

2.3 CBR Testing Procedure

California Bearing Ratio (CBR) tests were carried out on both natural and lime-treated soils to assess the improvement in bearing capacity due to chemical stabilization. All specimens were compacted using the Modified Proctor method in accordance with ASTM D1557. The lime-treated specimens were prepared by mixing the soil with 3%. After curing for 7 days, these specimens were soaked in water for 4 days to simulate field saturation conditions. Untreated soil specimens were also soaked for 4 days, but without any curing. Two sets of CBR tests were conducted for each soil type. The first followed a monotonic loading protocol using standard CBR equipment to measure penetration resistance under static load. The second applied cyclic loading using a modified CBR apparatus to simulate traffic-induced stresses. This approach enabled the evaluation of both immediate and long-term strength behavior of the subgrade materials under different loading conditions.



Fig. 1. CBR machine

2.4 UCS Testing Procedure

Unconfined Compressive Strength (UCS) tests were conducted on both untreated and lime-treated soils using procedures similar to those employed in the CBR tests. Specimens were prepared with 3% lime, cured for 7 days, saturated for 2 days, and tested under monotonic and cyclic axial loading to assess strength improvement due to lime stabilization.



Fig. 2. UCS test

2.5 Triaxial Testing Procedure

To obtain a reliable reference value for the resilient modulus (MR) and to simulate realistic in-situ stress conditions, the cyclic triaxial test was conducted in the Consolidated Undrained (CU) mode. This test was selected due to its ability to apply controlled deviatoric stresses under confining pressure, which replicates the stress path experienced by subgrade soils under traffic loading. The CU configuration allows for full saturation and isotropic consolidation of the specimens prior to the application of cyclic deviator stress under undrained conditions, which is

especially relevant for subgrades subjected to seasonal water table fluctuations.

Lime-treated and untreated cylindrical specimens (diameter: 50 mm; height: 100 mm) were prepared in accordance with ASTM D7181-11, which outlines the procedure for resilient modulus testing of cohesive soils using cyclic triaxial loading. The specimens were saturated under back-pressure conditions to achieve a Skempton B-value greater than 0.95, ensuring full saturation before consolidation. Isotropic consolidation was then applied at confining pressures representative of field conditions (50, 100, and 200kPa).

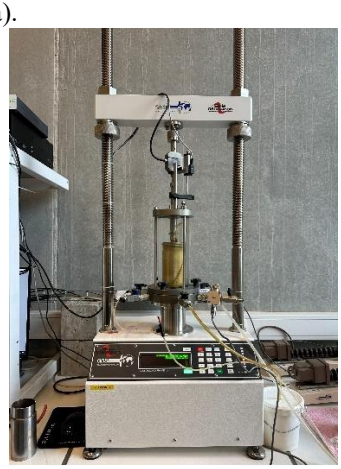


Fig. 3. Triaxial CU test

3. RESULTS AND DISCUSSION

3.1 CBR Untreated Soil and Treated soil

3.1.1 CBR Monotonic Loading

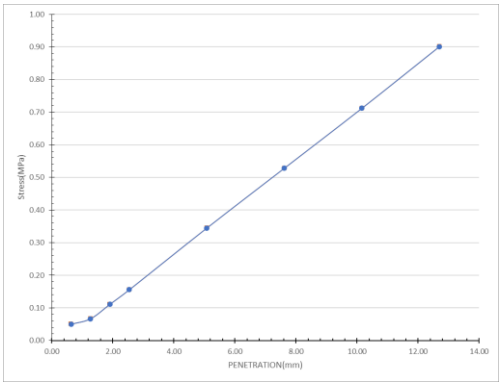


Fig. 4. CBR untreated soil monotonic loading

Fig. 4 presents the monotonic California Bearing Ratio (CBR) test results for untreated soil specimens. The corrected results from the CBR test under monotonic loading for untreated soil show a stress of 0.19 MPa at 2.5 mm penetration and 0.385 MPa at 5.0 mm penetration. These correspond to corrected CBR values of 2.76% and 3.73%, respectively. These low CBR values indicate that the untreated soil exhibits limited resistance to penetration under applied load, reflecting its weak strength and low stiffness characteristics. The nearly linear increase in stress with penetration also suggests a consistent but low load-bearing behavior throughout the loading process. Such mechanical response confirms that, in its natural state, the material is not capable of sustaining significant loads and would likely undergo considerable deformation under applied pressure. This highlights the need for improving its mechanical properties through appropriate treatment.

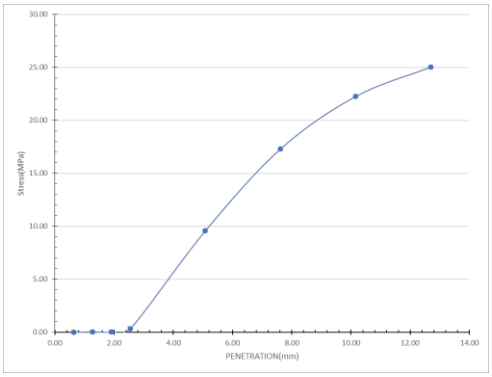


Fig. 5. CBR-treated soil under monotonic loading

The CBR test under monotonic loading for treated soil, as shown in Figure 5, demonstrates a substantial increase in load-bearing capacity compared to untreated soil. Initially, the stress remains low up to around 2.5 mm penetration, likely due to material seating or minor surface adjustment. Beyond this point, the stress increases rapidly and nonlinearly with depth. At 5 mm penetration, the stress reaches approximately 10 MPa, and continues to rise up to 25 MPa at 12.5 mm. This significant increase in resistance under load indicates a marked improvement in the mechanical behavior of the soil following treatment. The curve shape reflects a strong and stiff material response, capable of withstanding much higher loads before failure or excessive deformation occurs.

3.1.2 CBR Repeated loading

Figure 6 shows the stress–displacement behavior of untreated soil under repeated loading in a CBR test. The curve initially rises steadily up to about 0.43 MPa at around 7 mm displacement, indicating soil compaction and plastic deformation during initial loading. Beyond this point, the curve enters a cyclic loading phase, characterized by multiple hysteresis loops, which reflect plastic strain accumulation and energy dissipation during repeated loading. The loops indicate that the soil undergoes progressive deformation, failing to recover after each load cycle. This behavior demonstrates the limited strength and poor resilience of

untreated soil under repeated loads, highlighting its unsuitability for pavement layers without stabilization.

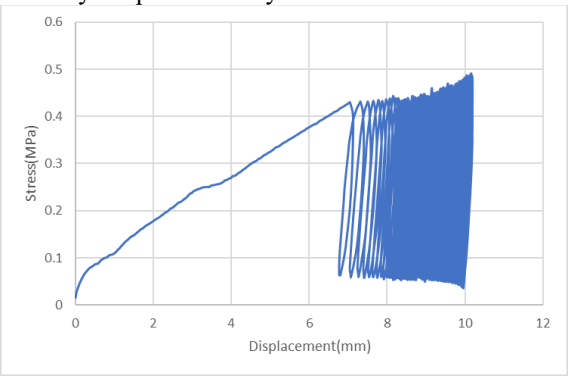


Fig. 6. CBR untreated soil repeated loading

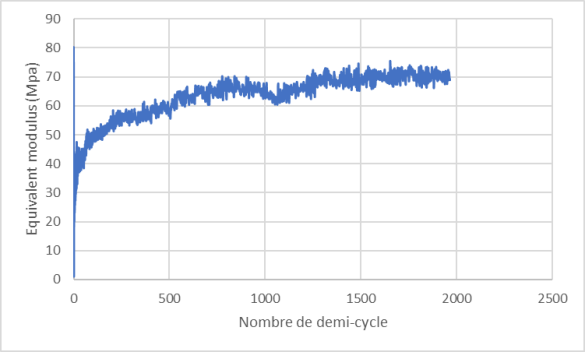


Fig. 7. CBR untreated Equivalent modulus graph

Figure 7 presents the variation of the equivalent resilient modulus of untreated soil with the number of half-cycles under repeated loading. The modulus was calculated using the empirical formula proposed by Haghighi et al. [18]. The curve shows a rapid increase in modulus within the first 200 cycles, indicating early compaction and stiffness gain due to particle rearrangement. After this initial phase, the modulus gradually stabilizes around 70–75 MPa beyond approximately 1500 cycles, suggesting the soil reaches a quasi-resilient state under continued loading. This plateau reflects a steady mechanical response with minimal further change in stiffness. The results confirm that untreated soils, despite low initial stiffness, can develop some level of resilient behavior under repetitive loading. However, the final modulus remains relatively low, highlighting the need for stabilization in pavement applications. These findings are consistent with the staged repeated load CBR methodology outlined in [1].

3.2 UCS Untreated Soil and Treated soil

3.2.1 UCS Monotonic Loading

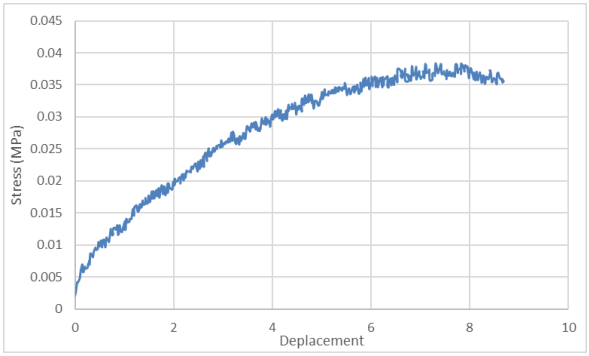


Fig. 8. UCS untreated soil monotonic loading

The results of the unconfined compressive strength (UCS) tests under monotonic loading conditions for both untreated and lime-treated soils are shown in Figures 8 and 9. The untreated soil exhibits a gradual increase in stress with displacement, reaching a maximum stress of approximately 0.037 MPa at a displacement of around 7 mm. This behavior reflects a ductile and low-strength material, with significant deformation prior to failure. The post-peak response shows a slight drop in stress, indicating strain-softening typically observed in weak clayey soils.

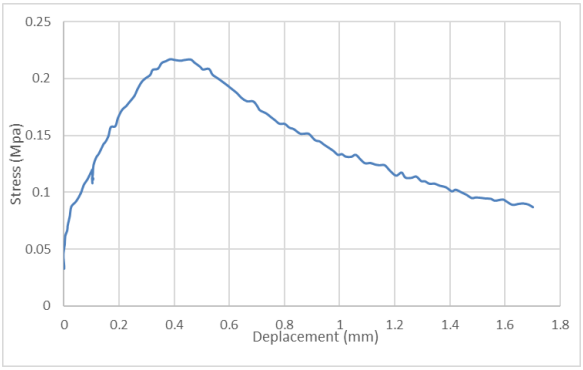


Fig. 9. UCS treated soil monotonic loading

In contrast, the lime-treated soil demonstrates a substantial improvement in strength and stiffness. The treated sample reached a peak stress of approximately 0.22 MPa at only 0.35 mm displacement, representing nearly a sixfold increase in compressive strength compared to the untreated sample. The stress–displacement curve of the treated soil shows a steep initial slope, indicating enhanced stiffness, followed by a pronounced post-peak softening, which is characteristic of brittle failure in cemented soils.

The significant difference in mechanical performance highlights the effectiveness of lime treatment in enhancing soil strength. Due to the very low strength of the untreated soil, it was not possible to perform reliable cyclic loading tests, as the specimen could not sustain repeated loading without failing prematurely. On the other hand, the improved load-bearing capacity of the treated soil allowed for further

cyclic loading tests, enabling the evaluation of its resilient modulus and long-term deformation characteristics.

3.2.2 UCS repeated Loading

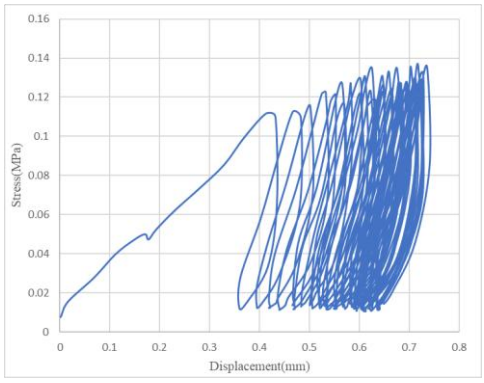


Fig. 10. UCS treated cyclic loading

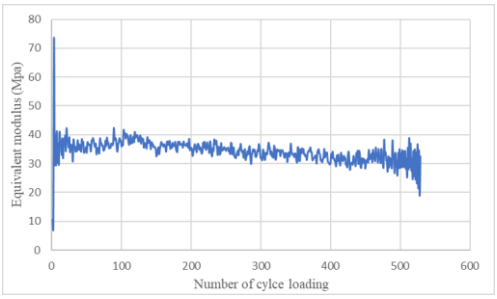


Fig. 11. UCS treated Equivalent modulus graph

The repeated UCS test on lime-treated soil reveals improved performance under cyclic loading. As shown in the Figure 10, the soil initially exhibits a monotonic increase in stress up to approximately 0.11 MPa at 0.4 mm displacement, followed by clear hysteresis loops under repeated loading. These loops indicate progressive plastic deformation and energy dissipation, reflecting the material's ability to withstand repeated stress without immediate failure.

Simultaneously, the equivalent modulus (Figure 11) stabilizes between 30–40 MPa after initial fluctuations and remains relatively consistent throughout most of the test. A slight decline beyond 400 half-cycles and a sharper drop after 500 cycles suggest gradual fatigue. Overall, the results confirm that lime treatment significantly enhances the soil's resilient stiffness and cyclic durability, making it suitable for subgrade applications in pavement design

Table 2. Comparison of MR from CBR and UCS Cyclic Tests

Test type	Soil Condition	Max MR(MPa)	Stabilized MR Range (MPa)
CBR	Untreated	83.3	70-75
UCS	Treated	10.5	30-40

3.3 Triaxial Untreated Soil and Treated soil

3.3.1 Triaxial Monotonic Loading

Stress Path Analysis

In figure 12 and 13 showing the stress path plots in the $p, p'–q, q'$ space reveal clear differences between untreated and lime-treated soils. Untreated soils follow curved, shallow paths, indicating progressive pore pressure build-up and limited shear resistance. In contrast, treated soils exhibit steeper, more linear paths with higher peak q, q' values. These steeper paths reflect enhanced shear strength and reduced contractive behavior due to lime-induced structure and cementation. The regression slopes of peak points confirm a higher effective friction angle for the treated soil.

Stress–Strain and Pore Pressure Response

In figure 14 and 16 show that treated soils display significantly higher peak deviator stresses and initial stiffness compared to untreated soils. At 200 kPa confinement, treated specimens reach over 300 kPa in deviator stress, while untreated samples peak much lower and show post-peak softening. Pore pressure response also differs: untreated soils generate rapid, positive excess pore pressures, while treated soils show slower build-up and, in some cases, negative Δu—indicating dilation and increased particle interlock.

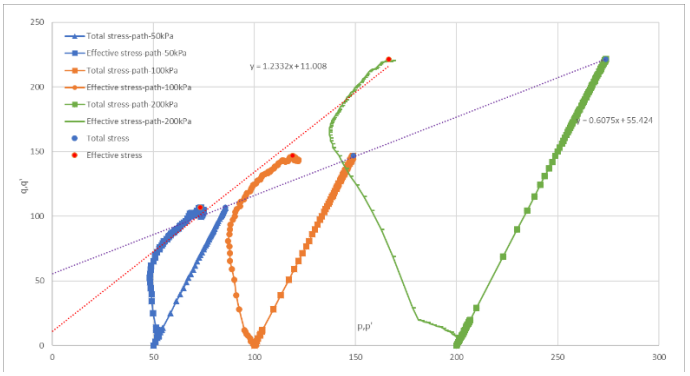


Fig. 12. Graph p, p′–q, q′ of untreated soil

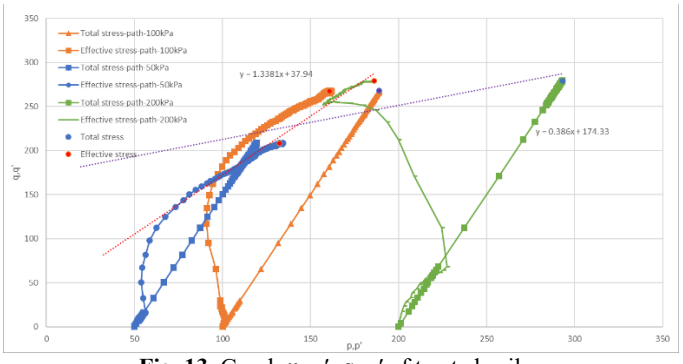
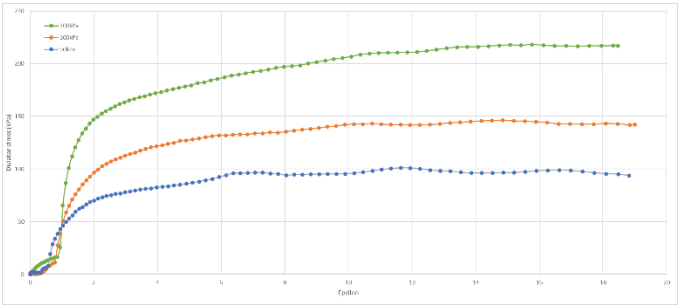


Fig. 13. Graph p, p′–q, q′ of treated soil



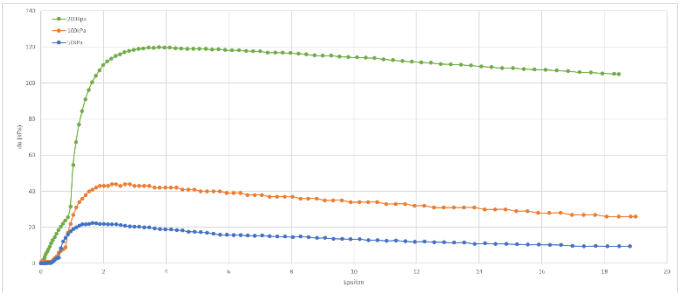
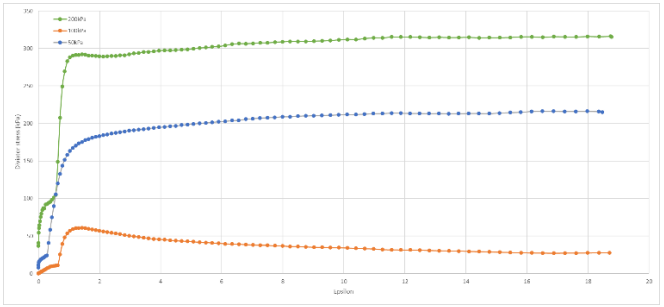


Fig. 14. Stress–Strain and Pore Pressure Response untreated soil



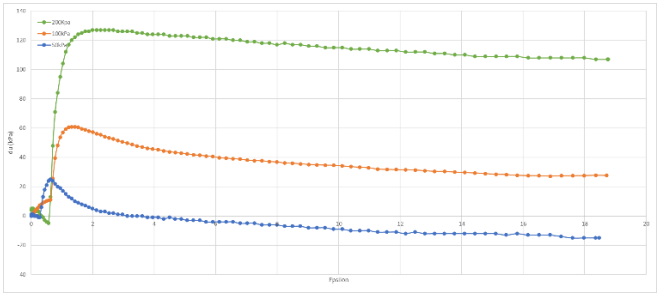


Fig. 16. Stress–Strain and Pore Pressure Response treated soil

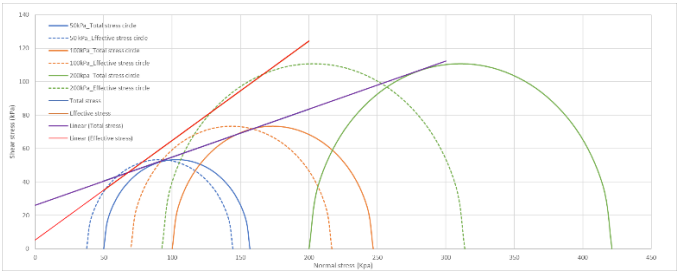


Fig. 15. Mohr's cycle untreated soil

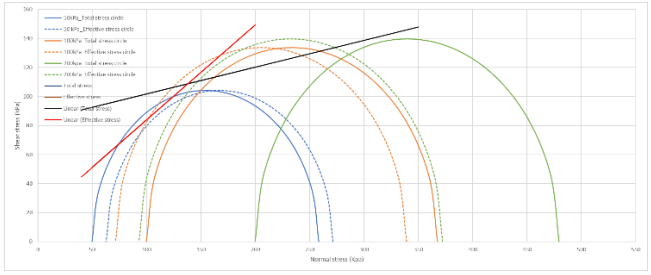


Fig. 17. Mohr's cycle treated soil

Mohr's Circles and Strength Envelopes

Mohr's circles for treated soils demonstrate larger effective stress states and a steeper failure envelope, indicating both increased cohesion and friction angle. The effective stress circles shift left due to pore pressure effects, but treated soil maintains higher strength even in undrained conditions. The untreated soil, by contrast, shows smaller circles and a flatter effective stress envelope, confirming its lower resistance to shear.

Table 3. Variation of cohesion and friction angle of soils from consolidated-undrained triaxial tests

Cambridge	Untreated soil	Treated soil
C(kPa)	26.2	83.3
φ (°)	16.1	10.5
C'(kPa)	5.3	18.5
φ′ (°)	30.8	33.2

4. CONCLUSIONS

This study demonstrates that lime stabilization greatly enhances the mechanical behavior of subgrade soils in Cambodia. A 3% lime content with a 7-day curing period significantly increased the resilient modulus and compressive strength of the soil, while also reducing permanent deformation under repeated loading. Treated soils showed marked improvements in both monotonic and cyclic CBR and UCS tests, and exhibited stronger, more stable responses in triaxial testing, including higher shear strength and more favorable stress paths.

Results from cyclic CBR and UCS tests correlated well with those from the triaxial CU test, confirming their suitability as practical alternatives for estimating resilient modulus in field conditions. Overall, lime treatment offers a reliable, cost-effective method to improve subgrade performance, supporting more durable pavement structures under Cambodia's challenging environmental conditions. Further studies should evaluate the long-term behavior of lime-treated soils under actual traffic loads and varying environmental conditions.

ACKNOWLEDGMENTS

The authors would like to sincerely thank Dr. Kan Kuchvichea and Prof. Bertrand François for their valuable guidance and continuous support throughout this study. Special appreciation is extended to the laboratory teams at the Institute of Technology of Cambodia (ITC) and the Université de Liège for providing the facilities and technical assistance necessary for completing the experimental work.

REFERENCES

[1] S. Bhuvaneshwari, R. G. Robinson, and S. R. Gandhi, "Resilient Modulus of Lime Treated Expansive Soil," *Geotech. Geol. Eng.*, vol. 37, no. 1, pp. 305–315, Jan. 2019, doi: 10.1007/s10706-018-0610-z.

[2] D. N. Little, "EVALUATION OF STRUCTURAL PROPERTIES OF LIME STABILIZED SOILS AND AGGREGATES VOLUME 1: SUMMARY OF FINDINGS PREPARED FOR THE NATIONAL LIME ASSOCIATION".

[3] J. L. Eades and R. E. Grim, "Reaction of Hydrated Lime With Pure Clay Minerals in Soil Stabilization".

[3] D. N. Little, Handbook for stabilization of pavement subgrades and base courses with lime, National Lime Association, 1995.

[5] *AASHTO, Standard Method of Test for Determining the Resilient Modulus of Soils and Aggregate Materials (T 307-99)*. Washington, D.C.: American Association of State Highway and Transportation Officials, 2007.

[6] W Sas, A Głuchowski, K Gabryś, and A Szymański, "Application of cyclic CBR test for the estimation of resilient modulus in the pavement construction Application du critère CBR cyclique pour l'estimation de module d'élasticité dans la construction de la chaussée," 2015, doi: 10.13140/RG.2.1.4391.2161.

[7] Camargo, F.F., Wen, H., Edil, T. & Son, Y.H., "Comperative assessment of crushed aggregates and bound/unbound recycled asphalt pavement as base materials," *Int. J. Pavement Eng.*, pp. 1–8, 2012.

[8] M. Budhu, *Soil Mechanics and Foundations*, 3rd ed. Hoboken, NJ: Wiley. 2010.

[9] S. Hossain, R. L. Parsons, and R. Han, "Evaluation of subgrade resilient modulus from repeated load CBR test," *J Mater Civ*, vol. 27, nos. 10, 04014269, 2015, doi: 10.1061/(ASCE)MT.1943-5533.0001230.

[10] K. Gnanendran and H. Woodburn, "Use of repeated load CBR test for characterizing granular pavement materials," *Can Geotech J*, vol. 40, no. 3, pp. 575–587, 2003, doi: 10.1139/t03-016.

[11] *ASTM, D2166/D2166M-16:Standard Test Method for Unconfined Compressive Strength of Cohesive Soil*. ASTM International: West Conshohocken, PA, 2016.

[12] T. B. Edil and J. L. Mickelson, "Durability and strength characteristics of lime-stabilized soils under cyclic loading," *Geotech. Spec. Publ.*, no. 130–142, pp. 1–10, 2005.

[13] D18 Committee, *Test Method for Particle-Size Analysis of Soils*. doi: 10.1520/D0422-63R07E02.

[14] D18 Committee, *Test Methods for Specific Gravity of Soil Solids by Water Pycnometer*. doi: 10.1520/D0854-14.

[15] *ASTM D4318-17, Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soil*, ASTM International. West Conshohocken: PA, 2017.

[16] D18 Committee, *Test Methods for Laboratory Compaction Characteristics of Soil Using Modified Effort (56,000 ft-lbf/ft3 (2,700 kN-m/m3))*. doi: 10.1520/D1557-12.

[17] H. Haghighi, A. Arulrajah, A. Mohammadinia, and S. Horpibulsuk, "A new approach for determining resilient moduli of marginal pavement base materials using the staged repeated load CBR test method," *Road Mater. Pavement Des.*, vol. 19, no. 8, pp. 1848–1867, Nov. 2018, doi: 10.1080/14680629.2017.1352532.

[18] H. Haghighi, A. Arulrajah, A. Mohammadinia, and S. Horpibulsuk, "A new approach for determining resilient moduli of marginal pavement base materials using the staged repeated load CBR test method," *Road Mater. Pavement Des.*, vol. 19, no. 8, pp. 1848–1867, Nov. 2018, doi: 10.1080/14680629.2017.1352532.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Deploying ANPR Technology to Detect Traffic Violation at Intersections in Phnom Penh

Sambath KEO [1*], Vengkheang PHUN [1], Yat YEN [1]

[1]*Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: keo_sambath@gsc.itc.edu.kh*

Abstract: *Road traffic accidents are a major public health concern worldwide (WHO). One of the key contributors to road traffic accidents is traffic rule violations In Phnom Penh City, traffic violation hinders the intra-city transportation by causing traffic accidents and traffic congestion. Types of traffic violation include red-light running, over speeding, tail-gating, not following the directional sign, driving in wrong direction, driving without seat-belt, talking on phone while driving, driving without helmet, failing to yield for pedestrian, and so on. This study aims to evaluate the capacity and functionality of the existing ANPR cameras in public video surveillance systems in Phnom Penh; and to apply algorithms for violations detection. Another objective is to have a clear understanding about traffic violations at intersections in Phnom Penh. The study is conducted at selected intersections based on traffic volume and potential number of traffic violations. Firstly, the site survey was done and traffic police officers at the sites were asked to share information about traffic violations at their location. Secondly, with approval from PPCA, researcher analyzes the video footage from public CCTV cameras at one site to understand the characteristics of traffic violations. Then, we compare the result from manual counting with report about traffic violations generated by the ANPR system. From this comparison, the approximate accuracy and recall rate for traffic violation detection by the ANPR system was determined. The study shows that number of traffic violations is high; and drivers of different vehicle types violate the traffic. It also shows that accuracy and recall of the violation detection by the existing ANPR cameras varies between violation types.*

Keywords: Traffic Violation; ANPR

1. INTRODUCTION

Road traffic accidents are a major public health concern worldwide. The World Health Organization reports that approximately 1.35 million people die each year due to road traffic accidents, making it the leading cause of death among young people aged 5–29 year. Moreover, road traffic accidents are estimated to cause economic losses of up to 3% of a country's gross domestic product [1]. Therefore, it is essential to prioritize measures that aim to reduce the incidence and severity of road traffic accidents.

One of the key contributors to road traffic accidents is traffic rule violations such as speeding, running red lights, and reckless driving that increase the risk of accidents and fatalities on the roads [2]. However, traffic violations not only lead to traffic accidents but also another big problem for transportation system in developing cities, that is traffic congestion. When traffic violation occurs, in most cases, the traffic flow will be interrupted even though there is not an accident.

Automatic Number Plate Recognition (ANPR) is an image processing technology that can capture images of vehicles and from the images extract information about their number (license) plates, and translate them into machine-readable formats, such as text string, which can then be processed and indexed into a database. A typical ANPR system often incorporates a camera device, a processing engine, a recognition engine, and a number plate library. Whilst the system can be designed with various packages, components, and slightly different functions, the ultimate functionality remains the same – to identify vehicles and record their movement information. In an age of rapid urbanization with fast-growing car ownership, the ANPR technology has been used as one of the key sources of big data by urban analysts, planners, and managers.

Since 2016, Phnom Penh Capital Administration (PPCA) has installed more than a thousand of CCTV cameras along the streets to monitor the traffic. Among those cameras, some cameras are ANPR cameras that can automatically detect and capture the license plate of the vehicles.

This study aims to evaluate the capacity and functionality of the existing ANPR cameras in public video surveillance systems in Phnom Penh; and to apply algorithms for violations detection. Another objective is to have a clear

[1] * Corresponding author: *Sambath KEO*
E-mail: keo_sambath@gsc.itc.edu.kh; Mobile: +855-12842880

understanding about characteristic of traffic violations at signalized intersection.

2. METHODOLOGY

2.1 Field of Study

The study is mainly done at the signalized intersection between Win-Win Blvd (or Ly Yongphat Street) and New Road in Phnom Penh (Map is attached in Fig.1). At this location (referred to as site ANPR-01), we applied algorithms for violations detection onto the existing ANPR cameras which were installed by PPCA.



Fig.1. Study location on Google Maps

Besides, we also obtained some more data from site survey and interview with traffic police officers at site ANPR-01 and four other intersections.

2.2 Method

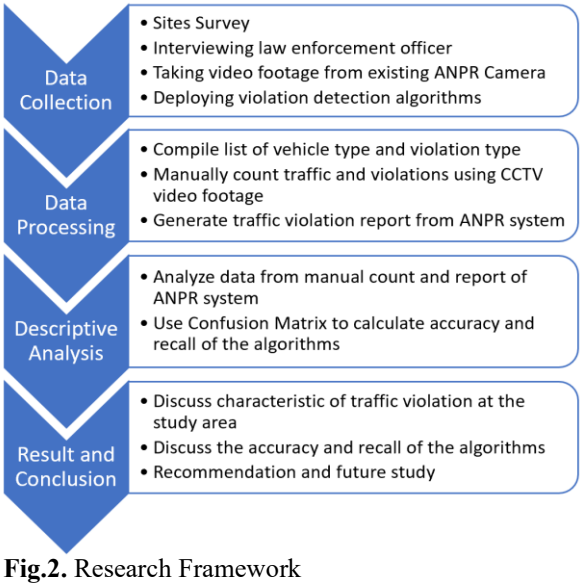


Fig.2. Research Framework

Data collection started with site survey at 5 intersections: 5 Makara (Techno) Overpass, Neak'Voin, Camko, 598 corner 1800, and Win Win Blvd. corner new road (ANPR-01: the site where ANPR cameras were installed). An interview was conducted with traffic police officers standing by at each site to understand their experience in dealing with traffic violations at their site. Moreover, we also looked for secondary data from reliable sources.

Most importantly, video footage from cameras at site ANPR-01 was analyzed. There are 2 cameras available at this site that were previously used only to capture plate number of passing by vehicles. Total number of traffic and number of violations of different types during 4 time-frames were manually counted. Meanwhile, algorithms for violations detection had been applied to existing ANPR cameras at site so that traffic violations were automatically detected and recorded by the system. The ANPR systems that are used as a tool for this study consist of ANPR cameras, supplement lights, terminal server, network infrastructure and backend platform servers. It can generate report in Excel file that we can use for data processing. All the component of the systems is developed by HIKVISION Technology, which is a company specialised in CCTV systems.

Using the statistics, we compare the result from manually counting the violations with the result generated by the ANPR system. Confusion Matrix is used to evaluate the performance of the system. By doing this, we could estimate the accuracy and recall rate of the algorithms when being used in the context of road and traffic condition in Phnom Penh. Having estimated the accuracy and recall rate of the ANPR systems, we could use it to estimate the approximate number of violations in other time frames.

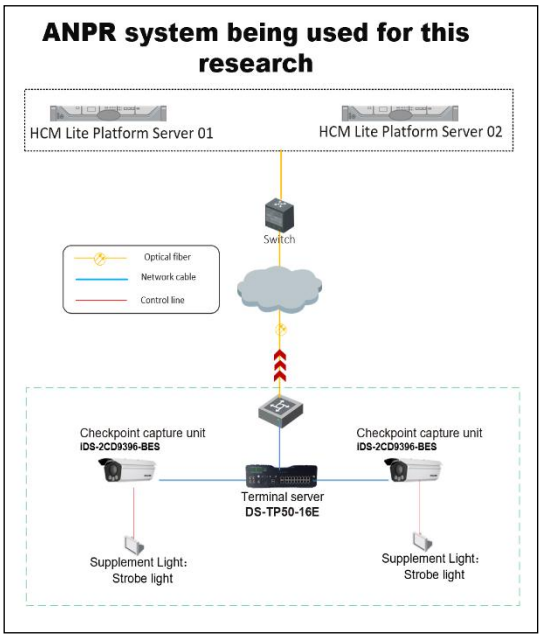


Fig.3. Diagram of the ANPR systems

Fig.3 describes the component of the ANPR system that is used for the research. The system consists of platform servers at the backend, network infrastructure, and on-site devices including terminal server, capture unit, and supplement strobe light.



Fig.4. Layout of camera installation at site ANPR-01

Fig.4 describes the layout of camera installation at the study location. There are 2 cameras on-site: one camera is installed on the traffic light pole arm to capture from the front of incoming vehicles, and another camera is installed on street light pole at the median of the street to capture from the rear of vehicles.

Table 1. Confusion Matrix

		Actual Values	
		Positive	Negative
Predicted Values	Positive	TP	FP
	Negative	FN	TN

Table 1 shows the general structure of the confusion matrix. There are two types of prediction: correct and incorrect(errors).
True Positive (TP): Both the actual values and the prediction are positive.
False Positive (FP): Although the prediction is positive, the actual value is negative. This is called a "Type I error".
True Negative (TN): The actual value is negative, and the prediction is negative.
False Negative (FN): Although predicted to be negative, the sample is positive. This is also called "Type II error".

The first part of TP, TN, FP, and FN, the terms True or False, concerns whether the prediction is correct or not.

Accuracy is the fraction of correct predictions among all predictions or how often a prediction is correct

Accuracy = (Number of correctly classified instances) / (Total number of instances)

$$\text{Accuracy} = (TP + TN) / (TP + FP + TN + FN) \quad \text{(Eq. 1)}$$

Recall is the proportion of actual positives predicted correctly, or how accurately the model predicts positive cases. It is calculated as:

$$\text{Recall} = TP / (TP + FN) \quad \text{(Eq. 2)}$$

3. RESULTS AND DISCUSSION

From our site survey and interview with police officers at intersections, common traffic violation types include car driving without seatbelt, using phone while driving, driving/riding motorbike without helmet, violation of traffic light, violation of lane control sign or marking, crosswalk violation(failing to yield to pedestrian), turns from the wrong lane, sudden lane change, driving in opposite direction, driving on sidewalk, violations by pedestrian (red-light crossing, outside crosswalk), expired driving license, expired vehicle inspection certificate, and vehicle with not sufficient technical specifications.

According to Cambodia Road Traffic Law [3], there are more types of traffic violation such as driving under the influence of alcohol, over speeding, driving over solid lines, using high beam when encounter other vehicles coming from the opposite direction, overloading passengers, increase speed while other vehicle is overtaking, driving without leaving sufficient safety space (tailgating), carelessly driving disturbing other road users, failure to give signal when starting to leave, not following traffic police signals, illegal parking, etc.

In this study we focus on more frequent violation types that occurs at the site ANPR-01. And due to the current technology limitation of the algorithms and the hardware of the cameras on-site, not all types of violation can be detected automatically. Thus, our data is more concentrated on violation types such as: Red-light running (going straight), Yellow-light running (going straight), Violation of lane control sign or marking (which is No U-turn sign), driving/riding motorcycle without helmet, using phone while driving, and car driving without seatbelt, which are more frequent and can be detected by the ANPR systems.

Table 2. Violation Type

Code	Violation Type
RR-ST	Red-light running (going straight)
YR-ST	Yellow-light running (going straight)
VLC-MC	Violations of lane control sign or marking by motorcycle
VLC-OV	Violations of lane control sign or marking by other vehicles
MWH	Driving/riding motorcycle without helmet
UPD	Using phone while driving
DWB	Car driving without seatbelt

In Cambodia, once the traffic light turns Yellow, all vehicles must stop unless the front wheel had already crossed the stop line. If the driver goes on otherwise, the traffic police will call the driver to pull over, and there will be punishment. However, the ANPR systems which were deployed do not detect yellow-light running as a violation. Meanwhile, the ANPR systems cannot detect violation of lane control sign or marking by motorcycle. So, from the ANPR systems, YR-ST and VLC-MC are not reported.

Table 3. Vehicle Type

Code	Vehicle Type
MC	Motorcycle
Car	Sedan,SUV, Van, Wagon, Pickup, Bus, Truck, others
3W	3-Wheeler (Rickshaw)
4W	4-Wheeler (Moto'reumork)

Table 3 describes the vehicle types being classified into different code in this research. Based on site observation and video footage from the CCTV camera, these vehicles in the table are the vehicles that were seen at the location. Sedan, SUV, Pickup, Van, Wagon, Bus, Truck and other vehicles of similar criteria are being recognized as Car in this study.



Fig.5. Example of violation being detected

Fig.5 shows an example of a motorcycle with plate number 1H0-15XX being detected by the ANPR camera for Red Light Running.

The working principle of the ANPR system is simple. Once a vehicle enters detection zone of the camera, it will automatically detect, capture and recognize the plate number of the vehicle. At the same time, the violation detection algorithms observe the movement of the vehicle, and trigger to the terminal server when traffic violation is detected. The terminal server sends all the data to platform server for generating reports.

Table 4. Manual count of traffic volume at site ANPR-01 on 11-March-2024 (during 00h:00 to 01h:00, 07h:00 to 09h:00, 11h:00 to 13h:00, and 17h:00 to 19h:00)

Vehicle Type Code	Volume by time-frame				
	00-01	07-09	11-13	17-19	Total
MC	24	1,961	724	1,213	*3,922*
Car	64	1,168	803	1,041	*3,076*
3W	2	134	110	85	*331*
4W	2	85	50	38	*175*
Total	92	3,348	1,687	2,377	7,504

Table 4 shows the result of manual count of traffic volume at site ANPR-01 during 7 hours of 11-March-2024. Motorcycle is the most common vehicle type with the count of 3,922.

Table 5. Comparing result from manual count and system report of traffic violation at site ANPR-01 on 11-March-2024 (during 00h:00 to 01h:00, 07h:00 to 09h:00, 11h:00 to 13h:00, and 17h:00 to 19h:00 combined)

Violation Code	Manual Count	System Report(=TP)	Difference (=FN)
RR-ST	383	352	31
YR-ST	85	none	n/a
VLC-MC	142	none	n/a
VLC-OV	177	166	11
MWH	209	208	1
UPD	41	10	31
DWB	115	55	60
Total	1,152	791	361

Counting error in manual traffic count is small, usually less than 1% [4]. Number shows that the systems tend to detect a smaller number of violation than manual count.

Within the studied time-frames, there were 1,152 violations while the total traffic volume was 7,504 vehicles. It means that about 15.35% of the vehicles that went past this location violated the traffic.

Table 6. Accuracy and Recall of the algorithms

Violation Code	TP	FN	Accuracy	Recall
RR-ST	352	31	0.995	0.919
VLC-OV	166	11	0.998	0.937
MWH	208	1	0.999	0.995
UPD	10	31	0.995	0.243
DWB	55	60	0.992	0.478

Because the traffic count was conducted in 7 hours when there was big traffic volume, the accuracy of the algorithms is not affected by the number of False Negative. To evaluate the performance of the algorithms in this study, Recall is a more reliable measure.

UPD and DWB detection have very low recall rate. These two violation types mostly occur during the night time; and the installed component of the ANPR systems did not include supplement flash light that can enable the camera to see clearly through the front windshield of the car. It is the reason that the ANPR systems cannot detect UPD and DWB well.

Using violation report generated from the systems and the recall rate of the system, we could see further into the number of traffic violations at this intersection during broader time frame.

We extracted number of three type of violations (RR-ST, VLC-OV, MWH) from report of violation on 11-Mar-2024 the whole day for further discussion. On this particular day, there were huge number of violations: RR-ST: 1,039 cases, VLC-OV:415 cases, and MWH: 756 cases. Only three types of violations in one day at one intersection account for 2,210 cases of traffic violation. Meanwhile, based on the 2023 annual report by Phnom Penh Capital Administration, traffic police and authorities have recorded only 40,438 traffic violations throughout the year [5]. Considering the number of intersections in the whole capital, this amount of 40,438 violations is relatively very low. However, it is worth mentioning about working hour of traffic police officers that standby at intersections. Normally, they standby from 7am to 11am and 1pm to 5pm. So, any violations out of this working hour mostly are not recorded.

4. CONCLUSIONS

The study shows that the existing ANPR cameras in Phnom Penh can be upgraded with more intelligent technology and algorithms to detect traffic violations at intersection. With traffic violation detection algorithms, these ANPR cameras can detect Red-light Running, Violations of lane control sign or marking, and Driving/riding motorbike without helmet at high recall rate of more than 90%. On the other hand, there should be more improvement on algorithms to detect other types of violation to make them applicable in Phnom Penh.

Number of traffic violations at intersection ANPR-01 is worryingly high. Red-light Running, Violation of lane control sign or marking, and Driving/riding motorcycle without helmet are top of the list. They are dangerous and can lead to fatal accidents. Thus, there should be more countermeasure against these violations especially at intersections.

Since this study was mostly done at sub-urban area, it would be useful to have future study at other intersections inside the city for the complexity of traffic inside the urban area may be much different from the sub-urban area. Researcher with computer coding capability might also be needed if more types of violation are to be automatically detected by the ANPR systems and algorithms.

REFERENCES

[1] World Health Organization (2023). Global status report on road safety

[2] El Mehdi B. Laoula, Omar Elfahim, Marouane El Midaoui, Mohamed Youssfi & Omar Bouattane. (2023) Traffic violations analysis: Identifying risky areas and common violations, Heliyon,Volume 9, Issue 9

[3] Cambodia Law on Road Traffic 2017

[4] P. Zheng and M. Mike, “An Investigation on the Manual Traffic Count Accuracy,” *Procedia - Social and Behavioral Sciences*, vol. 43, pp. 226–231, 2012, doi: 10.1016/j.sbspro.2012.04.095.

[5] Phnom Penh Capital Administration (2023). 2023 Annual Report on performance of PPCA

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

3D Road Surface Reconstruction: A Case Study of Indoor Corridor Using RPLiDAR 360° A2M8

Phearun MAO [1*], Veng Kheang PHUN [1], Narith SAUM [1]

[1] *Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: phearun_mao@gsc.itc.edu.kh*

Abstract: *Accurate and cost-effective road surface reconstruction is vital for infrastructure assessment in developing regions, where budget constraints limit access to advanced technologies. This study presents an affordable method for 3D road surface reconstruction using a 2D RPLIDAR A2M8 LiDAR sensor on a movable platform, validated in a controlled indoor environment to ensure precise data collection. Sequential 2D scans were acquired at 0.1-meter intervals along a marked floor simulating a road segment. A custom Python-based pipeline synchronized scans, corrected alignment errors, filtered noise, and generated a 3D point cloud. Surface reconstruction in MeshLab produced a 3D mesh capturing the test environment's geometry. The controlled indoor setup minimized variables like terrain irregularities, establishing a proof-of-concept for affordable 3D road surface modeling. Validation against reference objects demonstrated high accuracy, with (98.9% accuracy) for width and (98.1% accuracy) for height, highlighting exceptional vertical precision These findings underscore the potential of low-cost 2D LiDAR sensors for detailed reconstruction, offering a scalable solution for infrastructure monitoring. Future work will extend this methodology to outdoor road applications, integrating GPS and IMU for enhanced real-world applicability.*

Keywords: 2D LiDAR, Affordable Road surface reconstruction, 3D road surface modeling, Mobile mapping, Point cloud processing

1. INTRODUCTION

Accurate and affordable road surface reconstruction is critical for infrastructure monitoring, particularly in developing regions where access to advanced survey tools is limited. Conventional methods, such as Differential Global Positioning Systems (DGPS), total stations, or drone-based photogrammetry, offer high accuracy but are hindered by significant costs, specialized equipment, trained personnel, and logistical challenges [1]. Notably, total stations face significant drawbacks in road surface assessment, especially in undulating or mountainous terrains. Their systematic scanning, typically at 25-meter intervals, results in low spatial resolution, capturing sparse spot heights that fail to represent complex road geometries [2]. This coarse data leads to challenges in designing curved roadways, inaccurate earthwork volume estimates, and potential oversight of surface irregularities. Additionally, total stations are prone to missing data in hard-to-access areas, with measurements vulnerable to human error, poor instrument conditions, or obstacles like steep slopes [2]. Low-cost 2D LiDAR sensors, such as the RPLIDAR A2M8, offer a promising alternative for high-resolution data collection when paired with intelligent acquisition strategies. However, transforming 2D scans into accurate 3D models requires meticulous planning, calibration, and validation. This study validates a low-cost 2D LiDAR-based methodology in a controlled indoor environment, establishing a proof-of-concept for affordable 3D road surface reconstruction suitable for future outdoor deployment.

This study presents a methodology for 3D surface reconstruction using a low-cost 2D LiDAR sensor (RPLiDAR A2M8) on a movable platform. The experiment was

[1*] Corresponding author: Phearun MAO
Email: phearun_mao@gsc.itc.edu.kh; Tel: +855-95 817 917

conducted indoors to ensure controlled conditions for validating the methodology, minimizing variables like terrain irregularities, weather, or dynamic obstacles that could affect initial testing. By employing fixed-interval manual movement, calibration, and validation against known reference objects, this approach aims to provide a reliable, cost-effective solution for road surface profiling, suitable for future field deployment.

2. METHODOLOGY

This section provides a detailed overview of the process for system setup, sensor calibration, data collection, and 3D surface reconstruction from 2D LiDAR scans, optimized for controlled, cost-effective indoor testing to simulate 3d road surface reconstruction with precise movement intervals and sensor alignment, building on established practices for LiDAR-based mobile mapping systems.

2.1 Hardware and Device Installation

The data acquisition system was designed for accuracy and repeatability in an indoor setting (Fig. 1). The core sensing component is the RPLiDAR A2M8—a high-performance 2D planar LiDAR offering 360° scanning at 10 Hz, a range of 0.15–12 meters, and an angular resolution of 0.9°, making it capable of capturing fine surface variations[3]. Its compact size and low power consumption make it ideal for portable setups. In this configuration, the LiDAR is mounted on a tripod fixed at the front edge of a movable platform, oriented vertically to scan perpendicular to the floor. The platform itself is a wheeled desk measuring approximately 1.2 × 0.6 m, with a fixed height of 0.8 m. The LiDAR is positioned at a height of 0.86 m above the ground and is connected via USB to a laptop, which manages real-time data acquisition, processing, and storage. The platform moves smoothly along a tiled floor marked at 0.1 m intervals, providing consistent reference points for spatial positioning. The sensor's orientation is leveled using a digital spirit level and documented to enable accurate frame transformations during post-processing.



Fig. 1. Experimental setup

2.2 Device calibration

To ensure accurate and reliable data acquisition, a simple calibration procedure is performed before data collection session. The calibration process begins by allowing the LiDAR sensor to spin continuously for approximately 30 seconds. This warm-up period stabilizes the sensor's internal mechanisms and ensures consistent scanning performance [3]. Following this, a reference measurement is taken using a known calibration object. The object has a precisely measured height of 0.08 m and is placed at a fixed distance of 0.78 m from the LiDAR sensor (Fig. 2). The purpose of this step is to verify that the sensor accurately detects the object's distance and elevation within the expected range. The measured scan data is compared to the known dimensions to confirm proper sensor alignment and scanning accuracy. If discrepancies are found, adjustments are made before proceeding with data collection.

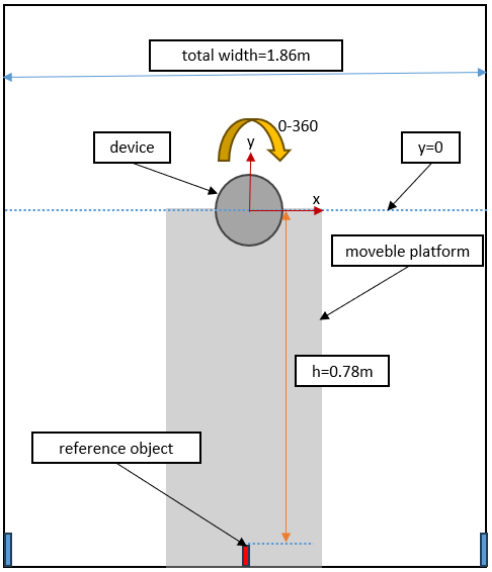


Fig. 2. Initial calibration with known object

2.3 Software and algorithm

A Python-based data acquisition system was developed to efficiently capture, process, and log LiDAR data in an indoor setup [4]. Using WSL with Ubuntu 22.04, ROS2 Humble was installed for seamless integration with the RPLIDAR A2M8 [5]. The system utilizes PyRPLiDAR for real-time data transfer, logging scans with timestamps, angles, distances, and intensity values [6]. Data is stored in CSV format for easy analysis using tools like MESHLAB, Cloud Compare, or Python-based libraries.

2.4 Data collection

Data collection was conducted in a controlled indoor environment to simulate road surface reconstruction, ensuring

precise and consistent measurements. A systematic stop-and-capture method was employed, with the LiDAR-equipped platform moved at fixed intervals along a marked corridor. A custom graphical interface facilitated scan initiation, data logging, and real-time monitoring (Fig. 3). Reference objects of known dimensions were placed along the corridor for validation purposes (Fig. 4). The process involved capturing sequential 2D LiDAR scans, maintaining a clear path to establish a baseline dataset. Each session was documented with photographs and metadata

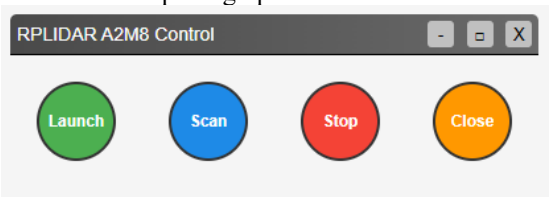


Fig 3. Custom LiDAR control GUI

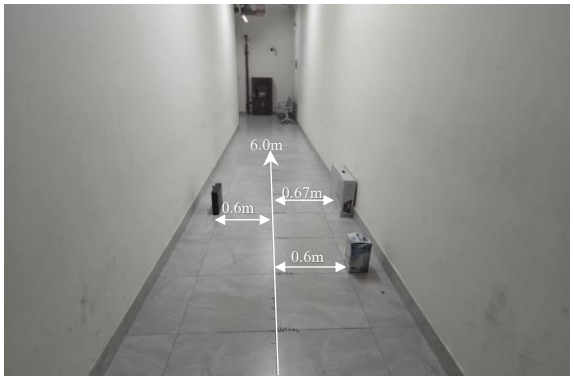


Fig. 4. Corridor data collection: known object placement

2.5 Data processing

The data processing pipeline converts raw two-dimensional LiDAR scans into a three-dimensional point cloud, thereby providing a digital representation of the corridor surface in our indoor experiment. This process is integral to the study, as it transforms raw sensor data into a format suitable for analyzing surface characteristics and performing geometric alignment. The process begins with importing raw scan data—comprising angle and distance measurements—from CSV files generated during the data collection phase. These files capture the spatial layout of the corridor, but due to the sensor's physical mounting orientation, slight angular misalignments may occur. To rectify this, angle correction is applied by adjusting the raw angles using a fixed offset. The corrected angle is computed as:

$$\theta = \theta.raw + \theta.adj \qquad (Eq.\ 1)$$

Where:

- θ is the corrected angle in degree
- $\theta.raw$ is the angle from sensor in degree
- $\theta.adj$ is the offset angle in degree, apply to align with the corridor layout

In addition to angle correction, the data is filtered to remove noisy or unreliable distance readings. Filtering may include thresholding to exclude physically implausible distances or statistical methods to eliminate outliers based on surrounding values (Fig. 5). Subsequent to the filtration and angular correction, the next step is to extract a representative distance value for each angle from the multiple LiDAR readings available in each scan. Each scan point consists multiple radial measurements, which are utilized to calculate a single distance value based on a specified percentile. This approach serves to mitigate the impact of residual noise and establishes a consistent foundation for transformation (Fig 5). The representative distance d for each angle is calculated using percentile base in excel formular with refer to:

$$dP = d\lfloor r \rfloor + (r - \lfloor r \rfloor).(d\lfloor r \rfloor + 1 - d\lfloor r \rfloor) \qquad (Eq.\ 2)$$

Where:

- d is the calculated representative distance.
- P is the desired percentile (ranging from 0 to 100).
- N is the total number of sorted radial distance measurements for that specific angle.
- $r = \frac{P}{100} x(N - 1) + 1$ the calculated rank for the P-the percentile.
- d⌊r⌋ is the distance value at the position corresponding to the integer part of the rank r within the sorted radial measurements.
- d⌊r⌋+1 is the distance value at the next position after d⌊r⌋ in the sorted radial measurements

Once the representative distance is computed, it is converted from polar to Cartesian coordinates, taking into account both the rotation and translation offsets defined in the parameter setup. The transformed coordinate is converted using:

$$x.lidar = d.(\cos(\theta)) + xo \qquad (Eq.\ 3)$$

$$y.lidar = d.(\sin(\theta)) + yo \qquad (Eq.\ 4)$$

Where:

- xo is the horizontal different error with actual point
- yo is the vertical different error with actual point

After performing these transformations for all target points, each transformed LiDAR point compared with its corresponding actual reference point to define the optimal error value. The alignment error for each point is calculated using the Euclidean distance and average the all-errors distance to get the mean error follow by:

$$Ei = \sqrt{x.lidar - x.actual)^2 + (y.lidar - y.actual)^2} \quad (Eq.\ 5)$$

Where:

- $x.actual$ is the actual x value or reference point
- $y.actual$ is the actual y value or reference point

In this optimization process, the Evolutionary algorithm in Excel Solver (Fig. 6) is employed due to its effectiveness in handling complex, nonlinear problems. This method utilizes a genetic algorithm-inspired approach, exploring the solution space through mechanisms such as mutation and population evolution. The objective is to minimize the mean Euclidean error between derived and reference coordinates by optimizing key transformation parameters. These parameters include a percentile value, bounded within a defined feasible range, and horizontal and vertical translations, constrained to physically realistic limits.

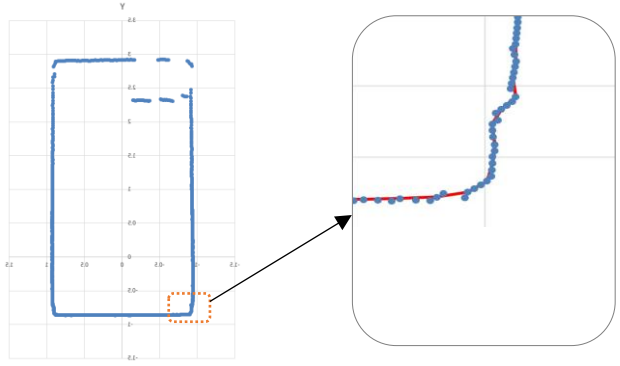


Fig. 5. Scatter plot compares before and after correcting angle with corner detection

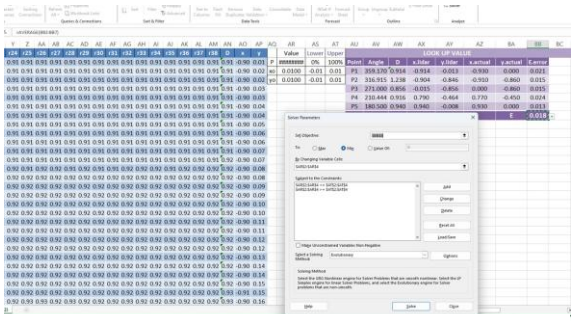


Fig. 6. Excel solver dialog

nize the average Euclidean error between LiDAR and reference points, a centering alignment step was implemented. This involves computing the centroids of both the transformed LiDAR points and the actual coordinates, calculating the offset between them, and shifting the entire LiDAR point cloud accordingly. This correction ensures that the transformed data is not only close to the actual measurements but also centrally aligned, improving spatial accuracy and reducing systematic offset (Fig. 7).

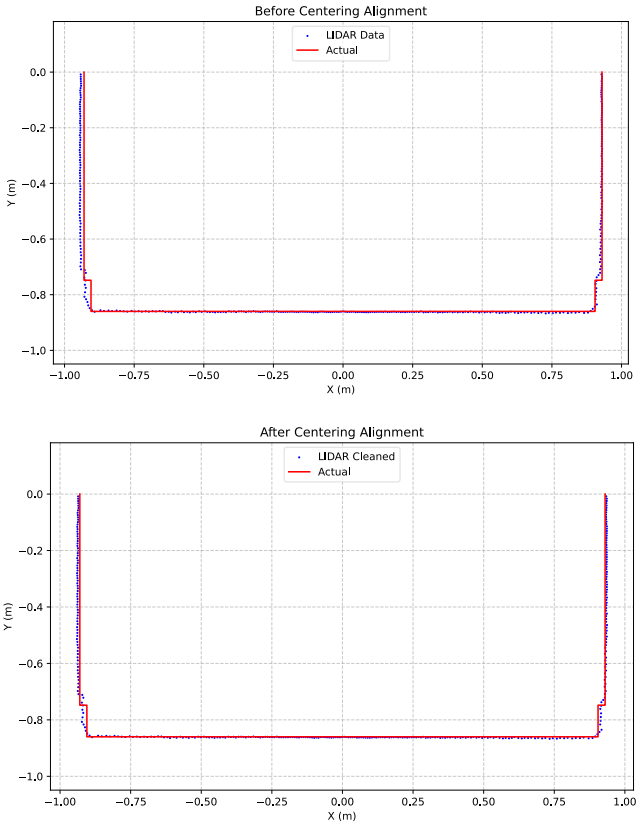


Fig. 7. Scatter plot compares between before and after alignment adjustment, (top) before and (bot) after alignment adjustment

3. RESULTS AND DISCUSSION

3.1 Experimental Setup

The data validation process ensures the integrity and reliability of sensor data used to measure corridor width at different 5 stations, Z=0.0m, Z=0.10m, Z=0.20m, Z = 3.0 m (two boxes, height=600 mm, 600 mm from center for the left-hand side, height=450 mm, 770 mm from center) and the last is Z=3.60m (one box, height=450mm), consider 5points for each station for the analysis, where objects are placed for detection were measured relative with Y=0. Each CSV file is checked for required columns (distance, angle) with numeric values, no missing or null entries, and plausible ranges

(distance > 0, angle within [0, 360°)). Z-position metadata is validated to align with the expected stations (Z = 1.90 m and Z = 3.0 m), flagging deviations exceeding 0.1 m to detect potential misalignment or metadata errors. After coordinate conversion, X (corridor width) and Y (height) values are verified to be physically realistic, with X consistent with corridor dimensions and Y within [-1.0, 1.0], and X-width measurements are checked against known object dimensions for consistency.

3.2 Error metric comparison

The accuracy of LiDAR measurements was evaluated across five stations by comparing the 2D point coordinates against actual measured positions. Two sets of results are presented raw error metrics and improved metrics using an Excel-based solver for angle correction. The table below is the summarizes of the result for this experiment:

Table 1. Error Metrics Comparison of Lidar and Lidar-Solver

Station	Ei	Ei-Solver
0.0	0.0166	0.0159
0.1	0.0074	0.0074
0.2	0.0222	0.0218
3.0	0.0133	0.0114
3.4	0.0139	0.0128
MAE	0.0147	0.0139

The LiDAR system consistently demonstrated high accuracy across all 25 measured points. With an overall Mean Absolute Error of 0.0141 for 2D position, the system achieved an impressive overall 2D accuracy. The X-coordinate demonstrated 98.90% accuracy, while the Y-coordinate showed 98.10% accuracy. These results confirm the LiDAR's strong capability in providing precise spatial measurements, making it well-suited for applications requiring reliable position data.

3.3 3D surface reconstruction

Surface reconstruction transforms the processed and validated LiDAR data into a 3D model of the corridor using MeshLab, an open-source tool for 3D mesh processing, to support road surface profiling applications [9]. The validated data, consisting of X (corridor width), Y (height), and Z (distance along the corridor) coordinates from CSV files, is first converted into a point cloud format (PLY) using Python scripts. The point cloud is imported into MeshLab, where preprocessing steps remove residual outliers beyond the. The Poisson Surface Reconstruction algorithm is applied to generate a 3D mesh, chosen for its robustness in handling noisy LiDAR data and producing a smooth, watertight surface that accurately represents the corridor geometry, including the positions of reference objects at Z = 1.90m, Z = 3.0m and Z = 3.6m. The reconstructed mesh is visualized as a wireframe (See Fig. 8), with axes labeled in meters (X: [-1.0, 1.0], Y: [0.0, 0.8], Z: [0.0, 6.0]), confirming the corridor's geometry and object placements. The mesh is exported for further analysis, such as integration with road design software, ensuring the 3D model accurately reflects the validated LiDAR measurements.

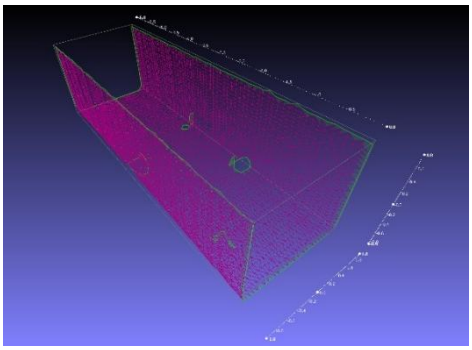


Fig.8. 3D surface reconstruction from MESHLAB

Applicability to Road Surface Issues

For road surface surveying, the LiDAR system can effectively detect irregularities such as potholes and rutting with depths or widths greater than 25 mm, supported by the sensor's accuracy and sufficient point density. Unevenness like bumps and dips are also well-captured due to the high vertical resolution. However, the detection of finer cracks (less than 10 mm) is limited by the sensor's 0.9° angular resolution and 0.1-m scan interval, suggesting that higher-resolution sensors may be necessary for such detailed features. For optimal data acquisition during road surveying, a recommended speed of approximately 3.6 km/h (1 m/s) should be maintained to ensure a 0.1-m resolution, as higher speeds would compromise point density. Furthermore, for outdoor deployment, integrating stabilization mechanisms and Inertial Measurement Units (IMU) is crucial to mitigate vibrations and maintain accurate alignment.

CONCLUSIONS

This study validates a low-cost 2D LiDAR (RPLiDAR A2M8) for 3D road surface reconstruction, achieving 98.90% (width) and 98.10% (height) accuracy in a controlled indoor environment. The methodology, tested indoors to ensure precise validation, demonstrates potential for detecting potholes, unevenness, and rutting, though fine cracks require higher-resolution sensors. The RPLiDAR A2M8 is recommended for its affordability (~$300–$400) and performance, though 3D LiDARs like Livox Mid-40

(~$1,000) offer advantages for outdoor applications due to direct 3D scanning and higher resolution. The indoor setup limits direct applicability to real roads, necessitating future outdoor tests (e.g., on ITC campus roads) to address terrain and environmental variables. Future work should automate data collection, integrate GPS/IMU.

ACKNOWLEDGMENTS

The author sincerely appreciates the invaluable guidance, continuous support, and insightful feedback received throughout the course, and extends special thanks to GTI laboratory for their contributions. Additionally, heartfelt gratitude is extended to the TSU unit for supplying the RPLIDAR device, which was instrumental in assisting this research

REFERENCES

[1] F. Hatta Antah, M. A. Khoiry, K. N. Abdul Maulud, and A. Abdullah, "Perceived Usefulness of Airborne LiDAR Technology in Road Design and Management: A Review," *Sustainability*, vol. 13, no. 21, Art. no. 21, Jan. 2021, doi: 10.3390/su132111773.

[2] M. Ono, "Comparative Analysis of DGPS and Total Station Accuracies for Static Deformation Monitoring Of Engineering Structures," Jun. 2018, Accessed: May 01, 2025. [Online]. Available: https://www.academia.edu/79876736/Comparative_Analysis_of_DGPS_and_Total_Station_Accuracies_for_Static_Deformation_Monitoring_Of_Engineering_Structures

[3] "RPLIDAR-A2 Laser Range Scanner_ Solid Laser Range Scanner|SLAMTEC." Accessed: Apr. 16, 2025. [Online]. Available: https://www.slamtec.com/en/Lidar/A2/

[4] "Python for Robotics - Full Course for Beginners," The Construct. Accessed: Apr. 25, 2025. [Online]. Available: https://www.theconstruct.ai/robotigniteacademy_learnros/ros-courses-library/python-robotics/

[5] "ROS 2 Documentation — ROS 2 Documentation: Humble documentation." Accessed: Apr. 25, 2025. [Online]. Available: https://docs.ros.org/en/humble/index.html

[6] "GitHub - Slamtec/rplidar_ros at dev-ros2," GitHub. Accessed: Apr. 25, 2025. [Online]. Available: https://github.com/Slamtec/rplidar_ros

[7] "tkinter — Python interface to Tcl/Tk," Python documentation. Accessed: Apr. 25, 2025. [Online]. Available: https://docs.python.org/3/library/tkinter.html

[8] "User Guide — pandas 2.2.3 documentation." Accessed: May 02, 2025. [Online]. Available: https://pandas.pydata.org/docs/user_guide/index.html#user-guide

[9] P. Cignoni, M. Callieri, M. Corsini, M. Dellepiane, F. Ganovelli, and G. Ranzuglia, "MeshLab: an Open-Source Mesh Processing Tool," 2008, *The Eurographics Association*. doi: 10.2312/LOCALCHAPTEREVENTS/ITALCHAP/ITALIANCHAPCONF2008/129-136.

[10] R. Ravi, A. Habib, and D. Bullock, "Pothole Mapping and Patching Quantity Estimates using LiDAR-Based Mobile Mapping Systems," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2674, no. 9, pp. 124–134, Sep. 2020, doi: 10.1177/0361198120927006.

[11] C. Koch, G. Jog, and I. Brilakis, "Automated Pothole Distress Assessment Using Asphalt Pavement Video Data," *Journal of Computing in Civil Engineering*, vol. 27, pp. 370–378, Jul. 2013, doi: 10.1061/(ASCE)CP.1943-5487.0000232.

[12] S. Chaithavee and T. Chayakul, "Classification of 3D Point Cloud Data from Mobile Mapping System for Detecting Road Surfaces and Potholes using Convolution Neural Networks," *International Journal of Geoinformatics*, vol. 18, no. 6, Art. no. 6, Dec. 2022, doi: 10.52939/ijg.v18i6.2455.

[13] H. Xin *et al.*, "Sustainable Road Pothole Detection: A Crowdsourcing Based Multi-Sensors Fusion Approach," *Sustainability*, vol. 15, no. 8, Art. no. 8, Jan. 2023, doi: 10.3390/su15086610.

[14] "Mid-40 lidar sensor - Livox." Accessed: May 26, 2025. [Online]. Available: https://www.livoxtech.com/mid-40-and-mid-100

[15] C. Cahalane, C. Mc Elhinney, and T. Mccarthy, *Mobile mapping system performance - an analysis of the effect of laser scanner configuration and vehicle velocity on scan profiles* . 2010.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessing the Impact of Motorcycle Emissions on Urban Air Quality: A Case Study at the Institute of Technology of Cambodia

Lihong PHINY [1*], Veng Kheang PHUN [1], Narith SAUM [1], Sokhit PLACK [1,2]

[1] *Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: Lihong_phiny@gsc.itc.edu.kh*

Abstract: *Air pollution poses a significant public health challenge in urban environments, particularly in rapidly developing cities like Phnom Penh, Cambodia. Motorcycles, which constitute over 80% of registered vehicles in the city, contribute substantially to air pollution through the emission of particulate matter (PM) and other harmful pollutants. However, research on motorcycle emissions in Cambodia remains limited, creating a gap in understanding their specific impact on air quality. This study investigates the influence of key motorcycle characteristics—fuel type, vehicle age, and model—on particulate matter (PM) emissions within the campus of the Institute of Technology of Cambodia (ITC), where motorcycles constitute the predominant mode of transportation. Emission data were collected from 100 motorcycles under idle conditions using a ZJSJ-G portable laser particle counter. Statistical analyses, including independent t-tests, one-way ANOVA, and Tukey's HSD post-hoc tests, were applied to determine significant differences in PM emissions among categories. Results revealed that motorcycles using Simple gasoline emitted significantly higher PM concentrations than those using Super gasoline ($p < 0.05$). Motorcycles aged 4–7 years exhibited greater emissions than those aged 0–3 years ($p = 0.016$). Furthermore, Honda Dream and Zoomer models recorded the highest PM levels. Although engine size appeared to influence emissions descriptively, inferential analysis found no statistically significant differences. These findings underscore the importance of promoting high-grade fuels, encouraging the replacement of mid-aged motorcycles, and regulating high-emission models. The study contributes to urban air quality management strategies in Southeast Asia, providing localized evidence to inform policy development. However, vehicle age alone may not fully capture the mechanical condition or emission performance of motorcycles, highlighting the need for future research to consider maintenance and operational history.*

Keywords: Air pollution, Cold-Start Emissions, Motorcycle emissions, Particulate matter (PM), Sustainable urban mobility

1. INTRODUCTION

Air pollution is an escalating environmental and public health challenge in rapidly urbanizing cities such as Phnom Penh, Cambodia. Among the primary contributors to this problem are emissions from motorized vehicles, particularly motorcycles, which dominate the urban transport system. In Phnom Penh, motorcycles represent over 80% of all registered vehicles, making them a dominant mode of urban transport [1]. Yet their emission characteristics remain poorly studied compared to other vehicle types, and their emissions remain under-researched in Cambodia, leaving a critical gap in local air quality management efforts [2]. Numerous studies across Southeast Asia have underscored the significant role motorcycles play in urban air pollution. In Bangkok, [3] demonstrated that two-wheelers emit large quantities of fine particulate matter (PM2.5) and hydrocarbons, particularly when older or poorly maintained. Similar patterns were observed in cities such as Jakarta and Ho Chi Minh City, where outdated motorcycle fleets and the use of low-quality fuels contribute to elevated levels of carbon monoxide (CO), nitrogen oxides (NO_x), and PM [4,5,6]. Health-related research has also established strong links between fine particulate matter and adverse health outcomes. Brook et al. (2010) found that PM exposure is closely associated with increased risks of respiratory and cardiovascular diseases [7,8]. More recently, Ting et al. (2023) reported that students and young adults, especially in traffic-heavy educational

[1*] Corresponding author: Lihong PHINY
E-mail: Lihong_phiny@gsc.itc.edu.kh; Tel: +855-68 602 747

environments, are increasingly vulnerable to PM-related respiratory symptoms due to frequent exposure [9]. Another important factor is the cold-start phenomenon. Borhani et al. (2021) emphasized that motorcycles emit significantly higher levels of pollutants during the first few minutes after ignition, primarily due to incomplete combustion at lower engine temperatures [10]. These emissions are especially concentrated in enclosed or semi-enclosed environments like motorcycle parking lots [11], where engines are frequently started and stopped. In response, studies such as Mahallana et al. (2022) have recommended targeted interventions, including phasing out older vehicles, improving fuel quality, and enforcing routine emission testing [2,12]. However, the practical implementation of such policies remains limited in Cambodia due to enforcement barriers, limited infrastructure, and a lack of localized emission data particularly from motorcycles.

Despite regional studies highlighting the impact of motorcycles on emissions, localized research in Cambodia remains scarce on the impact of motorcycle characteristics, specifically fuel type (simple vs. super), age, and model, on PM emissions under idle conditions at the Institute of Technology of Cambodia (ITC). Using a dataset of 100 motorcycles, the study applies t-test, one-way ANOVA, and Tukey's post-hoc analysis to identify significant differences in PM emissions among various subgroups. The objectives are to: (1) assess the influence of key motorcycle attributes on PM emission levels; (2) determine statistically significant group differences; and (3) provide evidence-based recommendations to support emission reduction strategies in urban academic environments. By providing empirical evidence from a high-use urban campus, this study contributes to the limited but growing body of literature on motorcycle emissions in Southeast Asia and supports data-driven policymaking for sustainable urban transport and air quality management in Cambodia.

2. METHODOLOGY

2.1 Study area

This study was conducted at the Institute of Technology of Cambodia (ITC), located along Russian Federation Boulevard in Phnom Penh, Cambodia. As one of the country's leading engineering and technology universities, ITC serves a large population of students, faculty, and staff who primarily commute by motorcycle. The campus includes over 2,800 designated motorcycle parking spaces, making it a high-density setting for two-wheeler traffic. Figure 1 shows the layout of the ITC campus, highlighting the designated motorcycle parking areas where emission measurements were conducted. The combination of frequent motorcycle activity, limited ventilation in parking areas, and high cold-start frequency creates ideal conditions to study localized air pollution from motorcycle emissions. These characteristics make the ITC campus a representative microenvironment for assessing the impact of motorcycle-related air pollution in urban academic institutions across Cambodia and similar Southeast Asian contexts.



Fig 1. Study Area at Institute of Technology of Cambodia (ITC).

2.2 Data Collection

A ZJSJ-G portable laser particle counter was used to measure particulate matter emissions. This device is capable of detecting airborne particles in six size ranges (0.3–10.0 μm) at a flow rate of 28.3 L/min. Emission measurements were recorded at one-minute intervals over a three-minute idling period, simulating cold-start conditions, with the motorcycle placed under a semi-enclosed hood to capture emissions efficiently and consistently. A total of 100 motorcycles were selected for PM emission testing. For each motorcycle, key information was recorded, including the fuel type (categorized as either Simple or Super), the year of manufacture (used to classify motorcycles into age groups: 0–3 years, 4–7 years, and over 8 years), and the motorcycle model (Honda Dream, Honda Scoopy, Honda Beat, or Honda Click). Emissions were measured while the motorcycles were idling, simulating real-world parking conditions where cold-start emissions typically peak.

2.3 Data Analysis

All emission data were cleaned and processed using Microsoft Excel tools. Descriptive statistics were first calculated to explore the distribution of PM emissions across different motorcycle characteristics. For inferential analysis, an independent samples t-test was conducted to compare the mean PM emissions between motorcycles using Simple and Super gasoline. A one-way ANOVA was applied to assess differences in emissions across motorcycle age groups and model types. Additionally, a two-way ANOVA with interaction was conducted to evaluate whether the effect of

fuel type on emissions varied across different motorcycle models. All statistical tests were conducted at a 95% confidence level ($\alpha = 0.05$). To support interpretation, the results were visualized using box plots and summary tables.

3. RESULTS AND DISCUSSION

3.1 Descriptive Statistics

Descriptive analysis revealed clear patterns in PM emission levels across motorcycle characteristics. Table 1 summarizes the descriptive statistics of PM emissions according to engine size, motorcycle age group, fuel type, and model. Engine size showed a general trend where motorcycles with smaller engine measurements emitted higher levels of PM. Motorcycles with engine sizes 100–110cc and 111–125cc recorded average PM concentrations of 347.79 µg/m³ and 424.11 µg/m³, respectively. In contrast, motorcycles in the 126–150cc group emitted lower PM levels, averaging 314.72 µg/m³. This trend suggests that newer or higher-capacity engines may be more efficient in combustion, although this factor did not yield statistically significant results in inferential tests. Motorcycle age had a stronger influence on emissions. Vehicles aged 0–3 years emitted the lowest average PM at 273.35 µg/m³. Emissions increased significantly for motorcycles aged 4–7 years and over 8 years, with averages of 436.97 µg/m³ and 440.04 µg/m³, respectively. These findings suggest a rapid decline in emission performance after the third year, likely due to engine wear and aging components. Fuel type also significantly affected emission levels.

Table 1. Descriptive Statistics of PM Emissions by Engine Size, Age, Fuel Type, and Model.

Factor	N =100	N =100	N =100	N =100
	1 Min (µg/m³)	2 Min (µg/m³)	3 Min (g/m³)	Average
Engine Size Group				
100–110cc	622.49	244.59	176.29	347.79
111-125cc	692.27	333.83	246.23	424.11
126-150cc	513.14	234.56	196.47	314.72
Age (Years)				
0–3 yrs	465.37	199.86	154.83	273.35
4–7 yrs	764.28	320.38	226.25	436.97
8+ yrs	705.65	352.4	262.06	440.04
Gasoline Type				
Simple	653.95	282.62	207.95	381.51
Super	673.53	324.45	233.17	419.39
Type of Motorcycle				
Honda Zoomer	806.8	301.09	210.01	439.30
Honda Dream	767.19	369.58	273.14	469.97
Honda Scoopy	638.29	263.33	192.03	364.55
Honda Click	469.76	222.3	166.04	286.03
Honda Beat	395.01	156.89	114.27	222.06

Motorcycles using Simple gasoline emitted an average of 381.51 µg/m³, whereas those using Super gasoline recorded a higher average of 419.39 µg/m³. Despite the higher average, inferential analysis later confirmed that Simple gasoline is associated with significantly greater PM emissions, likely due to lower combustion efficiency and fuel quality. When comparing motorcycle models, the Honda Dream and Zoomer exhibited the highest average emissions at 469.97 µg/m³ and 439.30 µg/m³, respectively. Meanwhile, the Beat model showed the lowest emissions, averaging only 222.06 µg/m³. These variations may stem from differences in engine design, fuel system efficiency, or maintenance status. Additionally, PM emissions were observed to be highest during the first minute of idling, averaging 658.02 µg/m³. Emissions declined to 291.32 µg/m³ at two minutes and further to 213.19 µg/m³ at three minutes. This trend reflects the typical cold-start emission pattern, where pollutant output is highest before engine temperature stabilizes. Overall, the descriptive statistics indicate that motorcycle emissions are influenced by multiple factors, including vehicle age, fuel type, and model design. These findings provided the basis for further statistical testing using ANOVA and t-tests to evaluate the significance of these effects. The pattern of emissions over the idling period was also evaluated. In figure 2 illustrates the trend of average PM emissions across three minutes of engine idling, showing that the highest emissions occurred within the first minute and progressively decreased in the following minutes, consistent with the typical cold-start behavior of motorcycles.

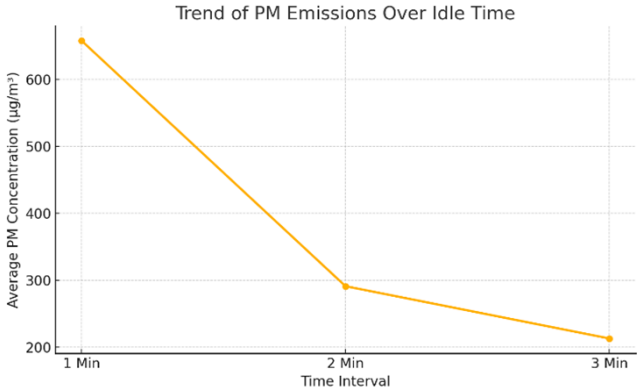


Fig 2. shows the trend of average PM emissions over three minutes of motorcycle engine idling.

3.2 Independent T-Test

To assess the influence of fuel type on particulate matter (PM) emissions, an independent sample t-test was conducted comparing motorcycles using Simple gasoline (n = 80) with those using Super gasoline (n = 20). Although the sample sizes were unequal, Welch's t-test was applied to account for differences in variance and ensure a robust comparison. The results, summarized in table 2, revealed a statistically significant difference between the two groups. Motorcycles fueled by Simple gasoline emitted an average PM concentration of 224.40 μg/m³, whereas those using Super gasoline recorded a significantly lower mean of 17.12 μg/m³. The t-test yielded a t-statistic of 25.88 and a p-value < 0.05, indicating a highly significant difference at the 95% confidence level. These findings confirm that fuel quality has a substantial impact on emission performance, with higher-grade Super gasoline leading to notably lower PM emissions.

Table 2. T-Test Results for Fuel Type

Fuel Type	Mean PM (μg/m³)	Standard Deviation	Sample Size (n)	T-Statistic	P-Value
Simple	224.4	70.77	80	25.88	< 0.001
Super	17.12	5.73	20	25.88	< 0.001

Note: The t-test confirms a statistically significant difference in average PM emissions between motorcycles using Simple and Super gasoline at the 95% confidence level ($p < 0.05$).

3.3 ANOVA and Tukey HSD Test

To assess the statistical significance of the observed differences, a one-way Analysis of Variance (ANOVA) was performed for three categorical factors: motorcycle age group, engine size group, and motorcycle model. The results indicated that motorcycle age group, and motorcycle model all had statistically significant effects on PM emissions. Motorcycle type also demonstrated a significant impact on emissions ($F = 4.88$, $p = 0.0013$). Post-hoc comparison showed that Honda Zoomer models emitted 170.14 μg/m³ more than Beat models, a difference that was statistically significant ($p = 0.001$). Similarly, the effect of motorcycle age was statistically significant ($F = 4.39$, $p = 0.015$). Tukey's test indicated that motorcycles aged 4–7 years emitted significantly more PM than those aged 0–3 years, with a mean difference of 67.19 μg/m³ ($p = 0.016$). However, no significant differences were found between other age group pairs. In contrast, engine size group was not a significant factor in PM emissions ($F = 1.75$, $p = 0.179$), suggesting that engine displacement alone does not sufficiently explain emission variability within the sample. The ANOVA results are summarized in table 3, highlighting that both motorcycle age and model had statistically significant impacts, while engine size did not show a significant difference.

Table 3. One-Way ANOVA Summary

Factor	F-value	p-value	Tukey Comparison	Mean Diff (μg/m³)	p-value
Age Group	4.39	0.015	0–3 yrs vs 4–7 yrs	67.19	0.016
Engine Size Group	1.75	0.179	-	-	-
Motorcycle Type	4.88	0.0013	Beat vs Zoomer	170.14	0.001

Note: Results show statistically significant differences in PM emissions at a 95% confidence level ($p < 0.05$).

These results emphasize the importance of fuel quality, vehicle age, and model type in determining motorcycle emissions and suggest that targeted interventions should prioritize these factors over engine size alone. To further explore the relationship between fuel type and motorcycle model, a two-way ANOVA with interaction was performed. The analysis revealed significant main effects of fuel type ($p < 0.001$) and motorcycle model ($p < 0.01$) on PM emissions. However, the interaction between fuel type and motorcycle model was not significant ($p > 0.05$). In table 4 provides a summary of the two-way ANOVA results, demonstrating that fuel type and model type individually had significant effects on emissions. This suggests that the influence of fuel quality on PM emissions was consistent across different motorcycle models and was not dependent on the specific model type. These results provide important insights into the factors influencing motorcycle emissions in urban academic environments. The following section discusses the broader implications of these findings.

Table 4. Two-Way ANOVA Summary

Source of Variation	F-value	p-value	Significance ($\alpha = 0.05$)
Fuel Type	45.23	< 0.001	Significant
Motorcycle Model	3.67	0.009	Significant
Fuel Type × Motorcycle Model Interaction	1.12	0.312	Not Significant
Residual	–	–	–

4. DISCUSSION

This study investigated the impact of motorcycle characteristics specifically fuel type, vehicle age, engine size, and model type on particulate matter (PM) emissions within the urban academic environment of the Institute of Technology of Cambodia (ITC). The results reveal several significant findings that have implications for air quality management and sustainable transport strategies in Cambodia. Among the analyzed factors, fuel type exhibited

the strongest and most statistically significant influence on PM emissions. The independent t-test confirmed that motorcycles fueled with Simple gasoline emitted significantly higher PM concentrations than those using Super gasoline ($p < 0.001$). This finding verifies previous studies across Southeast Asia, which associate lower-grade fuels with elevated emissions resulting from incomplete combustion [2,8]. These results highlight the critical need for stricter fuel quality regulations and the promotion of higher-grade fuels to mitigate urban air pollution, as also recommended in ADB's regional transport policy review [6]. Motorcycle age also showed a significant effect on emissions, as demonstrated by the one-way ANOVA results. Motorcycles aged 4–7 years emitted significantly more PM compared to newer motorcycles aged 0–3 years ($p = 0.016$). Although emissions from motorcycles older than 8 years did not differ significantly from the 4–7-year group, the overall trend indicates that emissions increase sharply after three years of use, likely due to engine wear and deterioration of fuel system components. This underscores the importance of implementing regular emissions inspections and encouraging the retirement of older motorcycles. Motorcycle model type significantly influenced PM emissions as well.

The one-way ANOVA and Tukey HSD tests revealed that Honda Zoomer and Honda Dream models emitted the highest levels of PM, while the Beat model exhibited the lowest emissions. This suggests that design differences, maintenance standards, and fuel system technologies may contribute to variability in emissions across motorcycle types. Interestingly, engine size group (100–110cc, 111–125cc, 126–150cc) did not significantly affect PM emissions ($p = 0.179$), despite initial expectations based on descriptive statistics. This indicates that engine displacement alone may not be a reliable predictor of emissions without considering additional factors such as engine technology, maintenance history, and riding patterns.

Furthermore, the two-way ANOVA confirmed significant main effects for both fuel type and motorcycle model on PM emissions but found no significant interaction effect between them. This suggests that the influence of fuel type on emissions is consistent across different motorcycle models, reinforcing the broader policy recommendation that fuel quality improvements will benefit the entire motorcycle fleet, regardless of specific model. While the analysis demonstrated a significant relationship between vehicle age and PM emissions, it is important to note that vehicle age may not always reflect the actual mechanical condition or emission performance of a motorcycle. For instance, regular engine maintenance, such as timely oil changes, tuning, and part replacement, can substantially reduce emissions even in older vehicles. Therefore, using age as a proxy for mechanical condition may introduce limitations. Future research should consider including maintenance records, cumulative mileage, or engine diagnostics to better capture real-world vehicle performance and its environmental impact.

Finally, the trend analysis of emissions over idle time indicated that PM emissions were highest during the first minute of engine operation (cold-start phase) and declined significantly thereafter. This behavior emphasizes the importance of minimizing engine idling, particularly in confined areas like university parking lots. Collectively, the findings support a multi-pronged strategy for reducing urban air pollution from motorcycles. Recommended actions include promoting the use of higher-grade fuels, encouraging the replacement of mid-aged motorcycles, targeting specific high-emission models for upgrades or replacement, and raising awareness about the harmful effects of prolonged idling. Implementing these measures could lead to substantial improvements in urban air quality, particularly in dense environments such as educational campuses.

5. CONCLUSIONS

This study examined the effects of key motorcycle characteristics—specifically fuel type, vehicle age, engine size, and model—on particulate matter (PM) emissions within the context of an urban university environment at the Institute of Technology of Cambodia (ITC). Based on data collected from 100 motorcycles under idle conditions, the study employed both descriptive and inferential statistical analyses, including independent t-tests, one-way ANOVA, and Tukey's HSD post-hoc tests, to determine the significance of each factor. The results indicated that fuel type, motorcycle age, and model type had statistically significant effects on PM emission levels. Motorcycles fueled with lower-grade Simple gasoline emitted significantly higher levels of PM compared to those using higher-grade Super gasoline, highlighting the critical role of fuel quality in emission control strategies. Similarly, motorcycles aged 4–7 years exhibited substantially higher emissions than those aged 0–3 years, underscoring the influence of aging on engine performance and combustion efficiency. In addition, certain motorcycle models—particularly the Honda Zoomer and Honda Dream—were identified as high-emission vehicles, suggesting that model-specific engineering and operational characteristics contribute significantly to pollutant output.

The study also emphasizes that while vehicle age emerged as a statistically significant variable, it may not fully capture the actual mechanical condition of a motorcycle. A well-maintained older vehicle may emit fewer pollutants than a newer one that is poorly maintained. This underscores the need for future studies to incorporate additional variables—such as maintenance history, oil change frequency, and total mileage—to better represent real-world emission variability. In contrast, engine size did not exhibit a statistically significant effect on PM emissions, suggesting that displacement alone is not a sufficient predictor of pollution levels. This finding highlights the importance of considering

a broader range of technological and operational factors when assessing motorcycle emissions.

Although the study provides meaningful insights, it is limited by a relatively small sample size (n = 100) and a focus solely on idle conditions. To improve the generalizability and policy relevance of future research, it is recommended that subsequent studies expand the sample size, include a wider range of motorcycle types and brands, and assess emissions under dynamic driving conditions. These enhancements would offer a more comprehensive understanding of real-world motorcycle emissions and support more effective air quality management strategies in urban environments.

Based on these insights; the study recommends a combination of policy measures aimed at reducing urban motorcycle emissions. These include:

- Promoting the use of higher-grade fuels through fuel quality standards and pricing incentives;
- Implementing fleet renewal programs and mandatory emission inspections for mid-aged motorcycles;
- Identifying and regulating high-emission models through targeted technical upgrades or replacement programs.

Future research should incorporate additional variables such as engine maintenance history, cumulative mileage, and fuel system technology (e.g., carbureted versus fuel-injected engines) to enhance the robustness and generalizability of the findings. Expanding the sample size and examining emissions under dynamic driving conditions rather than solely idling would provide a more comprehensive understanding of real-world motorcycle emission patterns in urban settings. Furthermore, investigating the influence of rider behavior, maintenance practices, and technological differences across motorcycle models could yield deeper insights into emission variability and inform more targeted mitigation strategies.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the Institute of Technology of Cambodia (ITC) for providing the opportunity, resources, and institutional support necessary to conduct this research. Special thanks are extended to the faculty members and research assistants whose dedication and expertise were invaluable during the data collection and analysis phases. The authors also wish to acknowledge the students and staff of ITC for their enthusiastic participation in the motorcycle usage survey. Their cooperation and input were essential to the success of this study.

REFERENCES

[1] Ministry of Public Works and Transport, "Urban Transport Developments in Cambodia," presented at the ASEAN Regional Transport Conference, Ha Long, Vietnam, Oct. 2023.

[2] A. D. Owen and J. Y. Tao, 'EMISSIONS AND FUEL STANDARDS: IMPLEMENTATION POSSIBILITIES FOR ASEAN', Asia Pac. J. Energy Environ., vol. 2, no. 1, 2014.

[3] N. T. Kim Oanh, M. T. Thuy Phuong, and D. A. Permadi, 'Analysis of motorcycle fleet in Hanoi for estimation of air pollution emission and climate mitigation co-benefit of technology implementation', Atmos. Environ., vol. 59, pp. 438–448, Nov. 2012, doi: 10.1016/j.atmosenv.2012.04.057.

[4] J.-H. Tsai, P.-H. Huang, and H.-L. Chiang, 'Air pollutants and toxic emissions of various mileage motorcycles for ECE driving cycles', Atmos. Environ., vol. 153, pp. 126–134, Mar. 2017, doi: 10.1016/j.atmosenv.2017.01.019.

[5] N. Zhang et al., 'Source profile and excess cancer risk evaluation of environmental tobacco smoking under real conditions, China', Atmospheric Pollut. Res., vol. 10, no. 6, pp. 1994–1999, Nov. 2019, doi: 10.1016/j.apr.2019.09.006.

[6] S. Gota and C. Huizenga, 'THE CONTOURS OF A NET ZERO EMISSION TRANSPORT SECTOR IN ASIA'.

[7] R. D. Brook et al., 'Particulate Matter Air Pollution and Cardiovascular Disease: An Update to the Scientific Statement from the American Heart Association', Circulation, vol. 121, no. 21, pp. 2331–2378, Jun. 2010, doi: 10.1161/CIR.0b013e3181dbece1.

[8] World Health Organization, WHO Global Air Quality Guidelines: Particulate Matter (PM2.5 and PM10), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide, Geneva, Switzerland: WHO, 2021.

[9] Y.-C. Ting, P.-K. Chang, P.-C. Hung, C. Chou, K. Chi, and T.-C. Hsiao, 'Characterizing emission factors and oxidative potential of motorcycle emissions in a real-world tunnel environment', Environ. Res., p. 116601, Jul. 2023, doi: 10.1016/j.envres.2023.116601.

[10] F. Borhani, M. Shafiepour Motlagh, A. Stohl, Y. Rashidi, and A. H. Ehsani, 'Changes in short-lived climate pollutants during the COVID-19 pandemic in Tehran, Iran', Environ. Monit. Assess., vol. 193, no. 6, p. 331, Jun. 2021, doi: 10.1007/s10661-021-09096-w.

[11] S. R. Kim, F. Dominici, and T. J. Buckley, 'Concentrations of vehicle-related air pollutants in an urban parking garage', Environ. Res., vol. 105, no. 3, pp. 291–299, Nov. 2007, doi: 10.1016/j.envres.2007.05.019.

[12] A. Mahalana, L. Yang, T. Dallmann, P. Lestari, K. Maulana, and N. Kusuma, 'Measurement of real-world motor vehicle emissions in Jakarta', 2022.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Correlation Between Dynamic Cone Penetrometer (DCP) and California Bearing Ratio (CBR) for Subgrade Soil Material in Cambodia

Bopheaktra PHO [1&2*], Narith SAUM [1], Tetsya SOK [2], Lyhour CHHAY[2]

[1] *Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.*

[2] *Public Infrastructure Department, General Directorate of Technics, Ministry of Public Works and Transport, Sangkat Chrang Chamres 2, Khan Russey Keo, Phnom Penh, Cambodia.*

[*]*Corresponding author: bopheaktra_pho@gsc.itc.edu.kh*

Abstract: *This research study investigates the correlation between the Dynamic Cone Penetrometer (DCP) and the California Bearing Ratio (CBR) for subgrade soil materials in Cambodia, which is essential for optimizing road pavement design. In Cambodia, there is a lack of a localized standard correlation equation customized to its unique soil and environmental conditions for estimating CBR value using DCP. This gap is particularly relevant with the recent requirement to use the DCP-9kg, replacing the previously used DCP-8kg, as specified in the updated Cambodian design specification D3 102-2024. The study aims to establish and evaluate a relationship of in-situ CBR resistance obtained from the DCP-9kg, develop a correlation equation between DCP-8kg and DCP-9kg, and formulate predictive equations. Field experiments were conducted across Cambodia using the DCP-8kg, DCP-9kg, and in-situ CBR equipment on various soil types, excluding gravel. The collected data was analyzed to formulate predictive equations. Experimental data were collected from 90 test locations across all 25 provinces and the capital city, covering diverse subgrade soils. DCP penetration depth measurements at increments of 50 mm to 250 mm were correlated with in-situ CBR tests. Results indicate a strong correlation at a 50mm depth, which aligns closely with the Australian standard. However, there were deviations of 5-7% for CBR values of 50-60% and 0.35-0.5% for 1-5%. Variations at deeper depths showed moderate to good correlations, with sandy soils showing weaker consistency. A correlation between DCP-8kg and DCP-9kg was also developed, noting that DCP-8kg produces higher CBR values above 30%. Field validation confirmed the reliability of the proposed equation for non-sandy soils. The findings contribute to the development of a localized standard for using DCP in Cambodia, which will enhance pavement design, road construction, and maintenance quality by providing a strong, site-specific CBR prediction model under local conditions.*

Keywords: Dynamic Cone Penetrometer (DCP), California Bearing Ratio (CBR), In-situ CBR, Correlation equation, Subgrade Soil

1. INTRODUCTION

In pavement design, investigating the strength of subgrade soil material is necessary to optimize structural safety and economic aspects of road infrastructures. The Dynamic Cone Penetrometer (DCP) test and the California Bearing Ratio (CBR) test are the key methods to evaluate subgrade material strength throughout the site investigation [1].

The Dynamic Cone Penetrometer (DCP), introduced by Scala (1956) [2] in Australia, is a device used to evaluate the resistance of soil at the actual site by driving a steel cone (20-30mm diameter) into the ground using a series of hammer blows to determine the in-situ CBR values, shear strength of soils, strata thickness, and bearing capacity of foundations and buildings [3]. The California Bearing Ratio (CBR) measures the strength of subgrade or base soil materials and its ability to support loads, commonly used in road design. It is obtained as the load ratio required to penetrate the compacted soil sample at a given moisture content and density compared to a standard crushed materials sample [4].

Since the laboratory CBR test is time-consuming and it cannot be readily determined in the field and the in-situ CBR test is also expensive, slow to conduct, DCP, being light and portable, takes a very short time to analyze and interpret, and cost-effectiveness with higher repeatability make it preferable for field use [1,5]. The DCP test quickly generates a

[1]* Corresponding author: Bopheaktra Pho
E-mail: bopheaktra_pho@gsc.itc.edu.kh; Tel: +855-16 672 147

continuous profile of in-situ subgrade and base strength measurements [6].

The Dynamic Cone Penetration (DCP) is utilized globally, including in Cambodia for road construction, however there is no official local standard or any study that clearly defines the correlation equation for CBR resistance using DCP. Previously, Cambodia used DCP-8kg equations, in accordance with Overseas Road Note 31 [7], which is derived from research in neighboring countries and other countries where environmental conditions, weather, soil, humidity, and materials are different, and the accuracy is limited compared to that used in the local. According to the new road design specification (D3 102-2024) [8] of the Ministry of Public Works and Transport (MPWT), which was updated in 2024 in accordance with the Australian standard [9], requires the use DCP-9kg instead of the previously used DCP-8kg, which differs in the In-situ CBR resistance correlation equation, analysis, and application.

To respond to new road design specification with the actual Cambodia's condition of road materials and soil conditions combined with hot and humid weather factors, new research is required. This paper objectively develops a correlation equation between DCP-8kg and DCP-9kg, both featuring a 16mm rod diameter and a cone tip with 60° apex angle and 20mm diameter The devices differ in drop height (575mm for DCP-8kg and 510mm for DCP-9kg). Additionally, the research establishes and evaluates a relationship of in-situ CBR resistance using a Dynamic Cone Penetration (DCP-9kg) based on the new road design specification (D3 102-2024) using a non-linear regression model.

Table 1. Summary of previous relationship between CBR and PI (DCP penetration index) [6]

Author	Correlation	Field or laboratory-based study	Material tested
Kleyn (1975) [10]	log (CBR) = 2.62 – 1.27 x log (PI)	Laboratory	Unknown
Harison (1987) [11]	log (CBR) = 2.56 – 1.16 x log (PI)	Laboratory	Cohesive
Harison (1987) [11]	log (CBR) = 3.03 – 1.51 x log (PI)	Laboratory	Granular
Livneh et al. (1995) [12]	log (CBR) = 2.46 – 1.12 x log (PI)	Field and Laboratory	Granular and cohesive
Ese et al. (1994) [13]	log (CBR) = 2.44 – 1.07 x log (PI)	Field and Laboratory	Aggregate base course
NCDOT (1998) [14]	log (CBR) = 2.60 – 1.07 x log (PI)	Field and Laboratory	Cohesive
Coonse (1999) [15]	log (CBR) = 2.53 – 1.14 x log (PI)	Laboratory	Piedmont residual soil
Gabr (2000) [16]	log (CBR) = 1.40 – 0.55 x log (PI)	Field and Laboratory	Aggregate base course

2. METHODOLOGY

2.1 Materials

The study will be conducted on different type of subgrade soils such as road subgrade, rural road, natural soil, field soil etc., except gravel soil, located in all 25 provinces and capital city of Cambodia. The data obtained in the experiment is 90 points. For these data, we have taken 80% for modeling and 20% for field validation on the model established. At each selected site, the in-situ CBR tests and both of DCP-8kg and DCP-9kg tests were carried out at the same site within $1m^2$ in area. Two to three tests in-situ CBR were performed in line at 150mm spacing. The DCP-8kg and DCP-9kg tests were performed adjacent to each other. Furthermore, two or three DCP penetrations were conducted approximately 1m apart with the in-situ CBR test location as the central point.

The in-situ CBR test was carried out in accordance with the American Standard ASTM D4429-09a [17] and was set up includes a loading machine, a 50mm diameter CBR plunger, load and penetration dial gauges, and surcharge plates. After setting up, perform the compression by press the plunger into the soil at a controlled rate, measuring the load at penetration depths of 2.54mm and 5.08mm. Calculate CBR is the ratio of the compressive strength of the soil at 2.54mm and 5.08mm compared to the standard compressive strength, which is shown in the formula below:

$$\text{CBR}(\%) = \frac{\text{Test Unit Load(kN)}}{\text{Standard Unit Load (kN)}} \times 100 \qquad \text{(Eq. 1)}$$

Where the standard compressive strength at 2.54mm is 13.44kN and the standard compressive strength at 5.08mm is 20.15kN.

The DCP-8kg consists of an 8kg hammer weight, a dropping height of 575mm, a calibrated 16mm stainless steel penetration rod, and replaceable hardened steel cone with 60° apex angle, 20mm diameter. This device was test according with Overseas Road Note 31 standard [7], which has the correlation equation is shown below:

$$\log(\text{CBR}) = 2.48 - 1.057\log(\text{mm/blow}) \quad \text{(Eq. 2)}$$

For DCP-9kg test was carried out in compliance with the Australian Standard AS 1289.6.3.2 [9], which consists of a 9kg hammer mass with a dropping height of 510mm, a 16mm steel rod, steel cone with a 20mm base diameter and 60° cone tip is attached. The standard correlation equation for predicting CBR from the DCP-9kg is shown below:

$$log\ (CBR) = 2.5117 - 1.1456log(mm/blow) \quad \text{(Eq. 3)}$$

2.2 CBR-DCPI correlation equation

After conducting field experiments between the depth of DCP and the in-situ CBR, the data obtained were analyzed by observing their relationship using the regression method to establish the relationship equation between DCP and in-situ CBR as well as the relationship between DCP-8kg and DCP-9kg. According to many international studies, as in Table 1, the data between DCP and in-situ CBR have a nonlinear relationship. Therefore, in this study, the equation to be established uses the Non-Linear Regression Model, which is expressed by the following equation:

$$log(CBR) = A * \log(DCPI) + B \quad \text{(Eq. 4)}$$

Where CBR is the predicted value of soil resistance at the site (%), DCPI is the value of the depth of penetration per hammer (mm/blow), and A, B are constant coefficients analyzed from the data to predict the CBR value.

The MATLAB software was used to calculate the coefficients A, B in the proposed equation model by evaluating the coefficient of determination (R^2). To verify that the proposed equation is acceptable, field validation was conducted. It should be noted that the development of this equation is to compare it with existing equations to evaluate whether the previous equation can be used on actual sites.

3. RESULTS AND DISCUSSION

3.1 Correlation Between DCP-9kg and In-situ CBR

The variation of penetration depth of the DCP was analyzed at depths of 50 mm (Fig. 1), 100 mm (Fig. 2), 150 mm (Fig. 3), 200 mm (Fig. 4), and 250 mm (Fig. 5), with the purpose of studying the relationship of the influence of the compressive strength of the In-situ CBR. Furthermore, the variation of penetration depth with the In-situ CBR was also analyzed by comparing it with the Australian standard.

Following scatter plots show how well the regression equations (lines) describe the observed relationship between the variation of settlement depth of the DCP and In-situ CBR for the tested soils, and to compare the local Cambodian equation (EQ) with the Australian standard equation.

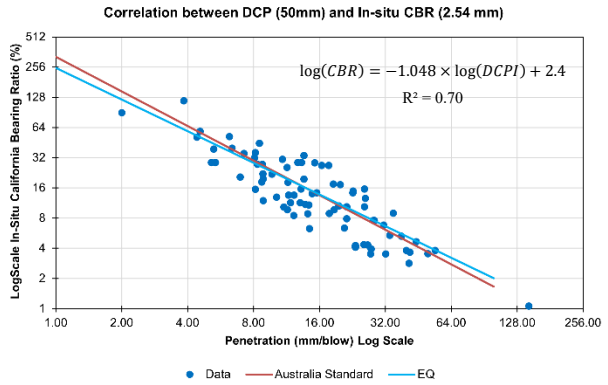


Fig. 1. Correlation between DCP at depth of 50mm and in-situ CBR

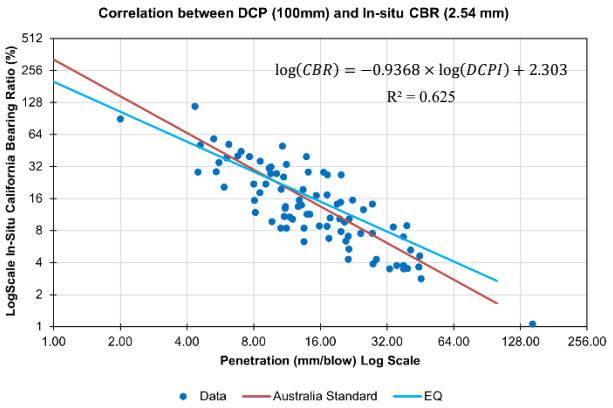


Fig. 2. Correlation between DCP at depth of 100mm and in-situ CBR

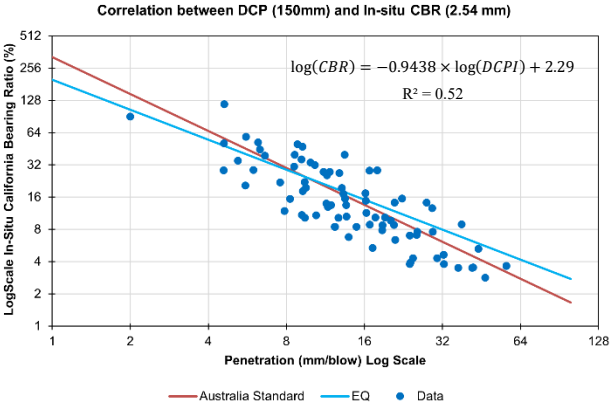


Fig. 3. Correlation between DCP at depth of 150mm and in-situ CBR

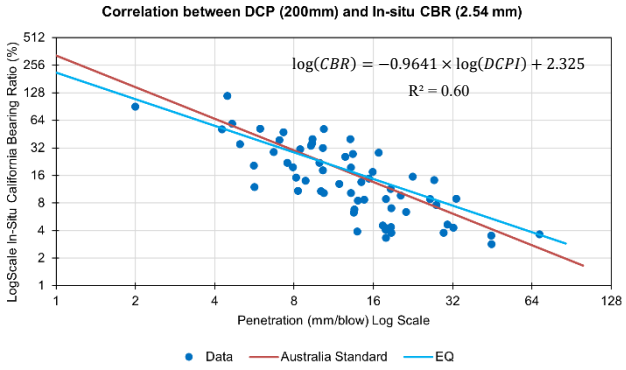


Fig. 4. Correlation between DCP at depth of 200mm and in-situ CBR

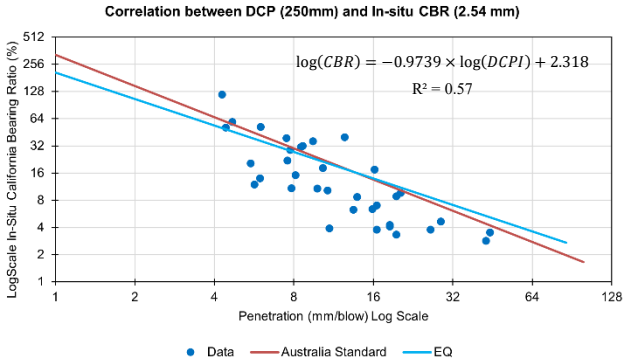


Fig. 5. Correlation between DCP at depth of 250mm and in-situ CBR

Based on the data analysis, we observed that the correlation equation between the variation of 50mm settlement depth and In-situ CBR exhibits a strong correlation with the highest R^2 coefficient (R^2=0.70), where the calculated equation coefficient is approximately similar to the equation from the Australian standard (R^2=0.85). However, upon examining the relationship between these two technical parameters, it is evident that the 50mm settlement depth shows the strongest correlation, as demonstrated in Table 2 and Fig. 1. Nevertheless, some data points exhibit a degree of scattering and do not respond consistently. In this experiment, we investigated various soil types, but a noteworthy observation is that sandy soil showed results that did not strongly correspond to the correlation between these two technical parameters. Therefore, we can preliminarily propose an equation for predicting CBR strength at the site based on the relationship with 50mm settlement depth, applicable to soil types other than sandy soil. This specification aligns with the Australian standard. When the CBR value ranges between 50% to 60%, the difference in CBR values between the Australian standard (Derived from equation Eq. 3) and the local experimental results (Proposed equation DCP50) is approximately 5% to 7% (of the CBR value). Meanwhile, when the CBR value ranges between 1% and 5%, the difference in CBR values between the Australian standard and the local experimental results is approximately 0.35% to 0.5% (of the CBR value). However, the Australian standard equation aligns well with the equation derived from the local experiment in cases where CBR ranges between 8% and 32%.

Table 2. Summary of the calculated equation coefficients

Description	Correlation Equation	R^2
Equation Australian Standard [9]	$\log(CBR) = -1.1456 \times \log(DCPI) + 2.512$	0.85
Equation in this study DCP50 - CBR	$\log(CBR) = -1.048 \times \log(DCPI) + 2.4$	0.70
Equation in this study DCP100 - CBR	$\log(CBR) = -0.9368 \times \log(DCPI) + 2.303$	0.625
Equation in this study DCP150 - CBR	$\log(CBR) = -0.9438 \times \log(DCPI) + 2.29$	0.52
Equation in this study DCP200 - CBR	$\log(CBR) = -0.9641 \times \log(DCPI) + 2.325$	0.60
Equation in this study DCP250 - CBR	$\log(CBR) = -0.9739 \times \log(DCPI) + 2.318$	0.57

The validation of the proposed equations, as in Table 2, has been evaluated through the above discussion. Fig. 6. illustrates the field validation between the predicted in-situ CBR derived from equation DCP50 and the in-situ CBR measured at the site. We observed that the validation of these two values is quite similar that the majority of predicted CBR values closely matched the measured in-situ CBR values, confirming the strongness and practical applicability of the proposed equation for Cambodian subgrade soils, as the data points fall within the line of equity. However, the proposed equation also has certain limitations under specific assumed conditions, especially for certain soil types (e.g., sandy soils), where the model's predictions may be less accurate.

The newly derived equation was shown to be reliable for predicting in-situ CBR from DCP results in most Cambodian soil conditions, as evidenced by the close alignment of predicted and measured values during field validation.

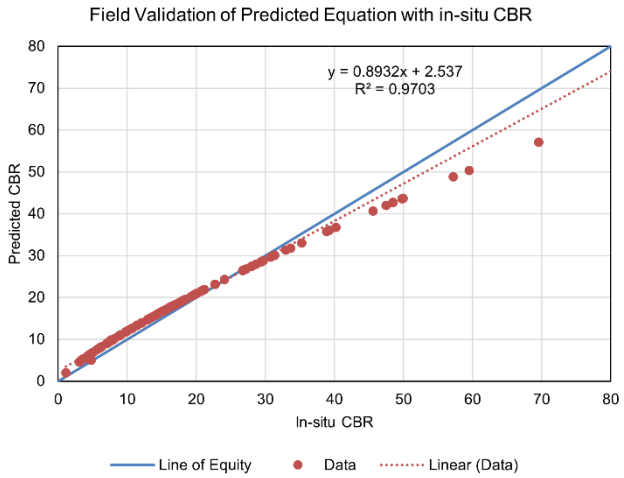


Fig. 6. Field validation of the predicted CBR with in-situ CBR

3.2 Correlation Between DCP-9kg and DCP-8kg

Taking into account the previous discussion and the experiment's aim of exploring the relationship between the DCP-9kg and DCP-8kg, this section describes the relationship between these two instruments. Fig. 7. illustrates the correlation between the CBR values of DCP-9kg derived from equation DCP50 and the CBR values of DCP-8kg derived from equation Eq. 2. We observe that the data from these two instruments exhibit a correlation with each other, which can be preliminarily used to predict In-situ CBR values by using the DCP-8kg to estimate the predicted In-situ CBR values of the DCP-9kg. However, based on the correlation equation, we note that the CBR from the DCP-8kg yields higher values than the DCP-9kg when CBR exceeds 30%.

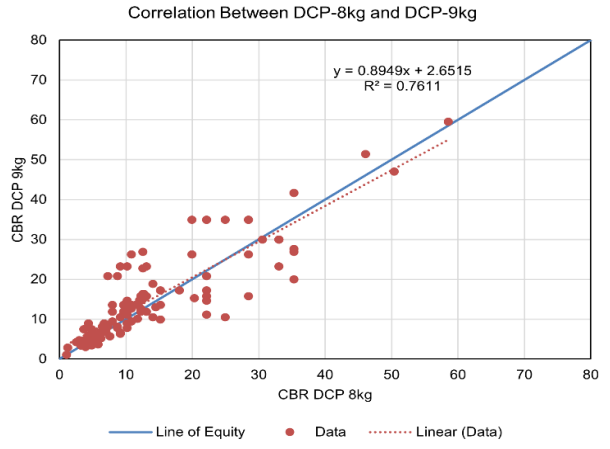


Fig. 7. Correlation Between DCP-8kg and DCP-9kg

4. CONCLUSIONS

The output of this study is a significant part of contributing to the evaluation of correlation equations between Dynamic Cone Penetrometer (DCP) and California Bearing Ratio (CBR) for subgrade soil materials used under Cambodian conditions, with the following conclusions:

- This research was conducted to confirm and establish standards for the use of the DCP (Dynamic Cone Penetrometer) tool to assess the quality of road layers, aiming to ensure the quality of road construction and maintenance. It provides concrete information about the condition and load-bearing capacity of soil layers.
- The research demonstrates the use of the DCP tool in analyzing CBR (California Bearing Ratio) values. It is observed that in-situ CBR values correlate with DCP penetration depth, which can be applied to predict actual CBR values at the site.
- The correlation equation between the DCP-9kg and in-situ CBR is effective at penetration depths ranging from 50mm to 150mm.
- The equation for the variation of 50mm penetration depth with in-situ CBR has the highest R^2 coefficient, with the equation's coefficient being approximately similar to the equation from the Australian standard.
- Based on the data analysis, the correlation equation between the DCP-9kg and In-situ CBR is:

$$\mathbf{Log}(\boldsymbol{CBR}) = -\mathbf{1.048} \times \mathbf{Log}(\boldsymbol{DCPI}) + \mathbf{2.4} \quad \text{(Eq. 5)}$$

- Sandy soil types (sandy soil and fine-grained soil) cannot be applied, as the experimental results at the site show errors consistent with the D3 102-2024 standard.
- Soil properties differ somewhat when compared to Australia, when CBR values range between 50% and 60%, the difference in CBR values between Australian standard and the local experimental results is approximately 5% to 7% (of the CBR value). When CBR values range between 1% and 5%, the difference in CBR values between Australia standard and the local experimental results is approximately 0.35% to 0.5% (of the CBR value). However, the Australian standard equation aligns with the equation derived from the local experiment in cases where CBR ranges between 8% and 32%.

ACKNOWLEDGMENTS

The first author wishes to thank his family for giving the majority of financial and emotional assistance throughout the project. The authors are also grateful to the Transport Study Unit (TSU) and the Ministry of Public Works and Transport (MPWT) for providing extra funds and technical support.

REFERENCES

[1] Feleke, G. G., & Araya, A. A. (2016). Prediction of CBR using DCP for local subgrade materials. International Conference on Transportation and Road Research, Mombasa, Kenya.

[2] Scala, A. J. (1956). Simple methods of flexible pavement design using cone penetrometers. New Zealand Engineering, 11(2), 34–44.

[3] Livneh, M. (1987). The use of dynamic cone Penetrometer in determining the strength of existing pavements and subgrades. Southeast Asian Geotechnical Conference. 9, 1–10.

[4] Montejo Fonseca, A. (2006). Ingeniería de pavimentos (3a ed). Universidad Católica de Colombia.

[5] Al-Refeai, T., & Al-Suhaibani, A. (1997). Prediction of CBR Using Dynamic Cone Penetrometer. Journal of King Saud University - Engineering Sciences, 9(2), 191–203. https://doi.org/10.1016/S1018-3639(18)30676-7

[6] Aldawi, M. (2016). Correlation Between the Standard Penetration Test and The Dynamic Cone Penetration Test for Sandy Soil [PhD Thesis, University of Tripoli].

[7] Rolt, J. et al., (2023). Road Note 31: A Guide to the Structural Design of Surfaced Roads in Tropical and Sub-Tropical Regions (5th ed.). Transport Research Laboratory (TRL).

[8] MPWT D3 102-2024. (2024). Road Design Specification, Part 2, Pavement, 1st ed.; Ministry of Public Works and Transport.

[9] Australian Standard AS 1289.6.3.2. (1997). Methods of testing soil for engineering purposes- Soil strength and consolidation tests - Determination of the penetration resistance of a soil – 9 kg dynamic cone penetrometer test, Sydney, New South Wales, Australia.

[10] Kleyn, E. G. (1975). The Use of the Dynamic Cone Penetrometer (DCP). Transvaal Provincial Administration.

[11] Harison, J. (1987). TECHNICAL NOTE. CORRELATION BETWEEN CALIFORNIA BEARING RATIO AND DYNAMIC CONE PENETROM ETER STRENGTH MEASUREMENT OF SOILS. TECHNICAL NOTE 463. Proceedings of the Institution of Civil Engineers, 83(4), 833–844. https://doi.org/10.1680/iicep.1987.204

[12] Livneh, M., Ishai, I., & Livneh, N. A. (1995). Effect of vertical confinement on dynamic cone penetrometer strength values in pavement and subgrade evaluations. Transportation Research Record, 1–1.

[13] Ese, D., Myre, J., Noss, P., & Vaernes, E. (1994). The use of dynamic cone penetrometer (DCP) for road strengthening design in Norway. 4th International Conference, Bearing Capacity of Roads and AirfieldsFHWA, U of Minnesota, Army Corps of Engineers, NRC Canada, FAA, 1. https://trid.trb.org/View/469210

[14] NCDOT. (1998). Pavement Design Manual. The North Carolina Department of Transportation.

[15] Coonse, J. W. (1999). Estimating California bearing ration of cohesive Piedmont residual soil using the Scala dynamic cone penetrometer [PhD Thesis]. North Carolina State University.

[16] Gabr, M. A., Hopkins, K., Coonse, J., & Hearne, T. (2000). DCP Criteria for Performance Evaluation of Pavement Layers. Journal of Performance of Constructed Facilities, 14(4), 141–148. https://doi.org/10.1061/(ASCE)0887-3828(2000)14:4(141)

[17] ASTM International. (2009). ASTM D4429-09a: Standard Test Method for CBR (California Bearing Ratio) of Soils in Place (Withdrawn 2018). ASTM International.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Geotechnical Properties of Clay Soil with River Sand for Sustainable Road Embankment Applications

Samnang Sok [1], Mankong Yin [1], Sarong Ri [1], Kean Thai Chhun [1, 2*]

[1] *Department of Research and Development, Techo Sen Institute of Public Works and Transport, St. 598, Phnom Penh, 120705, Cambodia.*

[2] *Ministry of Public Works and Transport, St. 598, Phnom Penh, 120705, Cambodia.*

[*] *Corresponding author: kt_chhun@yahoo.com*

Abstract: *Fluctuations in clay soil moisture lead to expansion and contraction, causing damage to foundations and road surfaces. This paper experimentally investigated the effect of river sand on the geotechnical properties of clay soil for sustainable road embankment applications. Numerous laboratory tests were adopted to examine the effect of river sand (0, 10, 30, and 50% by weight of dry clay) on the geotechnical properties of clay soil. Results indicated that increasing river sand content significantly increased maximum dry density (MDD) and California bearing ratio (CBR) and also led to a noticeable reduction in the plasticity, optimum moisture content (OMC), and swelling percent of the clay soil. From this, it was suggested that at least 50% of river sand needs to be mixed with clay soil to obtain ≥ 2 % of the CBR (i.e., requirement of the embankment), thereby reducing plasticity and swelling, and providing a cost-effective and environmentally friendly approach for road construction.*

Keywords: Clay; River Sand; Embankment; Geotechnical Properties

1. INTRODUCTION

Road embankments are critical components of transportation infrastructure, providing stable foundations for roads in diverse geotechnical conditions. However, in many regions of Cambodia, the prevalence of clay soils poses significant challenges for embankment construction due to their low bearing capacity, high compressibility, and poor drainage. These limitations often result in settlement, slope instability, and higher maintenance costs, impeding sustainable infrastructure development [1-2].

Soil stabilization techniques have been widely explored to address these challenges, with the incorporation of sand emerging as a cost-effective and environmentally friendly solution [3-4]. Previous studies have demonstrated that adding sand to clayey soils can significantly increase the engineering properties. Alnmr and Ray [5] investigated the effect of sand concentration on the physical properties of expansive clayey soil, and it found that incorporating sand into clay reduced plasticity and improved compaction density, making the soil more workable for construction. Endaryanta and Wibowo [6] reported that beach sand-clay mixtures demonstrated the increase both soaked and unsoaked California Bearing Ratio (CBR) values and reduced swelling. Moreover, the authors concluded that the optimum beach sand content of the beach sand-clay was about 85%. Al-Badran and Al-Ameri [7] conducted a laboratory study on the effect of adding sand on the shear strength characteristics of clay. They found that the friction angle of the soil increased with higher sand concentrations, while cohesion decreased as the sand percentage increased. The study proposed a correlation between the shear strength properties of clay and sand content. On the other hand, river sand, which is abundantly available in Cambodia, can also improve the shear strength, compaction characteristics, and permeability of clayey soils, making them more suitable for embankment construction. However, the optimal proportion of river sand-clay mixtures and their long-term performance under Cambodian climatic conditions remain areas requiring further investigation.

Therefore, in this study examines the geotechnical properties

[1*] Corresponding author: Kean Thai CHHUN
E-mail kt_chhun@yahoo.com; Tel: +855-95 935 466

of clayey soil mixed with river sand to assess its feasibility for sustainable road embankments in Cambodia. Key parameters such as Atterberg limits, compaction characteristics, and California Bearing Ratio (CBR) are evaluated to determine the most effective mix ratio. The findings aim to contribute to cost-effective and eco-friendly road construction practices, aligning with Cambodia's growing infrastructure needs while minimizing environmental impact.

2. METHODOLOGY

2.1 Material Used

In this study, both disturbed soil samples (clayed soil and river sand) were collected from Kampung Os, Ponhea Lueu, Kandal, Cambodia, as shown in Fig. 1. The clayed soil was collected at a depth of about 0.5 to 1 m. The soil specimens were oven-dried for 24 hours and the laboratory tests were then conducted to examine the basic properties. The basic properties of the soil samples were listed in Table 1.



Fig. 1. Location of the sample collected.

Table 1. Basic properties of the material used.

Parameters		Soil Sample	
		Clayed Soil	River Sand
Specific Gravity		2.661	2.632
Sieve Analysis	Fine (%)	99.27	5.24
	Sand (%)	0.73	94.76
	Gravel (%)	0	0
Atterberg Limit	Liquid Limit (%)	57.09	-
	Plastic Limit (%)	17.97	N.P
	Plastic Index (%)	39.12	-
Soil Classification		CL	SM
Compaction	Max. Dry Density (g/cm^3)	1.585	1.571
	Optimum Moisture Content (%)	21.08	15.22

2.2 Laboratory testing program

Numerous laboratory tests, including the Atterberg limits, compaction, and California Bearing Ratio (CBR), were conducted to evaluate the engineering properties of the mixed soil samples, as listed in Table 2. Additionally, five types of clay-river sand mixtures were prepared with varying clay content by weight. The river sand and clay soil were manually mixed for approximately 10 minutes to ensure homogeneity, and the resulting soil samples were directly stored in a curing chamber maintained at a temperature of 20 ± 2 °C.

In this study, laboratory tests were conducted following the American Society for Testing and Materials (ASTM) standards. The Atterberg limits (liquid limit, plastic limit, and plasticity index) were determined in accordance with ASTM D4318 for all mixture samples. The dry clay-river sand mixture was prepared by passing it through a No. 40 sieve (0.425 mm). Also, compaction tests were performed using the Standard Proctor method (ASTM D698) with a 24.4 N hammer and a mold volume of 943 cm³. The mixed soil was compacted in 3 layers, with each layer obtaining 25 blows from hammer dropped at a height of about 30 cm. From this test, the MDD and OMC was determined based on the relationship between the moisture content and density. To evaluate bearing capacity of the dry clay-river sand mixtures, the CBR test was also performed by following ASTM D1883. Based on the result of OMC, the dry clay-river sand mixtures were mixed with a predetermined amount of water, and the mixed soil was reconstructed in 3 different compactive efforts (10, 25, and 56 blows per layer). The soil samples were then soaked for 96 hours.

Table 2. Laboratory testing program of clay soil with river sand

No.	Test condition	Test	Standard
1	Clay + 0%, 10%, 30%, and 50% River Sand	Atterberg Limit	ASTM D-4318 [8]
2		Compaction	ASTM D-698 [9]
3		CBR	ASTM D-1883 [10]

3. RESULTS AND DISCUSSION

3.1 Atterberg Limit

The result of the Atterberg limit test with various test conditions was shown in Fig. 2. From this, it was noted that the value of liquid limit (LL), plastic limit (PL), and plastic index (PI) of clayey soil significantly decreased with an increase in the river sand content. This is due to a rise in sand content induces the reduction of the plasticity, surface area, and water retention of the clay-river sand mixture. This is in good agreement with the previous study by Alnmr and Ray [5], Aziz [11], and Qusai et al. [12].

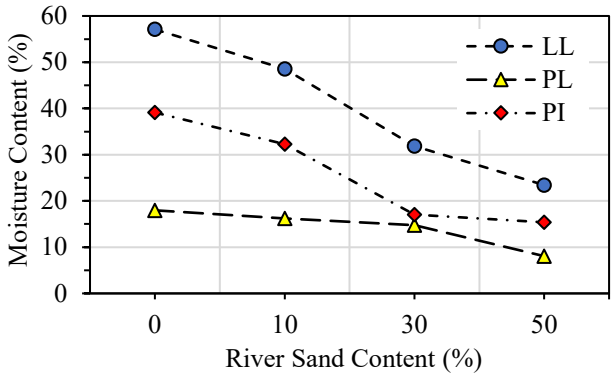


Fig. 2. Result of the Atterberg limit

3.2 Compaction Characteristics

Figure 3 presents the results of the compaction test. Based on the testing results, it was noted that the MDD of the soil sample increased and the OMC of the soil sample decreased as the percentage of river sand increased from 0 to 50%. This increase can be attributed to the river sand particles filling the voids between clay particles, leading to a denser structure. While the decrease of OMC can be due to the sand particles having a lower surface area and water retention than clay particles, and thus cannot hold as much water against gravity. Overall, the increase of MMD at lower OMC is often observed with sand addition and this result was consistent with the previous study by Alnmr and Ray [5] and Alkroosh et al. [13].

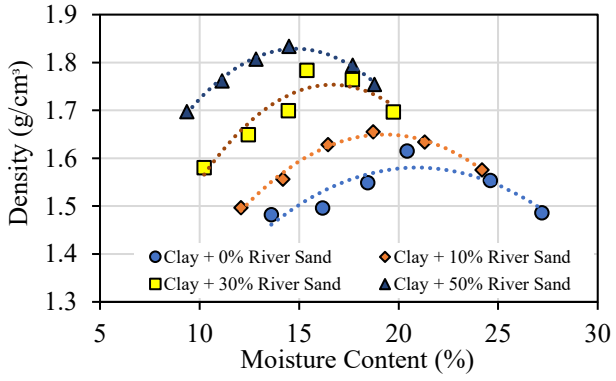


Fig. 3. Result of the compaction test

3.3 California Bearing Ratio

Figure 4 shows the result of the CBR test. From this, it was obvious that the value of CBR increased from 0.95 to 2.03% while the swelling percentage decreased from 1.204 to 0.711% with the increase of river sand concentration of the clay-sand mixture. Moreover, the enhancement of CBR value due to the river sand particles provides better bonding, increasing the mixture's shear strength and load-bearing capacity [6 & 14-15].

To evaluate the effectiveness of river sand and clay for soil stabilization, it is essential to compare the CBR value with the Cambodian standard criteria for the embankment layer [16]. The test results indicate that a mixture of at least 50% river sand and clayey soil achieves a high CBR value and a low swelling percentage, meeting the Cambodian standard criteria for the embankment layer.

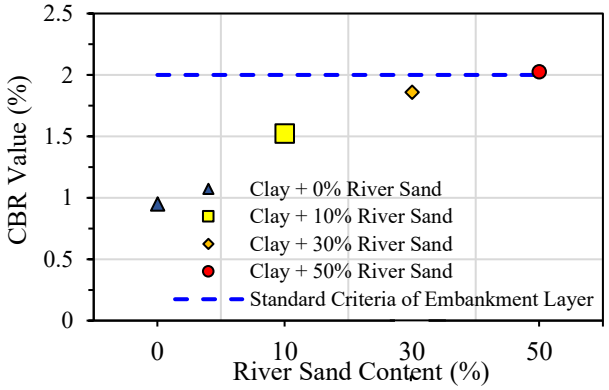


Fig. 4. Result of the CBR test

4. CONCLUSIONS

This study investigates the effect of river sand on the geotechnical properties of clay soil for sustainable road embankment applications. Laboratory tests, including Atterberg limits, compaction, and CBR were conducted on clay-river sand mixtures. The results demonstrate that incorporating river sand improves the soil's engineering properties by reducing plasticity, optimum moisture content, and swelling while increasing maximum dry density and strength. To meet the Cambodian embankment criteria (CBR $\geq$ 2% and swelling $\leq$ 4%), a 50% river sand content is recommended. In summary, the stabilized clay-sand mixture enhances durability and cost-effectiveness for road construction.

ACKNOWLEDGMENTS

This work was granted aid from Government of Republic of Korea through Ministry of Land, Infrastructure, and Transport of Republic of Korea (MOLIT) under Renewal Project of Road Construction and Management System in Cambodia. The author gratefully acknowledges Dr. Lok Vanno for his invaluable support and contributions to this project.

REFERENCES

[1] Nguyen, T. H., & Tran, Q. D. (2021). Stabilization of soft clay soils for road construction using granular materials: A case study in Mekong Delta. Geotechnical Engineering Journal, 52(3), pp.145-160.

[2] Smith, J. R., Kumar, A., & Lee, W. F. (2020). Geotechnical challenges in embankment construction on clay-rich soils: A review. Journal of Sustainable Infrastructure, 8(2), pp.78-95.

[3] Alnmr, A. and Ray, R., 2024. Investigating the impact of varying sand content on the physical characteristics of expansive clay soils from Syria. Geotechnical and Geological Engineering, 42(4), pp.2675-2691.

[4] Zhang, L., Wang, H., & Singh, R. M. (2019). Improving the mechanical properties of clayey soils using river sand: Laboratory and field evaluations. Transportation Geotechnics, 21, 100281.

[5] Alnmr, A. and Ray, R., 2024. Investigating the impact of varying sand content on the physical characteristics of expansive clay soils from Syria. Geotechnical and Geological Engineering, 42(4), pp.2675-2691.

[6] Endaryanta and Wibowo, D.E., 2020, September. Strengthening and Supporting Efforts to Reduce Swelling of Soil by Using Beach Sands Through CBR Test. In Journal of Physics: Conference Series (Vol. 1625, No. 1, p. 012008). IOP Publishing.

[7] Al-Badran, Y.M. and Al-Ameri, A.F., 2020, June. Effect of Adding Sand on Clayey Soil Shear Strength. In IOP Conference Series: Materials Science and Engineering (Vol. 870, No. 1, p. 012079). IOP Publishing.

[8] ASTM D4318, 2010. Standard test methods for liquid limit, plastic limit, and plasticity index of soils. Annual Book of ASTM Standards, American Society for Testing and Materials, West Conshohocken, Pennsylvania, U.S.A., pp.1-14.

[9] ASTM D698, 2007. Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Standard Effort (12,400 Ft-lbf/ft3 (600 kN-m/m^3)). ASTM International. Annual Book of ASTM Standards, American Society for Testing and Materials, West Conshohocken, Pennsylvania, U.S.A., pp.1-14.

[10] ASTM D1883, Standard Test Method for California Bearing Ratio (CBR) of Laboratory-Compacted Soils. Annual Book of ASTM Standards, American Society for Testing and Materials, West Conshohocken, Pennsylvania, U.S.A., pp.1-8.

[11] Aziz, M., 2023. Mechanical properties of a high plasticity clay mixed with sand and low-plastic silt. Materials Today: Proceedings.

[12] Qusai, A., Szendefy, J. and Vásárhelyi, B., 2025. Effect of Changing Sand Content on Liquid Limit and Plasticity Index of Clay. Geotechnics, 5(1), pp.1-11.

[13] Alkroosh, I., Al-Robay, A., Sarker, P. and Alzabeebee, S., 2021, July. Effect of sand percentage on the compaction properties and undrained shear strength of low plasticity soft clay. In *ACES* (pp. 157-161). Cihan University-Erbil.

[14] Roy, T.K., 2013. Influence of sand on strength characteristics of cohesive soil for using as subgrade of road. Procedia-Social and Behavioral Sciences, 104, pp.218-224.

[15] Cabalar, A.F. and Mustafa, W.S., 2017. Behaviour of sand–clay mixtures for road pavement subgrade. International Journal of Pavement Engineering, 18(8), pp.714-726.

[16] Cabalar, A.F. and Mustafa, W.S., 2017. Behaviour of sand–clay mixtures for road pavement subgrade. International Journal of Pavement Engineering, 18(8), pp.714-726.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Machine Learning-Based Prediction of Public Bus Ridership: A Case Study in Phnom Penh City

Piseth SRORN [1*], Veng Kheang PHUN [1], Narith SAUM [1]

[1] *Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.*

[*]*Corresponding author: pisethsrorn168@gmail.com*

Abstract: *The public bus system plays a crucial role in improving urban mobility, particularly in developing cities like Phnom Penh, Cambodia, where it is operated by the City Bus Authority (CBA). However, the lack of alignment between bus capacity and real-time ridership demand undermines service effectiveness, making the accurate prediction of public bus ridership essential for improving operational efficiency, service quality, and ridership satisfaction. To address this issue, this study applied machine learning techniques to predict public bus ridership using historical data, including temperature, temporal features (e.g., time of day and day of the week), and weather conditions. The predictive models in this research utilized a traditional statistical method like Linear Regression, and machine learning (ML) models such as Random Forest Regression (RF) and eXtreme Gradient Boosting (XGBoost) to enhance predictive accuracy. The effectiveness of these models is evaluated through standard performance metrics by Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE), providing a comprehensive analysis of predictive performance. The results demonstrate that the XGBoost model outperforms Linear regression and Random Forest Regression in capturing complex ridership patterns, achieving an MAE of 18.93 and an RMSE of 26.37, representing a 6.79% and 2.72% improvement over the Linear Regression and Random Forest, respectively. These key findings highlight the significance of accurate ridership prediction in optimizing bus frequency and adjusting bus capacity to better align with passenger demand. This research underscores the importance of machine learning in supporting data-driven decision-making for public transportation systems. Furthermore, future research may explore deep learning models to further improve the accuracy of ridership predictions, contributing to more efficient and sustainable urban mobility solutions.*

Keywords: Public Bus Ridership, Machine Learning, Prediction, Urban Transportation, Phnom Penh

1. INTRODUCTION[1]

Public transportation systems are essential for urban mobility, providing affordable and accessible travel for people to work, school, attend events, and other locations. Public bus transportation services are impacted by traffic congestion, weather conditions, and traffic signals [1]. Predicting demand accurately also allow operators to efficiently allocate resource [2]. Actually, machine learning has obtained significant for accurately modeling of complex relationships between inputs and outputs variables [3]. The application of machine learning algorithms is preferred over traditional methods for their accurate prediction performance [5, 6]. Despite increasing ridership for public transport in Phnom Penh, no exiting study has applied machine learning models to predict the public bus ridership. However, several international research have studied the effectiveness of advance models in predicting ridership of public transit. According to Palacio (2018), XGBoost outperformed better results based on error measurement [4]. Zhao et al. (2011) noted that traditional predictive methods always lead to inaccurate results [5]. Ruiz et al. (2024) found that Random Forest performed better in predicting ridership [6]. Wood et al. (2023) demonstrated that both Random Forest and Linear Regression models can effectively predict future bus ridership [7]. Mei et al.(2020) further illustrated that the XGBoost model can predict ridership volume more accurately [8]. Therefore. The objective of this research is to conduct a statistical analysis of public bus ridership to understand underlying patterns and explore key features such as departure, headway, day of week, time of day, and temperature, weather condition [9] that contribute to demand prediction. Then, all these features are used as inputs for ridership prediction models Linear Regression, Random

[1*] Corresponding author: Piseth Srorn
E-mail: pisethsrorn168@gmail.com; Tel: +855-97 407 6959

Forest, XGBoost, while their influence are presented as importance features.

2. METHODOLOGY

2.1 Research Framework

This stage outlines a machine learning framework for predicting city bus ridership. It starts with data collection, incorporating key features such as historical bus ridership, departure time, temperature, temporal characteristics, and weather condition. Then, exploratory data analysis (EDA) and descriptive statistics are explored to understand data and uncover patterns with the data. Next, data preprocessing for ML ensures the dataset was cleaned and properly structured for training the model. Feature selection assists in identifying relevant attributes. Furthermore, the data was split into training sets and testing set, and machine learning models such as Linear regression, Random Forest, and XGBoost are trained. To optimize the performance of the models, Optuna was the hyperparameter optimization techniques, and the model needs to be evaluated by using metrics such as mean absolute error (MAE) and root mean square error (RMSE). Finally, the best-performing model was selected for accurate prediction. This simulation environment was implemented using the Python programming language and executed on Google Colab.

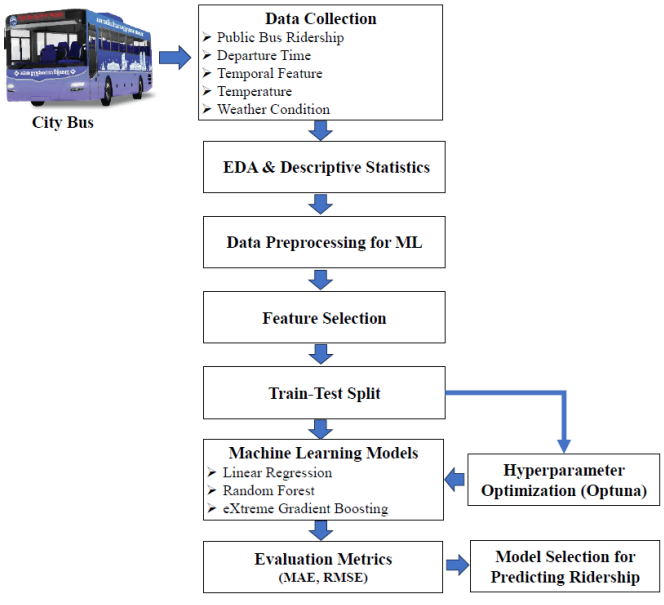


Fig.1. Research Framework

2.2 Raw Datasets

Nowadays, there are 13 bus lines serving in Phnom Penh City. Among these, line1, line 2, line 3 and line 4 have relatively high ridership compare to other lines, with line 3 having the highest amount of ridership. According to the dataset, it was manual data collection consists of three components such as departure time, arrival time, amount of ridership counted at each bus stop along the route, from the starting terminal to the final terminal by each trip, which operates from km9 to the Borey Santepheap 2 terminal, covering a total distance of 27.3 km. The route consists of 24 bus operations and 60 bus stops. Due to the dataset is limited, which was collected over a short period from October 28, 2024, to December 22, 2024, with a total of 3,329 samples.

Table 1. Raw Datasets

Departure	Arrival	Amount of ridership
5:30	6:45	38
5:40	6:54	21
5:50	7:10	25
6:00	7:21	49
6:10	7:30	15

2.3 Data Preprocessing

From the raw dataset, ridership was recorded by trips as the total amount of ridership per trip, not specific station during each predict interval (30 minute in this study). From existing features, new variables were considered based on previous studies. However, most machine learning requires numerical input and one-hot encoding of categorical variables to 0 or 1 [10,11]. Input: feature selections included departure time, departure 30-interval, headway, time of day, day of the week, holiday indicator, Am peak (7AM to 9AM), Pm peak (5PM to 7PM), temperature, and weather conditions. Process: The dataset was divided into features and the target variable, where the ridership count served as the target. Then, the dataset was split into 70% for training and 30% for testing to evaluate model performance. Output: The short-term public bus ridership prediction.

2.4 Prediction Models

The prediction models examined in this study are linear regression, random forest, and XGBoost. The model processes historical data to train and test these models, generating predictions based on key input variables. Linear regression provides a baseline trend analysis, while random forest improves model prediction by capturing non-linear relationships, while XGBoost further improve model performance by learning complex patterns in the dataset.

2.4.1 Linear regression

Linear regression is a basic regression model in statistics and machine learning that predicts the output variable based on a linear relationship with one or more input variables. This model creates mathematical equation to estimate unknown values. This model is easy to understand, fast to compute, and simple to use [12]. The equation for linear regression is:

$$y_i = \sum_{i=1}^{n} \beta_i x_i + \beta_o + \varepsilon \quad \text{(Eq.1)}$$

Where, ε is the error term, which has a mean of 0 and is distributed normally, β_i is calculated using the Ordinary Least Squares method, β_0 is intercept, x_i is the symbolizes the independent variables that were used to predict the dependent result, y_i is gives the set of independent variables and the expected value [7].

2.4.2 Random Forest Regression

In machine learning, random forest is an ensemble method that combine multiple decision trees and known as bootstrap and aggregation to accomplish both regression and classification tasks [11,12]. In Python, Random Forest can implement using libraries such as Pandas, numpy, scikit-learn.

$$\hat{y}_{rf}^{n}(x) = \frac{1}{n}\sum_{i=1}^{n} T_i(x) \quad \text{(Eq.2)}$$

Where, n is the total number of trees in the forest, $T_i(x)$ is the prediction of the i-th tree for the input feature vector x, and $\hat{y}_{rf}^{n}(x)$ is the prediction value.

2.4.3 XGBoost

XGBoost is one sort of boosting tree model aiming to create certain tree models for prediction [15]. The machine learning technique for both regression and classification tasks is gradient boosting, and it builds a predictive model by combining multiple weak learners, typically decision trees into a stronger ensemble model [14]. The XGBoost model integrates K tree, the predict ridership can be express as:

$$\hat{y}_i = \sum_{k=1}^{k} f_k(x_i) \qquad f_k \in \mathcal{F} \quad \text{(Eq.3)}$$

Where, f represented the regression tree in $\mathcal{F}$, $\mathcal{F} = f(x) = \omega_{q(x)}$ (also known as CART). The model aims to minimize an objective function:

$$L^{(t)} = \sum_{i=1}^{n} \left[l\left(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)\right)\right] + \sum_{k=1}^{K} \Omega(f_k) \quad \text{(Eq.4)}$$

$$\Omega(f_k) = \frac{1}{2}\eta \sum_{j=1}^{T} \omega_j^2 + \gamma T,$$

$\sum_{i=1}^{n} l\left(y_i, \hat{y}_i^{(t-1)}\right)$ represented the loss function, which measures the prediction error, whereas $\sum_{k=1}^{K} \Omega(f_k)$ represented of regularization terms, which reflect the complexity of model. ω_j is the weight assigned to the j-th leaf nodes. $\eta\Sigma_{j=1}^{T}\omega_j^2$ is the regularization term of L2, γ is the coefficient to control the complexity of model by setting the number of leaf nodes. j is the leaf nodes number, $f_t(x_i)$ is the model at the t-th iteration. Once second-order Tylor expansion applied to the error term of the objective function, and $l\left(y_i, \hat{y}_i^{(t-1)}\right)$ is the content that was no effect on the optimization of the objective function. The algorithm can be simplified objective function:

$$L^{(t)} = \sum_{i=1}^{n} \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)\right] + \Omega(f_t) \quad \text{(Eq.5)}$$

Where, $g_i = \partial_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$: first derivation(gradient), and $h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\left(y_i, \hat{y}_i^{(t-1)}\right)$: second derivation (Hessian). In the regression tree, each simple will be ends up in one specific leaf node. Thus, the final equation of objective function can choose the best tree:

$$L^{(t)} = \sum_{j=1}^{T} \left[G_j \omega_j + \frac{1}{2}\left(H_j + \eta\right)\omega_j^2\right] + \gamma T \quad \text{(Eq.6)}$$

Where, $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$,

2.5 Hyperparameter Tunning

A machine learning model consists of parameters that learns from data as known as hyperparameters. These settings can be adjusted to enhance the model's performance by training it on existing data. The process of selecting the optimal values before the training phase is referred to as hyperparameter tunning. In this study, hyperparameter tunning was implemented for both machine learning models. For random forest's parameters such as number of decision tree (n_estimators), the maximum depth of each tree (max_depth) were tuned. The setting of XGBoost model were optimized such as n_estimators, max_depth, learning_rate, subsample, colsample_bytree, gamma, reg_lambda (L2 regularization term), reg_alpha (L1 regularization term). Both models were used Optuna for tunning hyperparameters. it was a modern optimization, which created a probabilistic model of the objective function and utilized it to select hyperparameters that improve performance. This method was utilized to find parameters that can produce the most optimal value performance in the models [16].

2.6 Evaluation Metrics

The mean absolute error and root mean square error has been frequently used as evaluate criteria in previous studies to measures the quality of prediction models [15].

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n} | y_{actual,i} - \hat{y}_{prediction,i} | \quad \text{(Eq.7)}$$

$$\text{RMSE}= \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_{actual,i} - \hat{y}_{prediction,i}\right)^2} \quad \text{(Eq.8)}$$

Where, y_{actual} is the actual value of the observation of number of passengers, $\hat{y}_{prediction}$ is the prediction value of the i-th observation, n is the number of observations.

3. RESULTS AND DISCUSSION

3.1 Descriptive statistics

Table 2 presents the descriptive statistics of ridership of line 3. The result demonstrated that the average ridership per trip is approximately 45, with range from 0 to 180 and a standard deviation of 33.42, illustrating high variability in ridership demand each trip. Furthermore, the average of headway from 13.31min. In terms of mean of vehicle capacity of 33.69 ridership seats, while the bus size ranges from 14 to 41 seats in the network with three type of buses such as Japanese bus with 41 seats, Korean bus with 14 seats, and Chinese bus with 35 seats, comparing to maximum of ridership in some case could be overcrowding of ridership. Lastly, the average of temperature of 29.17 ℃ (±2.56℃) might influence in ridership pattern.

Table 2. Descriptive statistics (n=3,329 samples)

Variable	Mean	Std	Min	Max
Ridership (pax/trip)	44.95	33.42	0	180
Departure (0-1)	0.49	0.16	0.22	0.84
Headway (min)	13.31	10.09	2	113
Vehicle Capacity (size)	33.69	10.34	14	41
Time (hr)	11.52	3.91	5	20
Temperature (℃)	29.17	2.56	23	34

Noted: (0-1) is a fraction of the day

3.2 Result of hyperparameters

Table 4 shows the optimal hyperparameter value of models of machine learning models such as Random Forest and XGBoost were optimized by Optuna optimization within a defined search space to set for a parameter that gives highest model performance.

Table 3. The Optimal value of hyperparameters of ML models

ML	Hyper parameters	Search Space	Optimal value
RF	max_depth	[1,13]	8
	n_estimator	[10,300]	250
XGBoost	max_depth	[3,12]	4
	n_estimators	[100,1000]	400
	learning_rate	[0.01,0.3]	0.013
	subsample	[0.5,1]	0.959
	colsample_bytree	[0.5,1]	0.975
	gamma	[0,5]	0.003
	reg_lambda	[1e-3,10]	0.002
	reg_alpha	[1e-3,10]	0.792

3.3 Comparison of models

Table 4 demonstrates the performance metrics of machine learning models for ridership prediction, based on the hyperparameter tuning in Table 3. XGBoost model achieves the best performance, with an MAE and RMSE compared to LR, RF. As showed in Fig.2, XGBoost model performed well than linear regression (LR) owing to blue dots is closely group around the ideal line.

Table 4. Model results of testing set with hyperparameter tunning

Models	MAE	RMSE
Linear Regression (LR)	20.30	27.85
Random Forest (RF)	19.45	26.93
XGBoost	18.92	26.37

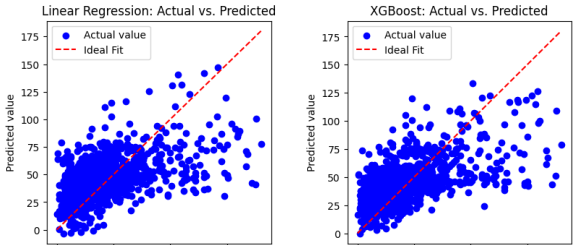


Fig.2. Actual Versus Predicted for LR, XGBoost

Fig.3 illustrates the RMSE value accross multiple iterations for Random Forest and XGBoost models, comparing training and testing performance. The blue line shows the training RMSE, which steadily decreased, indicating that the learning and fitting well of the model to the training data. However, the orange line representing testing remains relatively high and fluctuates, suggesting that the model struggles to generalize to unseen data. As show in Table 4, the RMSE of 26.93 for Random Forest was achieved at iteration 63, the model has reached the optimal point between training and testing performance. Beyond iteration 63, suggests the model

is no improvement its ability to generalize to unseen data. For comparison, the XGBoost model reached the RMSE of 26.37 at iteration 400.

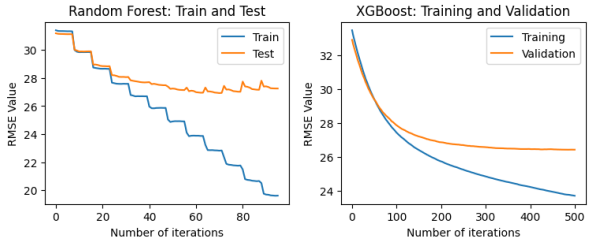


Fig.3. Comparison of training and testing of models

In Fig.4 presents about residual histograms for both training and testing sets are approximately bell-shaped and centered around zero, demonstrating that the predictions of models are generally unbiased. The similarity between the two distributions suggests good generalization with no significant overfitting.

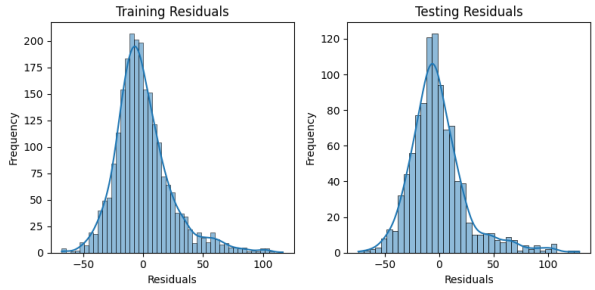


Fig.4. Residuals for training and testing datasets

3.4. Feature Importance

The feature importance analysis in Fig.5 shows that departure time, headway, and temperature are the most influential variables on public bus ridership. Together, these features togethers for over 66% of model's decision-making. In contrast, categorical variable such as the day of the week, peak hour, and weather conditions also contribute less significantly with the remaining 34% Thus, variables related to time and service characteristic play a crucial role in predicting ridership in this context.

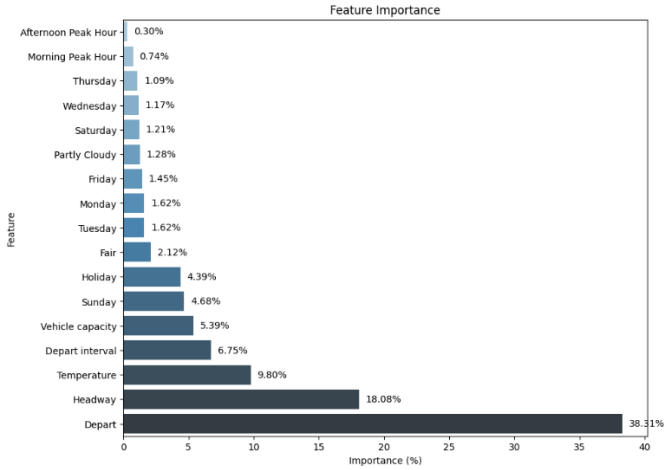


Fig.5. Relative importance of input features

3. CONCLUSIONS

This research explores machine learning techniques for predicting public bus ridership in Phnom Penh City, operated by City Bus Authority (CBA), with a using manually recorded data by dispatchers and drivers. As the data manual collection was limited and challenge, this study concentrated only single line 3. Furthermore, the data analysis showed that average of ridership approximately 45 ridership per trip with a high standard deviation of 33.42, indicating an irregular in the demand whereas headway and vehicle capacity seemed significant variation, indicating the need for flexible and efficient scheduling techniques. The study used for both traditional methods, and machine learning models. The remarkable accuracy achieved by XGBoost can be attributed to its ability to identify intricate patterns and temporal dependencies within the dataset. By including external variable such as temperature, with a mean of 29.17℃ and the time of the day, the model performed better for the prediction ridership. Furthermore, the key prediction features of ridership included departure, headway, and temperature, and residual analysis expressed good model generalization, indicating the model is suitable for reliable ridership prediction in public bus to optimize bus frequency to reduce user waiting time. These efforts contribute to improving service reliability, optimizing resource allocation.

Future studies will extend this research to all bus lines and may explore advanced learning models such as Long Short-Term Memory (LSTM) networks, for time-series prediction

ACKNOWLEDGMENTS

The author would like to thank the City Bus Authority (CBA) for providing the datasets.

REFERENCES

[1] J. Siswanto, D. Manongga, I. Sembiring, and S. Wijono, 'Deep Learning Based LSTM Model for Predicting the Number of Passengers for Public Transport Bus Operators', *J. Online Inform.*, vol. 9, no. 1, pp. 18–28, Apr. 2024, doi: 10.15575/join.v9i1.1245.

[2] A. Tirachini, D. A. Hensher, and J. M. Rose, 'Crowding in public transport systems: Effects on users, operation and implications for the estimation of demand', *Transp. Res. Part Policy Pract.*, vol. 53, pp. 36–52, Jul. 2013, doi: 10.1016/j.tra.2013.06.005.

[3] A. H. AlKhereibi, T. G. Wakjira, M. Kucukvar, and N. C. Onat, 'Predictive Machine Learning Algorithms for Metro Ridership Based on Urban Land Use Policies in Support of Transit-Oriented Development', *Sustainability*, vol. 15, no. 2, p. 1718, Jan. 2023, doi: 10.3390/su15021718.

[4] S. M. Palacio, 'Machine Learning Forecasts of Public Transport Demand: A Comparative Analysis of Supervised Algorithms Using Smart Card Data', *SSRN Electron. J.*, 2018, doi: 10.2139/ssrn.3165303.

[5] S.-Z. Zhao, T.-H. Ni, Y. Wang, and X.-T. Gao, 'A new approach to the prediction of passenger flow in a transit system', *Comput. Math. Appl.*, vol. 61, no. 8, pp. 1968–1974, Apr. 2011, doi: 10.1016/j.camwa.2010.08.023.

[6] E. Ruiz, W. F. Yushimito, L. Aburto, and R. De La Cruz, 'Predicting passenger satisfaction in public transportation using machine learning models', *Transp. Res. Part Policy Pract.*, vol. 181, p. 103995, Mar. 2024, doi: 10.1016/j.tra.2024.103995.

[7] J. Wood, Z. Yu, and V. V. Gayah, 'Development and evaluation of frameworks for real-time bus passenger occupancy prediction', *Int. J. Transp. Sci. Technol.*, vol. 12, no. 2, pp. 399–413, Jun. 2023, doi: 10.1016/j.ijtst.2022.03.005.

[8] Z. Mei, J. Yu, W. Ding, L. Kong, and J. Zhao, 'Bus Passenger Volume Forecasting Model Based on XGBoost Integrated Learning Algorithm', in *CICTP 2020*, Xi'an, China (Conference Cancelled): American Society of Civil Engineers, Dec. 2020, pp. 3100–3113. doi: 10.1061/9780784483053.261.

[9] 'Phnom Penh, Cambodia Weather History | Weather Underground'. Accessed: Apr. 28, 2025. [Online]. Available: https://www.wunderground.com/history/daily/VDPP/date/2025-4-28

[10] F. Pargent, F. Pfisterer, J. Thomas, and B. Bischl, 'Regularized target encoding outperforms traditional methods in supervised machine learning with high cardinality features', *Comput. Stat.*, vol. 37, no. 5, pp. 2671–2692, Nov. 2022, doi: 10.1007/s00180-022-01207-6.

[11] H. Lu, 'Quasi-orthonormal Encoding for Machine Learning Applications', May 29, 2020, *arXiv*: arXiv:2006.00038. doi: 10.48550/arXiv.2006.00038.

[12] B. P. Ashwini, R. Sumathi, and H. S. Sudhira, 'Bus Travel Time Prediction: A Comparative Study of Linear and Non-Linear Machine Learning Models', *J. Phys. Conf. Ser.*, vol. 2161, no. 1, p. 012053, Jan. 2022, doi: 10.1088/1742-6596/2161/1/012053.

[13] L. Breiman, 'Random Forests', *Mach. Learn.*, vol. 45, pp. 5–32, Oct. 2001, doi: 10.1023/A:1010950718922.

[14] N. Gaikwad and S. Varma, 'Performance Analysis of Bus Arrival Time Prediction Using Machine Learning Based Ensemble Technique', *SSRN Electron. J.*, 2019, doi: 10.2139/ssrn.3358828.

[15] W. Lv, Y. Lv, Q. Ouyang, and Y. Ren, 'A Bus Passenger Flow Prediction Model Fused with Point-of-Interest Data Based on Extreme Gradient Boosting', *Appl. Sci.*, vol. 12, no. 3, p. 940, Jan. 2022, doi: 10.3390/app12030940.

[16] R. Muzayanah, D. A. A. Pertiwi, M. Ali, and M. A. Muslim, 'Comparison of gridsearchcv and bayesian hyperparameter optimization in random forest algorithm for diabetes prediction', *J. Soft Comput. Explor.*, vol. 5, no. 1, Art. no. 1, Apr. 2024, doi: 10.52465/joscex.v5i1.308.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Evaluation of Modified Asphalt Binders Incorporating Recycled Plastic Waste for Enhanced Pavement Performance

Chy TENG [1*], Kuchvichea KAN [1], Phanny YOS [2]

Faculty of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

[2] *Research and Innovation Centre, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: chy_teng@gsc.itc.edu.kh*

Abstract: *The increasing accumulation of plastic waste poses a significant environmental challenge, with millions of tons being discarded annually. Simultaneously, conventional asphalt binders used in road construction often exhibit inadequate performance under high temperatures, leading to rutting and deformation. To address both concerns, this study explores the potential of incorporating recycled plastic waste into asphalt binders as a sustainable solution for improving pavement durability.*

The objective of this research is to evaluate the mechanical properties, thermal stability, and overall performance of modified asphalt binders incorporated with recycled plastic waste. This study focuses on modifying asphalt binders using only two materials: 60/70 penetration-grade bitumen and recycled plastic bottle caps composed primarily of HDPE.

The methodology involves laboratory experiments where asphalt binders are modified with different percentages of plastic waste (0%, 1%, 3%, 5%, 7% and 9%) and subjected to standard performance tests, including penetration, softening point, and ductility tests based on ASTM standards. The results indicate that increasing plastic content enhances the binder's resistance to deformation, increases softening point temperatures, and reduces penetration values, signifying improved binder hardness and durability. However, a higher plastic percentage reduces ductility, necessitating optimization for practical applications.

This study demonstrates that incorporating recycled plastic waste into asphalt binders can enhance binder performance while mitigating plastic pollution. The findings suggest that plastic-modified asphalt is a viable alternative for sustainable road construction, with economic and environmental benefits. Further research is recommended to assess the long-term field performance and environmental impacts of plastic-modified asphalt binders.

Keywords: Recycled Plastic Waste; Asphalt Binder Modification; Pavement Performance; Sustainable Road Construction; Plastic Pollution Reduction.

1 INTRODUCTION

The global surge in plastic waste has become a major environmental concern due to its synthetic and non-biodegradable nature, which leads to long-term pollution of land and marine ecosystems [1]. With over 430 million metric tons of plastic produced annually, much of which ends up in landfills or oceans, innovative reuse strategies are urgently needed [2]. One such approach is incorporating recycled plastic waste into asphalt binders for road construction, which can improve mechanical and thermal performance while contributing to sustainability [3]. In Cambodia, particularly Phnom Penh, plastic waste management faces significant challenges due to urbanization, population growth, and limited infrastructure. The city generates around 10,000 tons of plastic waste annually, much of which is not properly segregated or recycled, leading to increased environmental stress [4]. Past reports, such as Stemming the Tide [5], originally emphasized the contribution of Asian rivers like the Mekong to ocean plastic pollution. However, this report was later retracted by Ocean Conservancy for inaccuracies and bias.

Although many studies have explored the use of recycled plastic in asphalt mixtures, most focus on general performance enhancements or compare different types of polymers without considering regional material conditions or binder-specific effects. In Cambodia, waste plastic remains a growing environmental issue, while bitumen performance under tropical conditions is a critical concern for road durability. However, limited research has been conducted on the effects of locally sourced recycled HDPE specifically from bottle caps on the performance of 60/70-grade bitumen using the wet process. Furthermore, previous studies often do not define an optimal dosage range based on standardized performance thresholds.

2 METHODOLOGY

This study focuses on the wet process for modifying asphalt binders with recycled plastic waste, specifically targeting the reuse of plastic bottle caps as a sustainable additive. In this method, shredded plastic is blended directly into hot 60/70 penetration-grade bitumen at elevated temperatures (typically between 160 °C and 170 °C) to form a plastic-modified binder. The plastic is incorporated at varying dosages 1%, 3%, 5%, 7% and 9% by weight of the binder to evaluate the influence of plastic content on the binder's physical and mechanical properties.

The objective is to assess how the direct incorporation of recycled plastic into the binder affects critical characteristics such as penetration, softening point, and ductility. This method promotes uniform dispersion of plastic within the asphalt matrix and is considered effective for enhancing binder stiffness, high-temperature performance, and overall durability. The findings aim to support the development of environmentally responsible pavement materials while addressing the issue of plastic waste through practical reuse strategies.

2.1 Materials

2.1.1 Bitumen

Bitumen 60/70 was selected as the bituminous material in this research. The test performance of the laboratory to evaluate the properties were ductility, softening and penetration, and density.

Table 1: Physical Property of Bitumen 60/70

Test Item	*Test method*	*Spec.*	*Result*	*Unit*
Ductility	ASTM D113	>1000	1500	mm
Softening point	ASTM D36	45-55	45.5	°C
Penetration	ASTM D5	60-70	66.87	0.1mm
Density	ASTM D70	-	1.032	g/cm^3

Table *1* indicates the properties of the bitumen 60/70.

2.1.2 Recycle plastic

The recycled plastic utilized in this study was collected from post-consumer plastic bottle caps, which primarily consist high-density polyethylene (HDPE). Prior to use, a heating test at 180 °C for 90 minutes was performed to assess the thermal behavior of the materials. This test confirmed that plastics with melting points below 200 °C, particularly PP with a melting point of approximately 160 °C, were suitable for asphalt binder modification without undergoing degradation. Recycled HDPE plastics could be incorporated into asphalt through the wet method due to their relatively low melting temperature (~135°C) [7].

2.2 Testing Method

The performance of the modified asphalt binders was evaluated through standard laboratory tests in accordance with ASTM specifications. Penetration (ASTM D5), softening point (ASTM D36), and ductility (ASTM D113) tests were conducted to assess the consistency, thermal resistance, and flexibility of the binders, respectively.

2.2.1 Sample preparation

In this study, sample preparation involved modifying 1 kg batches of 60/70 penetration-grade asphalt binder with varying percentages of recycled Hight Density polyethylene (HDPE) plastic waste. The plastic was introduced at 0%, 1%, 3%, 5%, 7%, and 9% by weight of the binder. (see Table *2*)

Table 2: Plastic content for each sample preparation

Plastic content (%)	*Plastic weight (g)*	*Asphalt weight (g)*
0	0	1000
1	10	1000
3	30	1000
5	50	1000
7	70	1000
9	90	1000

2.2.2 Laboratory test

Recycled (HDPE) sourced from plastic bottle caps was cleaned, shredded into 2–5 mm particles, and added to 60/70 penetration-grade bitumen in varying proportions (0%, 3%, 5%, 7%, and 9% by binder weight). The bitumen was heated to 170 °C–180 °C, and the HDPE was gradually introduced using manual stirring to ensure uniform dispersion. Each batch was mixed for 5–6 hours, then poured into standardized molds for penetration, softening point, and ductility tests. The

prepared samples were cooled at room temperature and stored in sealed containers for subsequent testing.

3 RESULTS AND DISCUSSION

This section discusses the effects of recycled high-Density Polyethylene (HDPE) on the physical properties of asphalt binders, focusing on penetration, softening point, and ductility. The results for each plastic content level (0% to 9%) are compared against the Japan Road Association (JRA) standards to evaluate performance suitability. Key trends and trade-offs between stiffness and flexibility are highlighted to determine the optimal modification range.

3.1 Penetration (ASTM D5)

According to *Fig*. 1 the penetration values of the asphalt binder decreased consistently with the increasing proportion of recycled HDPE, confirming a significant stiffening effect. The unmodified binder exhibited a penetration value of approximately 66–67 1/10mm, characteristic of 60/70-grade bitumen. Upon adding 3% PP, the value dropped to approximately 45 1/10mm. Further increases in plastic content to 5%, 7%, and 9% resulted in penetration values of 39 1/10mm, 22 1/10mm, and 15 1/10mm, respectively. These results demonstrate that the addition of HDPE reduces binder penetration, enhancing its stiffness and resistance to deformation under high temperatures. However, excessive stiffness may lead to increased brittleness, which must be addressed in practical applications.

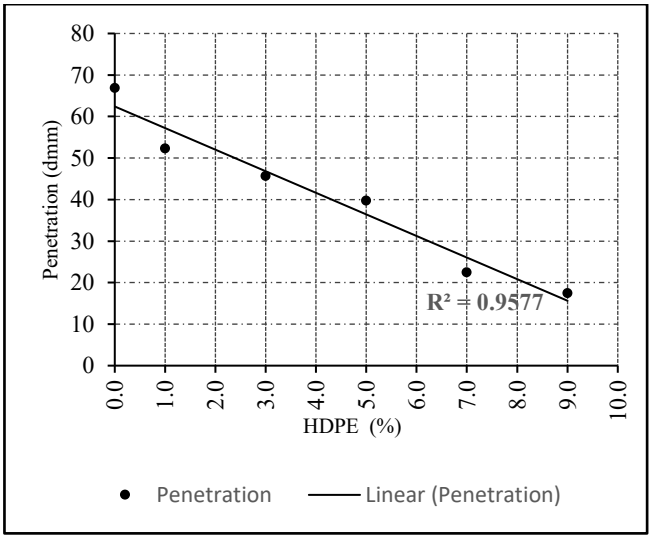


Fig. 1: Penetration affected by plastic percentage

These findings are consistent with previous studies that report a significant reduction in penetration as plastic content increases. The behaviour observed in this experiment confirms that recycled HDPE acts as a plastomeric modifier that reinforces the binder structure, making it suitable for pavements exposed to high temperatures and heavy loads.

3.2 Softening Point (ASTM D36)

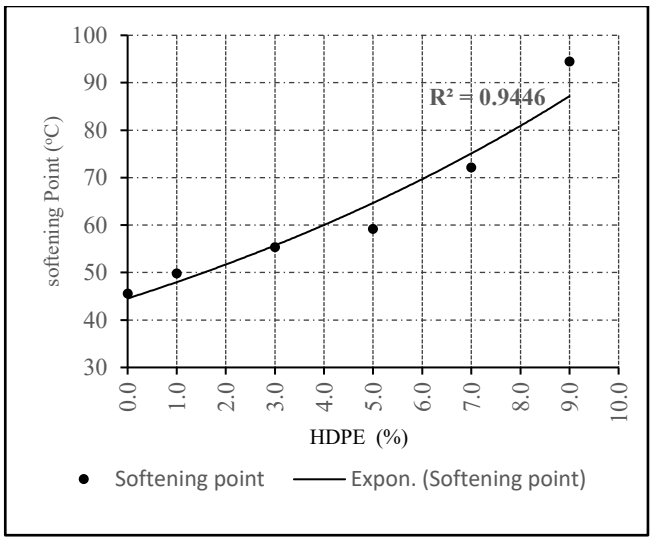


Fig. 2 Softening Point affected by plastic percentage

According to *Fig*. *2* The softening point of the bitumen increased with higher percentages of HDPE, indicating improved resistance to thermal deformation. The control sample (0% HDPE) had a softening point of approximately 45.5 °C. With 3% HDPE, the value rose to about 50 °C, while 5% and 7% contents led to softening points of 55 °C and 61 °C, respectively. At the highest modification level (9% HDPE), the softening point exceeded 70 °C. This trend demonstrates that plastic-modified binders can endure higher temperatures without losing structural integrity, which is beneficial for hot-climate applications.

These results are supported by literature where similar increases in softening points have been reported with the addition of polyolefin plastics like HDPE [8]. The higher thermal stability observed in this study indicates better rutting resistance, a critical performance metric for asphalt binders used in high-traffic roads.

3.3 Ductility (ASTM D103)

While stiffness and thermal stability improved, ductility decreased sharply with increasing HDPE content. The control binder had a ductility value exceeding 1500 mm, but this dropped to 700 mm at 3% HDPE and 500 mm at 5%. At 7% and 9% HDPE, ductility values were recorded at 123 mm and 82 mm, respectively. This decline highlights a key trade-off: as the binder becomes more resistant to deformation, it simultaneously loses flexibility, making it more prone to cracking under tensile stresses or low temperatures. (see *Fig*. *3*)

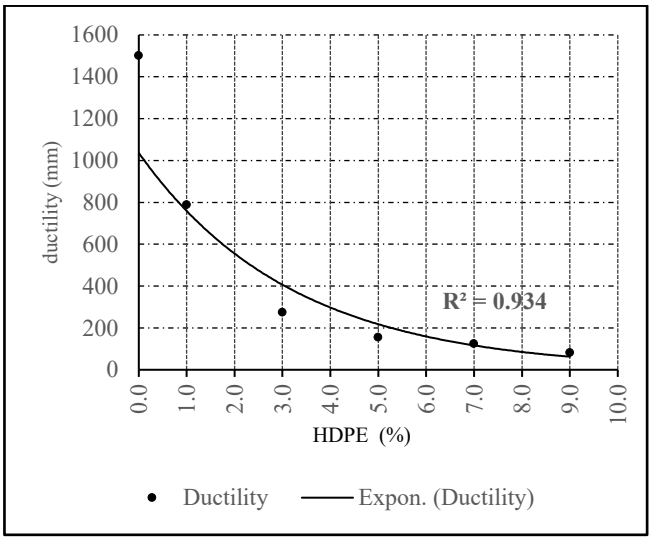


Fig. 3:Ductility affected by plastic percentage

This reduction in ductility is consistent with previous research, which notes that thermoplastic additives like HDPE tend to embrittle the binder at high concentrations [9]. Therefore, while high-temperature performance is enhanced, care must be taken to balance stiffness with sufficient flexibility to prevent cracking.

3.4 Overall Performance Implication

. To evaluate the suitability of the recycled plastic-modified asphalt binder, the experimental results for penetration, softening point, and ductility were compared against the standard requirements for polymer-modified asphalt established by the Japan Road Association as shown in Table *3*. These standards provide benchmark values that ensure the modified binder possesses sufficient stiffness, high-temperature stability, and flexibility needed for durable pavement performance.

Table 3: Standard Property of Polymer Modified asphalt (According to Japan Road Association)

Test item	*Value*
Softening Point ºC	min.56
Ductility (15ºC) , mm	min 300
Penetration (25ºC), 1/10mm	min 40

Softening point of modified asphalt binders compared to JRA standard as shown in Fig. *4*. The graph illustrates that the softening point rises with increasing plastic content. The binders with 3% to 9% PP and higher surpass the JRA minimum softening point standard (≥56°C), indicating improved resistance to thermal deformation.

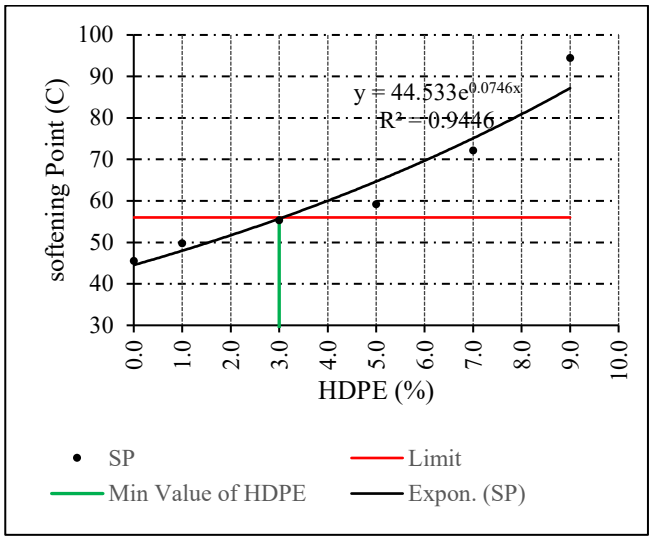


Fig. 4: Softening Point Results Compared to JRA Standards

Penetration values of modified asphalt binders compared to Japan Road Association (JRA) standard as shown in Fig. *5*. This figure shows how penetration decreases with increasing recycled HDPE content. Binders with 0% to 4.3% HDPE meet the JRA minimum penetration requirement (≥40 1/10 mm), while higher percentages result in excessively stiff binders.

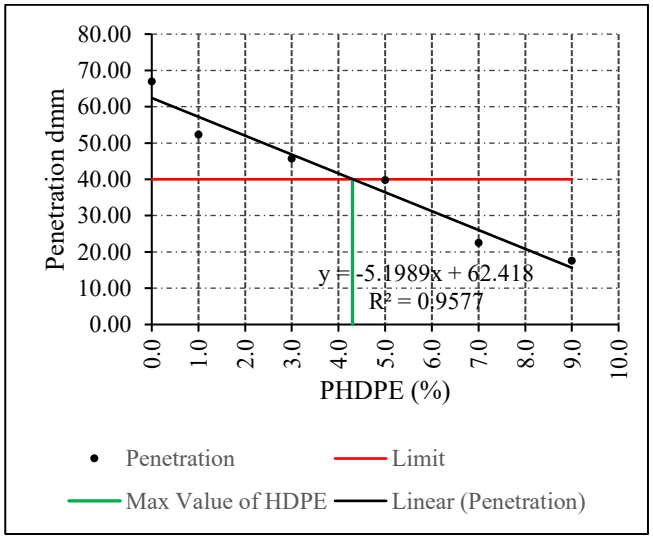


Fig. 5: Penetration Results Compared to JRA Standards

Ductility of modified asphalt binders compared to JRA standard as shown in Fig. *6* This figure depicts the decrease in ductility as HDPE content increases. Only binders with up to 4% HDPE maintain ductility above the JRA minimum standard (≥300 mm at 15°C).

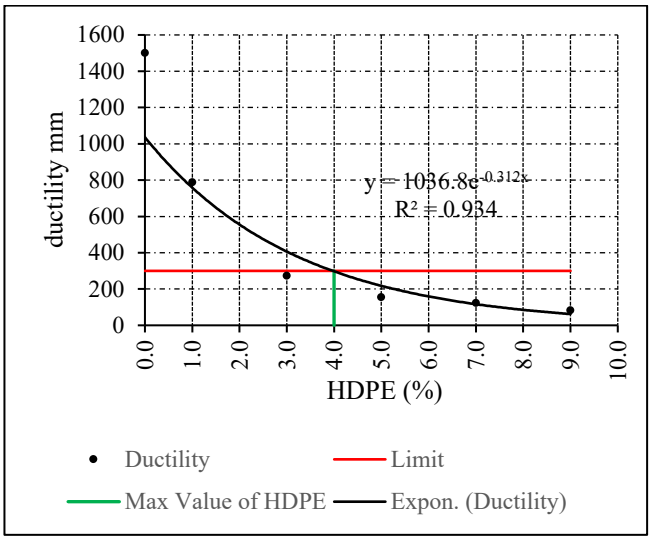


Fig. 6: Ductility Results Compared to JRA Standards

Overall performance evaluation of modified asphalt binders relative to JRA standards as shown in Fig. *7*.

The figure summarizes how penetration, softening point, and ductility of the plastic-modified binders align with JRA benchmarks, indicating that 3-4% PP provides the optimal balance between mechanical performance and flexibility.

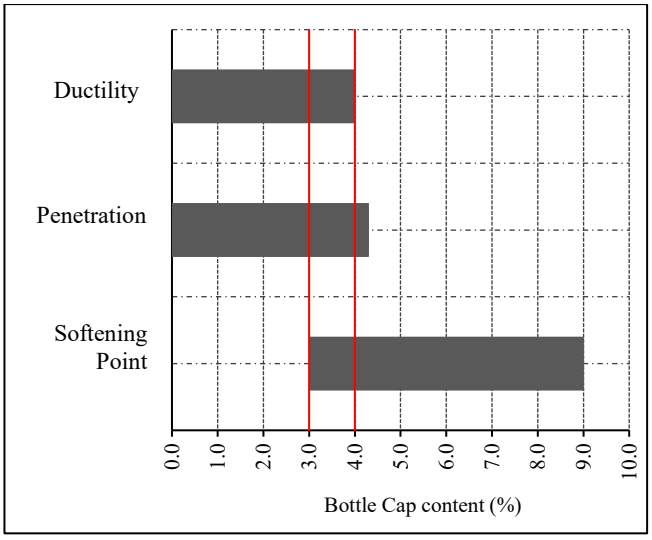


Fig. 7: Overall Performance Assessment Relative to JRA Standards

The results clearly show that HDPE modification improves binder stiffness and thermal resistance, both of which contribute to enhanced rutting performance. However, the trade-off in flexibility indicates that a moderate HDPE content around 3–4% by weight of binder is likely to offer the most balanced performance. Excessive HDPE dosage, while improving softening point and penetration resistance, significantly reduces ductility, potentially compromising long-term durability due to increased cracking risk

4 CONCLUSIONS

This study demonstrates that modifying asphalt binder with recycled High-density polyethylene (HDPE) significantly enhances high-temperature performance by increasing hardness and softening point while reducing penetration. However, higher HDPE content also leads to a notable decrease in ductility, indicating reduced flexibility. The optimal binder performance was observed at 3%-4% HDPE, which offers a balanced improvement in thermal resistance without severely compromising elasticity. These findings support the use of recycled HDPE as a sustainable asphalt modifier, particularly for hot climates like Cambodia, and contribute to the circular economy by repurposing plastic waste into durable infrastructure materials. Further research, including Marshall stability testing, is recommended to evaluate the full mechanical performance of HDPE-modified asphalt mixtures.

ACKNOWLEDGMENTS

This research would not have been possible without the generous financial support and valuable contributions from all involved. The author gratefully acknowledges the finance and material support provided by IKEE Bitumen Chemical (Cambodia) Co., Ltd. Sincere appreciation is also extended to the advisor and co-advisor for their dedicated supervision, professional guidance, and continuous encouragement. Their expertise, timely recommendations, and support in maintaining research progress were instrumental to the successful completion of this work.

REFERENCES

[1] A. L. Andrady and M. A. Neal, “Applications and societal benefits of plastics,” *Phil. Trans. R. Soc. B*, vol. 364, no. 1526, pp. 1977–1984, Jul. 2009, doi: 10.1098/rstb.2008.0304.

[2] UNDP, “Advancing circularity in Cambodia: Focus on plastic pollution | UNDP Climate Promise.” Accessed: Apr. 19, 2025. [Online]. Available: https://climatepromise.undp.org/news-and-stories/advancing-circularity-cambodia-focus-plastic-pollution

[3] T. Baghaee Moghaddam, M. R. Karim, and T. Syammaun, “Dynamic properties of stone mastic asphalt mixtures containing waste plastic bottles,” *Construction and Building Materials*, vol. 34, pp. 236–242, Sep. 2012, doi: 10.1016/j.conbuildmat.2012.02.054.

[4] World Bank, *Urban Development In Phnom Penh*. 2017.

[5] Q. Megan, “Ocean Conservancy rescinds 2015 ocean plastics report | Waste Dive.” Accessed: Apr. 19, 2025. [Online]. Available:

https://www.wastedive.com/news/ocean-conservancy-rescinds-ocean-plastic-report-asia/627368/

[6] L. You *et al.*, “Review of recycling waste plastics in asphalt paving materials,” *Journal of Traffic and Transportation Engineering (English Edition)*, vol. 9, Oct. 2022, doi: 10.1016/j.jtte.2022.07.002.

[7] A. P. F. Giustozzi and Y. J. Boom, “Use of Road-grade Recycled Plastics for Sustainable Asphalt Pavements: Overview of the Recycled Plastic Industry and Recycled Plastic Types”.

[8] L. Jenny Liu, W. Yizhuang David Wang, and Z. Anyou Zhu, *Feasibility Study of Incorporation of Recycled High-Density Polyethylene in Asphalt*. 2022.

[9] A. Oyelere and S. Wu, “State of the art review on the principles of compatibility and chemical compatibilizers for recycled plastic-modified asphalt binders,” *Journal of Cleaner Production*, vol. 492, p. 144895, 2025, doi: https://doi.org/10.1016/j.jclepro.2025.144895.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

IMPROVING THE DURABILITY AND STRENGTH OF EXPANSIVE SOILS STABILIZED WITH EMULSION ASPHALT AND CEMENT

Vanntheng Yoeunthy [1*], Kuchvichea Kan [1], Bertrand François [2], Kean Thai Chhun [3]

[1] *Department of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *ArGEnCo Departement, Université de Liège, Liège, Belgium*

[3] *Ministry of Public Works and Transport, Phnom Penh, Cambodia*

[*] *Corresponding author: vanntheng_yoeunthy@gsc.itc.edu.kh*

Abstract: *Expansive soils present major challenges in geotechnical engineering due to their high shrink-swell potential, which often results in pavement failures and structural damage. This study investigates the stabilization of expansive soils using a combination of 3% Portland cement and 3% asphalt emulsion. Soil specimens were compacted at optimum moisture content and cured for 7 and 14 days. To simulate environmental stresses, a set of specimens was also exposed to elevated temperature (70 °C for 24 hours) prior to testing. Mechanical performance was evaluated through Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS) tests. The results showed significant improvements in both compressive and tensile strength after stabilization. The difference between 7-day and 14-day curing was minor, indicating that most of the strength developed early. Although high-temperature exposure caused some reduction in strength, the stabilized soils still performed considerably better than untreated soils. Overall, the combined use of cement and asphalt emulsion proved effective in enhancing the strength and thermal stability of expansive soils, making it a practical solution for improving subgrade performance in hot climate conditions.*

Keywords: Expansive soil, Emulsion asphalt, Mechanical properties, Stabilized soil, Subgrade.

1. INTRODUCTION

In road and highway constructions, not only the pavement or premix quality is given serious scrutiny, but the substructure below the pavement is also equally vital. The stability of the underlying soils needs serious attention so as to ensure that the pavement structures that has been constructed can enhance the durability of the pavements. It is important to provide the optimum performance for the pavements because the pavement structures are significantly impacted by the direct loading of the traffic.

An expansive soil is a type of soil that takes up water, expands, and then shrinks as it dries out. The unique swelling and shrinking (Figure 1) of the soil are the primary engineering problem with expansive soils. The pressure from the volume change causes cracks on floors, pavements and roads, and wall movement as well as deformations in engineering constructions (Jones et al., n.d.).

Various soil stabilization methods have been proposed to mitigate the detrimental effects of expansive soils. Among these, chemical stabilization using binders such as lime and cement remains one of the most widely adopted practices. These additives react with the clay minerals through pozzolanic and cementitious processes, leading to a reduction in soil plasticity, an increase in mechanical strength, and an improvement in volumetric stability (Little, 1995).

Table 1. Potential soil volume changes as related to Plastic Index (PI) and Liquid Limit (LL) [28]

Potential for volume change	Plastic Index (PI%)	Shrinkage (SL%)	Liquid Limit (LL%)
Low	<18	>15	20-35
Medium	15-28	10-15	35-50
High	25-41	7-12	50-70
Very High	>35	<11	>70

In addition to traditional cementitious treatments, asphalt emulsions have been employed as alternative stabilizing agents. Emulsified asphalt acts by coating soil particles, thereby reducing permeability and enhancing resistance to water infiltration. Several studies have demonstrated that soils treated with asphalt emulsion exhibit improved moisture durability, increased cohesion, and greater mechanical integrity (Terrel and Wang, 1971).

Recent investigations highlight the potential of combining cement and asphalt emulsion to achieve synergistic stabilization effects. Cement primarily contributes

[1*] Corresponding author: Vanntheng Yoeunthy
E-mail: vanntheng_yoeunthy@gsc.itc.edu.kh; Tel: +855-93 285 017

to strength and stiffness enhancement through pozzolanic bonding, whereas asphalt emulsion improves ductility, mitigates brittleness, and imparts greater moisture resistance to the stabilized matrix (Mackiewicz and Zak, 2003). This hybrid stabilization technique is particularly attractive for applications where both strength performance and long-term environmental durability are critical.

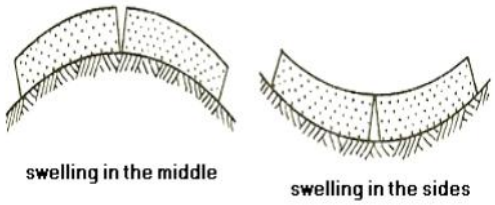


Figure 1 Possible pavement buckling [28]

Despite these promising developments, relatively few studies have focused on how this dual-stabilization method performs under extreme environmental conditions, particularly elevated temperatures. In hot climates, pavement surface temperatures can reach or exceed 70 °C during peak summer periods, especially in regions with high solar radiation. Studies by Li et al. (2013) and Moghaddam et al. (2015) have reported surface temperatures ranging from 60 °C to over 70 °C in tropical and arid environments. Because the temperature gradient decreases from the pavement surface downward, the upper portion of the subgrade—where stabilization is most critical—is still subject to elevated thermal exposure.

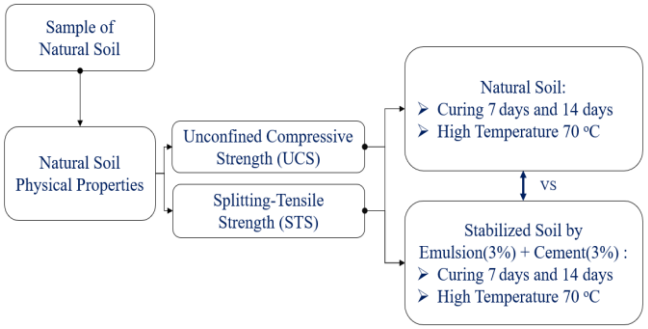
This is especially relevant because asphalt emulsion is temperature-sensitive; at high temperatures, it may soften or lose cohesion, potentially compromising the stability of the treated soil matrix. Therefore, evaluating the mechanical behavior of cement–asphalt stabilized soils under thermal loading is necessary to determine their suitability for long-term use in hot environments.

Additionally, limited research has examined the effect of elevated temperatures on settlement behavior, which is a key design parameter for subgrades and foundations. Understanding how stabilized soils deform under load at high temperatures is essential for designing resilient pavement systems.

Accordingly, this study investigates the combined use of 3% cement and 3% asphalt emulsion to stabilize expansive soils, focusing on compressive strength, tensile strength, and settlement behavior. The experimental program includes standard curing durations and controlled exposure to elevated temperatures (70 °C) to simulate harsh field conditions. The findings aim to support the development of practical and thermally resilient stabilization strategies for use in road and foundation engineering.

2. METHODOLOGY

This section details the experimental setup and procedure used to assess the deformation behavior of soil subject to curing time, high temperature, wet-dry under of applied stress. It describes the experimental apparatus and the specific steps involved in the sample preparation and testing procedure.



2.1 Experimental apparatus

To investigate the mechanical behavior of this expansive soil, three types of mechanical tests were conducted: Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS). The experimental apparatus triaxial machine (Figure 2) for each test was designed to accurately capture the relevant strength and deformation characteristics of the specimens under controlled conditions.



Figure 2: Testing Machine



Figure 3 Specimen's preparation

2.1.1 Unconfined Compressive Strength (UCS)

This test was done to determine the unconfined compressive strength of the soils. The tests were performed on cylindrical samples with 36mm in diameter and 72mm in height compacted in mold at the modified proctor optimum conditions. For the stabilized soil, the cement and emulsion asphalt content used were 3% by weight of air-dried soil. The uniaxial compression strength at failure is given in the equation:

$$\sigma_c = \frac{4P}{\pi d^2} \qquad \text{(Eq. 1)}$$

Where σ_c is the compressive strength in kPa, P is the load at failure in N and d is diameter of specimen in mm

2.1.2 Splitting Tensile Strength (STS)

The specimens for the indirect tensile strength test were prepared identically to those used in the unconfined compressive strength test, with cylindrical dimensions of 18 mm in height and 36 mm in diameter. The test, commonly known as the Brazilian or diametrical compression test, is widely used to assess the tensile strength of rocks, concrete, and also soils. In this method, a cylindrical specimen is loaded horizontally between two rigid platens along its diameter. The applied compressive force (P) generates a maximum tensile stress perpendicular to the loading direction.

Figures 4 and 5 illustrate the stress distribution along the x- and y-axes, respectively. At y = 0, both compressive and tensile stresses reach peak values at the specimen's center (Figure 4), while in Figure 5, the tensile stress is uniformly distributed along the y-axis, and the minimum compressive stress matches that along the x-axis [19].

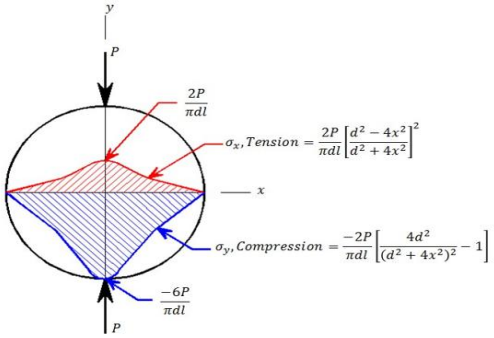


Figure 4 Stress distribution on x-axis (Hudson and Kennedy, 1968)

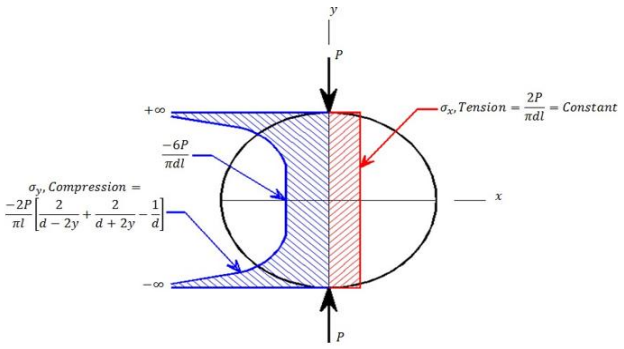


Figure 5 Stress distribution on y-axis (Hudson and Kennedy, 1968)

The indirect tensile strength at failure is given in the equation:

$$\sigma_t = \frac{2P}{\pi dl} \qquad \text{(Eq. 2)}$$

where σ_t is the tensile strength in kPa, P is a load at failure in N, d is diameter of specimen in mm and l is thickness of specimen in mm.

2.2 Test Material and sample preparation

The soil was obtained from a site construction in ITC Phnom Pehn, Cambodia. The soil was pre-treated by oven drying to obtain uniform moisture content.

Atterberg limits, specific gravity, hydrometer analysis and modified compaction were performed using the relevant tests according to the American Society for Testing and Materials (ASTM) Standard Test Method.

Table 2. Physical properties of the expansive soil

Properties	Symbol	Unit	Value
Liquid Limit	LL	%	47.42
Plastic Limit	PL	%	17.08
Plasticity Index	PI	%	30.34
Specific Gravity	Gs	-	2.7
Sand (4.75 to 0.075mm)	-	%	8.89
Silt (0.074 to 0.005mm)	-	%	40.57
Clay (Smaller than 0.005mm)	-	%	50.54
Unified Soil Classification System	USCS	-	CL (Clay of Low Plasticity)
Maximum Dry Density	MDD	g/cm^3	1.895
Optimum Moisture Content	OMC	%	15.01

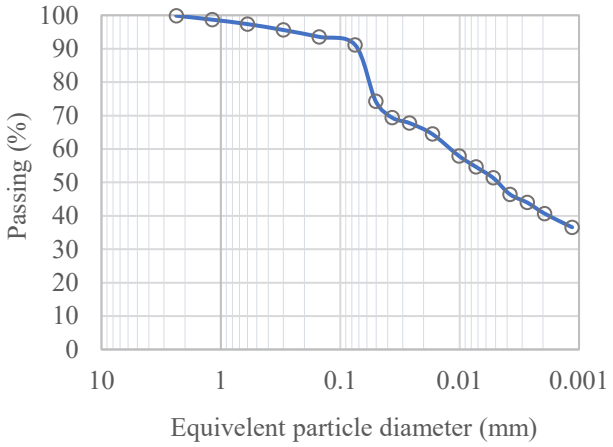


Figure 6 The particle size distribution curve of the studied soil

2.3 Experimental protocol

In this study, all mechanical tests were conducted on unsaturated specimens compacted at their respective optimum moisture contents, as determined from the Modified Proctor test. Figure 7 presents the maximum dry densities obtained for the untreated and stabilized soils, with the stabilization carried out using 3% Portland cement and 3% emulsion asphalt by dry weight. The soil samples were prepared by dynamic compaction using Modified Proctor energy levels to ensure consistency and field-representative conditions prior to mechanical testing.

From a geotechnical perspective, particularly in the context of pavement subgrades, the strength and elastic stiffness (modulus) of the soil are critical parameters influencing load-bearing capacity and predicted settlement behavior. Therefore, the experimental program was designed to assess these properties systematically through a series of strength and deformation tests on the compacted specimens, following standardized procedures.

In this study, a temperature of 70 °C was selected to represent extreme heat conditions that can occur at or near the ground surface, especially in hot and arid regions. Pavement surfaces are known to reach temperatures between 60 °C and 70 °C under strong solar radiation, and the heat can also transfer into the subgrade soils. Testing at this temperature helps to evaluate how stabilized soils behave when exposed to severe heat, particularly in terms of strength loss and material softening. Since asphalt emulsions are sensitive to high temperatures, 70 °C provides a critical point to observe any potential reduction in bonding quality and mechanical performance. In this study, the specimens were placed in a controlled environment and maintained at 70 °C for 24 hours before mechanical testing to ensure full temperature exposure. This approach allows a realistic assessment of how stabilized soils perform under sustained thermal conditions.

2.3.1 Step 1: Specimen curing time and temperature

Table 3: Summary data of the testing specimens

Experi-mental	Dimension of Specimen (mm)		Number of Specimens
	Height (H) or Thickness (l) in mm	Diameter (D) in mm	
UCS	72	36	18
SLS	18	36	18
Total			36

To ensure the reliability of the results, three specimens were tested for each condition. Table 3 presents the number of specimens used for each test factor—natural soil and high-temperature exposure—for the Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS) tests. This approach helped reduce variability and supported a more accurate comparison of the mechanical behavior under different environmental conditions.

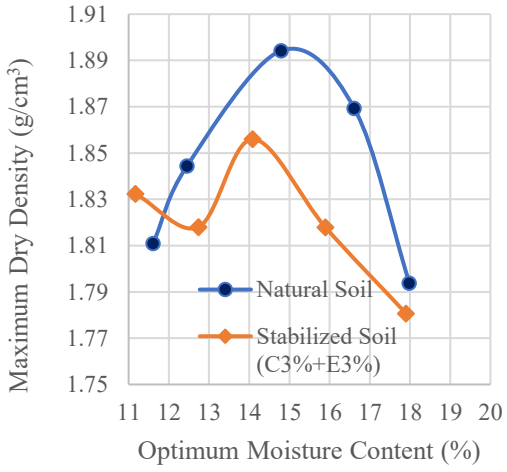


Figure 7 Modified Proctor curves for natural soil and stabilized soil with 3% of cement + 3% of emulsion.

2.3.2 Step 2: Load application

Following specimen preparation, load application was performed under controlled environmental conditions to evaluate the mechanical response of both untreated and stabilized soils. The mechanical tests included Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS), conducted using a loading frame with displacement control.

For the untreated soils, loading was applied immediately after compaction. Specimens were tested under two environmental conditions: (1) at ambient laboratory temperature (natural condition), and (2) after exposure to an elevated temperature of 70 °C for 24 hours.

For the stabilized soils, which were treated with 3% cement and 3% asphalt emulsion by dry weight, loading was applied after 7 and 14 days of curing. Each curing group was tested under two conditions: (1) following standard curing at ambient conditions, and (2) after thermal exposure at 70 °C for 24 hours.

The compressive load in the UCS tests was applied vertically at a constant strain rate of 0.5% per minute, in accordance with ASTM D2166. For STS tests, the diametral compressive load was applied at the same rate using rigid platens until tensile failure occurred. The loading continued steadily until the maximum load was reached; no unloading was applied during either UCS or STS tests, as both tests are destructive and conducted to failure.

All tests were performed under monotonic loading conditions, and the maximum axial load and corresponding deformation were recorded to determine compressive strength, tensile strength, and axial strain at peak.

3. RESULTS AND DISCUSSION

This section presents and discusses the experimental results obtained from the Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS) tests on both natural and stabilized expansive soils under various conditions. The key factors considered include curing duration and exposure to elevated temperature (70 °C). Settlement behavior was interpreted through axial strain at peak strength.

3.1 Effect of Curing Time on Strength and Deformation Behavior

Figure 8 and Figure 9 demonstrate that the UCS and STS values for specimens stabilized with 3% cement and 3% asphalt emulsion after 7 and 14 days of curing, compared to untreated natural soil. The results demonstrate a significant improvement in both compressive and tensile strength due to stabilization.

The UCS of natural soil was 702 kPa, while stabilized soils achieved 3114 kPa after 7 days and 2923 kPa after 14 days. Similarly, STS increased from 101.7 kPa (natural) to 346.4 kPa (7-day cure) and 339.9 kPa (14-day cure). These results indicate that most of the strength gain occurred within the first 7 days, with only marginal improvements observed at 14 days. This is beneficial from a practical perspective, as it allows earlier use of treated subgrades without sacrificing performance.

Axial strain at peak load also decreased with stabilization—from 1.62% (natural) to 1.47% and 1.57% for 7- and 14-day curing, respectively—indicating increased stiffness and reduced deformation. These results suggest the cement–emulsion treatment not only improves strength but also enhances the soil's resistance to settlement under load.

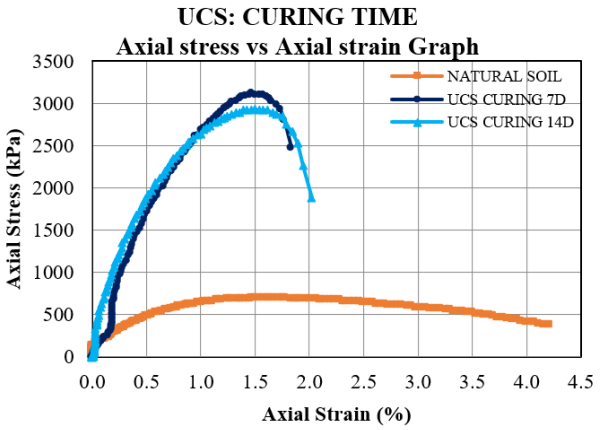


Figure 8 Stress-strain curves of UCS test at different curing times

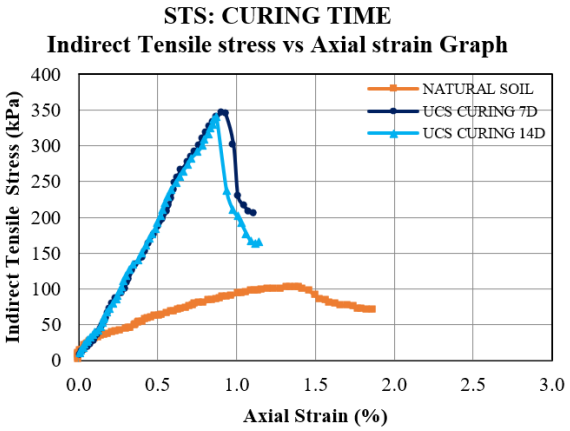


Figure 9 Stress-strain curves of STS test at different curing times

3.2 Effect of High Temperature on Mechanical Behavior

Under high temperature, Figure 10 shows that the UCS of untreated soil increased slightly to 1070 kPa, possibly due to minor moisture evaporation and densification. However, the STS in Figure 11 shows the remained low at 164 kPa, and axial strain was relatively high (2.32%), indicating limited structural integrity.

In contrast, stabilized soils maintained considerably higher strength even after heat exposure. UCS values were 2170 kPa (7-day) and 2382 kPa (14-day), while STS reached 348.2 kPa and 381.8 kPa, respectively. Although slightly lower than values under standard curing, these results confirm that the stabilized matrix retains good mechanical performance under thermal stress. Axial strain also remained lower in stabilized soils (1.42–1.71%) compared to the natural soil, further indicating reduced settlement risk.

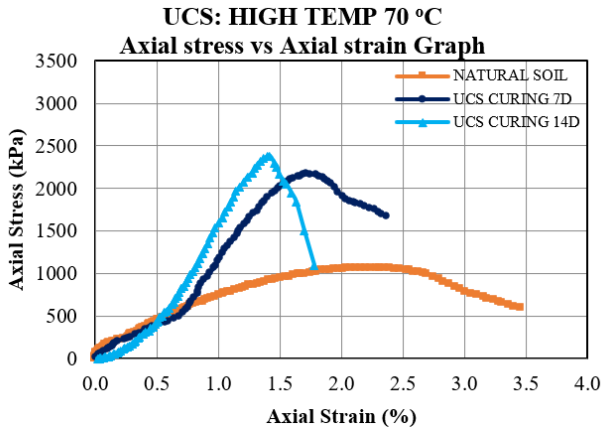


Figure 10 Stress-strain curves of UCS test at different curing times with the effect of High Temp 70 °C 24h before testing

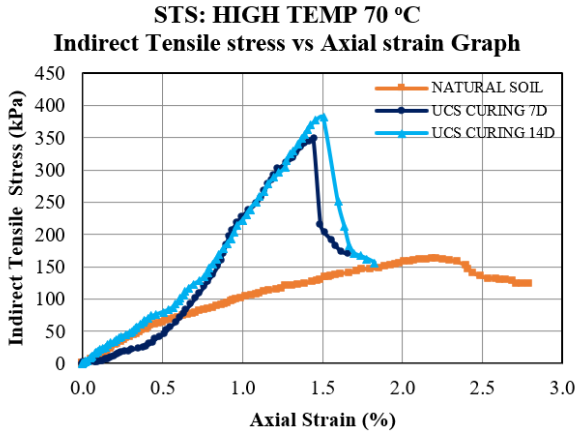


Figure 11 Stress-strain curves of STS test at different curing times with the effect of High Temp 70 °C 24h before testing

3.3 Correlation Between Compressive and Tensile Strength

A strong correlation was observed between compressive and tensile strength across all test conditions. Specimens with higher UCS consistently showed higher STS values, indicating that the treatment improved both failure modes proportionally. This balanced improvement is essential for real-world applications where soils are subjected to a combination of compressive and tensile stresses due to wheel loads, thermal movements, and settlement.

The ratio of UCS to STS remained relatively consistent, reinforcing the idea that the stabilized matrix responds uniformly under both loading conditions.

4. CONCLUSIONS

This study investigates the mechanical behavior of natural expansive soil and stabilized soil treated with 3% Portland cement and 3% asphalt emulsion. Three different factors were considered: curing time (7 days and 14 days), and high-temperature exposure at 70 °C. The aim is to explore changes in compressive strength and tensile strength under different environmental conditions, focusing on the strength and deformation behavior of the treated soils. To achieve this objective, Unconfined Compressive Strength (UCS) and Splitting Tensile Strength (STS) tests were conducted. The applied load was controlled to produce an axial strain at a minimum rate of 0.5% per minute for both UCS and STS tests.

The experimental results lead to the following conclusions:

- The difference between 7-day and 14-day curing periods had negligible influence on the strength and settlement characteristics of the stabilized soil.
- Exposure to high temperatures (70 °C) resulted in a noticeable reduction in strength and an increase in settlement, indicating thermal sensitivity.
- Tensile strength also improved notably, suggesting that the treatment enhanced the soil's resistance to cracking and contributed to overall structural integrity.
- A consistent relationship was observed between compressive and tensile strength gains, pointing to a well-balanced mechanical response after stabilization

These findings underscore the effectiveness of cement and asphalt emulsion stabilization in enhancing the mechanical properties of expansive soils, particularly under environmental and loading conditions relevant to pavement subgrades.

To advance our understanding, future investigations should examine the effect of varying stabilizer content by testing a broader range of cement and asphalt emulsion percentages. This would help optimize the mix design for improved strength and durability while maintaining cost efficiency. Additionally, more research is needed on the long-term performance of stabilized soils, particularly under extended wetting and drying cycles, to better understand how they behave in real-world conditions over time. Finally, developing numerical models that can simulate the combined effects of temperature, moisture, and mechanical loading would be valuable for improving the accuracy and reliability of pavement and subgrade design.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the Institute of Technology of Cambodia for hosting this research as part of its Master's program. Special thanks to Dr. Kuchvichea Kan

and Prof. Bertrand François for their invaluable guidance and support. We also thank the Soil Mechanics Laboratory at ITC, the Ministry of Public Works and Transport, and the JICA-LBE project for their essential contributions. Appreciation is extended to IKEE Company for providing the asphalt emulsion used in this study.

REFERENCES

[1] ASTM International, Standard Practice for Classification of Soils for Engineering Purposes (Unified Soil Classification System), ASTM D2487, 1st ed., West Conshohocken, PA, USA: ASTM International, Year Unknown.

[2] ASTM International, Standard Test Method for One-Dimensional Consolidation Properties of Saturated Cohesive Soils Using Controlled-Strain Loading, ASTM D4186-re98, 1st ed., West Conshohocken, PA, USA: ASTM International, 1998.

[3] ASTM International, Standard Test Method for Splitting Tensile Strength of Intact Rock Core Specimens, ASTM D3967-16, 1st ed., West Conshohocken, PA, USA: ASTM International, 2016.

[4] ASTM International, Standard Test Method for Unconfined Compressive Strength of Cohesive Soil, ASTM D2166, 1st ed., West Conshohocken, PA, USA: ASTM International, 2000.

[5] ASTM International, Standard Test Methods for Laboratory Compaction Characteristics of Soil Using Modified Effort (2,700 kN-m/m³), ASTM D1557, 1st ed., West Conshohocken, PA, USA: ASTM International, 2012.

[6] ASTM International, Standard Test Methods for Liquid Limit, Plastic Limit, and Plasticity Index of Soils, ASTM D4318, 1st ed., West Conshohocken, PA, USA: ASTM International, 2000.

[7] ASTM International, Standard Test Methods for Specific Gravity of Soil Solids by Water Pycnometer, ASTM D854-00, 1st ed., West Conshohocken, PA, USA: ASTM International, 2000.

[8] Alexandre Simas de Medeiros, M. Hugo Sant'Anna Cardoso, and M. Aurélio Vieira da Silva, Evaluation of the Mechanical Behavior of Soil Stabilized with Asphalt Emulsion Using Multi-Stage Loading, 1st ed., Rio de Janeiro, RJ, Brazil: Civil Engineering Journal, 2024.

[9] Ali Akbar Firoozi, C. Guney Olgun, Ali Asghar Firoozi, and Mojtaba Shojaei Baghini, Fundamentals of Soil Stabilization, 1st ed., Seoul, South Korea: International Journal of Geo-Engineering, 2017.

[10] D.I. Boardman, S. Glendinning, and C.D.F. Rogers, Development of Stabilisation and Solidification in Lime–Clay Mixes, Géotechnique, vol. 51, no. 6, pp. 533–543, 2001.

[12] F.G. Bell, Lime Stabilization of Clay Minerals and Soils, Engineering Geology, vol. 42, no. 4, pp. 223–237, 1996.

[13] F.H. Chen, Foundations on Expansive Soils, 1st ed., Amsterdam, Netherlands: Elsevier, 1988.

[14] H. Fred Waller Jr., Emulsion Mix Design Methods: An Overview, 1st ed., Washington, DC, USA: Transportation Research Board, 1980.

[16] Jawad K. Thajeel, Hussein A. Shaia, Sohaib K. Al-Mamoori, and Alaa D. Almurshedi, Effect of Emulsified Asphalt on Expansive Soil Strength and Swelling, 1st ed., Baghdad, Iraq: E3S Web of Conferences, 2023.

[17] Jaspreet Singh Pooni, Filippo Giustozzi, Dilan Robert, Sujeeva Setunge, and Brian O'Donnell, Durability of Enzyme Stabilized Expansive Soil in Road Pavements Subjected to Moisture Degradation, 1st ed., Amsterdam, Netherlands: Elsevier, 2019.

[18] J.D. Nelson and D.J. Miller, Expansive Soils: Problems and Practice in Foundation and Pavement Engineering, New York, NY, USA: Wiley, 1992.

[19] K. Kan and B. François, Triaxial Tension and Compression Tests on Saturated Lime-Treated Plastic Clay upon Consolidated Undrained Conditions, 1st ed., Brussels, Belgium: Journal of Rock Mechanics and Geotechnical Engineering, 2023.

[20] Muchui Loyford Mugambi, Julius Ratumo Toeri, Ismael Kithinji Kinoti, Kidist Dereje Bedada, and Joseph Mwiti Marangu, A Comprehensive Review on Methods, Agents and Durability Factors for Stabilization of Expansive Soils, 1st ed., Istanbul, Turkey: Journal of Sustainable Construction Materials and Technologies, 2023.

[21] Nilo Cesar Consoli, Hugo Carlos Scheuermann Filho, Luis Segadaes, and Nuno Miguel Cordeiro Cristelo, Effect of Wet-Dry Cycles on the Durability, Strength, and Stiffness of Granite Residual Soil Stabilized with Portland Cement, 1st ed., London, United Kingdom: International Society for Soil Mechanics and Geotechnical Engineering, 2019.

[22] Noraida binti Razali, Norazzlina binti M. Sa'don, and Abdul Razak bin Abdul Karim, Strength and Durability Effect on Stabilized Subgrade Soil, 1st ed., Kota Samarahan, Sarawak, Malaysia: University Malaysia Sarawak, 2016.

[23] R.D. Holtz and W.D. Kovacs, An Introduction to Geotechnical Engineering, Englewood Cliffs, NJ, USA: Prentice Hall, 1981.

[24] R.L. Terrel and C.K. Wang, Early Curing Behavior of Emulsified Asphalt Stabilized Soils, Highway Research Record, no. 351, pp. 28–38, 1971.

[25] S.M. Mackiewicz and J.D. Zak, Cement Stabilization of Native Soils for Pavement Applications, Ames, IA, USA: Iowa State University, 2003.

[27] D.N. Little, Stabilization of Pavement Subgrades and Base Courses with Lime, Arlington, VA, USA: National Lime Association, 1995.

[28] Behzad Kalantari, Foundations on Expansive Soils: A Review, University of Hormozgan, Bandar Abbas, Iran, 2012.

Water and Environment (WAE)

3 Sub-parallel Sessions

"WAE-1: Special Session for Project for Establishment of Risk Management Platform for Air Pollution"

Chair: Prof. Furuuchi MASAMI; Co-chair: Assist. Prof. PENG Chanthol

No.	Topic
1	Preliminary Assessment of Air Quality in the Workplace in Phnom Penh *Authored by: Sereyvath YOEUN, Vuthy RUOS, Channmuny THANH, Kimleang KHOEURN, Muhammad AMIN, Chanvorleak PHAT*
2	Variations in Particulate Matter (PM2.5) Concentration in Phnom Penh: The Role of Wind Speed and Wind Direction During Daytime and Nighttime *Authored by: Sochan Sao, Saksoleakhena SAMBATH, Seyha DOEURN, Chanthol PENG, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHI, Chanmoly OR*
3	Gravimetric Analysis of PM2.5 Mass Concentration: Daily Behavior in Early Dry season in Phnom Penh, Cambodia *Authored by: Vahet SARY, Seyha DOEURN, Sreyoun SORN, Sochan SAO, Chanthol PENG, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHi, Chanmoly OR*
4	Driving Factors of Particulate Matter (PM2.5 and PM10) Dynamics During Heavy Air Pollution in Phnom Penh, Cambodia *Authored by: Sokkeav SRONG, Sochan SAO, Seyha DOEURN, Chanthol PENg, Saksoleakhena SAMBATH, Muhammad AMIN, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHI, Chanmoly OR*
5	Preliminary Assessment of Indoor Air Quality in Residences in Phnom Penh *Authored by: Channmuny THANH1, Bophavry SARAV, Sereyvath YOEUN, Kimleang KHOEURN, Muhammad AMIN, Chanvorleak PHAT*
6	Daytime and Nighttime Monitoring of Ultrafine Particulate Matter (PM0.1) during Dry Season in Phnom Penh, Cambodia *Authored by: Chanrat VORNG, Seyha DOEURN, Sokuntheary SOK, Sreyvich SIENG, Sochan SAO, Chanthol PENG, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHI, Chanmoly OR*
7	Variability of PM2.5 Distribution in Relation to Gaseous Pollutants Concentrations in Phnom Penh, Cambodia *Authored by: Titsothearna YOU, Sochan SAO, Seyha DOEURN, Chanthol PENG, Saksoleakhena SAMBATH, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHI, Chanmoly OR*
8	Characterization of Bacterial Communities in Aquatic Environments of Phnom Penh as Potential Bioaerosol Sources Using Nanopore Sequencing *Authored by: Titsothearna YOU, Sochan SAO, Seyha DOEURN, Chanthol PENG, Saksoleakhena SAMBATH, Fumikazu IKEMORI, Mitsuhiko HATA, Masami FURUUCHI, Chanmoly OR*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Preliminary Assessment of Air Quality in the Workplace in the Institute of Technology of Cambodia

Sereyvath Yoeun1,2*, Vuthy Ruos[1,2], Channmuny Thanh[1,2], Kimleang Khoeurn[1], Muhammad Amin[3], Chanvorleak Phat[1,2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Faculty of Geoscience and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kanazawa 920-1192, Ishikawa, Japan*

[*] *Corresponding author: svath@itc.edu.kh*

Abstract: *Air pollution is a pressing global concern with far-reaching impacts on both human health and environmental sustainability. This study focuses on the assessment of particulate matter (PM2.5) and polycyclic aromatic hydrocarbons (PAHs) in workplace environments within the Institute of Technology of Cambodia (ITC), located in Phnom Penh, a rapidly urbanizing city. Air quality monitoring was conducted at six distinct sites, with PM2.5 samples collected over one weekday per site. High-Performance Liquid Chromatography (HPLC) was employed to analyze PAHs concentrations in these samples. The findings revealed that PM2.5 levels at certain locations significantly exceeded the World Health Organization's Air Quality Guidelines, with some readings alarmingly reaching 300–400 µg/m³. PAHs concentrations, particularly in enclosed areas such as Canteen-2, were markedly elevated due to poor ventilation. Notably, carcinogenic compounds such as Benzo[a]pyrene (BaP) were detected at higher levels compared to outdoor settings. These elevated levels pose significant health risks, potentially comparable to or even exceeding the risks associated with activities such as smoking. This study underscores the urgent need for improved workplace air quality management interventions, including the implementation of more effective ventilation systems and the monitoring of indoor air pollutants. As a preliminary investigation constrained by limited resources, the study provides valuable insights while emphasizing the necessity for further comprehensive research to corroborate the findings and guide the development of robust mitigation strategies to safeguard public health in rapidly urbanizing regions like Phnom Penh.*

Keywords: Air Pollution; Workplace; PAHs; PM2.5

1. INTRODUCTION

Air pollution is a global issue that deteriorates air quality, negatively affecting human health and the environment, and is increasingly recognized by businesses and policymakers as one of the most pressing challenges worldwide. Many epidemiological studies have identified a link between air pollution and various negative health outcomes in the general population, spanning from minor subclinical impacts to premature death[1].

One of the well-known parameters to consider in air quality observation is particulate matter (PM), where PM2.5 and PM10 refer to particles with diameters smaller than 2.5 µm and 10 µm, respectively [1]. Within these PM, polycyclic aromatic hydrocarbons (PAHs) are organic pollutants made up of two or more fused aromatic rings of carbon and hydrogen. These colorless, white, or pale-yellow solid compounds naturally occur in substances like coal, crude oil, and gasoline [2]. PAHs with lower molecular weights, consisting of 2–3 aromatic rings, are primarily found in the gas phase due to their higher vapor pressures. In contrast, PAHs with higher molecular weights, containing 5–6 aromatic rings, are mainly associated with particulate matter (PM) [3]. PAH sources can be from different emission sources, including industrial, mobile, domestic, agricultural, and natural[4] [5].

Each PAH compound has distinct health effects and does not exhibit identical impacts. Numerous PAHs are known to be mutagenic, carcinogenic, teratogenic, and harmful to the immune systems of various living organisms, including microorganisms, animals, and humans [2], [6]. Exposure to PAHs and their impact on the human reproductive system was studied to assess health outcomes such as polycystic ovary syndrome, fertility issues, miscarriage, and preterm birth [6], [5].

[1]* Corresponding author: Sereyvath Yoeun
E-mail: svath@itc.edu.kh; Tel: +855-10 348 588

Cambodia, particularly Phnom Penh, has undergone significant development with new construction projects, a rise in vehicle numbers, and a growing population, all of which may contribute to air pollution. Previous data indicate that ambient air monitoring revealed ultrafine particulate matter levels in Phnom Penh to be slightly higher than in other cities in Southeast Asia [5]. However, information on air pollution in workplace environments remains unavailable. Therefore, in this study, we conducted preliminary measurements of PM2.5 and performed PAH analysis in specific workplaces at the Institute of Technology of Cambodia.

2. METHODOLOGY

2.1 Site description

The sampling site in this study is depicted in Fig. 1. The Sample was collected at 6 different sites inside the campus of Institute of Technology of Cambodia (ITC), which is positioned at coordinates 11°34′14.8″N 104°53′53.4″E in the city of Phnom Penh. Those sampling sites include a printing room (photocopy service), a staff office, two cooking sites (canteen), and two security-check sites (campus gate).



Fig. 1 Sampling site in Institute of Technology of Cambodia (ITC). The yellow rectangles showed the six specific site.

2.2 Sampling methods and PM concentration

Samples were collected using a personal air sampler (Sibata Scientific Technology Ltd.) for one weekday at each sampling site (five working days per site) from 8:00 AM to 4:00 PM in May 2024. The air samplers were placed on a stand at a height of 1.0–1.5 m at the sampling sites [7]. Preconditioned Teflon-membrane filters with a diameter of 25 mm were used to collect PM 2.5 mass concentrations and PAHs in the air. The filters were connected to a mini pump (MP-∑100HNII) with a flow rate of 1.5 L per hour. After sampling, the filters were carefully packed in aluminum foil, sealed in a zip-lock polyethylene bag, and stored in a freezer at -20℃ until chemical analysis[8]. PM2.5 concentration was calculated by measuring the change in mass before and after sampling by microbalance (Satorius Cubis MCA6.6S-2S01-F)and dividing it by the volume of air sampled.

2.3 PAHs analysis

This study was the preliminary testing of the PAHs in the Institute of Technology of Cambodia, the standard calibration mix was obtained from Professor Akira Toriba's laboratory, Nagasaki University, containing 10 PAHs (10 µg/mL in acetonitrile), included fluoranthene (Flu), pyrene (Pyr), benz[a]anthracene (BaA), benzo[b]fluoranthene (BbF), benzo[k]fluoranthene (BkF), benzo[a]pyrene (BaP), dibenz[a,h]anthracene (DBA), benzo[ghi]perylene (BghiP), and indeno[1,2,3-cd]pyrene (IDP). The PAHs standard was diluted to concentrations of 2 ng/mL, 5 ng/mL, and 10 ng/mL for injection into the HPLC system (Shimadzu, Japan) was equipped with dual LC-30AD pumps, a SIL-30AC autosampler injector, a degasser (DGU-20A5R), a CMB-20A system controller, an integrator (LCsolution software), a CTO-20AC column oven, and an RF-20Axs fluorescence detector. The mobile phase used was an isocratic phase consisting of 100% acetonitrile (ACN) at a flow rate of 1.8 mL/min. Sample separation was achieved using tandem ODS columns (Inertsil ODS-P, 250 × 4.6 mm i.d., 5 µm particle size, GL Sciences, Tokyo, Japan) maintained at 20°C. Dual data sets were generated from two channels of the fluorescence detector with a scheduled wavelength change, as previously described [5].

The filters were extracted twice with 10 mL of dichloromethane (DCM) by sonication for 10 minutes, and the extracts were combined. Subsequently, 100 µL of DMSO was added, and the mixture was evaporated using a centrifugal evaporator (CVE-3100, EYELA) at 40°C until only the DMSO remained. Acetonitrile (900 µL) was added to the extracted samples and mixed well by sonication for 2 minutes. The samples were then filtered using a 0.45 µm syringe filter and injected into the HPLC.

3. RESULTS AND DISCUSSION

3.1 Concentration of PMs

The data showed that the mass concentration at most sites was less than 100 µg/m³, with two sites exhibiting very high concentrations of approximately 300 and 400 µg/m³ (Fig. 2). Although this study focuses on the workplace air environment,

most of the sampling sites were located in open spaces, and similar results were observed when combining PM2.5 and ultrafine particulate data from the Ambient Nanosampler (ANS) sampler [5].This finding was significantly higher than the WHO Air Quality Guidelines (AQG), which state that the annual average concentration should not exceed 5 µg/m³, while short-term exposures (24-hour average) should not exceed 15 µg/m³ for more than 3–4 days per year [9]. However, this data was based on preliminary testing, which could introduce some errors related to the weighting instrument used for measuring PM.

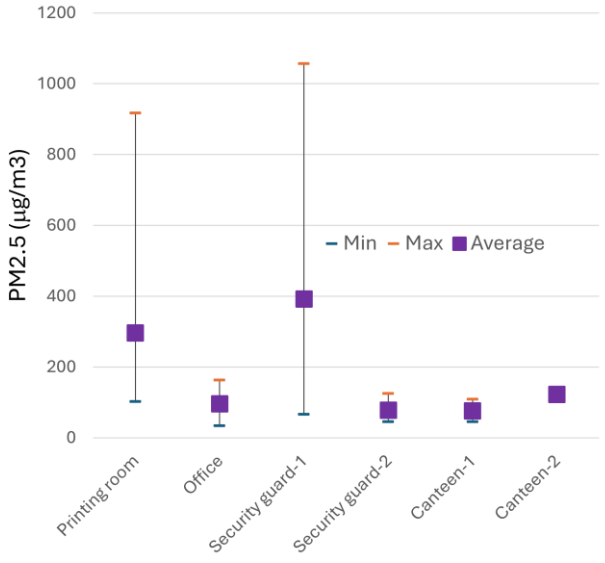


Fig. 2 Particle mass concentration of six sites in the ITC campus.

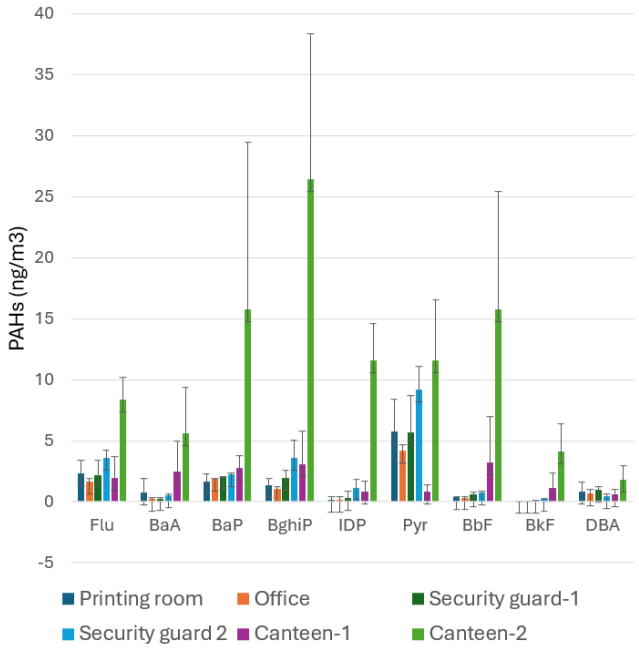


Fig. 3 PAHs component in different sampling sites in the ITC campus.

3.2 PAHs concentration

Around 70–90% of atmospheric PAHs are attached to particulate matter, particularly PM2.5[10]. The International Agency for Research on Cancer (IARC) estimates that the average levels of individual PAHs in urban ambient air typically range between 1 ng/m³ and 30 ng/m³ [10]. In this study, with one-weekday sampling (8 h/day), the data showed a significant difference, with the Canteen-2 site exhibiting very high concentrations of PAHs (Fig. 3 and Fig. 4), although similar mass concentrations were observed (Fig. 2). This high concentration of PAHs was attributed to the site being enclosed, as well as the limited ventilation and confined space during the cooking process, which likely led to reduced dispersion of the emitted pollutants. Many PAHs are recognized as carcinogenic or potentially carcinogenic to humans. Benzo[a]pyrene (BaP) is among the most studied PAHs and is classified as a Group 1 carcinogen by the International Agency for Research on Cancer (IARC) [11]. Smoking one cigarette can result in the absorption of 20–40 ng of BaP [12]. Consuming one pack of unfiltered cigarettes daily leads to a BaP exposure of 0.7 µg/day, while using filtered cigarettes reduces this amount to 0.4 µg/day [13]. An average healthy adult male, with normal body weight and engaging in moderate activity for 16 hours while resting for 8 hours, breathes approximately 22.8 cubic meters of air daily [14]—equivalent to nearly 1 cubic meter per hour. During the 8-hour work shift at Canteen-2 (Fig. 3), the staff are exposed to BaP levels of approximately 2 ng/h, which is significantly higher compared to other sites where the exposure is less than 0.4 ng/h. However, these findings are based on a preliminary study with limited resources, particularly the addition of internal standards during sample preparation.

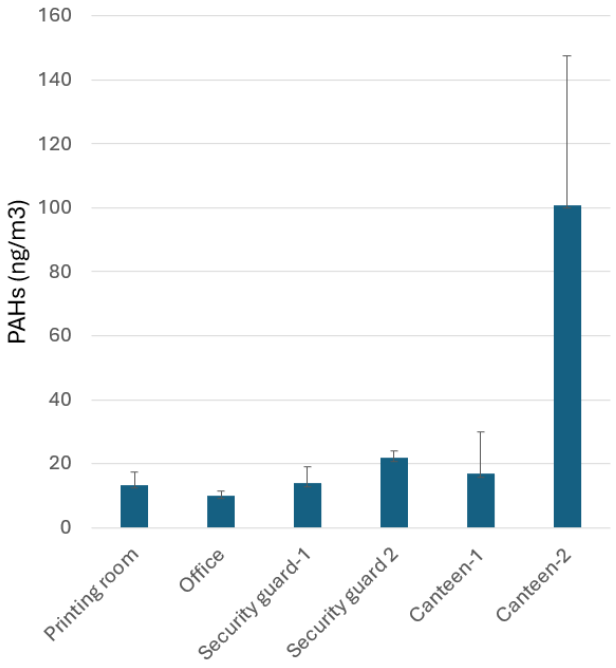


Fig. 4 Total PAHs concentration from sampling sites in the ITC campus.

4. CONCLUSIONS

The study highlights the alarming levels of air pollution, specifically PM2.5 and PAH concentrations, in workplace environments at the Institute of Technology of Cambodia in Phnom Penh. The results showed that PM2.5 concentrations at certain sites significantly exceeded the WHO Air Quality Guidelines, underscoring substantial concerns regarding air quality and associated health risks. Elevated levels of PAHs, particularly in enclosed spaces like Canteen-2, where limited ventilation and confined cooking activities take place, present serious carcinogenic risks. Notably, one of the most significant PAHs, Benzo[a]pyrene (BaP), which serves as a key indicator, was detected at levels considerably higher than those in other areas, thereby exposing workers to health risks that surpass those typically linked to activities such as smoking.

The findings underscore the urgent need for workplace air quality monitoring and improved ventilation systems to reduce exposure to hazardous particles and compounds. However, as the study was preliminary and constrained by limited resources and measurement methodology, further comprehensive studies are essential to validate the risks and develop targeted mitigation strategies.

ACKNOWLEDGMENTS

This research was financially supported by the JICA-JST SATREPS. Equipment was supported by the East Asia Nanoparticle Monitoring Network (EA-NanoNet).

REFERENCES

[1] WHO, “Health risk assessment of air pollution - general principle,” 2016.

[2] H. I. Abdel-Shafy and M. S. M. Mansour, “A review on polycyclic aromatic hydrocarbons: Source, environmental impact, effect on human health and remediation,” 2016. doi: 10.1016/j.ejpe.2015.03.011.

[3] H. Zhang, X. Zhang, and A. Geng, “Expression of a novel manganese peroxidase from Cerrena unicolor BBP6 in Pichia pastoris and its application in dye decolorization and PAH degradation,” *Biochem Eng J*, vol. 153, 2020, doi: 10.1016/j.bej.2019.107402.

[4] A. B. Patel, S. Shaikh, K. R. Jain, C. Desai, and D. Madamwar, “Polycyclic Aromatic Hydrocarbons: Sources, Toxicity, and Remediation Approaches,” Nov. 05, 2020, *Frontiers Media S.A.* doi: 10.3389/fmicb.2020.562813.

[5] M. Amin *et al.*, “Characterization of carbonaceous components and PAHs on ultrafine particles in Phnom Penh, Cambodia,” *Environ Monit Assess*, vol. 196, no. 10, Oct. 2024, doi: 10.1007/s10661-024-13045-8.

[6] A. L. Bolden, J. R. Rochester, K. Schultz, and C. F. Kwiatkowski, “Polycyclic aromatic hydrocarbons and female reproductive health: A scoping review,” 2017. doi: 10.1016/j.reprotox.2017.07.012.

[7] Y. Yoda, K. Tamura, and M. Shima, “Airborne endotoxin concentrations in indoor and outdoor particulate matter and their predictors in an urban city,” *Indoor Air*, vol. 27, no. 5, 2017, doi: 10.1111/ina.12370.

[8] A. McCarron, S. Semple, C. F. Braban, C. Gillespie, V. Swanson, and H. D. Price, “Personal exposure to fine particulate matter (PM2.5) and self-reported asthma-related health,” *Soc Sci Med*, vol. 337, 2023, doi: 10.1016/j.socscimed.2023.116293.

[9] WHO, “WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide,” Sep. 2021, Accessed: May 29, 2025. [Online]. Available: https://iris.who.int/bitstream/handle/10665/345329/9789240034228-eng.pdf

[10] WHO, “Some non-heterocyclic polycyclic aromatic hydrocarbons and some related occupational exposures.,” 2010.

[11] WHO, *Human health effects of polycyclic aromatic hydrocarbons as ambient air pollutants*. 2021. [Online]. Available: https://apps.who.int/bookorders/.

[12] P. Markowicz, J. Löndahl, A. Wierzbicka, R. Suleiman, A. Shihadeh, and L. Larsson, "A study on particles and some microbial markers in waterpipe tobacco smoke," *Science of the Total Environment*, vol. 499, pp. 107–113, Nov. 2014, doi: 10.1016/j.scitotenv.2014.08.055.

[13] H. J. Haussmann, "Use of hazard indices for a theoretical evaluation of cigarette smoke composition," Apr. 16, 2012, *American Chemical Society*. doi: 10.1021/tx200536w.

[14] J. D. Pleil, M. Ariel Geer Wallace, M. D. Davis, and C. M. Matty, "The physics of human breathing: Flow, timing, volume, and pressure parameters for normal, on-demand, and ventilator respiration," Oct. 01, 2021, *IOP Publishing Ltd*. doi: 10.1088/1752-7163/ac2589.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Variations in Particulate Matter ($PM_{2.5}$) Concentration in Phnom Penh: The Role of Wind Speed and Wind Direction During Daytime and Nighttime

Sochan Sao [1*], Saksoleakhena Sambath [1], Seyha Doeurn [1,2], Chanthol Peng [1,2], Fumikazu Ikemori [3], Mitsuhiko Hata [4], Masami Furuuchi [4], Chanmoly Or [1]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Graduate School of Integrated Science and Technology, Nagasaki University, 1-14, Bunkyo-machi, Nagasaki, 852-8521, Japan*

[4] *Faculty of Geosciences and civil Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

[*]*Corresponding author: saosochan2015@gmail.com; sochan.sao@itc.edu.kh*

Abstract: *Particulate matter ($PM_{2.5}$) is a major component of air pollution that poses significant impact on public health risks. In urban areas like Phnom Penh, Cambodia, $PM_{2.5}$ pollution has become a growing concern due to local emissions and transboundary influences. This study investigates the temporal distribution of $PM_{2.5}$ concentration in Phnom Penh and its relationship with wind speed and wind direction during daytime (7 am to 7 pm) and nighttime (7 pm to 7 am) in January 2025. $PM_{2.5}$ concentrations were monitored at the Institute of Technology of Cambodia using a sensor-based Ambient Air Monitor 11000 series. Wind conditions were measured with a Smart Weather Sensor, which was connected to the Air Quality Monitoring System. Sensor-based $PM_{2.5}$ measurements were validated against filter-based mass concentrations collected from a low-volume air sampler. Results revealed that average mass concentrations of $PM_{2.5}$ during daytime were 20 % lower than those during nighttime. Polar plot analysis showed that higher $PM_{2.5}$ concentrations during the daytime were associated with weak winds from north to east at speeds below 0.28 m/s. Similarly, at nighttime, elevated $PM_{2.5}$ concentrations occurred with weak winds from north to west at speeds below 0.14 m/s. Pearson's correlation analysis indicated that $PM_{2.5}$ concentration was weakly and negatively correlated with wind speed during both daytime ($n = 372$, $r = -0.45$, $p < 0.001$) and nighttime ($n = 372$, $r = -0.40$, $p < 0.001$), suggesting the effects of upwind sources emissions from traffic or biomass burning. Although this study provides insights into short-term $PM_{2.5}$ variations, its limited duration restricts analysis of seasonal trends and long-term risks. Further research on seasonal patterns, local emissions, and health impacts is needed for a comprehensive understanding.*

Keywords: Particulate matter; $PM_{2.5}$; Wind speed; Wind direction; Air quality

1. INTRODUCTION

Particulate matter (PM), typically classified by particle diameter, is a major air pollutant of global concern, associated with a wide range of environmental and health problems. Among its various forms, $PM_{2.5}$, particles with an aerodynamic diameter of 2.5 µm or less, is particularly harmful due to its ability to degrade air quality and pose serious risks to human health [1,2]. Because of its small size, $PM_{2.5}$ can penetrate deep into the respiratory tract and enter the bloodstream, with prolonged exposure linked to respiratory and cardiovascular diseases [3]. Given these severe health implications, the World Health Organization (WHO) has established strict air quality guidelines, recommending that $PM_{2.5}$ concentrations not exceed 15 µg/m³ for 24-hour exposure and 5 µg/m³ for annual averages [4]. However, a global analysis reveals that over 90 % of the world's population is exposed to $PM_{2.5}$ levels exceeding this

[1*] Corresponding author: Sochan Sao
E-mail: saosochan2015@gmail.com; Tel: +855-978445346

guideline, with even a complete elimination of anthropogenic emissions insufficient to meet the standard in many regions due to natural sources [5]. Among the key anthropogenic sources of $PM_{2.5}$ are vehicular emissions, industrial activities, and open biomass burning [6,7]. Biomass burning, including the combustion of agricultural residues and forest biomass, plays a significant role in $PM_{2.5}$ pollution across Asia [8].

In Southeast Asia, the air quality is heavily impacted by agricultural residue burning across both regional and local scales [7,9]. This practice is commonly adopted due to its low cost and effectiveness in field preparation, despite its considerable environmental consequences [10,11]. In Southeast Asia the open burning of rice straw/stubble during the dry season is widespread and contributes significantly to air pollution [7]. For example, a previous study found that biomass burning contributed approximately 57 % of $PM_{2.5}$ concentrations during a peak smog event in Southeast Asia [7]. In addition to $PM_{2.5}$, agricultural residue burning activities release large quantities of greenhouse gases (e.g., carbon dioxide (CO_2), carbon monoxide (CO), and hydrocarbons) and harmful pollutants (e.g., sulfur dioxide (SO_2) and nitrogen oxides (NO_x)) [7,12], compounding both air quality and climate challenges. This regional issue is exemplified by recent events in Cambodia.

In January 23rd and 24th, 2025, $PM_{2.5}$ concentrations in Phnom Penh increased dramatically to 109 µg/m³ between 4:00 pm and 7:00 am (Ministry of Environment; MoE). This unusually high concentration coincided with widespread rice straw/stubble burning, as reported by Phnom Penh Post, a local media [13], suggesting a potential link between agricultural burning and the increase in $PM_{2.5}$ concentrations. The incident drew significant public attention and underscored the urgent need for evidence-based strategies to manage air quality.

In addition to emission sources, meteorological factors such as wind speed, wind direction, temperature, and humidity also play a crucial role in influencing the dispersion, transport, and chemical transformation of air pollutants [14,15,16]. These atmospheric conditions significantly affect how pollutants spread and how long they persist in the environment. In tropical urban areas, such conditions can vary between daytime and nighttime, shaping the dynamics of $PM_{2.5}$ concentrations and affecting population exposure patterns.

In this context, the present study hypothesizes that a local emission and agricultural burning activities, in combination with meteorological variables, was a major contributor to the elevated $PM_{2.5}$ concentrations observed during the January 2025 episode. To test this hypothesis, we analyze the diurnal variations in $PM_{2.5}$ concentrations and examine their relationship with meteorological variables across daytime and nighttime periods. The goal of this short-term investigation is to improve our understanding of $PM_{2.5}$ pollution dynamics and support the development of targeted, evidence-based air quality interventions.

2. METHODOLOGY

2.1 Study area

The study was conducted at the Institute of Technology of Cambodia (ITC), located in Phnom Penh (11.5683° N, 104.8885° E). The campus is situated in a densely developed urban area with flat terrain and limited green space. The surrounding environment includes a mix of residential area, commercial activities, traffic congestion, and nearby industrial operations, all of which contribute to local emission sources of air pollution. Numerous research initiatives have been conducted at ITC, including the AIR-SATREPS Project (2022-2027). The present study highlights a specific component of this project, focusing on the variation in $PM_{2.5}$ concentrations, and its association with meteorological conditions.

2.1 Data collection and calibration

$PM_{2.5}$ concentrations in Phnom Penh, and across the country, increased dramatically in early 2025. In this study, variations in $PM_{2.5}$ concentrations were measured during the haze period, from January 1st to 30th, 2025, to better understand their dynamics in relation to meteorological variations.

$PM_{2.5}$ concentrations were continuously monitored at ITC using a sensor-based Ambient Air Monitor (Particles Plus, Model EM11000, Japan). Meteorological variables, including wind speed (m/s), wind direction, humidity (%), temperature (°C), and pressure (hPa), were measured using a Smart Weather Sensor (Lufft, WS700-UMB, Japan), which was connected with an Air Quality Monitoring System (AQMS). The sensor-based $PM_{2.5}$ monitoring data was validated against filter-based mass concentration collected using a low-volume (LV) air sampler (Sibata, Model LV-250R, Japan). Polytetrafluoroethylene (PTFE) filters were collected for two-month period (1st September to 31st October 2024; 47 filter samples), and their masses were measured under controlled humidity and temperature conditions in a weighing chamber (Satorius Cubis, MCA6.6S-2S01-F, Japan). The results were then compared, showing an average difference of less than 20 % (**Fig. 1**).

2.2 Statistical analysis

The data of one-hour measurement were selected and categorized into daytime and nighttime periods from January 1st to 30th, 2025 (372 data points for daytime and 372 for nighttime). Bivariate polar plots were adopted to determine the spatial variation of $PM_{2.5}$ concentration in relation to wind speed and wind direction. These polar plots were developed using the Openair R package (version 4.4.1), as described in the previous studies [17,18]. Pearson's correlation analysis was conducted to determine the relationships between $PM_{2.5}$ and meteorological variables during daytime and nighttime, using the vegan package (version 2.6.4). All the packages used in this study are available in R-studio [19].

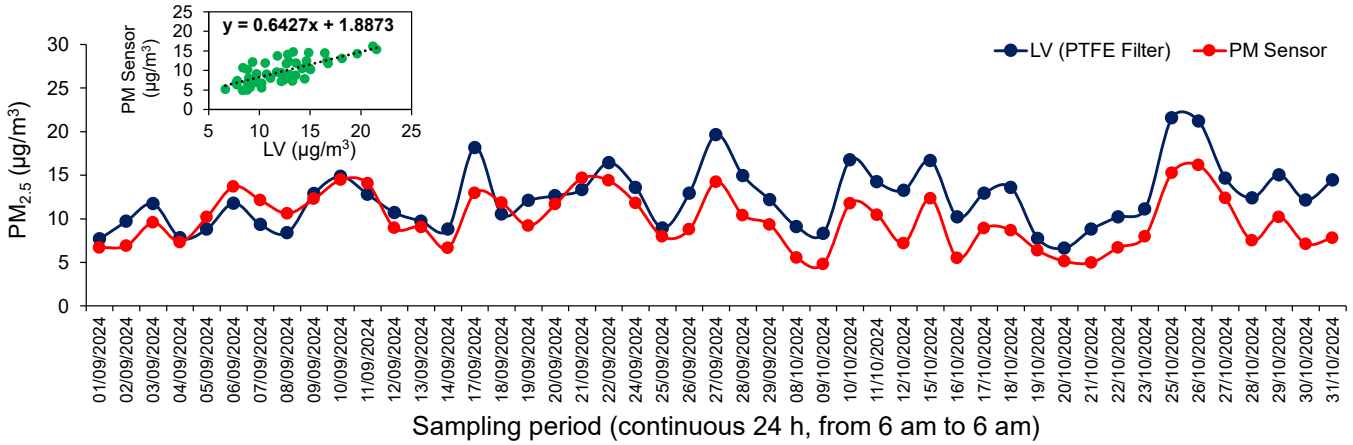


Fig. 1. PM2.5 Sensor calibration by comparing filter sampling (filter-based) and optical particle count (sensor-based). The sensor data was plotted from 24-hour average, along with the collection time of the filter sampling.

3. RESULTS AND DISCUSSION

$PM_{2.5}$ concentrations were monitored continuously for one month in January 2025, and the data were categorized into daytime (7 am to 7 pm) and nighttime (7 pm to 7 am) periods (**Fig. 2**). The data were further classified into two events: a high pollution event (22nd to 25th January) and a normal event (1st to 30th January, excluding 22nd to 25th January).

Table 1. Average $PM_{2.5}$ concentrations under the normal event (1st to 30th January, except 22nd to 25th January) and high pollution event (22nd to 25th January).

Variable \ Events		Normal		High pollution	
		Day (n=312)	Night (n=312)	Day (n=48)	Night (n=48)
$PM_{2.5}$ (µg/m³)	Min.	5.98	5.07	22.44	29.85
	Max.	43.90	45.45	81.40	89.39
	Aver.	18.69	22.01	40.06	57.47

Under the normal event, $PM_{2.5}$ concentrations ranged from 5.98 to 43.90 µg/m³ during the daytime and from 5.07 to 45.45 µg/m³ at night. In contrast, under the high pollution event, $PM_{2.5}$ concentrations ranged from 22.44 to 81.40 µg/m³ during the daytime and from 29.85 to 89.39 µg/m³ at night (**Table 1**). On average, $PM_{2.5}$ concentrations under high pollution event was 40.06 µg/m³ during the daytime and 57.47 µg/m³ at night—more than twice the concentration observed under normal event, which averaged 18.69 µg/m³ and 22.01 µg/m³ during the daytime and nighttime, respectively (**Table 1**). The elevated $PM_{2.5}$ concentrations under the high pollution event were likely caused by biomass burning. In Cambodia, farmers typically prepare fields by burning crop residues, especially rice straw/stubble, during the dry season from January to February [13]. Pollution levels during this period, 22nd to 25th January, were likely influenced by rice straw/stubble burning from agricultural areas such as Kampong Cham, Pursat, Banteay Meanchey, Siem Reap Provinces (e.g., $PM_{2.5}$ > 120 µg/m³, MoE). The wind direction during this event was came from the northeast part of Cambodia, passing Kampong Cham province (Trajectory model). Additionally, during the high pollution event, the concentrations of key gases such as CO and NO_x, originating from agricultural burning [7,12], were approximately more than two times higher than those during the normal event (unpublished results).

Wind speeds were observed to be higher during the daytime (ranging from 0.56 to 0.83 m/s) and gradually decreased throughout the evening, remaining below 0.42 m/s during the nighttime (**Fig. 2**). To better understand the relationship between $PM_{2.5}$ levels, wind speed, and wind direction, the polar plots were employed (**Fig. 3**). Red and blue colors represent high and low $PM_{2.5}$ concentrations, respectively. In this study, $PM_{2.5}$ concentrations during the daytime were lower than that at nighttime, regardless of wind speed and direction. The higher $PM_{2.5}$ during both daytime and nighttime were associated with weak winds speed at 0.28 m/s and 0.14 m/s, respectively, indicating that the pollution accrued at the sampling site due to local emission sources.

Pearson's correlation analysis (**Fig. 4**) indicated that $PM_{2.5}$ concentration was weakly and negatively correlated with wind speed during both daytime ($n = 372$, $r = -0.45$, $p < 0.001$) and nighttime ($n = 372$, $r = -0.40$, $p < 0.001$), supporting the results from the polar plots (**Fig. 3**). Higher wind speeds (up to 0.8 m/s) during daytime, predominantly from the northeast (riverside) (**Fig. 3**), likely enhanced the horizontal transport and dilution of $PM_{2.5}$ [14], which may have contributed to lower $PM_{2.5}$ concentrations during daytime compared to nighttime. Additionally, the elevated $PM_{2.5}$ concentrations observed at nighttime may be attributed to transformation processes. $PM_{2.5}$ concentrations are influenced by secondary processes such as chemical reactions that can sustain or increase concentrations [20], involving gaseous precursors (i.e., CO and NO_x, which showed a positive correlation with $PM_{2.5}$ levels; unpublished result).

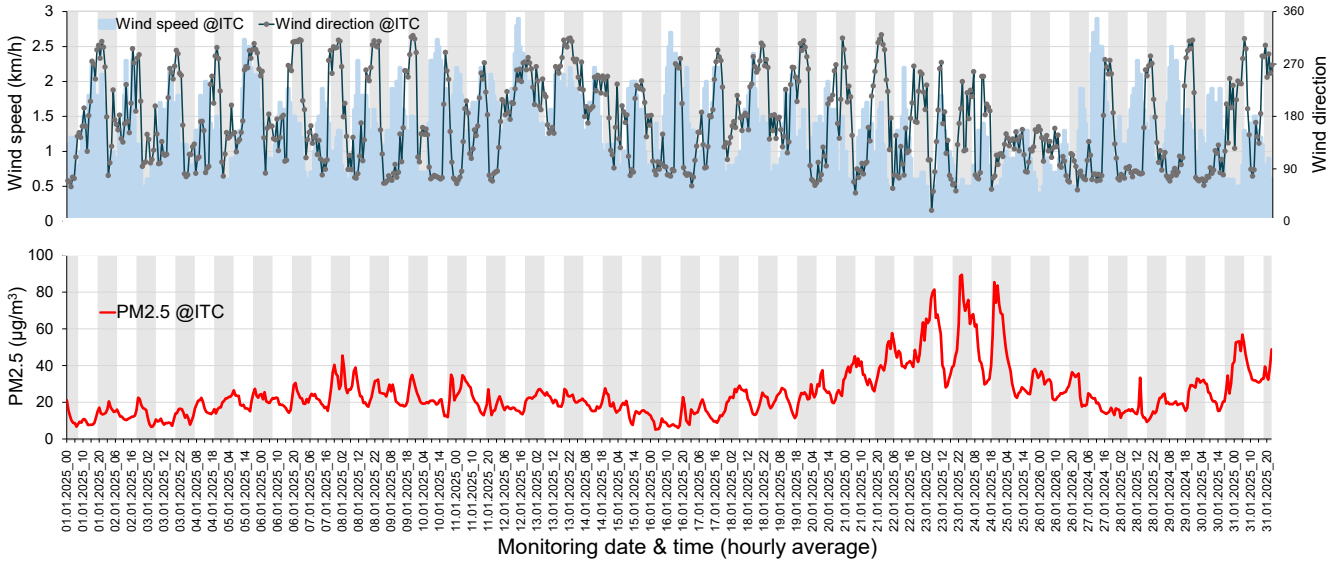


Fig. 2. $PM_{2.5}$ concentration, wind speed, and wind direction distribution at ITC from January 1st to 30th, 2025.

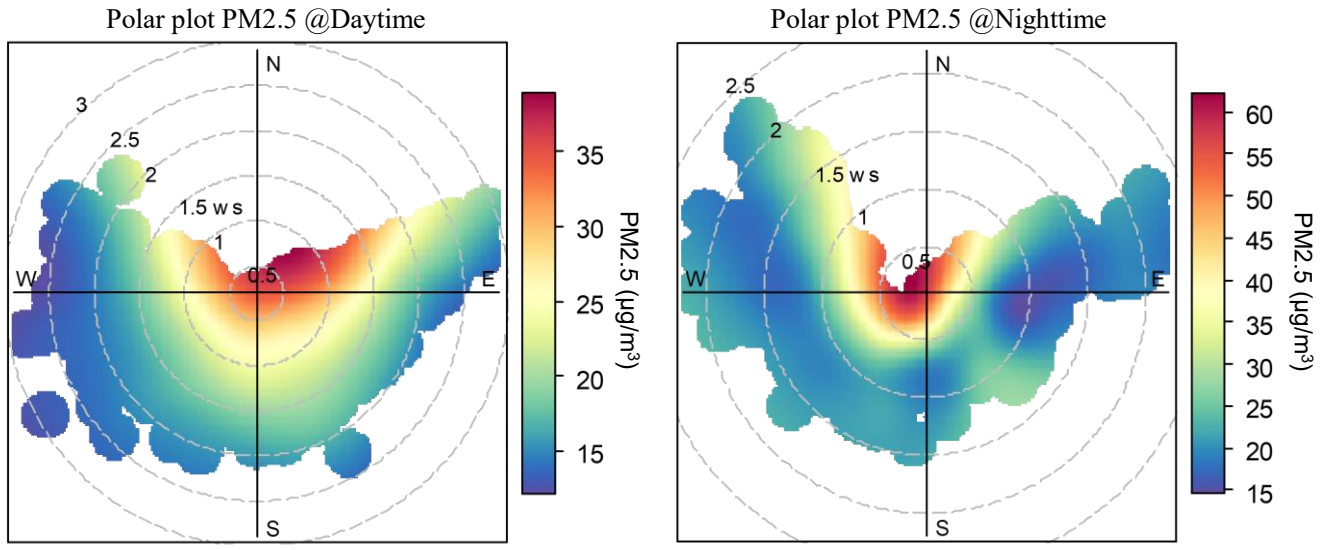


Fig. 3. Polar plots of daytime and nighttime $PM_{2.5}$ concentrations. The red and blue color represent the high and low concentrations of $PM_{2.5}$, respectively. ws stands for wind speed (km/h).

During the daytime, air temperature also showed a weak negative correlation with $PM_{2.5}$ concentration ($n = 372$, $r = -0.15$, $p < 0.01$), while no significant correlation was observed at nighttime ($n = 372$, $r = 0.06$, $p > 0.05$). This suggests that higher daytime temperatures may contribute slightly to lower $PM_{2.5}$ concentrations, likely due to enhanced atmospheric mixing. Temperature gradients, which measure the change in temperature with altitude, play a key role in atmospheric dispersion by influencing vertical movement and turbulent mixing, affecting atmospheric stability [21]. Additionally, photolysis processes during the daytime may contribute to lower $PM_{2.5}$ concentrations by reducing NO_x concentrations.

Throughout the observation period, humidity and atmospheric pressure did not show a significant influence on $PM_{2.5}$ variations during either daytime or nighttime ($P > 0.05$; **Fig. 4**). This may be attributed to the low humidity levels recorded, averaging 53 % during daytime and 63 % at nighttime. Under such dry conditions, the effect of humidity on $PM_{2.5}$ concentrations is limited, as reduced moisture availability suppresses the hygroscopic $PM_{2.5}$ growth, thereby minimizing their size and mass increase [22]. In contrast, higher humidity can enhance the hygroscopic $PM_{2.5}$ growth [22]. Since this study was conducted during the dry season, further investigation is needed to assess seasonal variations.

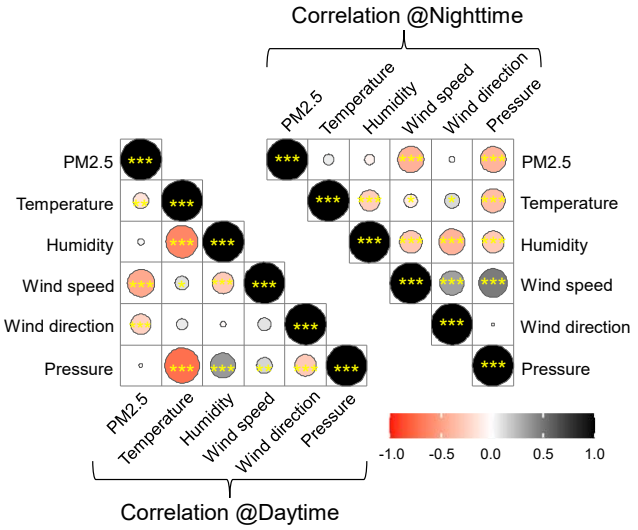


Fig. 4. Correlation between $PM_{2.5}$ and mereological data during daytime and nighttime. The negative and positive correlation coefficients are represented in the red and black, respectively (n = 372; *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$).

4. CONCLUSIONS

This study highlights how meteorological factors influence $PM_{2.5}$ variation in Phnom Penh. $PM_{2.5}$ concentrations showed a distinct diurnal pattern, with 22 % average at nighttime higher than daytime, largely due to the local emissions and the formation of a stable boundary layer that limits vertical dispersion. A significant pollution peak was observed from January 23rd to 24th (up to 85 µg/m^3), accompanied by wind from the northeast, which may have been influenced by transboundary pollution resulting from agricultural burning in other provinces, a common land preparation method (open burning) during the dry season. Polar plots and correlation analysis confirmed significant negative correlations between $PM_{2.5}$ and wind speed, and a daytime negative correlation with air temperature. The humidity averaged 53 % during the daytime and 63 % at nighttime, which limited the growth of hygroscopic $PM_{2.5}$ during the dry season. These findings emphasize the need for continued air quality monitoring, as well as further source-specific and seasonal analyses to support effective air quality management.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided by the **JICA–JST SATREPS** program (Grant No. JPMJSA2102) for this study. Sincere appreciation is also extended to Mr. **Tomoyuki Yamada**, the project coordinator, for his invaluable support in both administrative and technical aspects of the project.

REFERENCES

[1] E. E. McDuffie *et al.*, "Source sector and fuel contributions to ambient PM2.5 and attributable mortality across multiple spatial scales," *Nat. Commun.*, vol. 12, no. 1, pp. 1–12, 2021, doi: 10.1038/s41467-021-23853-y.

[2] J. Rentschler and N. Leonova, "Global air pollution exposure and poverty," *Nat. Commun.*, vol. 14, no. 1, pp. 1–11, 2023, doi: 10.1038/s41467-023-39797-4.

[3] D. W. Dockery *et al.*, "An Association between Air Pollution and Mortality in Six U.S. Cities," *N. Engl. J. Med.*, vol. 329, no. 24, pp. 1753–1759, 1993, doi: 10.1056/nejm199312093292401.

[4] WHO, "WHO global air quality guidelines." [Online]. Available: https://www.who.int/news-room/questions-and-answers/item/who-global-air-quality-guidelines (accessed on 9 April 2025)

[5] S. J. Pai, T. S. Carter, C. L. Heald, and J. H. Kroll, "Updated World Health Organization Air Quality Guidelines Highlight the Importance of Non-anthropogenic PM2.5," *Environ. Sci. Technol. Lett.*, vol. 9, no. 6, pp. 501–506, 2022, doi: 10.1021/acs.estlett.2c00203.

[6] F. Karagulian *et al.*, "Contributions to cities' ambient particulate matter (PM): A systematic review of local source contributions at global level," *Atmos. Environ.*, vol. 120, pp. 475–483, 2015, doi: 10.1016/j.atmosenv.2015.08.087.

[7] S. Lorn *et al.*, "Health Impact Assessment from Rice Straw Production in Cambodia," *Appl. Sci.*, vol. 12, no. 20, 2022, doi: 10.3390/app122010276.

[8] K. Uranishi, H. Shimadera, F. Ikemori, K. Takami, A. Nogami, and S. Sugata, "Modeling study of PM2.5 pollution episode of early spring 2019 in Hokkaido, Japan caused by biomass burning in Northeast China," *E3S Web Conf.*, vol. 530, pp. 1–8, 2024, doi: 10.1051/e3sconf/202453001002.

[9] S. Yin, X. Wang, X. Zhang, M. Guo, M. Miura, and Y. Xiao, "Influence of biomass burning on local air pollution in mainland Southeast Asia from 2001 to 2016," *Environ. Pollut.*, vol. 254, pp. 1–25, 2019, doi: 10.1016/j.envpol.2019.07.117.

[10] K. Lasko, K. P. Vadrevu, V. Bandaru, T. N. Thi Nguyen, and H. Q. Bui, "PM2.5 Emissions from Biomass Burning in South/Southeast Asia – Uncertainties and Trade-Offs," *Biomass Burn. South Southeast Asia*, no. June, pp. 149–169, 2021, doi: 10.1201/9780429022036-12.

[11] K. Lasko, K. P. Vadrevu, and T. T. N. Nguyen, "Analysis of air pollution over Hanoi, Vietnam using multi-satellite and MERRA reanalysis datasets," *PLoS One*, vol. 13, no. 5, pp. 1–21, 2018, doi: 10.1371/journal.pone.0196629.

[12] J. Chen *et al.*, "A review of biomass burning:

Emissions and impacts on air quality, health and climate in China," *Sci. Total Environ.*, vol. 579, no. November 2016, pp. 1000–1034, 2017, doi: 10.1016/j.scitotenv.2016.11.025.

[13] Phnom_Penh_Post, "Nearly 3K illegal burn sites identified across Kingdom." [Online]. Available: https://www.phnompenhpost.com/national/nearly-3k-illegal-burn-sites-identified-across-kingdom?utm_source=chatgpt.com (Accessed on 21 April 2025)

[14] Z. Chen *et al.*, "Influence of meteorological conditions on PM2.5 concentrations across China: A review of methodology and mechanism," *Environ. Int.*, vol. 139, no. February, p. 105558, 2020, doi: 10.1016/j.envint.2020.105558.

[15] C. Lou, H. Liu, Y. Li, Y. Peng, J. Wang, and L. Dai, "Relationships of relative humidity with PM2.5 and PM10 in the Yangtze River Delta, China," *Environ. Monit. Assess.*, vol. 189, no. 11, p. 10, 2017, doi: 10.1007/s10661-017-6281-z.

[16] P. Y. Yin, R. I. Chang, R. F. Day, Y. C. Lin, and C. Y. Hu, "Improving pm2.5 concentration forecast with the identification of temperature inversion," *Appl. Sci.*, vol. 12, no. 1, 2022, doi: 10.3390/app12010071.

[17] S. L. Wright, J. Ulke, A. Font, K. L. A. Chan, and F. J. Kelly, "Atmospheric microplastic deposition in an urban environment and an evaluation of transport," *Environ. Int.*, vol. 136, no. September 2019, 2020, doi: 10.1016/j.envint.2019.105411.

[18] F. Ikemori *et al.*, "Source Contribution Analysis of Polycyclic Aromatic Hydrocarbons in PM2.5 at Three Japanese Cities Using Positive Matrix Factorization with Organic Tracers," *Atmosphere (Basel).*, vol. 16, no. 2, 2025, doi: 10.3390/atmos16020175.

[19] R Core Team, "R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria."

[20] J. H. Seinfeld and S. N. Pandis, *Atmospheric Chemistry and Physics*, Third Edit. Wiley, 2016.

[21] J. Yang, L. Li, X. Zheng, H. Liu, F. Li, and Y. Xiao, "Pollutant Dispersion of Aircraft Exhaust Gas during the Landing and Takeoff Cycle with Improved Gaussian Diffusion Model," *Atmosphere (Basel).*, vol. 15, no. 10, 2024, doi: 10.3390/atmos15101256.

[22] W. S. Won, R. Oh, W. Lee, S. Ku, P. C. Su, and Y. J. Yoon, "Hygroscopic properties of particulate matter and effects of their interactions with weather on visibility," *Sci. Rep.*, vol. 11, no. 1, pp. 1–12, 2021, doi: 10.1038/s41598-021-95834-6.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Gravimetric Analysis of $PM_{2.5}$ Mass Concentration: Daily Behavior in Early Dry Season in Phnom Penh, Cambodia

Vahet Sary[2*], Seyha Doeurn [1,2*], Sreyoun Sorn[2], Sochan Sao[1], Chanthol Peng [1,2], Fumikazu Ikemori [3], Mitsuhiko Hata [4], Masami Furuuchi [4], Chanmoly Or [1]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Graduate School of Integrated Science and Technology, Nagasaki University, 1-14, Bunkyo-machi, Nagasaki, 852-8521, Japan*

[4] *Faculty of Geosciences and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

[*] *Corresponding author: vahetsary93@gmail.com , doeurn.seyha@itc.edu.kh*

Abstract: *Fine particulate matter ($PM_{2.5}$) is one of the most critical air pollutants due to its severe consequences to public health, environmental quality and climate change. Gravimetric analysis of $PM_{2.5}$ mass concentration is the most credible primary standard information for evaluation of pollution levels, assessing health risks, and developing effective mitigation strategies. To contribute to these efforts, this study observes the daily $PM_{2.5}$ mass concentration over a three-month period (November 2024 to January 2025) from the beginning to mid-dry season in Phnom Penh, Cambodia. $PM_{2.5}$ was collected daily using PTFE filters in a low-volume air sampler for 24 hours on the rooftop of a fourth-floor building at the Institute of Technology of Cambodia. Mass concentration was measured by weighing PTFE filters before and after sampling using a Cubic II Ultra Microbalance in the weighing chamber following the set guideline of the Japanese Ministry of Environment. The $PM_{2.5}$ concentration ranged from 7.4 to 81.2 µg/m3, with an average of 21.8±14.5 µg/m³. Though, majority of the samples were within the Cambodian ambient air quality standard (50 µg/m³), more than half of the samples exceeded the WHO 24-hr $PM_{2.5}$ recommended level for public health risks. Daily behavior of $PM_{2.5}$ across the 3 months depicted a slight rise in Novembers, followed by a decrease in December and back to sharp increase by late January. Friday recorded the highest average $PM_{2.5}$, followed by a weekend drop and relative stability early in the week, reflecting differences in emission activities between weekdays and weekends and PM formation process across the week. To gain a deeper understanding of air quality in Phnom Penh, further studies should be conducted over a longer period, including detailed analysis of PM2.5 components to identify specific pollution sources.*

Keywords: Fine particulate matter; $PM_{2.5}$; Daily behavior; Dry Season

1 INTRODUCTION

Fine particulate matter ($PM_{2.5}$), defined as airborne particles with an aerodynamic diameter of less than 2.5 micrometers, has been recognized as a significant environmental and public health concern over the past century [1]. Exposure to $PM_{2.5}$ poses serious health risks, contributing to approximately 4.2 million premature deaths globally in 2015, mainly due to cardiovascular and respiratory diseases [2] $PM_{2.5}$ affects environmental quality by reducing visibility, contributing to haze, and depositing harmful substances into soil and water. It also impacts climate: sulfate aerosols reflect sunlight and cool the atmosphere, while black carbon absorbs heat, accelerates glacier melt, and enhances global warming [1] Wildfire smoke, a major $PM_{2.5}$ source, worsens due to climate change, creating a feedback loop that increases both pollution and warming [3].

[1*] Corresponding author: Vahet Sary
E-mail: vahetsary93@gmail.com; Tel: +855-312 048 708

[2*] Corresponding author: Seyha Doeurn
E-mail: doeurn.seyha@itc.edu.kh; Tel: +855-11 248 886

Monitoring $PM_{2.5}$ is crucial for protecting public health, managing air quality, and supporting climate research. It provides essential data to identify pollution sources, assess compliance with air quality standards such as those set by the WHO or US EPA, and enforce regulatory measures when exceedances occur. Extensive monitoring of $PM_{2.5}$ mass concentrations has been conducted in many parts of the world to understand ambient pollution levels and assess potential risks to human health [4]–[7]. In Kitakyushu, Japan, long-term air sampling programs have been established to monitor $PM_{2.5}$ levels, characterize pollution patterns, and inform risk assessment and management strategies by using low-volume (LV) air sampler[8]. In Charlotte, North Carolina, $PM_{2.5}$ measurements were evaluated by comparing data from a low-cost optical particle counter (PA-II) with a Federal Equivalent Method (FEM) Beta Attenuation Monitor (BAM 1022) [9].

Gravimetric method is used as the basis reference method for PM mass evaluation by many institutions and regulatory agencies. The United States Environmental Protection Agency (U.S. EPA), and World Health Organization (WHO) recommended for using this method to regular monitoring and providing baseline data for health risk assessments and air pollution modeling [10], [11]. Gravimetric analysis is a quantitative method used to determine the mass of particulate matter (PM), such as $PM_{2.5}$, by collecting particles on a filter and then weighing the filter before and after sampling. The difference in mass gives the concentration of PM in the air, usually expressed in µg/m³[12]. Gravimetric analysis is considered the gold standard for $PM_{2.5}$ measurement due to its simplicity, accuracy, and reliability.

Phnom Penh, like many other urban centers in Southeast Asia, has experienced increasing air pollution levels due to its on-going development phase [13]–[15]. Amid growing concerns over $PM_{2.5}$ pollution across Southeast Asian cities, the availability of accurate monitoring data—particularly through gravimetric analysis—remains limited in the Cambodian context. Achieving true mass concentration and understanding the daily behavior of $PM_{2.5}$ are crucial for further analyzing its composition and estimating its impacts on air quality and public health.

This study aims to obtain $PM_{2.5}$ mass concentration in Phnom Penh, Cambodia, by employing the gravimetric analysis method. The daily behavior of $PM_{2.5}$ during the early dry season, from November to January, in relations to its meteorological condition such as temperature, relative humidity, and precipitation, including its day-of-the-week patterns will be investigated and analyzed.

2 MATERIALS AND METHODS

2.1. Sampling site and sampling period

Sampling site of this study is located on a rooftop of a fourth-floor building at the Institute of Technology of Cambodia (ITC), in Phnom Penh, Cambodia, with the coordination of 11° 34' 12.72" N for Latitude and 104° 53' 50.64" E for Longitude (**Fig. 1.)**. ITC is situated along Russian Federation Boulevard in Toul Kork District, a densely populated area with extensive traffic congestion.

As the study aims to observe the daily behavior of $PM_{2.5}$ in the early dry season, the study period also follows the typical dry season in Cambodia (November to April), yet our study period was from November 1st, 2024, and continued until the mid-dry season which was on January 31st, 2025.

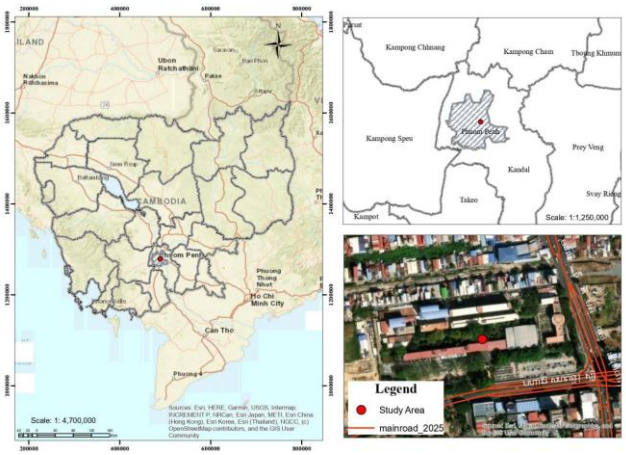


Fig. 1. Location of sampling site (ITC, Phnom Penh)

2.2 $PM_{2.5}$ samples collection and measurement

Fine particulate matter ($PM_{2.5}$) was collected and measured by gravimetric analysis following the standard measurement method of $PM_{2.5}$ collection of the Ministry of Environment, Japan (MOEJ) [16][17]. Polytetrafluoroethylene (PTFE) filter, used for PM mass concentration, was pre-conditioned for 48h in a chamber of temperature 21.5±0.5°C and relative humidity 35±5%, and weighed for pre-sampling mass (M_i), then employed in the low-volume air sampler (Model LV-250R, Sibata Scientific Technology Ltd., Japan). With this sampler, airborne particles in environmental air are aspirated in at a constant flow rate (16.7 L/min) via a suction port. $PM_{2.5}$ particles, passed through an impactor, are then collected on a filter for the period of 23.5 hours daily. After sampling, the filter was conditioned in same condition as the pre-sampling for 48h, then the filter was weighed (M_f). The $PM_{2.5}$ mass concentration was then calculated using the following formula:

$$PM = \frac{M_f - M_i}{Q \times t} \times 10^3 \quad \text{(Eq. 1)}$$

where PM is the particulate matter concentration (µg/m³); M_i is the mass of filters before sampling (mg), M_f is the mass of filters after sampling (mg), t is the sampling time (s), and Q is the volumetric flow rate (m³/s). The resulting PM

concentration is reported in units of micrograms per cubic meter (μg/m³).

2.2 Temperature, relative humidity, and precipitation data

Due to limited availability of full set data on meteorological information from November 2024 to January 2025 in ITC sampling site, the data of temperature, relative humidity, and precipitation used for this study were obtained from the NASA POWER database, accessible from https://power.larc.nasa.gov/data-access-viewer/. This open access database provided accurate meteorological information corresponding to the exact longitude and latitude uploaded. Daily data of temperature, relative humidity and precipitation between November 2024 and January 2025 were downloaded and analyzed in correlation to the daily $PM_{2.5}$ mass concentrations using Pearson correlation method.

3 RESULTS AND DISCUSSION

3.1. Descriptive Statistics of $PM_{2.5}$ Concentrations in early dry season

Fig. 2 depicted the daily $PM_{2.5}$ mass concentration with some meteorological conditions across the 3 months period of the early dry season in Phnom Penh city. **Table 1** summarized the minimum, maximum and average values of $PM_{2.5}$ mass concentration in each month of the early dry season in Phnom Penh city. The lowest concentration was observed in a day of November at 7.4 μg/m³, though most of the daily $PM_{2.5}$ concentration within this month was generally high compared to the later month (**Fig. 2**). The highest concentration was observed in January, the mid-dry season, with the value of 81.2 μg/m³. The overall three-months average $PM_{2.5}$ mass concentration for this early dry season period was 21.8±14.5 μg/m³. It should be noted that monitoring efforts of $PM_{2.5}$ in Phnom Penh had also been conducted by the Ministry of Environment in Cambodia using the low-cost air quality monitoring sensors in 2017, 2018 and 2019; their annual average concentrations were respectively increasing from 13.6 to 19.3, and to 21.2 μg/m³ [18]. Though the values were within the Cambodian ambient air quality standard for annual average of $PM_{2.5}$ (25 μg/m³) [19], they were far exceeded the WHO annual average (5 μg/m³) [20].

Similarly, for the current data, comparing to the Cambodian ambient air quality standard for 24h average of $PM_{2.5}$ (50 μg/m³) [19], majority of samples across these 3 months stayed within this standard, with only four peak episodes observed in the end of January. However, comparing to the WHO Air Quality Guideline for 24h $PM_{2.5}$ level [20], more than half of the total sampling days (61% of n=88) exceeded the recommended value of 15 μg/m³; with 55%, 45% and 83% exceedances occurred in November, December and January, respectively (**Fig. 2**). Growing evidences pointed out the associated risks of diseases and mortality due to exposure to $PM_{2.5}$. n United States, Canada, and Cuba, about 28,000 premature deaths were recorded due to the association of the exposure of urban $PM_{2.5}$ [2]. Long term exposure to high $PM_{2.5}$ concentrations could cause significant risks to public health. Effective implementation of air pollution control measures shall be taken to reduce $PM_{2.5}$ pollutant level and safe-guard the well-being of the population.

Table 1. Minimum, maximum and average $PM_{2.5}$ values in the 3 months of dry season in Phnom Penh

$PM_{2.5}$ Mass Concentration (μg/m³)	November (n=29)	December (n=29)	January (n=30)
Min	7.4	8.8	8.7
Max	39.9	27.4	81.2
Average	17.8	14.7	32.4
Std	7.3	4.3	19.4

3.2. $PM_{2.5}$ daily behavior in early dry season, in accordance with meteorological factors

Observing the daily behavior pattern of $PM_{2.5}$, as illustrated in **Fig. 2**, a slight rise was observed at the beginning to the middle of November, followed by a modest decline and relative stability throughout December, with a noticeable rise in January, reaching higher levels than those observed in November and December. From early to mid-November which is the transition from rainy season to the beginning of dry season, high humidity (79-89%) and residual precipitations (<8mm/day), along with slightly low temperature were observed. Such high humidity with light precipitation events tended to increase particulate matter concentrations due to hygroscopic growth of particles and insufficient removal efficiency, as reported by Zalakeviciute et al. (2018). However, despite the presence of similar condition in our study for these early weeks of November, correlation analysis for the whole month of November revealed almost none to weak relationships between $PM_{2.5}$ and temperature, humidity and precipitation (r=-0.02, r=0.02, r=0.17, respectively, **Table 2**). In December, rather weak correlations were also observed, suggesting minimal effects of temperature, humidity and precipitation to the $PM_{2.5}$ concentrations during this period. In January, only humidity indicated moderate negative correlations to the $PM_{2.5}$ levels (r=-0.57), with noticeable decrease of humidity (46-63%), and high $PM_{2.5}$ concentrations observed. Contrasting to above phenomenon where high humidity can increase PM concentrations, low humidity can also elevate its concentration [22]. Overall, the $PM_{2.5}$ variation in the early dry season period in Phnom Penh, as indicated by the correlation analysis in this study, were more likely to be less influenced by meteorological conditions, and may be more attributed to different emission sources. Further analysis for

$PM_{2.5}$ chemical components shall be conducted to identify specific sources contributing to the atmospheric $PM_{2.5}$ pollution in Phnom Penh.

Table 2. Correlation coefficients of Temperature, Relative humidity and Precipitations with $PM_{2.5}$ values across the 3 months of dry season in Phnom Penh

	November	December	January
Temperature(°C)	-0.02	-0.21	0.05
Relative Humidity (%)	0.02	0.16	-0.57*
Precipitation (mm/d)	0.17	-0.11	-0.14

**denote p-value <0.001*

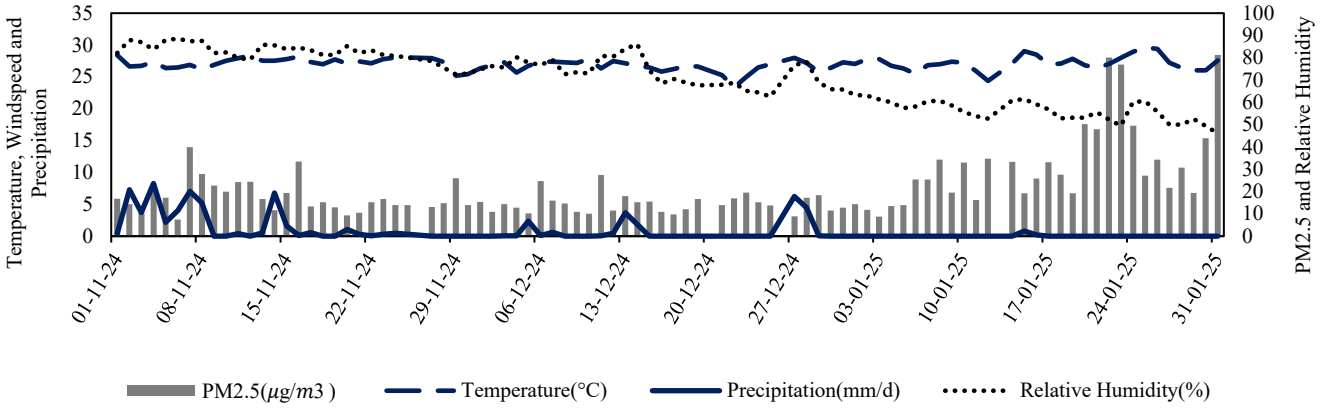
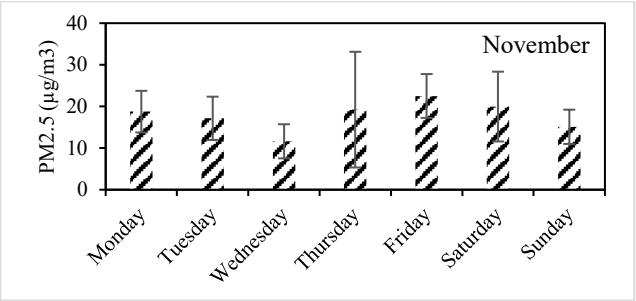


Fig. 2. $PM_{2.5}$ concentration, its daily behavior, and meteorological conditions in Phnom Penh in early dry season

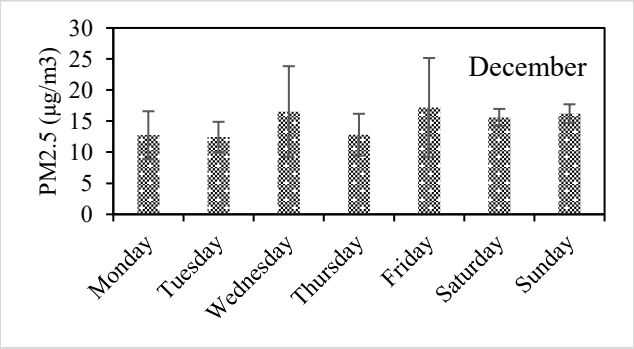
3.2. Day-of-week patterns of $PM_{2.5}$

Analysis of day-of-week pattern aims to investigate the variation of $PM_{2.5}$ throughout the week and to identify which specific day may exhibit higher or lower $PM_{2.5}$ during each month of the early dry season in Phnom Penh city. **Fig. 3 (a), (b)**, and **(c)** respectively shows the average $PM_{2.5}$ concentration of each day of the week in November, December and January. The highest averages were consistently observed during the late workweek, particularly on Fridays. At the beginning of the week, starting from Monday, average $PM_{2.5}$ level tended to remain relatively stable, followed by a slight increase toward Friday. Concentrations then generally declined over the weekend, on Saturday and Sunday. This pattern of variations in the average $PM_{2.5}$ level was similarly observed for all the three months across this early dry season period. This finding is consistent with previous studies by Motallebi et al. (2003), and Lough et al., (2006), which also reported that Friday exhibited the highest weekly $PM_{2.5}$ concentrations, while weekend, particularly Sunday, showed the lowest.

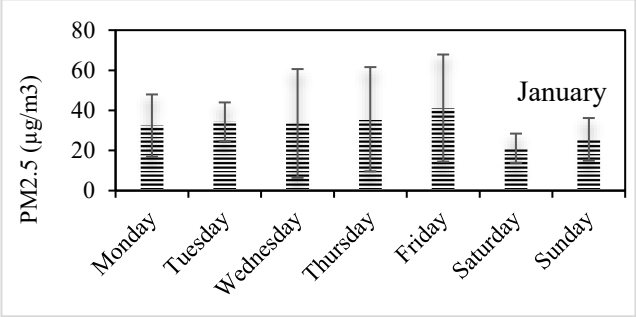
The relatively higher concentrations during weekdays, rising toward Friday, can be tentatively explained by emissions linked to motor vehicle activity and the formation processes of particulate matter. However, further analysis of chemical components is necessary to clarify this pattern more definitively. From actual observation, the sampling site in this study is located in urban area surrounding by major roads that experience heavy traffic from Monday to Friday. Furthermore, as PM consists of primary and secondary pollutants, formation of secondary pollutants may not occurred immediately after emission [25]. This delayed formation can lead to accumulation and higher concentrations toward the late workweek i.e., the Friday. The decreasing concentration patterns during weekends could be attributed to the lower numbers of motor vehicles compared to weekdays, as travel activities and industrial operations are typically reduced on weekends.



(a) Average $PM_{2.5}$ by days in November, 2024



(b) Average $PM_{2.5}$ by days in December, 2024



(c) Average $PM_{2.5}$ by days in January, 2025

Fig. 3. Days of week patterns of $PM_{2.5}$ in early dry season for (a), November (b), December (c), January. (Data Analysis: Average of each Days-of-the-week in November, December and January along its standard deviation were plotted. E.g. Monday of November, all data from 4 Mondays of the 4 weeks of November were calculated for average and standard deviation to plot the above data.)

4 CONCLUSIONS

This study presented the daily and monthly variation of $PM_{2.5}$ mass concentration in Phnom Penh during the early dry season (November 2024 to January 2025), using gravimetric analysis. The results revealed that $PM_{2.5}$ levels exhibited a decreasing trend from November to December, followed by a sharp increase in January, where the highest mass concentration reached 81.19 µg/m³. Although most days were within the Cambodian ambient air quality's 24-hour $PM_{2.5}$ limit (50 µg/m³), comparing to the WHO AQG recommended level for 24-hour $PM_{2.5}$ (15 µg/m³), more than half of the total sampling day were exceeded, suggesting potential public health risks during long term exposure. Furthermore, day-of-week pattern was observed for each month. Friday consistently showed the highest average concentration, followed by a decrease over the weekend and relatively stable levels at the beginning of the week, suggesting a varying pattern of $PM_{2.5}$ emission and formation between weekdays and weekends.

These results provide a valuable baseline for air quality assessment in Cambodia and highlight the need for continuous monitoring, advanced source identification, and targeted mitigation strategies to protect public health. To gain a deeper understanding of air quality in Phnom Penh, it is recommended that future research should include long-term monitoring across multiple seasons and integrate chemical composition analysis of $PM_{2.5}$ (e.g., organic carbon, elemental carbon, trace metals) to identify specific emission sources.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided by the **JICA–JST SATREPS** program (Grant No. JPMJSA2102) for this study. Sincere appreciation is also extended to Mr. **Tomoyuki Yamada**, the project coordinator, for his invaluable support in both administrative and technical aspects of the project.

REFERENCES

[1] N. R. Council, G. Affairs, T. for S. Program, and C. on I. S. in the U. S. E. P. Agency, *Sustainability and the US EPA*. National Academies Press, 2011.

[2] N. Fann, A. D. Lamson, S. C. Anenberg, K. Wesson, D. Risley, and B. J. Hubbell, "Estimating the National Public Health Burden Associated with Exposure to Ambient PM 2.5 and Ozone," *Risk Anal.*, vol. 32, no. 1, pp. 81–95, 2012, doi: 10.1111/j.1539-6924.2011.01630.x.

[3] WMO, "Vicious circle of climate change, wildfires and air pollution has major impacts," World Meteorological Organization.

[4] G. Snider *et al.*, "Variation in global chemical composition of PM 2.5: emerging results from SPARTAN," *Atmos. Chem. Phys.*, vol. 16, no. 15, pp. 9629–9653, 2016.

[5] R. Nakatsubo *et al.*, "Influence of marine vessel emissions on the atmospheric PM2. 5 in Japan's around the congested sea areas," *Sci. Total Environ.*, vol. 702, p. 134744, 2020.

[6] M. Shupler, W. Godwin, J. Frostad, P. Gustafson, R. E. Arku, and M. Brauer, "Global estimation of exposure to fine particulate matter (PM2. 5) from household air pollution," *Environ. Int.*, vol. 120, pp. 354–363, 2018.

[7] D. J. Lary *et al.*, "Estimating the global abundance of ground level presence of particulate matter (PM2. 5)," *Geospat. Health*, vol. 8, no. 3, p. S611, 2014.

[8] X. Zhang, Y. Eto, and M. Aikawa, "Risk assessment and management of PM2. 5-bound heavy metals in the urban area of Kitakyushu, Japan," *Sci. Total Environ.*, vol. 795, p. 148748, 2021.

[9] B. I. Magi, C. Cupini, J. Francis, M. Green, and C. Hauser, "Evaluation of PM2.5 measured in an urban setting using a low-cost optical particle counter and

a Federal Equivalent Method Beta Attenuation Monitor," *Aerosol Sci. Technol.*, vol. 54, no. 2, pp. 147–159, 2020, doi: 10.1080/02786826.2019.1619915.
[10] E. Triantafyllou, E. Diapouli, E. M. Tsilibari, A. D. Adamopoulos, G. Biskos, and K. Eleftheriadis, "Assessment of factors influencing PM mass concentration measured by gravimetric & beta attenuation techniques at a suburban site," *Atmos. Environ.*, vol. 131, pp. 409–417, 2016, doi: https://doi.org/10.1016/j.atmosenv.2016.02.010.
[11] S. O'Connor, P. F. O'Connor, H. A. Feng, and K. Ashley, "Gravimetric Analysis of Particulate Matter using Air Samplers Housing Internal Filtration Capsules.," *Gefahrstoffe, Reinhaltung der Luft = Air Qual. Control*, vol. 74, no. 10, pp. 403–410, Oct. 2014.
[12] E. Hanks, "SLAMS/NAMS/PAMS Network review guidance," *Monit. Qual. Assur.*, 1998.
[13] V. San, V. Spoann, and J. Schmidt, "Industrial pollution load assessment in Phnom Penh, Cambodia using an industrial pollution projection system.," *Sci. Total Environ.*, vol. 615, pp. 990–999, Feb. 2018, doi: 10.1016/j.scitotenv.2017.10.006.
[14] P. Sokharavuth *et al.*, "Air pollution mitigation assessment to inform Cambodia's first clean air plan.," *Environ. Res.*, vol. 220, p. 115230, Mar. 2023, doi: 10.1016/j.envres.2023.115230.
[15] M. Amin *et al.*, "Characterization of carbonaceous components and PAHs on ultrafine particles in Phnom Penh, Cambodia.," *Environ. Monit. Assess.*, vol. 196, no. 10, p. 895, Sep. 2024, doi: 10.1007/s10661-024-13045-8.
[16] Ministry of the Environment, "Analysis of fine particulate matter components | Manual for measuring fine particulate matter (PM2.5) components in the atmosphere | Air quality and automobile measures | Ministry of the Environment," Ministry of the Environment.
[17] Asia Center for Air Pollution Research, "Monitoring of PM2.5 - Experience in Japan," 2020.
[18] MOE, "Clean air plan of Cambodia," 2021.
[19] DAQNM, "Air quality management," UNESCAP.ORG.
[20] WHO, "WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.," 2021.
[21] R. Zalakeviciute, J. López-Villada, and Y. Rybarczyk, "Contrasted effects of relative humidity and precipitation on urban PM2.5 pollution in high elevation urban areas," *Sustain.*, vol. 10, no. 6, 2018, doi: 10.3390/su10062064.
[22] L. Zhang, Y. Cheng, Y. Zhang, Y. He, Z. Gu, and C. Yu, "Impact of air humidity fluctuation on the rise of PM mass concentration based on the high-resolution monitoring data," *Aerosol Air Qual. Res.*, vol. 17, no. 2, pp. 543–552, 2017, doi: 10.4209/aaqr.2016.07.0296.
[23] N. Motallebi, H. Tran, B. E. Croes, and L. C. Larsen, "Day-of-week patterns of particulate matter and its chemical components at selected sites in california?," *J. Air Waste Manag. Assoc.*, vol. 53, no. 7, pp. 876–888, 2003, doi: 10.1080/10473289.2003.10466229.
[24] G. C. Lough, J. J. Schauer, and D. R. Lawson, "Day-of-week trends in carbonaceous aerosol composition in the urban atmosphere," *Atmos. Environ.*, vol. 40, no. 22, pp. 4137–4149, 2006, doi: 10.1016/j.atmosenv.2006.03.009.
[25] N. Zhao, Y. Liu, J. K. Vanos, and G. Cao, "Day-of-week and seasonal patterns of PM2. 5 concentrations over the United States: Time-series analyses using the Prophet procedure," *Atmos. Environ.*, vol. 192, pp. 116–127, 2018.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Title: Driving Factors of Particulate Matter ($PM_{2.5}$ and PM_{10}) Dynamics During Heavy Air Pollution in Phnom Penh, Cambodia

Sokkeav Srong [1*], Sochan Sao [2*], Seyha Doeurn [1,2], Chanthol Peng [1,2], Saksoleakhena Sambath [2], Muhammad Amin [3], Fumikazu Ikemori [4], Mitsuhiko Hata [3], Masami Furuuchi [3], Chanmoly Or [2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Faculty of Geosciences and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

[4] *Graduate School of Integrated Science and Technology, Nagasaki University, 1-14, Bunkyo-machi, Nagasaki, 852-8521, Japan*

[*] *Corresponding author: srongsokkeav11@gmail.com; saosochan2015@gmail.com*

Abstract: *Frequent haze events at the beginning of 2025 worsened air pollution in Phnom Penh, Cambodia. Variation of daily particulate matter (PM) concentrations were observed, prompting this study to examine $PM_{2.5}$ and PM_{10} levels and their relationship with meteorological conditions. Data were collected using the Ambient Air Monitor 11000 series across four monitoring stations: Institute of Technology of Cambodia (ITC), Royal University of Agriculture (RUA), Prek Pnov District Hall (PPH), and Areyksat Commune Hall (ARK). Wind direction and wind speed were monitored at ITC. Daily PM data were collected from January 20th to 26th, 2025. The results at ITC showed that at the highest wind speed of 2.2 km/h, $PM_{2.5}$ and PM_{10} were 39.63 and 95.05 µg/m³, respectively, while at the lowest wind speed of 0.1 km/h, the values were 78.19 and 235.67 µg/m³, respectively. This indicated a negative correlation between wind speed and $PM_{2.5}$ and PM_{10} concentrations, with r = -0.47 and r = -0.45, respectively. The $PM_{2.5}/PM_{10}$ ratio was below 0.5 at all sites. When the wind is driven from the southeast, RUA had $PM_{2.5}$ (24.76–178.51 µg/m³) and PM_{10} (81.01– 1104.55 µg/m³). Similarly, PPH also recorded elevated concentrations, with $PM_{2.5}$ ranging from 26.49 to 149.20 µg/m³ and PM_{10} from 92.41 to 668.72 µg/m³, possibly influenced by the surrounding deserted lands. In contrast, strong easterly winds and better dispersion from the riverside, Areyksat experienced low concentration, with $PM_{2.5}$ ranging from 8.68 to 57.74 µg/m³ and PM_{10} between 56.46 and 97.21 µg/m³. This study underscores the role of wind speed, wind direction, and other sources during the haze periods, emphasizing that weather is not the only contributing factor. However, this study lacked the meteorological data and source apportionment. Future research should include meteorological data at all monitoring locations and source identification to improve the analysis of PM variations.*

Keywords: Particulate matter (PM); $PM_{2.5}$; PM_{10}; Wind direction; Wind speed

1 INTRODUCTION

Rapid urbanization and population growth have significantly contributed to worsening air quality worldwide [1]. In Cambodia, this trend is becoming increasingly evident, with air quality steadily deteriorating and raising concerns about its impacts on public health and the environment [2]. The severity of air pollution is not determined solely by emission volumes; it is also influenced by geographical features, weather conditions, and local environmental factors [3]. In early 2025, Cambodia experienced a notable spike in air pollution, indicated by elevated air quality index and a visible haze over the city [4]. Among the most concerning pollutants are fine particulate matter (i.e., $PM_{2.5}$ and PM_{10}), which are small enough to penetrate deep into the lungs and enter the bloodstream [5]. These particles are especially dangerous and linked to serious health problems like

[1*] Corresponding author: Srong Sokkeav
E-mail: srongsokkeav11@gmail.com; Tel: +855-85 620 115

[2*] Corresponding author: Sochan Sao.
E-mail: saosochan2015@gmail.com; Tel: +855-97 844 5346

respiratory infections, lung cancer, heart disease, and even premature death [6]. Beyond human health, they also reduce visibility, harm the environment, and contribute to global climate change [7], [8].

Geographical and urban factors exacerbate air pollution levels. Urban centers, with their dense traffic, higher energy consumption, and industrial activities, are particularly susceptible to elevated $PM_{2.5}$ and PM_{10} concentrations [9]. Brochu et al. (2011) reported that urban areas typically exhibit higher levels of these pollutants compared to rural regions due to these intensified human activities.

Meteorological conditions also play a pivotal role in shaping air pollution dynamics. Variables such as wind speed, wind direction, temperature, and humidity influence the dispersion, accumulation, and transformation of pollutants in the atmosphere. These meteorological parameters can also use to identify pollution sources. For instance, Ikemori et al. (2025) demonstrated that combining wind patterns with chemical analysis enabled the tracking of pollution pathways and source attribution. Therefore, understanding the interactions among particulate matter variations, emission sources, and accompanying meteorological conditions is essential for effective air quality management.

The present study aims to explore the relationship between meteorological variables and ambient $PM_{2.5}$ and PM_{10} concentrations during the haze episode in Phnom Penh City. The analysis includes average $PM_{2.5}/PM_{10}$ ratios and wind rose diagrams to better understand pollution trends at the four monitoring sites.

2 MATERIALS AND METHODS

2.1. Study area

This study was conducted in Phnom Penh City across four monitoring sites (**Figure 1**), including the Institute of Technology of Cambodia (ITC), the Royal University of Agriculture (RUA), Prek Pnov District Hall (PPH), and Areyksat Commune Hall (ARK). Among the monitoring sites, ITC is situated in a central urban area within Toul Kork District, in the western part of Phnom Penh. It is geographically distinct from the other three stations, located approximately 10 km from RUA in the southeast (Dangkor District), 13 km from ARK in the east, and 17 km from PPH in the far west of the city. The topography of the four monitoring sites selected for this study in Phnom Penh City varies slightly due to their location and surrounding features.

ITC is in a densely built-up, central urban area with flat terrain, surrounded by major roads, buildings, and limited green space. RUA, located in a suburban zone [11], features a mix of flat open fields and scattered residential areas. Its relatively open landscape offers better ventilation than the city center. PPH lies on the northern outskirts of Phnom Penh, in a semi-rural setting with low-rise buildings, open spaces, and nearby water bodies, contributing to more natural airflow. ARK, across the Mekong River from the city, is in a more rural area characterized by agricultural land, fewer structures, and open, flat terrain. These varying topographical features play a role in air pollutants' local distribution and accumulation.

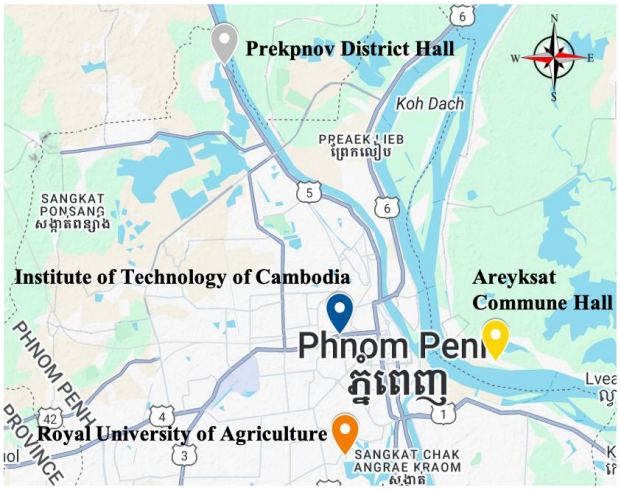


Figure 1. Four study areas representing different surrounding aspects of Phnom Penh City.

2.2. Data collection

In this study, an Ambient Air Monitor (Particles Plus, Model 11000, Japan), referred to as the PM sensor, was installed at four sites (i.e., ITC, RUA, PPH, and ARK), while meteorological parameters (i.e., wind speed and wind direction) were measured only at ITC. Hourly $PM_{2.5}$ and PM_{10} concentrations during the haze period, from January 20 to 26, 2025, were recorded using the PM sensor. Meteorological data were measured using a Smart Weather Sensor (WS700-UMB, Japan), which was connected to an Air Quality Monitoring System (AQMS).

2.3. Statistical analysis

The study employed Spearman rank correlation to examine how daily $PM_{2.5}$ and PM_{10} concentrations were associated with meteorological parameters during the one-week period in January. Furthermore, the $PM_{2.5}/PM_{10}$ ratios were averaged, and wind rose chart was created for ITC monitoring site to illustrate prevailing wind directions and assess the impact on particulate matter distribution.

3 RESULTS AND DISCUSSION

3.1. PM concentrations and meteorological parameters

The time-series analysis revealed that $PM_{2.5}$ and PM_{10} concentrations followed similar patterns throughout the haze episode (**Figure 2**). The highest pollution levels for both particulate types were consistently recorded between January 22nd and 24th. Among the monitoring sites, $PM_{2.5}$ and PM_{10} concentrations at RUA and PPH were notably higher compared to those at ITC and ARK.

Daily average $PM_{2.5}$ and PM_{10} concentrations (μg/m³) recorded at four monitoring sites, along with wind speed data recorded at ITC during the haze period from January 20th to 26th, 2025, are listed in **Table 1**. ITC, located in a central urban setting, showed $PM_{2.5}$ and PM_{10} levels of 40.91 ± 16.43 μg/m³ and 105.14 ± 43.42 μg/m³, respectively. This site is surrounded by construction zones and experiences heavy traffic. The ongoing construction activities significantly contribute to elevated particulate matter concentrations, as typical earthwork, demolition, and foundation operations release substantial amounts of both coarse and fine particles into the air [12]. Additionally, traffic emissions including exhaust gases and tire or brake wear are a significant source of fine particles ($PM_{2.5}$) [13].

Table 1. Summary statistic of $PM_{2.5}$ and PM_{10} concentrations (μg/m³), and wind speed parameter based on daily average at four sites during the haze period January from 20th to 26th.

Site		$PM_{2.5}$ (μg/m³)	PM_{10} (μg/m³)	WS (Km/h)
ITC	Average ± SD	40.91±16.43	105.14±43.42	0.85±0.43
	Min	19.52	50.10	0.1
	Max	89.39	294.56	2.2
RUA	Average ± SD	67.29±33.72	237.93±158.07	N/A
	Min	24.76	81.01	
	Max	178.51	1104.55	
PPH	Average ± SD	65.03±25.83	237.37±109.90	N/A
	Min	26.49	92.41	
	Max	149.20	668.72	
ARK	Average ± SD	23.98 ±11.16	69.97±32.94	N/A
	Min	8.68	29.53	
	Max	57.74	200.90	

Among the four sites, RUA, followed by PPH, reported the highest $PM_{2.5}$ and PM_{10} concentrations. The average $PM_{2.5}$ concentrations were 67.29 ± 33.72 μg/m³ at RUA and 65.03 ± 25.83 μg/m³ at PPH, while PM_{10} concentrations averaged 237.93 ± 158.07 μg/m³ at RUA and 237.37 ± 109.90 μg/m³ at PPH (**Table 1**). These elevated values reflect the influence of nearby emission sources, such as local industries, unpaved roads, and vehicles, which can significantly contribute to both fine and coarse particulate matter ($PM_{2.5}$ and PM_{10}) concentrations, as observed during the study period. Industrial activities often emit $PM_{2.5}$ through combustion processes [14], while unpaved roads contribute to PM_{10} due to dust resuspension caused by vehicular movement [15]. Collectively, these sources have a substantial impact on air quality. In contrast, ARK showed the lowest concentrations of $PM_{2.5}$ and PM_{10}, at 23.98 ± 11.16 μg/m³ and 69.97 ± 32.94 μg/m³, respectively. The lower pollution levels could be attributed to its riverside location as the wind data indicating higher average wind speeds near the riverside compared to urban areas (**Figure 4**), which may facilitate pollutant dispersion and potentially fewer localized emission sources than more urbanized areas [16].

3.2. $PM_{2.5}/PM_{10}$ ratio

This study also calculated the daily average $PM_{2.5}/PM_{10}$ concentration ratios (**Figure 3**). The results revealed that the average $PM_{2.5}/PM_{10}$ ratios for ITC, RUA, PPH, and ARK were 0.39, 0.30, 0.28, and 0.35, respectively. Similar trends were observed in the daily average ratios across all monitoring sites. The $PM_{2.5}/PM_{10}$ ratio at all locations remained below 0.5, indicating that coarse particles, likely from soil dust resuspension, were the dominant type [17].

The $PM_{2.5}/PM_{10}$ ratio variations could be attributed to differing emission sources, meteorological factors, and terrain features at each location. The results suggest that PM_{10} was the dominant particulate matter in all areas, with fine particles ($PM_{2.5}$) accounting for approximately 28% to 39% of the average PM10 mass concentration. ITC showed the highest ratio of 0.39, indicating that fine particulate matter ($PM_{2.5}$) constitutes a significant portion of the total pollution. The elevated presence of fine particles at ITC can be attributed to its urban setting, with dense traffic. These particles stay suspended in the air longer, especially under low wind speed conditions [18].

At RUA (0.30), the ratio was lower compared to ITC (0.39), reflecting a larger proportion of coarse particles (PM_{10}). Located further from the city center and surrounded by residential zones, the area is influenced by road dust, construction debris, and possibly landfills (observed results). These sources typically emit larger particles [15], explaining the lower ratio. Additionally, local wind patterns and topography may promote the settling of coarse particles more rapidly than fine ones [19].

PPH exhibited a similar trend with a ratio of 0.28, indicating that larger particles also dominate particulate matter here. Given PPH's proximity to mixed-use areas including the deserted lands, agricultural zones, the contribution of unpaved roads and farming practices (such as burning or tilling) is significant. The combination of local emissions and wind behavior likely contributes to the higher PM_{10} concentrations compared to $PM_{2.5}$. Meanwhile, ARK displayed a ratio of 0.35, higher than both RUA and PPH but still below ITC. This suggests a more balanced presence of fine and coarse particles. Located near residential areas and along the river, ARK may be affected by both local emissions and pollution transported from upwind sources.

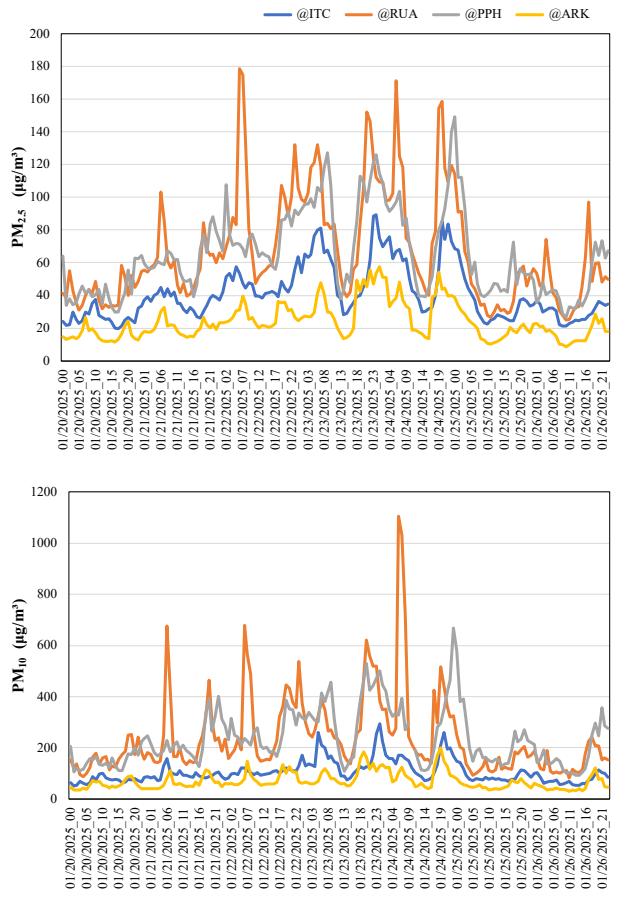


Figure 2. Trends of daily average concentrations of $PM_{2.5}$ and PM_{10} concentrations of the 4 sites during the haze period from January 20th to 26th, 2025.

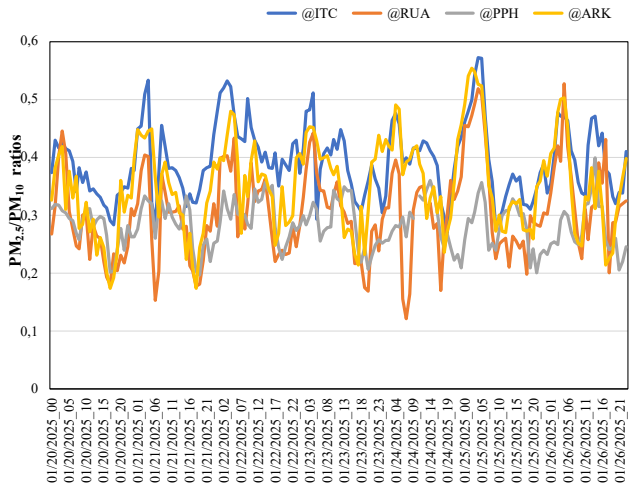


Figure 3. $PM_{2.5}/PM_{10}$ ratios of the 4 sites during the haze period from January 20th to 26th, 2025.

3.3. Relationship between PM2.5, PM10, and Wind Speed

Wind is critical in the horizontal dispersion of air pollutants and can significantly influence pollutant concentrations [20]. This study examined the relationship between wind speed and PM concentrations ($PM_{2.5}$ and PM_{10}) at ITC during the haze period (January 20–26) using the Spearman rank correlation with 168 data samples. The results showed a moderate negative correlation, with coefficients of r = -0.47 for $PM_{2.5}$ and r = -0.45 for PM_{10} (**Table 2**).

Table 2. Spearman-Rank Correlation Coefficients (r) and *p*-Values between wind speed and particulate matter in ITC during the haze period from 20th to 26th January, 2025 (n=168)

Parameter	$PM_{2.5}$	PM_{10}
Correlation Coefficient	-0.47	-0.45
Significance *p*-Value	<0.001	<0.001

During the study period, wind speed at ITC ranged from 0.1 to 2.2 km/h (0.028 to 0.61 m/s). Between January 20th and 21st, the average wind speed was 1.13 km/h (0.31 m/s), with $PM_{2.5}$ and PM_{10} concentrations recorded at 31.02 µg/m³ and 84.00 µg/m³, respectively. As wind speed dropped to 0.72 km/h (0.20 m/s) from January 22nd to 24th, $PM_{2.5}$ and PM_{10} concentrations increased to 53.19 µg/m³ and 134.27 µg/m³ (**Figure 2**). A slight increase in wind speed to 0.77 km/h (0.21 m/s) on January 25th and 26th corresponded with a decline in PM levels to 32.38 µg/m³ ($PM_{2.5}$) and 83.58 µg/m³ (PM_{10}). This inverse relationship highlights how lower wind speeds can lead to pollutant accumulation, while higher wind speeds facilitate pollutant dispersion [21].

3.4. Relationship between $PM_{2.5}$, PM_{10} and Wind conditions

The wind rose diagram offers a clear overview of the wind speed and direction patterns observed across the ITC monitoring site (**Figure 4**). At ITC, the prevailing winds were from the east (27.98%), southeast (27.38%), and south (12.50%). These patterns suggest that pollutant dispersion and transport were largely influenced by winds from these directions during the one week-period. Light winds during the haze period may contribute to limited pollutant dispersion at ITC and surrounding areas.

Pollution levels were notably higher when winds originated from the east (**Figure 4**), which likely carried pollutants from that direction. However, the presence of the Mekong River between ITC and Areyksat could have facilitated some dispersion, as the site's riverside location may have promoted better dispersion due to enhanced airflow [16]. Despite Areyksat being located to the east, ITC, which is situated in a more central area, is exposed to additional sources of pollution, such as traffic emissions and ongoing construction, which contribute to higher PM concentrations. This may explain why PM levels at Areyksat did not exceed those at ITC. Additionally, winds from the southeast played a significant role in influencing pollution

levels at ITC, likely transporting pollutants from the southeast and south, including areas around RUA. In contrast, winds from the northwest were scarcely observed (**Figure 4**), which corresponds to the PPH area.

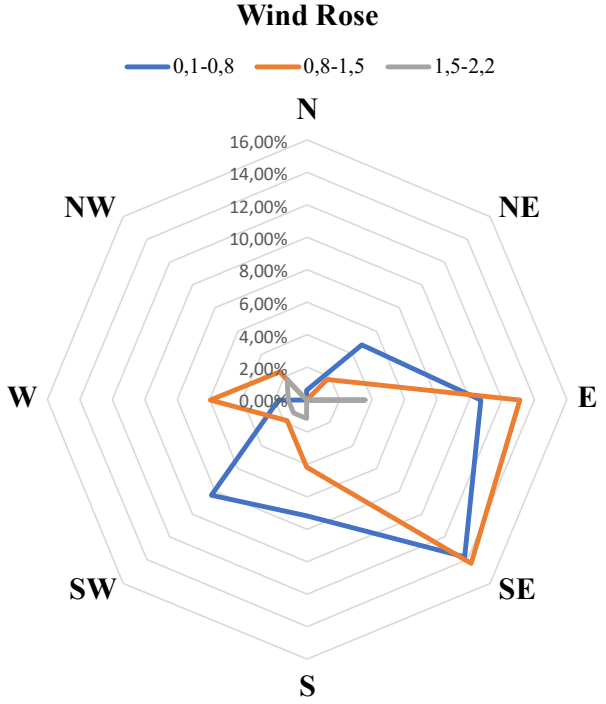


Figure 4. Wind rose diagram showing the hourly average wind direction and wind speed at ITC during the haze period from January 20th to 26th, 2025.

4 CONCLUSIONS

This study examined how wind speed and wind direction influenced $PM_{2.5}$ and PM_{10} levels at four monitoring sites in Phnom Penh during the haze period from January 20th to 26th, 2025. Result showed that both $PM_{2.5}$ and PM_{10} concentrations increased during the periods of low wind speed, indicating a negative correlation between wind speed and PM levels. The average $PM_{2.5}$ /PM_{10} ratios were 0.39 at ITC, 0.30 at RUA, 0.28 at PPH, and 0.35 at ARK, suggesting that ITC site was more affected by finer particles. The wind rose diagrams further revealed that winds from the east and southeast may have transported pollutants from Areyksat and RUA, potentially contributing to higher particulate levels at RUA. To effectively mitigate $PM_{2.5}$ and PM_{10} pollution in the future, more research is needed to identify specific emission sources at each monitoring site.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided by the **JICA–JST SATREPS** program (Grant No. JPMJSA2102) for this study. Sincere appreciation is also extended to Mr. **Tomoyuki Yamada**, the project coordinator, for his invaluable support in both administrative and technical aspects of the project.

REFERENCES

[1] C. Zhan *et al.*, "Impacts of urbanization on air quality and the related health risks in a city with complex terrain," *Atmospheric Chem. Phys.*, vol. 23, no. 1, pp. 771–788, Jan. 2023, doi: 10.5194/acp-23-771-2023.

[2] P. Sokharavuth, "Air pollution mitigation assessment to inform Cambodia's first clean air plan," *Environ. Res.*, 2023.

[3] "European Environment Agency." Accessed: Apr. 12, 2025. [Online]. Available: https://www.eea.europa.eu/en/analysis/maps-and-charts/many-factors-contribute-to-air-1

[4] G. GOH, "Review of Regional Haze Situation for January 2025," ASMC. Accessed: May 01, 2025. [Online]. Available: https://asmc.asean.org/review-of-regional-haze-situation-for-january-2025/

[5] WHO, "WHO global air quality guidelines: particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide." Accessed: Apr. 21, 2025. [Online]. Available: https://www.who.int/publications/i/item/9789240034228

[6] Y.-F. Xing, Y.-H. Xu, M.-H. Shi, and Y.-X. Lian, "The impact of PM2.5 on the human respiratory system," 2016.

[7] B. D. Pienkosz, R. K. Saari, E. Monier, and F. Garcia-Menendez, "Natural Variability in Projections of Climate Change Impacts on Fine Particulate Matter Pollution," *Earths Future*, vol. 7, no. 7, pp. 762–770, Jul. 2019, doi: 10.1029/2019EF001195.

[8] O. US EPA, "Particulate Matter (PM) Pollution." Accessed: Apr. 09, 2025. [Online]. Available: https://www.epa.gov/pm-pollution

[9] C. Chotamonsak and D. Lapyai, "Climate change impacts on air quality-related meteorological conditions in upper northern Thailand," 2020.

[10] P. J. Brochu *et al.*, "Particulate Air Pollution and Socioeconomic Position in Rural and Urban Areas of the Northeastern United States," *Am. J. Public Health*, vol. 101, no. S1, pp. S224–S230, Dec. 2011, doi: 10.2105/AJPH.2011.300232.

[11] S. Chim, S. Chhin, and R. Yim, "Impacts of Air Pollution on Richness and Abundance of Bird Species in Phnom Penh Urban Habitats," *Insight Cambodia J. Basic Appl. Res.*, vol. 6, no. 1, pp. 23–39, May 2024, doi: 10.61945/cjbar.2024.6.1.02.

[12] H. Yan, Q. Li, K. Feng, and L. Zhang, "The characteristics of PM emissions from construction sites during the earthwork and foundation stages: an empirical study evidence," *Environ. Sci. Pollut. Res.*,

vol. 30, no. 22, pp. 62716–62732, Mar. 2023, doi: 10.1007/s11356-023-26494-4.

[13] P. Pant and R. M. Harrison, "Estimation of the contribution of road traffic emissions to particulate matter concentrations from field measurements: A review," *Atmos. Environ.*, vol. 77, pp. 78–97, Oct. 2013, doi: 10.1016/j.atmosenv.2013.04.028.

[14] C. A. Belis, F. Karagulian, B. R. Larsen, and P. K. Hopke, "Clarifications and Corrigendum to 'Critical review and meta-analysis of ambient particulate matter source apportionment using receptor models in Europe' [Atmos. Environ. 69 (2013) 94–108]," *Atmos. Environ.*, vol. 85, pp. 275–276, Mar. 2014, doi: 10.1016/j.atmosenv.2013.12.027.

[15] J. A. Gillies, V. Etyemezian, H. Kuhns, D. Nikolic, and D. A. Gillette, "Effect of vehicle characteristics on unpaved road dust emissions," *Atmos. Environ.*, vol. 39, no. 13, pp. 2341–2347, Apr. 2005, doi: 10.1016/j.atmosenv.2004.05.064.

[16] R. M. Harrison and J. Yin, "Particulate matter in the atmosphere: which particle properties are important for its effects on health?," *Sci. Total Environ.*, vol. 249, no. 1–3, pp. 85–101, Apr. 2000, doi: 10.1016/S0048-9697(99)00513-6.

[17] O. A. Choobari, P. Zawar-Reza, and A. Sturman, " he global distribution of mineral dust and its impacts on the climate system: A review," *Atmospheric Res.*, vol. 138, pp. 152–165, Mar. 2014, doi: 10.1016/j.atmosres.2013.11.007.

[18] P. Kumar, P. Fennell, and R. Britter, "Effect of wind direction and speed on the dispersion of nucleation and accumulation mode particles in an urban street canyon," *Sci. Total Environ.*, vol. 402, no. 1, pp. 82–94, Aug. 2008, doi: 10.1016/j.scitotenv.2008.04.032.

[19] X. Chang *et al.*, "Windbreak efficiency in controlling wind erosion and particulate matter concentrations from farmlands," *Agric. Ecosyst. Environ.*, vol. 308, p. 107269, Mar. 2021, doi: 10.1016/j.agee.2020.107269.

[20] Q. Yang, Q. Yuan, T. Li, H. Shen, and L. Zhang, "The Relationships between PM2.5 and Meteorological Factors in China: Seasonal and Regional Variations," *Int. J. Environ. Res. Public. Health*, vol. 14, no. 12, p. 1510, Dec. 2017, doi: 10.3390/ijerph14121510.

[21] Z. Chen *et al.*, "Influence of meteorological conditions on PM2.5 concentrations across China: A review of methodology and mechanism," *Environ. Int.*, vol. 139, p. 105558, Jun. 2020, doi: 10.1016/j.envint.2020.105558.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Preliminary Assessment of Indoor Air Quality in Residences in Phnom Penh

Channmuny Thanh[1,2], Bophavry Sarav[1,2], Sereyvath Yoeun[1,2], Kimleang Khoeurn[1], Muhammad Amin[3], Chanvorleak Phat[1,2*]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Faculty of Geoscience and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kanazawa 920-1192, Ishikawa, Japan*

[*] *Corresponding author: phatchanvorleak@itc.edu.kh*

Abstract: *Ambient air pollution is still considered as a primary global health concern, resulting in numerous mortality and loss of life expectancy. Outdoor and indoor air pollution, continues to be a significant issue, resulting in around 4.2 million and 3.2 million deaths worldwide respectively per year. The majority of indoor air pollution originates from sources that emit gasses through biomass combustion or release particles into the atmosphere. This study aims to evaluate the indoor air quality related to particulate matter (PM) and polycyclic aromatic hydrocarbons (PAHs) in business households (BH) involved in the roasted duck and normal residential households (NH) in Phnom Penh. The sampling was done by using Ambient Nanoparticle sampler (ANS) within 24h for three days continuously for each location. PAHs were analyzed using High-Performance Liquid Chromatography (HPLC). The PM were found three time higher in BH comparing to NH in all particle size including $PM_{0.1}$, $PM_{0.5}$, PM_{1}, $PM_{2.5}$ and PM_{10}. PAHs levels were observed to be higher in BH compared to NH. The highest concentration of PHAs detected in BH was 197.06 ng/m³, whereas the lowest concentration was 51.01 ng/m³. At BH, one of the PAHs known as benzo[a]pyrene (BaP), an indicator of carcinogenic risk in ambient air, was detected at a concentration of 69, 95 ng/m³ whereas the European Union has established a limit of annual average no more than 1 ng/m³. NH contains below the threshold established by EU (2004/107/EC) for BaP. Increased levels of BaP are associated with several health risks, as stated by the National Institutes of Health. Moreover, the highest concentration of $PM_{2.5}$ was recorded at the BH, measuring 24.13 µg/m³, which exceeds the WHO limit by almost 1.6 times. The $PM_{2.5}$ levels in NH are half of the permissible limit (7.2 µg/m³). These results indicate that charcoal combustion in the BH generated substantial indoor air pollution, characterized by elevated $PM_{2.5}$ and PAH levels, posing significant health risks to residents and the surrounding population.*

Keywords: Air pollution; $PM_{2.5}$; PAHs; residential household; business household

1. INTRODUCTION

Air pollution refers to the contamination of the indoor or outdoor environment by any chemical, physical, or biological substance that alters the natural properties of the atmosphere [1], [2]. Household combustion appliances, automobiles, industrial establishments, and wildfires are prevalent causes of air pollution [3]. Significant pollutants of public health concern encompass particulate matter, carbon monoxide, ozone, nitrogen dioxide, and sulfur dioxide. Both outdoor and indoor air pollution contribute to respiratory and other diseases, serving as significant sources of morbidity and mortality [4]. Data from the World Health Organization (WHO) indicate that nearly the entire of global population (99%) inhales air that surpasses WHO guideline thresholds and contains elevated pollutant levels, particularly in low- and middle-income nations, which have the largest exposures [1].

High levels of outdoor air pollution remain a major concern, causing an estimated 4.2 million deaths globally each year. Indoor air pollution contributes to an additional 3.2 million deaths annually [5]. Globally, air pollution is responsible for about 9% of lung cancer fatalities, 17% of chronic obstructive pulmonary disease deaths, and over 30% of ischemic heart disease and stroke deaths [1], [2], [5].

[1,2*] Corresponding author: Chanvorleak Phat
E-mail: phatchanvorleak@itc.edu.kh; Tel: +855-93 418 000

Domestic cooking with combustion fuels is a significant source of harmful pollutants such as polycyclic aromatic hydrocarbons (PAHs) and their nitrated derivatives (NPAHs) in the environment. These pollutants adversely influence human health and substantially affect the indoor air quality [6]. Therefore, the depth study of chemical composition in particulate matter in each size-fractionated particle is important to understand the threat to human health. Notably, benzanthracene, benz[a]pyrene, and benzo[k]fluoranthene which were identified as the primary contributors and harmful indicators to the total PAHs content [7].

This study aims to evaluate indoor air quality associated with particulate matter (PM) and PAHs in the normal residences and business households.

2. METHODOLOGY

2.1 Chemical and reagents

Solvents and reagents were obtained from Sigma-Aldrich (Saint Louis, USA). The PAH mixture standard was supplied by Nagasaki University which contained the following compunds: dibenz[a,h]anthracene (DBA), indeno [1,2,3-cd] pyrene (IDP), benz[a]anthracene (BaA), chrysene (Chr), benzo[b]fluoranthene (BbF), benzo[k]fluoranthene (BkF), benzo[a]pyrene (BaP), and fluoranthene (Flu). The benzo[a]pyrene-d10 (BaP-d10) was utilized as internal standards which was provided by Wako Pure Chemicals (Osaka, Japan). All solvents utilized in this investigation were in the grade of high-performance liquid chromatography (HPLC) or analytical grade.

2.2 Sample collection

The sampling was done in two different locations including the business households (BH) involved in the roasted duck and the normal residential households (NH) in Phnom Penh. BH normally used charcoal for roasting duck around 9 hours continuously per day, while the NH representing a standard residential which primarily used gas for cooking daily. The sampling was collected by using Ambient Nanoparticle sampler (ANS) within 24h for three days continuously for each location. TX filters (Pall, 2500 QAT -UP, diameter 55mm) were used in this sample collection. ANS has four impactor stages (>10, 10-2.5, 2.5-1.0, 1.0-0.5 μm), an initial filter (IF) stage (0.5-0.1 μm), and a backup filter (<0.1 μm) located at downstream of IF. It was operated with the air inlet at 40L/min [8].

2.3 Sample preparation

Before and after sampling, TX filters were stored in the weighing chamber (PWS-PM 2.5, Tokyo Delec Corp., Japan) at a temperature of 25 ºC, with a relative humidity of 50%, for at least 48hours to study the characteristics of filters before weighing mass. Blank filters were equipped during sampling and transit to validate the possibility of contamination by adsorption gaseous carbon. Micro balance was used to weight the filter mass before and after sampling (Sartorius Cubis MSU2.7S-000-DF). The weighing process weas done in the weighing chamber. After sampling, collected samples were kept in aluminum paper with plastic bag filters. To prevent the sample from oxidation and deterioration of the other chemical components, all sample filters were placed in the refrigerator at -20 ºC until the analysis.

2.4 Particulate matter (PM) calculation

The collected filters were at once weighed by using the gravitational method. The mass concentration was calculated by the following formula:

$$PM = \{(mf - mi)/(t \times q)\} \times 10^6 \quad \textbf{(Eq.1.)}$$

Where:

PM = mass of particulate matter in micrograms per cubic meter ($\mu g/m^3$)

mi = initial mass of filter, in milligrams (mg)

mf = final mass of filter, in milligrams (mg)

t = sampling period (s)

q = volume flow rate (m^3/s)

10^6 = factor of conversion to get the final unit in $\mu g/m^3$.

2.5 PAHs analysis

PAHs were extracted from the TX filters by following [4]. The filter samples were cut into 4 pieces. Only one portion of the filter samples was deposited in a 50 mL Erlenmeyer flask, and 10 mL of dichloromethane (DCM) was added. The sample was sonicated at room temperature for 10 minutes, then filtered accordingly. The extracted sample was then concentrated by adding 20 μL of internal standard and 20 μL of dimethyl sulfoxide (DMSO), followed by evaporation using a centrifugal evaporator at 40°C for 40 minutes. The PAH-extracted sample was analyzed by using HPLC (Shimazu, Japan) coupled with 2 LC30AD pumps, a SIL-30AC auto sample injector, a degasser (DGU-20A5R), a CMB-20A system controller and an integrator (LCsolution software), a CTO-20AC column oven and a RF-20Axs fluorescence detector. The mobile phase was using an isocratic mobile phase with 100% Acetone nitrile (ACN) with a 1.8 mL per min flow rate. Two ODS columns were used (Inertsil ODS-P, 250 × 4.6 mm i.d., 5 μm, GL Sciences, Tokyo, Japan) with their guard column (10 ×4.6 mm i.d.) in tandem. The PAHs mixture comprised 10 previously described PAHs at a concentration of 0.1, 1, 2, 5, and 10 μg/mL were used to create standard curve for quantification.

3. RESULTS AND DISCUSSION

3.1 Particulate matter concentration

The mass concentration of PM ranging from PM 0.1 to PM10 at the two different residents (BH and NH) was shown in the **Fig.1**. The results were shown that the concentration of PM in each size was found as the higher in BH comparing to NH. The highest concentration was found in PM 0.1 in both BH and NH with the amount of 81.2 and 8.2 µg/m^3 respectively. The PM 2.5 and PM 10 concentration were found three times higher in BH comparing to NH. The high concentration of all PM in this study especially PM 2.5 and PM 10 content in the BH might influence by the primary emission sources of charcoal burning that was used for duck roasting. The most harmful particles to human health are fine and ultra fine particles (PM 2.5, PM 1, PM 0.5 and PM 0.1). Particles that are smaller have a higher chance of traveling to the farthest reaches of the respiratory tube and into the bloodstream, where they can damage more to the body [1]. Comparing with the findings of the earlier study conducted on the roof of the Institute of Technology of Cambodia, the concentration of PM 2.5 was found in lower concentration [4]. Moreover, the highest concentration of PM 2.5 was recorded at the BH, measuring 24.13 µg/m³, which exceeds the WHO limit by almost 1.6 times. The PM 2.5 levels in NH are half of the permissible limit (7.2 µg/m³) [1]. However, the limitation of PM 0.5 and PM 0.1 currently were not established yet.

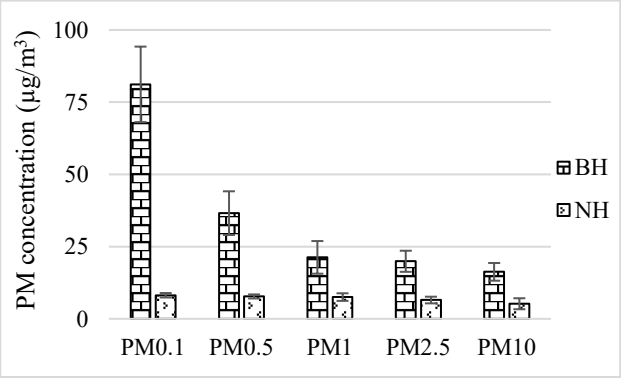


Fig. 1. PM concentration in BH and NH

Chronic obstructive pulmonary disease, lung cancer, acute lower respiratory infections, ischemic heart disease, and stroke are among the health effects of exposure to fine and ultra fine particulate matter in household air pollution [5]. However, the criterion was for general population which was not specific to any vulnerable groups, such as children, the elderly, or pregnant women. The limitation of fine and ultra fine PM should include the specific groups in the future. Nonetheless, the constraints of our research study suggest that the results may serve as a preliminary indicator for a more comprehensive investigation into the whole source profile in further analyses.

3.2 Polycyclic aromatic hydrocarbons (PAHs)

Total concentration of PAHs accumulated in different particle matter in both BH and NH were indicated in the Table 1. In general, the concentration of PAHs was found higher in BH and followed by NH. Highest accumulation of PAHs was found in PM 0.1 of BH with the value of 144.8 ng/m^3, following by PM 0.5, PM 0.1, PM 2.5 and PM 10 with the value of 81.8, 31.3, 18.4 and 11.6 ng/m^3 continuously. No significant different between the accumulation into different size of particle in NH, and the value was found in between 1.9 to 2.3 ng/m^3. The prior investigation of PAHs in the ambient air of Phnom Penh revealed concentrations of 0.18 ng/m3 during the wet season and 0.19 ng/m^3 during the dry season [4]. Comparative analysis of PAH concentrations in other Southeast Asian nations revealed that the amounts of PAH in BH and NH were much greater than those recorded in Hat Yai, Thailand, during the haze period (1.1–4.8 ng/m^3) and under normal conditions (0.3–0.4 ng/m^3) [9]. Consequently, the combustion of charcoal during duck roasting or for everyday cooking can be a significant contribution source of indoor air pollution dissemination.

Table 1. Concentration of total PAHs in both BH and NH

PAHs (ng/m^3)	PM 0.1	PM 0.5	PM 1	PM 2.5	PM 10
BH	144.8±81.4	81.8±39.1	31.3±8.5	18.4±5.8	11.6±2.6
NH	2.3±1.2	2.0±0.4	1.9±0.2	1.5±0.8	1.3±0.8

The concentrations of the PAHs component in the BH and NH were presented in **Fig 2.** and **Fig. 3** respectively. Nine main components of PAHs were detected in this study such as Flu, BaA, BaP, Pyr, Chr, BbF, BkF, DBA, and IDP.

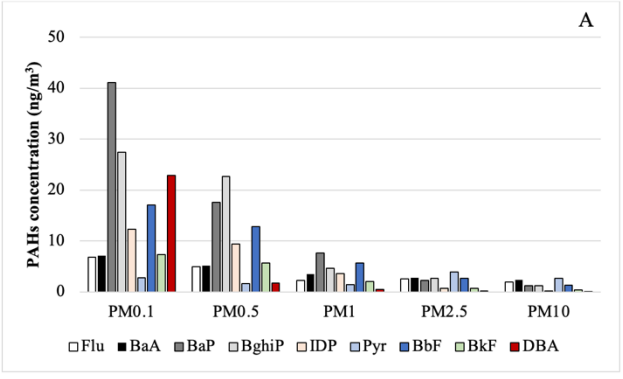


Fig. 2. Concentration of PAH components analysis in BH

The concentration of PAH compounds in BH was highest at PM 0.1. A smaller particle size correlates with a greater concentration of PAHs. Among the identified components of PAHs, BaP is considered as a pollution marker due to its associated with risk of cancer [10]. In the obtained data, it was observed that the smaller size of PM, the higher accumulation of BaP in PM 10, PM 2.5, PM 1, PM 0.5 and PM 0.1 respectively. European Union (2004/107/EC) published a guideline concentration of not exceed 1 ng/m³ for BaP in ambient air [11]. Comparing to our study, BaP accumulated in BH was almost 70 times higher than EU limitation.

BaP in coal is a principal source of organic pollutants and may be easily emitted during combustion, coking, pyrolysis, and various coal processing techniques. This chemical is also found in air, water, soil, sediments, and some foods [10]. According to the previous study, the authors suggested that a consistent source of PAH emissions were contributed by human activities, transportation emissions, or maybe agricultural burning, rather than seasonal change [4]. Hence, due to the limitation of our investigation, these findings are just preliminary data. Additional locations and extended sampling durations are necessary to provide an appropriate assessment of PAH emissions, particularly BaP.

BaP produced tumors in all species tested for which data were reported following exposure including oral, dermal, inhalation, intratracheal, intrabronchial, subcutaneous, intraperitoneal, and intravenous. This was confirmed by the IARC organization [12]. In conclusion, BH or charcoal combustion may be associated with an increased risk of cancer for Cambodian people.

The primary PAH component in NH was Pyr. Nevertheless, it still remained below the threshold, indicating that the maximum permissible concentration of Pyr in the air is typically lower than 15 mg/m^3. The concentration of BaP was determined to be below the EU limit. Therefore, residing in NH and utilizing a minimal quantity of charcoal for cooking will mitigate the risk of air pollution dispersion.

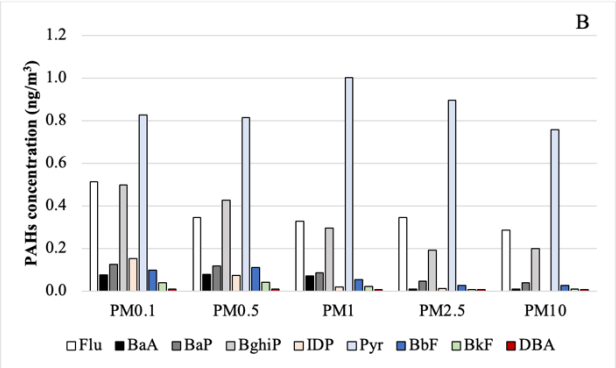


Fig. 3. Concentration of PAH components analysis in NH

4. CONCLUSIONS

The research examined indoor air quality was influences by the cooking method of the people. The roasted duck industry (BH) exhibited markedly elevated levels of particulate matter across all particle sizes. Polycyclic aromatic hydrocarbons, a deleterious air contaminant, were detected in significantly higher amounts at the BH compared to NH. BaP, a cancerogenic factor, were presented in the studied area. PM 2.5 levels were observed to be elevated in comparison to the previous study and in Asian countries. This indicates the trend of elevated air pollution levels in Phnom Penh city. This issue is a significant concern, particularly for vulnerable populations such as pregnant women, young children, and the elderly, who generally spend most of their time indoors.

Future investigations into the sources of PAHs emissions should be conducted to elucidate their origins and behavior, with the objective of developing more comprehensive health risk guidelines, as data for assessing the toxicity of these pollutants is limited.

ACKNOWLEDGMENTS

This study was financially supported by the JICA-JST SATREPS (Grant No. JPMJSA2102). Equipment was supported by the East Asia Nanoparticle Monitoring Network (EA-NanoNet).

REFERENCES

[1] WHO, “WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.,” World Health Organization, Geneva, Switzerland, 2021.

[2] H. M. Adly and S. A. Saleh, “Long-Term Trends in PM10, PM2. 5, and Trace Elements in Ambient Air: Environmental and Health Risks from 2020 to 2024,” *Atmosphere*, vol. 16, no. 4, p. 415, 2025.

[3] P. Suriyawong *et al.*, “Airborne particulate matter from biomass burning in Thailand: Recent issues, challenges, and options,” *Heliyon*, vol. 9, no. 3, 2023.

[4] M. Amin *et al.*, “Characterization of carbonaceous components and PAHs on ultrafine particles in Phnom Penh, Cambodia,” *Environmental Monitoring and Assessment*, vol. 196, no. 10, p. 895, 2024.

[5] WHO, “Introduction to household air pollution. Air pollution and health training toolkit for health workers (APHT),” World Health Organization, 2024.

[6] S. Lara, F. Villanueva, P. Martín, S. Salgado, A. Moreno, and P. Sánchez-Verdú, “Investigation of PAHs, nitrated PAHs and oxygenated PAHs in PM10 urban aerosols. A comprehensive data analysis,” *Chemosphere*, vol. 294, p.

133745, 2022, doi: https://doi.org/10.1016/j.chemosphere.2022.133745.
[7] Y. Zhang, Y. Wang, H. Zheng, and J. Wei, “Increased mortality risk from airborne exposure to polycyclic aromatic hydrocarbons,” *Journal of hazardous materials*, vol. 474, p. 134714, 2024.
[8] M. Furuuchi *et al.*, “Development and performance evaluation of air sampler with inertial filter for nanoparticle sampling,” *Aerosol and Air Quality Research*, vol. 10, no. 2, pp. 185–192, 2010.
[9] J. Chomanee, K. Thongboon, S. Tekasakul, M. Furuuchi, R. Dejchanchaiwong, and P. Tekasakul, “Physicochemical and toxicological characteristics of nanoparticles in aerosols in southern Thailand during recent haze episodes in lower southeast Asia,” *Journal of Environmental Sciences*, vol. 94, pp. 72–80, 2020.
[10] D. Wang, M. H. Rietdijk, L. Kamelia, P. J. Boogaard, and I. M. Rietjens, “Predicting the in vivo developmental toxicity of benzo [a] pyrene (BaP) in rats by an in vitro–in silico approach,” *Archives of Toxicology*, vol. 95, no. 10, pp. 3323–3340, 2021.
[11] EU, “Directive 2004/107/EC of the European Parliament and of the Council of 15 December 2004 Relating to Arsenic, Cadmium, Mercury, Nickel and PAH in Ambien Air,” European Union, 2004.
[12] IARC Working Group on the Evaluation of Carcinogenic Risks to Humans, “Some non-heterocyclic polycyclic aromatic hydrocarbons and some related exposures,” 92, 2010

.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Daytime and Nighttime Monitoring of Ultrafine Particulate Matter ($PM_{0.1}$) during Dry Season in Phnom Penh, Cambodia

Chanrat Vorng [2*], Seyha Doeurn [1,2*], Sokuntheary Sok[2], Sreyvich Sieng [1], Sochan Sao[1], Chanthol Peng [1,2], Fumikazu Ikemori [3], Mitsuhiko Hata [4], Masami Furuuchi [4], Chanmoly Or [1]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Graduate School of Integrated Science and Technology, Nagasaki University, 1-14, Bunkyo-machi, Nagasaki, 852-8521, Japan*

[4] *Faculty of Geosciences and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

[*] *Corresponding author: chanrathvorng145@gmail.com , doeurn.seyha@itc.edu.kh*

Abstract: *Air pollutants consist of various sizes of particulate matter (PM) such as PM_{10}, $PM_{2.5}$, and $PM_{0.1}$. Among these, $PM_{0.1}$, also known as ultrafine particulate matter with diameter of less than 100 nm, can penetrate alveolar regions of the lungs, making it a serious threat to human health. While many studies focus on fine particulate matter ($PM_{2.5}$), the study of ultrafine particles, particularly in Cambodia, is still limited. This study aims to monitor ultrafine particulate matter ($PM_{0.1}$) during daytime and nighttime of dry season in Phnom Penh, Cambodia. $PM_{0.1}$ was sampled on TX filters at the rooftop of a third-floor building in Institute of Technology of Cambodia, using Ambient Nanoparticle Samplers (ANS, Kanazawa University). Nighttime sampling started from 6 PM – 6 AM, and daytime sampling continued from 6 AM- 6 PM covering every Monday and Tuesday for 16 weeks during the dry season (November 2024–February 2025). The results indicated that the mass concentrations of $PM_{0.1}$ were slightly higher during the daytime (7.5±3.0 μg/m^3) than during nighttime (6.5±2.7 μg/m^3). One possible explanation is emissions from nearby motor vehicles during the daytime, though additional chemical component analysis is required to more clearly identify and quantify source contributions to $PM_{0.1}$ Meteorological factors such as temperature, precipitation and windspeed were found to be weakly correlated with the $PM_{0.1}$ mass concentration, except the relative humidity exhibiting significant negative correlation for both the daytime and nighttime $PM_{0.1}$.*

Keywords: Ultrafine particulate matter; $PM_{0.1}$; Mass concentration; Dry Season

1. INTRODUCTION

Ambient particulate matter (PM) is classified by particle size, ranging from larger to smaller particles as PM_{10}, $PM_{2.5}$, and $PM_{0.1}$, each with distinct physical and chemical properties [1]. However, nanoparticles or ultrafine particulate ($PM_{0.1}$; aerodynamic diameter ≤ 0.1 μm) has emerged as a pollutant of increasing concern due to its ability to penetrate alveolar regions of the lungs and translocate into the circulatory system, contributing to a range of adverse health effects including respiratory, cardiovascular, and neurological disorders [2][3][4]. Compared to larger particulate size such as $PM_{2.5}$ and PM_{10}, $PM_{0.1}$ is less studied; largely due to technical challenges associated with its measurement and a lack of routine monitoring infrastructure, particularly in low- and middle-income countries [5][6].

Morawska et al. [4] stated that motor vehicle emissions and biomass burning are among the dominant sources of ultrafine particulate matter ($PM_{0.1}$) in urban and peri-urban environments. Combustion engines, particularly diesel-powered vehicles, emit high concentrations of $PM_{0.1}$ due to incomplete fuel combustion and the formation of soot and volatile organic compounds that nucleate to form ultrafine

[1*] Corresponding author: Chanrat Vorng
Email: chanrathvorng145@gmail.com; Tel: +855-96 670 8707

[2*] Corresponding author: Seyha Doeurn
Email: doeurn.seyha@itc.edu.kh; Tel: +855-11-248-886

particles. Additionally, biomass burning includes open burning agricultural waste, domestic cooking with wood or charcoal, and seasonal forest fires significantly contribute to $PM_{0.1}$ level [7]. Both sources are particularly intensified during the dry season, when stagnant air and limited atmospheric dispersion further elevate $PM_{0.1}$ concentration in the urban atmosphere.

Phnom Penh is undergoing rapid population growth, with Cambodia ranking as the second-fastest expanding urban area in Asia [8][9]. This growth is accompanied by rising levels of vehicle traffic, booming construction, and widespread open burning, all of which are known to release fine and ultrafine particulate matter into the air. Seasonal changes, particularly during the dry months, tend to worsen air quality due to factors like low humidity, temperature inversions, and limited air circulation. These environmental conditions are likely to influence both the concentration and daily fluctuations of ultrafine particles ($PM_{0.1}$). Population growth and increased transportation in urban areas contribute to higher concentrations of PM in the atmosphere. However, published data available on $PM_{0.1}$ level or their temporal variation in Phnom Penh is still limited.

This study gave the focus on $PM_{0.1}$ concentration in Phnom Penh, with a focus on diurnal variation during the dry season. The sub-objectives were to compare daytime (D) and nighttime (N) $PM_{0.1}$ level, assess their temporal trends, and explore associations with meteorological parameters.

2. METHODOLOGY

2.1 Sampling site and Sampling period

The sampling site of this study was situated in the Institute of Technology of Cambodia (ITC), in Phnom Penh, Cambodia (Fig. 1). ITC is in Russian Federation Boulevard which is one of the busiest streets in Phnom Penh. Since it is in the city center, ITC is surrounded by major roads and commercial areas and experiences different vehicular traffic roads including trains, cars and motorbikes [10].

The monitoring of $PM_{0.1}$ in this study was conducted from the beginning to mid-dry season in Cambodia, starting from November 2024 to February 2025. The sampling was conducted weekly every Monday and Tuesday from 6PM to 6AM and 6AM to 6PM, respectively for nighttime and daytime.

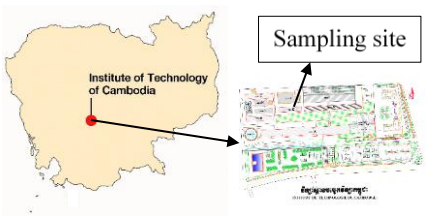


Fig. 1. Sampling location of this study

2.2 Particulate matter collection

Particulate matters in this study were collected using a cascade-type air sampler known as the Ambient Nano Sampler (ANS) (Kanazawa University, Japan), capable of capturing particles across multiple size ranges, including ultrafine particles (UFPs - the $PM_{0.1}$) [11]. The ANS consists of four impactor stages, an initial filter stage (IF) and a backup filter stage, that can collect PM in six distinct size ranges including those exceeding 10 µm, within the ranges of 10–2.5 µm, 2.5–1 µm, 1–0.5 µm, 0.5–0.1 µm, and those smaller than 100 nm ($PM_{0.1}$). The flowrate set for the sampling operation was 40 L/min.

A set of five TX filters (Pallflex), used for sampling and for calculation of PM mass concentration, were loaded to the ANS. The TX filters were conditioned in a controlled weighing chamber for 48 hours at a stable temperature of 21.5 ± 0.5 °C and a humidity of 35 ± 5%, both before and after the sampling process. Their mass before and after sampling were also recorded to calculate the PM mass concentration, following the formula as shown in **Eq. 1**, for the daytime and nighttime sampling. All filters were prepared before sampling for one week. To prevent external contamination, each filter was individually wrapped in aluminum foil and sealed in a zip lock plastic bag.

$$PM\ Concentration = \frac{(m_f - m_i)}{Q \times T} \times 10^{-3} \quad \textbf{(Eq. 1)}$$

Which m_i : mass of initial filter (mg)
m_f : mass after sampling (mg)
Q : volumetric flowrate (L/min)
T: sampling period (min)

2.3 Meteorological Data

The meteorological data in this study were obtained from the NASA Prediction of Worldwide Energy Resources (POWER) database, accessible from https://power.larc.nasa.gov/data-access-viewer/ . This open access database provided accurate meteorological information corresponding to the exact longitude and latitude uploaded. Temperature, relative humidity, precipitation and windspeed for the ITC sampling site were downloaded, and analyzed for the correlation with both the daytime and nighttime $PM_{0.1}$ mass concentration.

3. RESULTS AND DISCUSSION

3.1 Daytime & Nighttime $PM_{0.1}$ Variation

Fig. 2. presents the daytime and nighttime variation of $PM_{0.1}$ in Phnom Penh city during a period of dry season,

starting from November 2024 to February 2025. A total of 32 samples (Nighttime (N) – n=16; Daytime (D) – n=16) were collected for mass concentration. Daytime $PM_{0.1}$ were generally higher than those of the nighttime, as observed in 14 out of the 16 weeks studied. Daytime concentration ranged from 2.8 – 15.5 μg/m^3, with an average of 7.5±3.0 μg/m^3. While for nighttime, the average was lower (6.5±2.7 μg/m^3), and the concentrations ranged smaller, from 2.4 – 11.8 μg/m^3. The sampling site of this study was located in the city center, bordered by major roads with cars, trucks, and motorbikes in the front, and a railway line at the back (**Fig. 1**). During the daytime, these roads, particularly the Russian Federation Boulevard where the sampling site is located, tend to be busy with these fuel-combustion vehicles, mainly during its peak hours of morning and evening time [12]. Similar study by Kim et al. [13] investigated traffic volume in Korean roads and found that higher $PM_{0.1}$ concentrations during the daytime were associated with increased traffic compared to nighttime levels. This may suggest that higher $PM_{0.1}$ concentration in daytime in this current study might also be attributed to the emission from those motor vehicles, which is known as one among the major sources of $PM_{0.1}$ [4][6][14]. However, further analysis of chemical components is necessary to better clarify the sources and their contributions to the $PM_{0.1}$. Furthermore, from **Fig. 2**. temporal variation of $PM_{0.1}$, for both day and nighttime in this dry season period, exhibited a rising trend from early November, peaking in January, and then declining in February. During the dry season, particularly in January, agricultural burning activities by Cambodian farmers and subsequent forest fires were reported in Cambodia [15]. Biomass burning has also been identified as a major source of $PM_{0.1}$ [4][6][16], suggesting that the observed $PM_{0.1}$ peak may be linked to this source. However, further analysis of wind direction and chemical indicators is needed to provide stronger evidence.

From our review, the diurnal variation of $PM_{0.1}$ during the daytime and nighttime, particularly in Southeast Asian countries, were only reported by a few studies [17][18], and ours may be the first to report $PM_{0.1}$ mass concentration using TX filters, while the formers reported from the use of Quartz Fiber Filter (QFF). In Chiang Mai, the studies conducted between March and April, 2020, found the average nighttime $PM_{0.1}$ (31.7±15.4 μg/m^3) higher than average daytime (29.5±8.2 μg/m^3) [17]. Yet, its temporal trend showed fluctuations between nighttime and daytime data, with higher nighttime $PM_{0.1}$ level observed in early March, then declined and lower than daytime from mid-April—implying to the increased biomass burning activity in March and more unstable atmospheric conditions during daytime in April [17]. Both average daytime and nighttime concentrations showed a 4- to 5-fold difference from the average value in this current study, likely due to the type of filters used, or differences in study periods and locations.

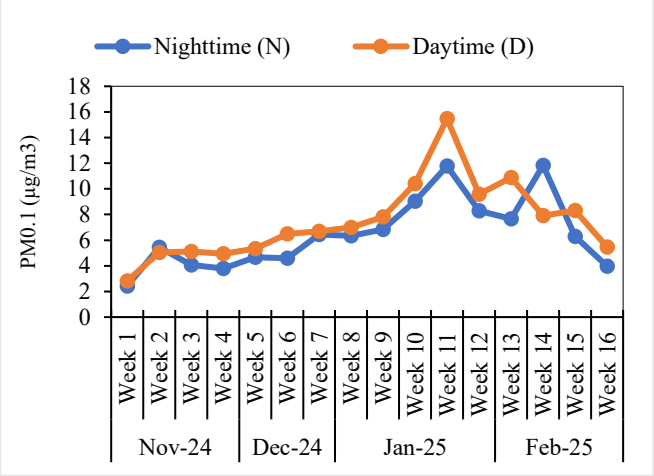


Fig. 2. $PM_{0.1}$ concentration during dry session from November 2024-February 2025

3.2 Correlations of meteorological factors with daytime & nighttime $PM_{0.1}$

The correlations of meteorological factors such as temperature (T), relative humidity (RH), precipitation (Pr) and wind speed (Ws) on mass concentrations of nighttime and daytime $PM_{0.1}$ were examined and presented in **Table 1**. Relative humidity exhibited a statistically significant negative correlations with $PM_{0.1}$ concentration during both daytime and nighttime (r=-0.65, p<0.01; r=-0.60, p<0.05, respectively. High humidity could cause the growth of ambient particles to larger size, enhance particle deposition, thus reducing $PM_{0.1}$ level [19]. Temperature showed a negative, but relatively weak, correlation with the $PM_{0.1}$ concentration during nighttime (r=-0.31, p>0.05). Cool temperatures are likely to accumulate PM to the surface, leading to high concentrations [4]. Conversely, during daytime, temperature showed weak positive correlation (r=0.20, p>0.05) suggesting that higher temperature may enhance photochemical reactions of pollutants during high temperature, leading to the formations of secondary particles, especially for the $PM_{0.1}$ [4]. Precipitations were negatively correlated to $PM_{0.1}$ for both nighttime and daytime, though no significant statistical relations were observed (r=-0.40, p>0.05 for nighttime; r=-0.30, p>0.05 for daytime). Rainfalls may remove airborne particulate matters, thus reducing its mass concentrations (Kurotsuchi et al., 2022). Windspeed showed weak and non-significant positive correlations, suggesting that dispersion effects were minimal on both the nighttime and daytime $PM_{0.1}$. Overall, among the meteorological parameters analyzed, relative humidity exhibited significant effect to the level of ultrafine particles.

Table 1. Correlation coefficients between meteorological factors and mass concentration

	T	RH	Pr	Ws
Nighttime $PM_{0.1}$	-0.31	-0.60*	-0.40	0.20
Daytime $PM_{0.1}$	0.20	-0.65**	-0.30	0.11

*T, RH, Pr, Ws denote the temperature, relative humidity, precipitation, and windspeed during the daytime and nighttime of the sample period; The numbers in the tables are correlation coefficient; and *&** denote the p-value <0.05 and <0.01, respectively.*

3.3 Overall $PM_{0.1}$ mass concentration compared with other studies

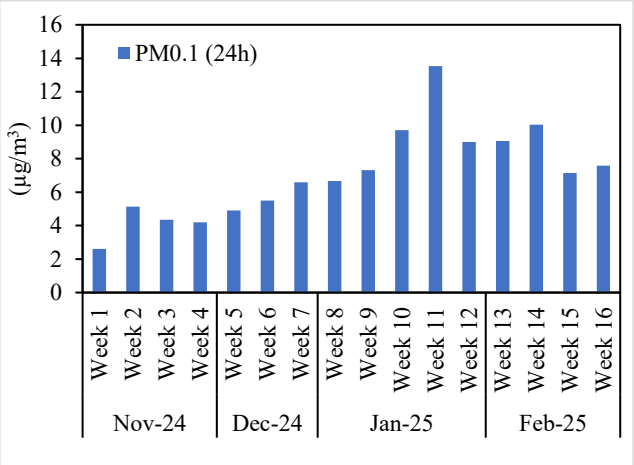


Fig. 3. $PM_{0.1}$ mass concentration of 24h in dry season in Phnom Penh city

Fig. 3 presented the $PM_{0.1}$ mass concentration across 4 months of dry season in Phnom Penh. These $PM_{0.1}$ concentrations were calculated from the daytime and nighttime $PM_{0.1}$ with the correcting factors corresponding to the sampling hours. On average, daily $PM_{0.1}$ was 7.1±0.8 µg/m^3, with the minimum of 2.6 µg/m^3 and the maximum of 13.5 µg/m^3. The current daily $PM_{0.1}$ data, sampling from the use of TX filters, exhibited lower mass concentrations comparing to similar studies [9][20], conducted in the same sampling sites, using QFF for their sampling techniques. Phairuang et al. [20] observed the $PM_{0.1}$ in ITC sampling site from 2011–2021, and reported their mass concentration ranging from 13.4 to 20.6 µg/m^3, which is about 2 to 3 times higher than the values reported by the current study. Amin et al. [9] conducted the sampling of daily $PM_{0.1}$ in 3 sampling sites in Phnom Penh city, including ITC, and found $PM_{0.1}$ mass was 10.5 µg/m^3, about 1 to 2 times differences to the current study. This suggests that besides different sampling periods or effect of meteorological conditions, the use of different sampling materials and techniques also resulted in different PM mass concentrations.

4. CONCLUSIONS

This study monitored ultrafine particulate matter ($PM_{0.1}$) in Phnom Penh during the dry season and found that concentrations generally increased over time, with daytime concentrations were higher than the nighttime. This may be attributed to human activities such as traffic and open burning, occurring within the daytime period, yet further analysis on its chemical components is needed to better clarify the sources and their contributions to the $PM_{0.1}$. Meteorological factors exhibited weak positive and negative correlations for temperature, precipitation and windspeed, yet a significant negative correlation was observed for relative humidity. 24-hour $PM_{0.1}$ in this current data exhibited lower values compared to similar previous studies, mainly due to the differences in sampling techniques i.e. the use of different sampling filters. Our study may be the first in Cambodia, to report mass concentrations of $PM_{0.1}$ from TX filters. Future research should explore seasonal variation, source identification, health impacts, and spatial distribution to better inform air quality management and public health strategies, particularly on the matters of this ultrafine particulate size.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the financial support provided by the JICA–JST SATREPS program (Grant No. JPMJSA2102) for this study. Sincere appreciation is also extended to **Mr. Tomoyuki Yamada**, the project coordinator, for his invaluable support in both administrative and technical aspects of the project.

REFERECNES

[1] WHO, "WHO global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide.," 2021.

[2] S. F. I. Abdillah and Y. F. Wang, "Ambient ultrafine particle (PM0.1): Sources, characteristics, measurements and exposure implications on human health," *Environ. Res.*, vol. 218, p. 115061, Feb. 2023.

[3] B. V. S. Chauhan, K. Corada, C. Young, K. L. Smallbone, and K. P. Wyche, "Review on Sampling Methods and Health Impacts of Fine (PM2.5, ≤2.5 µm) and Ultrafine (UFP, PM0.1, ≤0.1 µm) Particles," *Atmosphere (Basel).*, vol. 15, no. 5, 2024.

[4] L. Morawska, Z. Ristovski, E. R. Jayaratne, D. U. Keogh, and X. Ling, "Ambient nano and ultrafine particles from motor vehicle emissions: Characteristics, ambient processing and implications on human exposure," *Atmos. Environ.*, vol. 42, no. 35, pp. 8113–8138, 2008.

[5] P. Kumar *et al.*, “The rise of low-cost sensing for managing air pollution in cities,” *Environ. Int.*, vol. 75, pp. 199–205, Feb. 2015.
[6] W. Phairuang, M. Amin, M. Hata, and M. Furuuchi, “Airborne Nanoparticles (PM0.1) in Southeast Asian Cities: A Review,” *Sustain.*, vol. 14, no. 16, pp. 1–10, 2022.
[7] M. K. S. Tial *et al.*, “Characterization of PM0.1 mass concentrations and elemental and organic carbon in upper Southeast Asia,” *Atmos. Pollut. Res.*, vol. 15, no. 8, p. 102157, 2024.
[8] V. San, V. Spoann, and J. Schmidt, “Industrial pollution load assessment in Phnom Penh, Cambodia using an industrial pollution projection system.,” *Sci. Total Environ.*, vol. 615, pp. 990–999, Feb. 2018.
[9] M. Amin *et al.*, “Investigation of size-segregated particulate matter and carbonaceous components in Phnom Penh, Cambodia,” *Environ. Sci. Pollut. Res.*, vol. 31, no. 55, pp. 63993–64006, 2024.
[10] P. Ngoun *et al.*, “Monitoring Particulate Matters and Total Suspended Particles Along the Roadside and Public Area in Phnom Penh,” in *IOP Conference Series: Earth and Environmental Science*, 2023, vol. 1199, no. 1.
[11] M. Furuuchi *et al.*, “Development and performance evaluation of air sampler with inertial filter for nanoparticle sampling,” *Aerosol Air Qual. Res.*, vol. 10, no. 2, pp. 185–192, 2010.
[12] V. Chheng, P. Pheng, and V. K. Phun, “Improving Urban Traffic Flow at Congested Signalized Intersections in Phnom,” *Techno-Science Res. J.*, vol. 10, no. 2, pp. 43–49, 2022.
[13] K. H. Kim *et al.*, “Carbonaceous components in PM2.5 and PM0.1 with online measurements of gaseous and particulate pollutants: Implication of thermal-optical derived EC2 fraction as a component of ultrafine particles in the roadside environment,” *Aerosol Air Qual. Res.*, vol. 16, no. 2, pp. 361–372, 2016.
[14] A. Y. P. Wardoyo, I. Majid, A. Budianto, and E. T. P. Adi, “The influence of gasoline fuel additive on the motorcycle PM0.1emission factor,” *J. Phys. Conf. Ser.*, pp. 0–7, 2021.
[15] Phnom Penh Post, “Nearly 3K illegal burn sites identified across Kingdom | Phnom Penh Post,” *Phnom Penh Post*, 2021. [Online]. Available: https://www.phnompenhpost.com/national/nearly-3k-illegal-burn-sites-identified-across-kingdom. [Accessed: 25-May-2025].
[16] M. Hata *et al.*, “Characteristics of nanoparticles emitted from burning of biomass fuels,” *J. Environ. Sci. (China)*, vol. 26, no. 9, pp. 1913–1920, 2014.
[17] W. Phairuang *et al.*, “Daytime-nighttime variations in the concentration of PM0.1 carbonaceous particles during a biomass fire episode in Chiang Mai, Thailand,” *Particuology*, vol. 87, pp. 316–324, 2024.
[18] V. T. Le Ha *et al.*, “Occurrence of PM2.5 and PM0.1 at high polluting event days in Ha Noi and health implication,” *Vietnam J. Sci. Technol.*, vol. 62, no. 2, pp. 335–345, 2024.
[19] N. T. T. Thuy, N. T. Dung, K. Sekiguchi, L. B. Thuy, N. T. T. Hien, and R. Yamaguchi, “Mass concentrations and carbonaceous compositions of pm0.1, pm2.5, and pm10 at urban locations in hanoi, vietnam,” *Aerosol Air Qual. Res.*, vol. 18, no. 7, pp. 1591–1605, 2018.
[20] W. Phairuang *et al.*, “Ten Years Behavior of Carbonaceous Ultrafine Particulate Matter (PM0.1) in Phnom Penh, Cambodia,” in *IOP Conference Series: Earth and Environmental Science*, 2023, vol. 1199, no. 1, pp. 0–7.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Variability of $PM_{2.5}$ Distribution in Relation to Gaseous Pollutants Concentrations in Phnom Penh, Cambodia

Titsothearna You [1*], Sochan Sao [2*], Seyha Doeurn [1,2], Chanthol Peng [1,2], Saksoleakhena Sambath [2], Fumikazu Ikemori [4], Mitsuhiko Hata [3], Masami Furuuchi [3], Chanmoly Or [2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Faculty of Geosciences and Civil Engineering, Institute of Science and Engineering, Kanazawa University, Kakuma-machi, Kanazawa, Ishikawa 920-1192, Japan*

[4] *Graduate School of Integrated Science and Technology, Nagasaki University, 1-14, Bunkyo-machi, Nagasaki, 852-8521, Japan*

[*] *Corresponding author: yousotheana@gmail.com; saosochan2015@gmail.com*

Abstract: *Phnom Penh has recently experienced worsened air pollution, particularly fine particulate matter ($PM_{2.5}$), posing significant environmental and health risks. This study aimed to investigate the variability of $PM_{2.5}$ distribution within Phnom Penh, Cambodia, and to explore its relationship with key gaseous pollutants, specifically nitrogen dioxide (NO_2), sulfur dioxide (SO_2), carbon monoxide (CO), and ozone (O_3) using data from real-time air quality monitoring system (AQMS). AQMS showed that $PM_{2.5}$ concentrations were markedly higher between 21-24 January 2025 compared to other days. During this period, the daily average $PM_{2.5}$ levels surged from 36.085 µg/m³ on January 21 to 45.452 µg/m³ on January 22, followed by a sharp increase to 55.303 µg/m³ on January 23, before peaking at 58.824 µg/m³ on January 24. After this peak, concentrations gradually declined to 35.50 µg/m³ on January 25. Correlation analysis derived from 24-hourly measurements over the five-day period demonstrated a weak positive relationship between $PM_{2.5}$ and SO_2 with a correlation coefficient (r) of 0.355 ($p < 0.001$). On the other hand, $PM_{2.5}$ showed a moderate positive correlation with NO_x ($r = 0.432$ and $p < 0.001$) and a stronger correlation with CO ($r = 0.738$ and $p < 0.001$), as both exhibited hourly peaks typically between 7-8am and 5-7pm. Conversely, a negative correlation was observed between PM2.5 and O_3 ($r = -0.397$ and $p < 0.001$) as its concentrations tend to peak at midday when $PM_{2.5}$ levels are relatively low. These results suggest that combustion-related emissions, primarily from vehicular traffic during typical rush-hours, are the source of $PM_{2.5}$ pollution in Phnom Penh, while O_3 variability is primarily influenced by atmospheric chemistry. Beyond that, meteorological conditions also involved in $PM_{2.5}$ distribution. However, this study is limited by its short observation period and reliance on data from a single monitoring station. Further investigation is needed to validate these findings and verify this critical phenomenon.*

Keywords: Air pollution; $PM_{2.5}$; nitrogen dioxide; sulfur dioxide; carbon monoxide; Ozone

1. INTRODUCTION

Air pollution refers to any chemical, physical, or biological substances that contaminate indoor and outdoor air, degrading air quality and negatively impacting the environment [1]. It is a major contributor to respiratory diseases and other health complications, significantly increasing global mortality rates [2]. According to estimates from the WHO, approximately 7 million deaths annually, primarily from noncommunicable diseases, are attributed to the combined effects of ambient and household air pollution [3]. Furthermore, a global assessment of ambient air pollution alone shows an annual death rate ranging from 4 to 9 million, along with hundreds of millions of lost healthy life years. This burden is particularly severe in low- and middle-income countries [1], where exposure to pollutants is often higher and healthcare access is more limited.

Cambodia has been experiencing rapid economic development since the early 2000s. It is reported that urban expansion between 1973 and 2015 increased eightfold in Phnom Penh [4]. This urban growth has been accompanied by a noticeable decline in air quality. Urban modernization is

[1*] Corresponding author: Titsothearna YOU
E-mail: yousotheana@gmail.com; Tel: +855-69 455 352

[2*] Corresponding author: Sochan SAO
E-mail: saosochan2015@gmail.com; Tel: +855-97 844 5346

strongly linked to increased levels of atmospheric pollution, particularly fine particulate matter (i.e., $PM_{2.5}$) [5]. This trend is largely driven by population growth in the city, intensifying transportation, industrial activities, and energy consumption, all of which accelerate the emission of harmful air pollutants and contribute to environmental degradation [6].

Among various pollutants, $PM_{2.5}$ is of particular concern due to its small size and complex composition. It consists of a mixture of solid and liquid particles suspended in the atmosphere, originating from various sources such as diesel emissions, road dust, agricultural activities, and industrial processes [1]. $PM_{2.5}$ exhibits a wide range of characteristics in terms of shape, chemical composition, physical properties, and thermodynamic behaviors [7]. Its ability to penetrate deep into the human respiratory tract and its long atmospheric lifetime make it one of the most hazardous air pollutants. Its concentration in the atmosphere is influenced by both primary and secondary particulate components.

Primary $PM_{2.5}$ concentrations are directly emitted from sources like vehicle exhaust, while secondary particles are formed in the atmosphere through chemical reactions involving precursor gases such as sulfur dioxide (SO_2), nitrogen oxides (NOx), volatile organic compounds (VOCs), and ammonia (NH_3) [8]. Typically, the secondary fraction makes up the largest portion of $PM_{2.5}$ [9]. For instance, secondary inorganic aerosols, are produced via reactions involving SO_2, NO_x, and NH_3. A significant reduction in SO_2 emissions can lead to decreased sulfate production, allowing excess NH_3 to react and form ammonium nitrate [10].

Ozone (O_3) plays a crucial role in accelerating the oxidation of precursor gases such as SO_2, NO_x, and volatile organic compounds (VOCs), promoting the formation of both secondary organic aerosols (SOAs) and secondary inorganic aerosols (SIAs), such as sulfates and nitrates, which contribute to increased $PM_{2.5}$ concentrations [11]. Conversely, elevated $PM_{2.5}$ concentrations may influence atmospheric chemistry by reducing ground-level ultraviolet radiation through light extinction, thereby slowing photolysis rates that affect the formation and degradation of O_3 [12]. Additionally, particulate matter such as black carbon, dust, and sea salt can act as surfaces for heterogeneous chemical reactions, altering key gaseous (i.e., O_3, N_2O_5, NO_2, and SO_2) [13]. These interactions between $PM_{2.5}$ and gaseous pollutants exhibit positive, negative, and neutral correlations depending on local atmospheric conditions [14].

Given these complex dynamics, this study aims to assess the variability of $PM_{2.5}$ concentrations and their relationship with key gaseous pollutants, including NO_2, SO_2, CO, and O_3, in Phnom Penh, Cambodia. The analysis utilized data collected from a real-time Air Quality Monitoring System installed in the city. By doing so, this research seeks to provide a better understanding of the influence of combustion-related emissions and atmospheric chemical processes on $PM_{2.5}$ variability, ultimately contributing to evidence-based air quality management and mitigation strategies for Phnom Penh.

2. METHODOLOGY

2.1. Study area

Phnom Penh, the capital and largest city of Cambodia, has a tropical monsoon climate characterized by distinct dry and wet seasons [15]. Air quality monitoring was conducted at the Institute of Technology of Cambodia (ITC), located in Russey Keo district in the northern Phnom Penh. The monitoring station is situated at latitude 11.5682°N and longitude 104.8986°E, in a densely populated urban area surrounded by residential neighborhoods, commercial establishments, and construction sites. The surrounding environment is heavily influenced by vehicular emissions due to its proximity to major roads such as the Russian Federation Boulevard and National Road No.5.

2.2 Data collection

In early 2025, Cambodia experienced a sharp peak of $PM_{2.5}$ concentrations, raising concerns about air quality and public health. To investigate this event, data were collected using the real-time Air Quality Monitoring System (AQMS-1560 by HORIBA, Japan) installed at ITC in Phnom Penh. The AQMS provides continuous measurements of $PM_{2.5}$ and key gaseous pollutants, including NO_2, SO_2, CO, and O_3. Hourly average data from January 21 to 25, 2025, were used for analysis. This period was selected based on observed pollution peaks to explore the relationship between $PM_{2.5}$ and key gaseous pollutant concentrations, contributing to a better understanding of air pollution dynamics in Phnom Penh.

2.3 Statistical analysis

In this study, correlation analysis was applied to examine the relationship between $PM_{2.5}$ concentrations and key gaseous pollutants. Pearson's correlation coefficient (r) was used to measure the strength and direction of the linear relationship between $PM_{2.5}$ and each pollutant. The value of *r* ranges from -1 to +1, where values close to +1 indicate a strong positive correlation, values close to -1 indicate a strong negative correlation, and values around 0 indicate no correlation (**Table 1**). A significance level of *p* was used to determine the statistical significance of the results. Additionally, correlation graphs were plotted to visually observe the patterns and variations of $PM_{2.5}$ and key gaseous pollutants throughout the monitoring period.

3. RESULTS AND DISCUSSION

3.1 Variation in $PM_{2.5}$ and key gaseous pollutants

The daily average concentration of $PM_{2.5}$ increased from 36.085 $\mu g/m^3$ on January 21 to the highest peak of 58.824

µg/m^3 on January 24 before declining to 35.50 µg/m^3 on January 25. A similar trend was observed for CO, which rose from 0.757 ppm to 1.405 ppm before decreasing to 0.980 ppm on the same days.

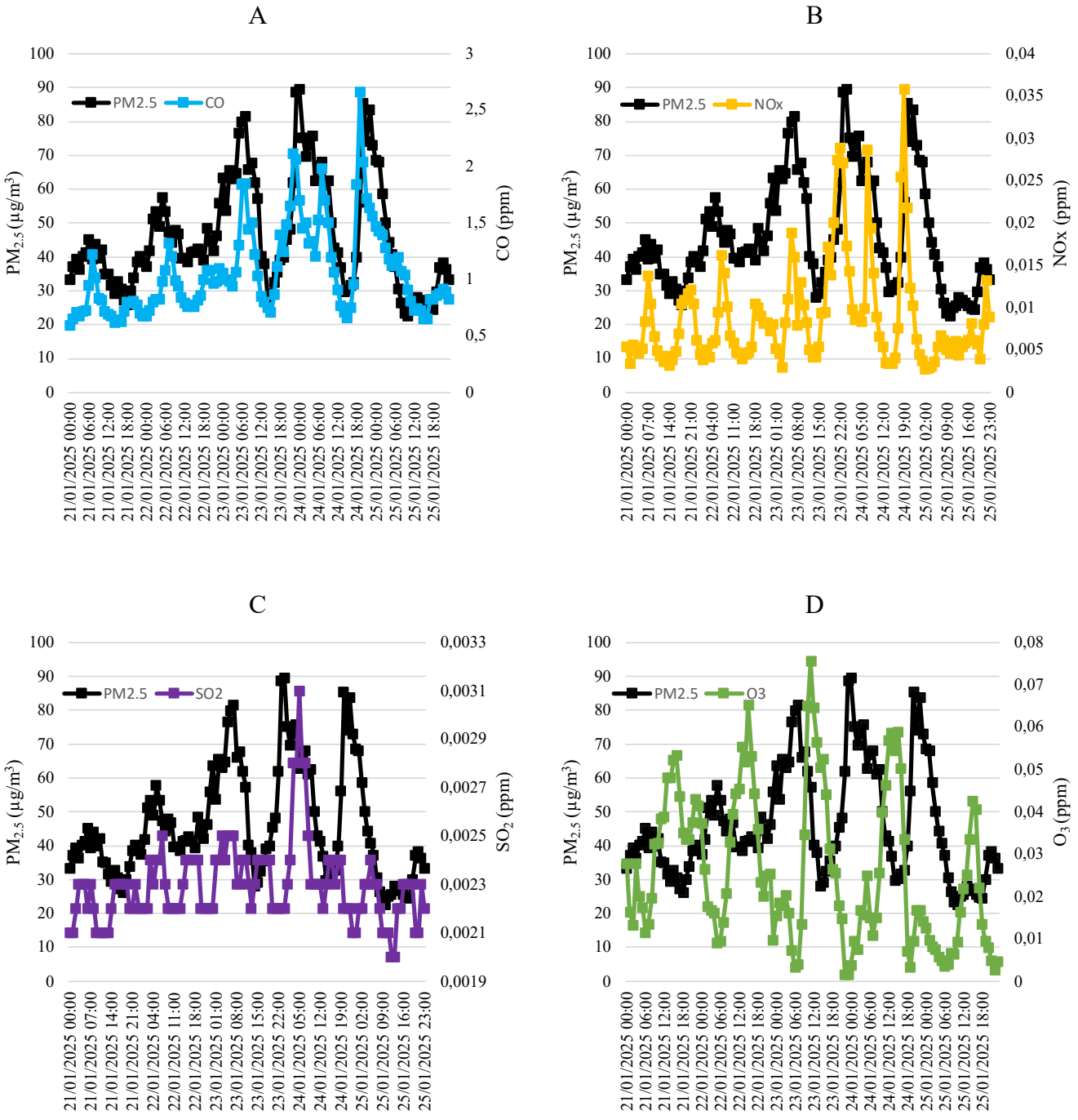


Fig.1. Correlation between $PM_{2.5}$ and key gaseous pollutants

NO_x concentrations also increased steadily from 0.0069 ppm on January 21 to 0.0125 ppm on January 24, before declining to 0.00566 ppm on January 25. The increase in $PM_{2.5}$ concentration was consistent with the fluctuation in CO and NO_x concentrations, aligning with a previous study by Wang et al., 2017. The average concentration of SO_2 slightly increased from 0.00221 ppm to 0.00241 ppm before declining to 0.00220 ppm. A previous study concluded that elevated

SO_2 concentration is likely to originate from coal burning and industrial activities. The relatively low SO_2 concentration observed in the current study compared to those typically reported in more industrialized Southeast Asian cities, where daily averages often exceed 0.005 ppm [17]. This is likely due to the ITC monitoring station's distance from heavy industrial activity. In contrast, the average concentration of O_3 displayed an opposite trend compared to $PM_{2.5}$ (**Fig. 1**). O_3 levels were highest on January 21 (0.032 ppm) and gradually decreased to the lowest level on January 25 (0.013 ppm). An increase in PM2.5 can significantly influence photochemical reactions [16].

3.2 Correlation between $PM_{2.5}$ and key gaseous pollutants

The correlation coefficient between the hourly average concentrations of $PM_{2.5}$ and gaseous pollutants is shown in Table 1. CO (r = 0.738) > NO_x (r = 0.432) > SO_2 (r = 0.355) > O_3 (r = –0.397), with all correlations statistically significant ($p < 0.001$). These results indicate a strong positive correlation between CO and $PM_{2.5}$ followed by a moderate positive relationship with NO_x, and a weak positive relationship with SO_2. In contrast**,** O_3 showed a statistically significant moderate negative correlation with $PM_{2.5}$.

Table. 1. Correlation coefficient between $PM_{2.5}$ and key gaseous pollutants (n =120).

Pollutants	r	p-value
CO	0.738	<0.001
NO_x	0.432	<0.001
SO_2	0.355	<0.001
O_3	-0.397	<0.001

The strong positive correlation between $PM_{2.5}$ and CO suggests a shared origin, predominantly from traffic emissions and other combustion-related activities [18]. Their temporal patterns showed aligned peaks during early morning and evening hours, such as at 08:00 on January 23 and 20:00 on January 24, pointing to shared combustion sources from traffic activities. These matching peaks indicate direct emissions as a common contributor. However, clear differences in their variation patterns are observed. CO peaks are sharp and decline quickly, consistent with short-lived combustion emissions.

Similarly, a moderate correlation between NO_x and $PM_{2.5}$ was shown by the distinct peaks during the typical morning (7-8AM) and evening hours (6-8PM). This pattern aligns with known traffic congestion times in Phnom Penh [19], suggesting vehicle emissions are a likely source of NO_x and a contributor to $PM_{2.5}$. NO_x also acts as a precursor for secondary particulate matter formation, particularly nitrate (NO_3^-) aerosols, through atmospheric reactions under favorable conditions such as the presence of ammonia (NH_3) [20]. Correlation analysis supports that CO (r = 0.738) and NO_x (r = 0.432) are positively correlated with $PM_{2.5}$, affirming a shared source in combustion processes. However, while CO and NOx display sharp, narrow peaks that drop quickly, $PM_{2.5}$ shows broader peaks with slower rises and falls. This difference suggests that, beyond direct combustion, $PM_{2.5}$ levels are influenced by secondary processes such as chemical reactions and atmospheric conditions [21] that sustain concentrations even after primary emissions decline. If $PM_{2.5}$ came only from combustion, its pattern would mirror the sharp peaks of CO and NOx more closely.

The weak correlation between SO_2 and $PM_{2.5}$ (r = -0.397) suggests that SO_2 concentrations do not strongly influence $PM_{2.5}$ levels in the study area. Although SO_2 is a recognized precursor to sulfate aerosols (SO_4^{2-}), which contribute to $PM_{2.5}$, its transformation into particulate matter requires complex atmospheric oxidation processes involving oxidants such as ozone (O_3) and hydroxyl radicals (OH) under specific meteorological conditions [21]. These secondary formation processes can delay or spatially displace the impact of SO_2 on $PM_{2.5}$ concentrations [22]. Furthermore, the prevalence of gasoline-powered vehicles, which emit relatively little SO_2 compared to diesel engines and industrial activities, combined with the study area's distance from major industrial sources, likely explains the limited influence of SO_2 on observed $PM_{2.5}$ variability [23]. Hence, while SO_2 contributes to secondary PM formation under favorable conditions, its immediate impact in this urban context appears minimal.

Conversely, a moderate negative correlation was found between $PM_{2.5}$ and O_3. Higher $PM_{2.5}$ concentrations typically occurred during early morning and evening hours when sunlight was limited and combustion activity peaked, while O_3 levels increased around midday (12:00–14:00) when solar radiation is strongest, promoting photochemical reactions. $PM_{2.5}$ levels declined due to atmospheric dispersion and reduced emission activity. This inverse relationship is consistent with previous studies that elevated $PM_{2.5}$ concentrations can attenuate solar radiation, suppressing the photolysis processes required for O_3 formation [24].

4. CONCLUSIONS

Based on the short-term monitoring results, the findings highlight that $PM_{2.5}$ concentrations significantly increased during the study period. Correlation analysis revealed that $PM_{2.5}$ had a weak to moderate positive correlation with SO_2, a moderate positive relation with NO_x, and a stronger positive correlation with CO. In contrast, a negative correlation was observed between $PM_{2.5}$ and O, reflecting the photochemical behavior of ozone, which tends to increase under strong midday solar radiation as $PM_{2.5}$ levels decline. However, $PM_{2.5}$ concentrations did not always follow the same sharp diurnal trends observed in gaseous pollutants. This suggests that, in addition to the main traffic-related emissions, secondary aerosol formation, atmospheric dispersion, and

meteorological conditions also contribute significantly to $PM_{2.5}$ variability. These findings provide valuable insights into the complex interactions between $PM_{2.5}$ and key gaseous pollutants in Phnom Penh's urban atmosphere. To strengthen these observations, future studies should consider long-term monitoring and multiple observation sites to better capture the spatiotemporal variations and chemical characteristics of $PM_{2.5}$.

ACKNOWLEDGMENTS

The author gratefully acknowledges the financial support provided by the **JICA–JST SATREPS** program (Grant No. JPMJSA2102) for this study. Sincere appreciation is also extended to Mr. **Tomoyuki Yamada**, the project coordinator, for his invaluable support in both administrative and technical aspects of the project.

REFERENCES

[1] WHO, "WHO global air quality guidelines," *Part. matter (PM2.5 PM10), ozone, nitrogen dioxide, sulfur dioxide carbon monoxide.*, pp. 1–360, 2021.

[2] A. K. Rana, E. Mostafavi, W. F. Alsanie, S. S. Siwal, and V. K. Thakur, "Cellulose-based materials for air purification: A review," *Ind. Crops Prod.*, vol. 194, no. June 2022, p. 116331, 2023, doi: 10.1016/j.indcrop.2023.116331.

[3] P. J. Landrigan *et al.*, "The Lancet Commission on pollution and health," *Lancet*, vol. 391, no. 10119, pp. 462–512, 2018, doi: 10.1016/S0140-6736(17)32345-0.

[4] F. Mialhe *et al.*, "Spatial growth of Phnom Penh, Cambodia (1973–2015): Patterns, rates, and socio-ecological consequences," *Land use policy*, vol. 87, no. January 2018, p. 104061, 2019, doi: 10.1016/j.landusepol.2019.104061.

[5] World Bank, "The cost of air pollution abatement," *Appl. Econ.*, vol. 29, no. 6, pp. 759–774, 2016, doi: 10.1080/000368497326688.

[6] P. Behinaein, H. Hutchings, T. Knapp, and I. C. Okereke, "The growing impact of air quality on lung-related illness: a narrative review," *J. Thorac. Dis.*, vol. 15, no. 9, pp. 5055–5063, 2023, doi: 10.21037/jtd-23-544.

[7] P. Thangavel, D. Park, and Y.-C. Lee, "Toxicity of airborne particles—established evidence, knowledge gaps and emerging areas of importance," *Int. J. Environ. Res. Public Heal.*, vol. 19, no. 12, p. 7511, 2022, [Online]. Available: https://www.mdpi.com/1660-4601/19/12/7511

[8] W. M. Hodan and W. R. Barnard, "Evaluating the Contribution of PM2. 5 Precursor Gases and Re-entrained Road Emissions to Mobile Source PM2. 5 Particulate Matter Emissions. MACTEC Under Contract to the Federal Highway Administration," *MACTEC Fed. Programs, Res. Triangle Park. NC*, p. 58 pp, 2004, [Online]. Available: https://www3.epa.gov/ttnchie1/conference/ei13/mobile/hodan.pdf

[9] S. Aksoyoglu, G. Ciarelli, I. El-Haddad, U. Baltensperger, and A. S. H. Prévôt, "Secondary inorganic aerosols in Europe: Sources and the significant influence of biogenic VOC emissions, especially on ammonium nitrate," *Atmos. Chem. Phys.*, vol. 17, no. 12, pp. 7757–7773, 2017, doi: 10.5194/acp-17-7757-2017.

[10] A. Clappier, P. Thunis, M. Beekmann, J. P. Putaud, and A. de Meij, "Impact of SOx, NOx and NH3 emission reductions on PM2.5 concentrations across Europe: Hints for future measure development," *Environ. Int.*, vol. 156, no. x, p. 106699, 2021, doi: 10.1016/j.envint.2021.106699.

[11] K. L. Chan, S. Wang, C. Liu, B. Zhou, M. O. Wenig, and A. Saiz-Lopez, "On the summertime air quality and related photochemical processes in the megacity Shanghai, China," *Sci. Total Environ.*, vol. 580, pp. 974–983, 2017, doi: 10.1016/j.scitotenv.2016.12.052.

[12] Y. Li *et al.*, "Spatiotemporal Variations of PM2.5 and O3 Relationship during 2014–2021 in Eastern China," *Aerosol Air Qual. Res.*, vol. 23, no. 10, pp. 1–16, 2023, doi: 10.4209/aaqr.230060.

[13] H. Lu, X. Lyu, H. Cheng, Z. Ling, and H. Guo, "Overview on the spatial-temporal characteristics of the ozone formation regime in China," *Environ. Sci. Process. Impacts*, vol. 21, no. 6, pp. 916–929, 2019, doi: 10.1039/c9em00098d.

[14] Xin Zhang, "Spatiotemporal distribution of pm2.5 and its correlation with other air pollutants in winter during 2016~2018 in Xi'an, China," *Polish J. Environ. Stud.*, vol. 30, no. 2, pp. 1457–1464, 2021, doi: 10.15244/pjoes/124755.

[15] World Bank, *Climate Risk Country Profile*. 2021. doi: 10.1596/36382.

[16] M. Wang, J. Cao, C. Gui, Z. Xu, and D. Song, "The characteristics of spatiotemporal distribution of PM2.5 in Henan province, China," *Polish J. Environ. Stud.*, vol. 26, no. 6, pp. 2785–2792, 2017, doi: 10.15244/pjoes/70631.

[17] N. D. Luong *et al.*, "Investigation of sources and processes influencing variation of PM2.5 and its chemical compositions during a summer period of 2020 in an urban area of Hanoi city, Vietnam," *Air Qual. Atmos. Heal.*, vol. 15, no. 2, pp. 235–253, 2022, doi: 10.1007/s11869-021-01100-z.

[18] L. Han, G. Zhuang, S. Cheng, Y. Wang, and J. Li, "Characteristics of re-suspended road dust and its impact on the atmospheric environment in Beijing,"

Atmos. Environ., vol. 41, no. 35, pp. 7485–7499, 2007, doi: 10.1016/j.atmosenv.2007.05.044.

[19] JICA, "The Kingdom of Cambodia Phnom Penh Capital City Ministry of Public Works and Transport the Project for Comprehensive Urban Transport Plan in Phnom Penh Capital City (Pputmp)," pp. 1–5, 2014.

[20] S. N. P. Jonh H.Seinfeld, "Atmosphere Chemistry and Physics: From Air pollution to Climate change," 2006. [Online]. Available: https://books.google.com.kh/books?id=YH2K9eWsZOcC&printsec=frontcover&redir_esc=y#v=onepage&q&f=false

[21] Asiva Noor Rachmayani, *Atmospheric Chemistry and Physics: From Air Pollution to Climate Change (3rd ed.)*. 2015.

[22] C. Cai *et al.*, "Incorporation of New Particle Formation and Early Growth Treatments into WRF/Chem: Model Improvement, Evaluation, and Impacts of Anthropogenic Aerosols over East Asia," *Atmos. Environ.*, vol. 124, May 2015, doi: 10.1016/j.atmosenv.2015.05.046.

[23] X. Yao, A. P. S. Lau, M. Fang, C. K. Chan, and M. Hu, "Size distributions and formation of ionic species in atmospheric particulate pollutants in Beijing, China: 1 - Inorganic ions," *Atmos. Environ.*, vol. 37, no. 21, pp. 2991–3000, 2003, doi: 10.1016/S1352-2310(03)00255-3.

[24] H. Yang, L. Chen, H. Liao, J. Zhu, W. Wang, and X. Li, "Impacts of aerosol-photolysis interaction and aerosol-radiation feedback on surface-layer ozone in North China during multi-pollutant air pollution episodes," *Atmos. Chem. Phys.*, vol. 22, no. 6, pp. 4101–4116, 2022, doi: 10.5194/acp-22-4101-2022.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Characterization of Bacterial Communities in Aquatic Environments of Phnom Penh as Potential Bioaerosol Sources Using Nanopore Sequencing

Saksoleakhena SAMBATH [1*], Chanthol PENG [1], Chan Leakhena PHOEUNG [2], Chanvorleak PHAT [1], Chhaipo CHHEANG [2], Tomoyuki YAMADA [3], Kimiko UCHII [4], Masao NASU [4]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *University of Health Sciences, Preah Monivong Blvd., Duan Penh District, Phnom Penh, Cambodia*
[3] *JICA Air SATREPS Project, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[4] *Osaka Ohtani University, 3-chōme-11 Nishikiorikita, Tondabayashi, Osaka, Japan*

[*] *Corresponding author: saksoleakhena@gmail.com*

Abstract: *Bioaerosols, comprising airborne microorganisms and their byproducts, pose growing concerns for urban air quality and public health. Aquatic environments, especially wastewater systems, can serve as reservoirs for diverse bacterial communities, some of which may become aerosolized and contribute to airborne microbial loads. However, the aquatic sources of bioaerosols in Phnom Penh remain largely uncharacterized. This study presents preliminary results on the composition of bacterial communities in Phnom Penh's aquatic environments, Cambodia, and their potential role as sources of bioaerosols. Water samples were collected from four sites, including two freshwater sites: Tonle Sap River (TSR) and the Royal University of Phnom Penh Pond (RUPP), and two wastewater sites: Tamok Lake (TML) and Tror Bek Pumping Station (TBS), during two sampling sessions on 28 August and 19 September 2023. Bacterial DNA was extracted and analyzed using full-length 16S rRNA gene sequencing with the Oxford Nanopore MinION platform, and cloud-based taxonomic classification was performed using the EPI2ME FASTQ 16S workflow. Phylum-level analysis revealed that Pseudomonadota dominated all sites, particularly in wastewater samples, with relative abundances of 35.7% at TBS and 21.8% at TML. Bacillota and Bacteroidota were also prevalent in wastewater samples. Campylobacterota, an indicator of fecal contamination associated with animal waste, was notably abundant at TML (31.7%) and TBS (17.9%), highlighting potential health risks if wastewater is aerosolized. In contrast, freshwater sites exhibited greater microbial diversity and distinct dominant phyla. Cyanobacteria, often linked to algal blooms, were more abundant in the stagnant pond RUPP (21.5%) compared to the flowing river TSR (4.8%). Planctomycetota, a phylum that includes nitrogen-cycling bacteria, was found at 5.4% in RUPP and 3.7% in TSR. At the genus level, wastewater samples were dominated by Aliarcobacter, which includes pathogenic species for gastrointestinal infections and respiratory diseases. RUPP pond was abundant with Planktothrix, a potentially harmful bloom-forming cyanobacterium, Legionella and Clostridium, waterborne pathogenic genera capable of airborne transmission. These preliminary findings reflect clear microbial distinctions between freshwater and wastewater environments and highlight wastewater and stagnant freshwater systems as potential reservoirs for bioaerosol-relevant taxa. The identification of genera with known aerosolization potential reinforces the need to investigate microbial transitions from water to air in flood-prone urban environments. Future research should expand sampling across locations and seasons, conduct species-level characterization, and investigate the microbial transitions from aquatic environments to the air to better assess bioaerosol dynamics and potential public health implications.*

Keywords: Bacterial communities; Aquatic environments; Bioaerosols; Nanopore sequencing

1. INTRODUCTION

Phnom Penh, the capital city of Cambodia, is experiencing rapid urbanization that has drastically transformed its natural landscapes and disrupted long-standing hydrological systems. Historically, the city was surrounded by extensive wetlands and aquatic ecosystems that played essential roles in flood mitigation, wastewater absorption, and ecological regulation. However, over 6,000 hectares of wetlands have been replaced by urban infrastructure in recent decades, significantly reducing the city's natural capacity to manage stormwater and contributing

[1*] Corresponding author: Sambath Saksoleakhena
E-mail: saksoleakhena@gmail.com; Tel: +855-16 967 172

to increased surface runoff and seasonal flooding [1]. This accelerated urban transformation poses serious environmental and public health risks, particularly urban flooding. The intensification of flood events not only causes infrastructural damage but also raises health concerns due to the spread of waterborne contaminants. One emerging concern is the production and dissemination of bioaerosols [2, 3].

Bioaerosols are airborne microscopic particles of biological origin, including bacteria, viruses, fungi, and their fragments [4]. In flood-prone cities like Phnom Penh, the overflow of contaminated surface waters, especially those connected to wastewater systems, may lead to the aerosolization of harmful microorganisms through wind, rainfall, or human disturbance [5, 6]. These bioaerosols can remain airborne, travel significant distances, and increase exposure risks in densely populated areas. Exposure has been associated with respiratory infections, gastrointestinal diseases, and allergic reactions, particularly affecting vulnerable groups such as children, the elderly, and immunocompromised individuals [7]. Their transport and infectivity are further influenced by environmental variables such as humidity, temperature, and solar radiation [8, 9].

Globally, research on bioaerosols has predominantly focused on high-income countries, often in the context of healthcare, agriculture, industry, and occupational settings. A bibliometric analysis of global bioaerosol research from 1989 to 2019 showed that the majority of studies originate from high-income countries [10]. By contrast, little is known about bioaerosol generation in rapidly urbanizing, flood-prone cities in Southeast Asia, including Phnom Penh. The combination of inadequate sanitation, informal settlements, and the direct discharge of untreated wastewater creates complex airborne microbial exposure pathways that remain poorly understood. Studies from neighboring countries such as China and India have detected fecal indicator bacteria, respiratory pathogens, and antibiotic resistance genes in urban air near wastewater channels and open drains [11]. These findings highlight the need for similar investigations in Cambodia to address growing environmental health concerns.

Recent advances in molecular technology, particularly Oxford Nanopore sequencing, offer a transformative approach for studying environmental microbial communities. This real-time, portable platform enables full-length 16S rRNA gene sequencing, providing high-resolution taxonomic profiling. Unlike conventional culture-based or short-read techniques, Nanopore sequencing can detect both dominant and rare taxa, including unculturable organisms [12]. This innovation remains underutilized in environmental research in Cambodia.

This study utilizes Oxford Nanopore sequencing to characterize bacterial communities in aquatic environments of Phnom Penh and assess their potential role in contributing to urban bioaerosols. The findings offer initial insights into microbial diversity within these aquatic environments and their relevance to airborne bioaerosols.

2. METHODOLOGY

2.1 Water Samples Collection

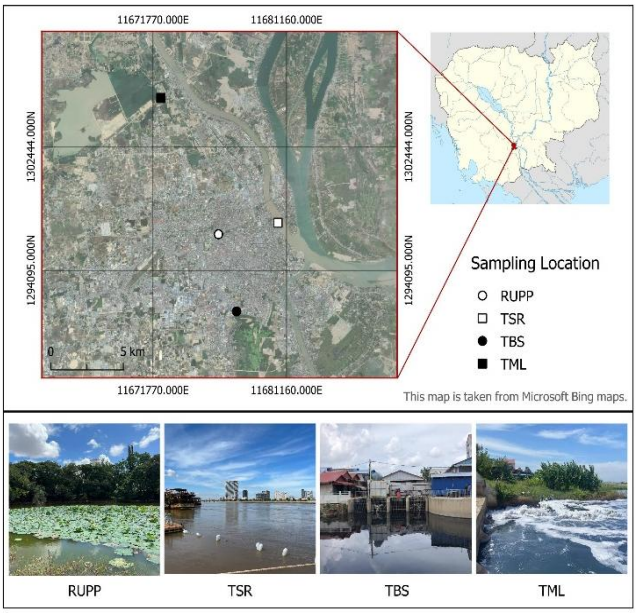


Fig. 1. Sampling locations.

Water samples were collected from four distinct aquatic environments in Phnom Penh, each representing different hydrological and contamination profiles. The sampling sites included two freshwater sites: a stagnant pond within the Royal University of Phnom Penh (RUPP) and a surface water sample from the Tonle Sap River (TSR), and two wastewater sites from urban drainage system: Tamok Lake (TML) and Tror Bek Pumping Station (TBS) as shown in **Fig. 1**. Sampling was conducted during the rainy season, from August to September, to capture microbial community profiles under flood-prone, high-moisture conditions. Two sampling sessions were selected to provide an initial characterization of microbial communities across diverse sites, as shown in **Table 1**.

Table 1. Sampling date and weather conditions.

Site	Sampling Date	Sample ID	Weather Conditions
RUPP	28 Aug 2023	RUPP-0828	Sunny
	19 Sep 2023	RUPP-0919	Sunny
TSR	28 Aug 2023	TSR-0828	Sunny
	19 Sep 2023	TSR-0919	Sunny
TBS	28 Aug 2023	TBS-0828	Sunny
	19 Sep 2023	TBS-0919	Sunny
TML	28 Aug 2023	TML-0828	Sunny
	19 Sep 2023	TML-0919	Sunny

At each site, 50-100 mL of surface water was collected using a gamma-ray sterilized polyethylene bag (Monotaro, Japan) and immediately filtered through a 0.22 μm Sterivex filter (Millipore, Germany) using a sterile syringe. Each filter was sealed with sterilized Luer fittings (VRSP6 and VRMP6), labeled, and stored in a cooler with ice packs. To monitor for potential contamination, a field blank was prepared during each sampling session by filtering 100 mL of sterilized ultrapure water under the same conditions. All samples were transported to the laboratory within 3 hours of collection and stored at −20°C until DNA extraction.

2.2 DNA Extraction

Genomic DNA was extracted from the filtered membranes using the QIAGEN AllPrep PowerViral DNA/RNA Kit (Qiagen, Germany), following the manufacturer's protocol optimized for environmental samples. The integrity and concentration of extracted DNA were assessed using a NanoDrop-1000 spectrophotometer (Thermo Fisher Scientific). All extracted DNA was stored at -20°C until further downstream analysis.

2.3 PCR Amplification and 16S rRNA Gene Sequencing

Full-length 16S rRNA gene amplification was carried out by double PCR protocol using the Oxford Nanopore 16S Barcoding Kit 24 V14 (SQK-16S114.24), with KAPA HiFi HotStart ReadyMix to ensure high-fidelity amplification. Barcoded primers targeting the V1–V9 regions were used under optimized thermocycling conditions. The resulting amplicons were quantified with a Qubit Fluorometer (Thermo Fisher Scientific). Sequencing was performed on MinION device using Flongle flow cells (R.10.4.1) using fast basecalling in MinKNOW software.

2.2 Bioinformatic Analysis

Raw sequencing reads were processed using the EPI2ME Fastq 16S workflow for basecalling and preliminary taxonomic assignment. Quality filtering was applied to retain only reads between 1,000 and 2,000 base pairs, corresponding to full-length 16S rRNA gene sequences. The resulting taxonomic classification data were exported as CSV files and imported into RStudio (version 4.4.3) for further analysis. Read counts were aggregated across taxonomic levels ranging from superkingdom to species. Taxa with low abundance (≤4 reads per sample) were removed to minimize background noise. Relative abundance values were then calculated. Microbial communities were characterized at the phylum and genus levels using heatmaps to compare patterns across different sampling sites.

3. RESULTS AND DISCUSSION

3.1 Bacterial Communities at the Phylum Level

Phylum-level taxonomic analysis revealed distinct bacterial community profiles across the four aquatic sites, reflecting differences in water type, pollution load, and environmental conditions, as illustrated in **Fig. 2**. *Pseudomonadota* was the most dominant phylum overall, with especially high proportions in wastewater samples (35.7% at TBS; 21.8% at TML), but also found in freshwater sites (32.9% at TSR; 21.0% at RUPP), suggesting a broad ecological presence across both urban and natural waters. Other prevalent phyla commonly associated with the wastewater environment include *Bacillota* and *Bacteroidota*, both of which are often linked to fecal pollution and organic matter decomposition. *Bacillota* contributed substantially to the bacterial communities, with the highest abundance observed at RUPP (25.2%), and considerable proportions at TML (18.6%) and TBS (21.7%). Likewise, *Bacteroidota* was found to be abundant in TSR ranges from 17.1% to 19.5%, reflecting the influence of organic matter in Phnom Penh large riverine systems. Notably, *Campylobacterota*, a phylum that includes many fecal-associated and pathogenic taxa, showed a more localized pattern, with high relative abundances in wastewater samples (31.7% at TML; 17.9% at TBS), moderate level in TSR (15.9%), and minimal presence in RUPP, suggesting minimal fecal contamination in the stagnant pond. Moderate presence of *Campylobacterota* in TSR likely reflects terrestrial runoff or sewage discharge from surrounding and upstream residential areas, contributing to contamination of fecal-associated bacteria [13, 14]. This sharp contrast highlights the need for further investigation into public health implications if such taxa were to become aerosolized, especially in urban settings. The freshwater sites displayed greater microbial diversity and a broader distribution of phyla. *Cyanobacteria*, a photosynthetic group often associated with algal blooms, were most abundant in the stagnant pond at RUPP (21.5%) but nearly absent from wastewater samples and low in TSR (4.8%), reflecting conditions favorable to phototrophic growth, such as light availability and nutrient accumulation in stagnant water. While generally non-pathogenic, *Cyanobacteria* are capable of forming harmful algal blooms (HABs) that can release toxins and contribute to bioaerosol emissions, potentially impacting air quality [15]. *Planctomycetota*, known for roles in nitrogen cycling, was also more prevalent in freshwater sites (5.4% at RUPP and 3.7% at TSR), indicating possible ecological differences in community function compared to wastewater samples. These distinctions point to different potentials for bioaerosol formation and environmental health risks, warranting further analysis at finer taxonomic and functional levels.

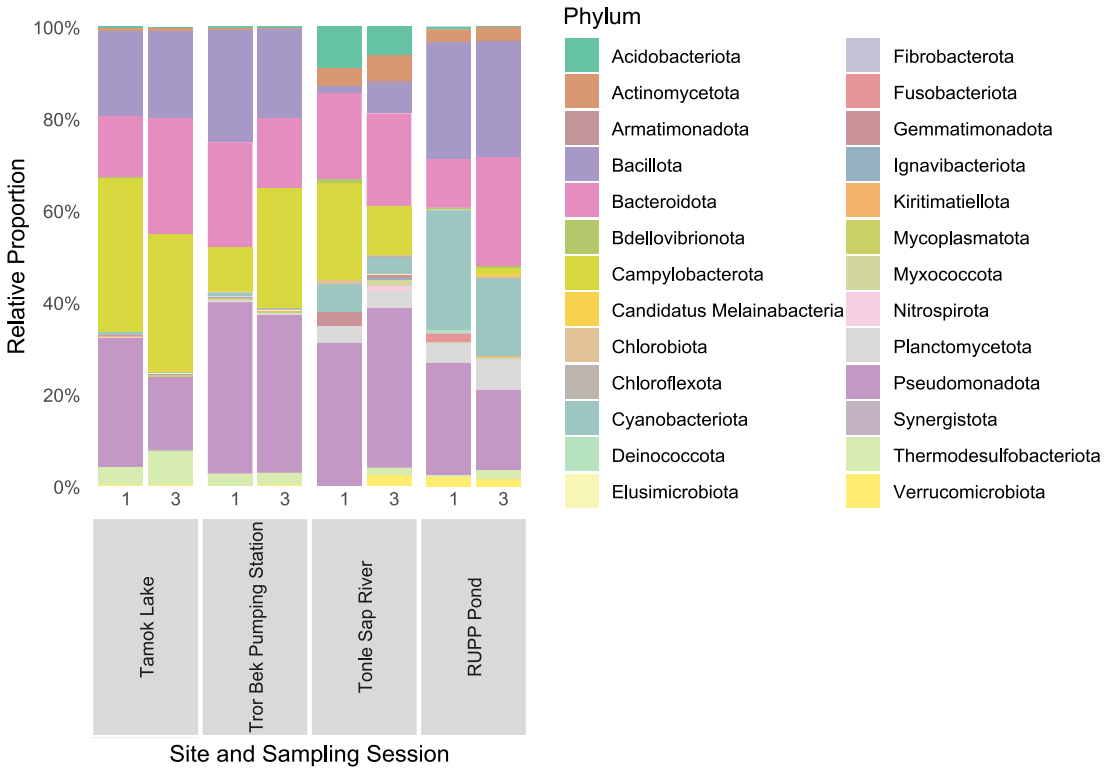


Fig. 2. Bacterial community in Phnom Penh aquatic environments at the phylum level, based on 16S rRNA gene amplicon sequencing using MinION. 1, 3 indicates different sampling sessions.

3.2 Heat Map at the Genus Level

At the genus level, notable variation was observed between freshwater and wastewater samples. After applying a 1% relative abundance cutoff, several dominant genera emerged that may be indicative of site-specific environmental conditions and potential health risks **Fig. 3**. Wastewater sites were consistently dominated by the genus *Aliarcobacter*. This genus contains pathogenic bacterial species known for gastrointestinal infections and respiratory diseases. Notably, *Aliarcobacter* reached over 16% relative abundance in both TML-0828 and TBS-0919, aligning with its characteristics as wastewater.

In contrast, the stagnant freshwater site at RUPP displayed a very different microbial profile. *Planktothrix*, a genus of cyanobacteria linked to harmful algal blooms and toxin production, was particularly dominant (up to 8.45%). Additionally, *Legionella* was detected at notable levels in RUPP samples (7.22% in RUPP-0828). This genus is known for aerosol-transmissible pathogens that cause legionellosis disease and Pontiac fever [16, 17]. The co-occurrence of these genera in stagnant, nutrient-rich water underscores the potential for bioaerosol generation in poorly circulated aquatic environments.

TSR exhibited a more diverse profile compared to other sites. In addition to fecal-indicating genera such as *Aliarcobacter* and *Arcobacter*, the genus *Terrimonas*, part of the phylum *Bacteroidota* that is commonly associated with soil and sediment, was relatively abundant at 7.80%, suggesting the deposition of terrestrial runoff. This further highlights the influence of land-based contamination on river microbial composition.

Temporal variation was also evident across all sites. Shifts in genus-level abundance between August and September suggest the influence of monsoonal rainfall on microbial dynamics. For example, reductions in *Arcobacter* or *Polynucleobacter,* and increases in *Alistipes,* or *Dechloromons* at certain sites may reflect microbial responses to changing hydrological and nutrient conditions. These patterns emphasize the importance of seasonal context in future monitoring efforts.

The detection of genera with known aerosolization potential and pathogenicity, such as Legionella, reinforces the need for downstream species-level characterization and air sampling. Such data could inform public health risk assessments and enhance microbial surveillance frameworks, especially in flood-prone, densely populated urban settings like Phnom Penh.

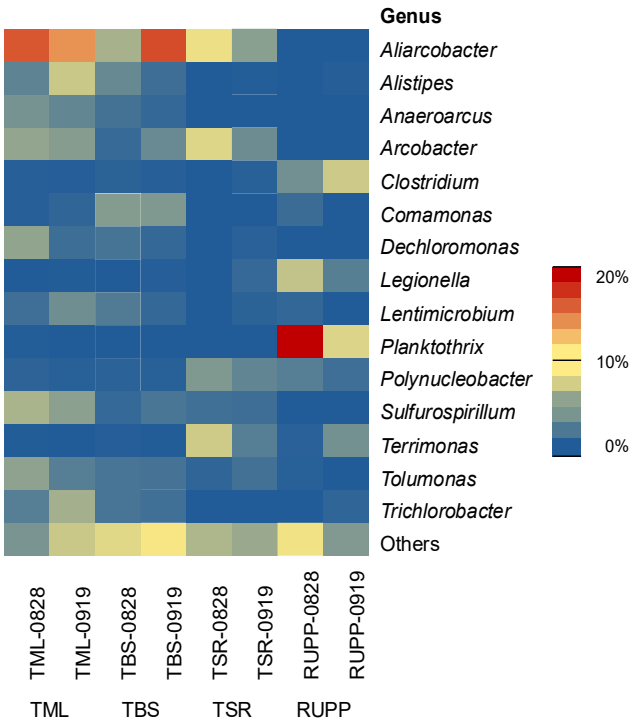


Fig. 3. Heat map of the dominant bacteria in freshwater and wastewater systems of Phnom Penh city at the genus level.

3.3 Limitation of the Study

This study provides a preliminary assessment of bacterial communities in Phnom Penh's aquatic systems and their potential implications for bioaerosol research. However, several limitations must be considered when interpreting the findings. The study was based on a limited number of samples collected during two sampling sessions within a single rainy season, which restricts both statistical integrity and temporal representativeness. Moreover, while the detection of genera with known aerosolization potential is noteworthy, the absence of direct air sampling limits the ability to confirm whether these taxa contribute to airborne exposure. Similarly, the use of a 1% relative abundance cutoff at the genus level, though effective in reducing analytical noise, may have excluded rare but clinically significant taxa. Species-level resolution was also beyond the scope of this analysis but is essential for accurately assessing public health risks associated with specific pathogenic organisms.

To advance upon these findings, future research should expand the spatial and temporal coverage of sampling and integrate air microbiome analysis to investigate water-to-air microbial transitions. Incorporating functional gene profiling, including markers for antibiotic resistance and virulence, would enhance the risk assessment framework.

4. CONCLUSIONS

This study provides preliminary insights into the bacterial communities present in Phnom Penh's urban aquatic environments and their potential role as sources of bioaerosols. Using full-length 16S rRNA gene sequencing via the Oxford Nanopore MinION platform, distinct differences were observed between freshwater and wastewater sites at both the phylum and genus levels. Wastewater environments were dominated by fecal-associated genus such as *Aliarcobacter*, while the stagnant pond at RUPP showed high relative abundances of *Planktothrix*, *Legionella*, and *Clostridium,* suggesting possible risks related to harmful algal blooms and aerosol-transmissible pathogens.

The detection of bacterial taxa with known pathogenicity or aerosolization potential highlights the need for further investigation into microbial transitions from water to air, particularly in flood-prone urban areas. Future studies should expand sampling across seasons and sites, incorporate species-level analysis, and integrate bioaerosol monitoring to better understand exposure pathways and support public health risk assessments. These findings contribute to a growing body of evidence emphasizing the importance of microbial water quality monitoring in rapidly urbanizing Southeast Asian cities.

ACKNOWLEDGMENTS

This study was financially supported by the Japan International Cooperation Agency (JICA) and Japan Science and Technology Agency (JST) under the SATREPS program (Grant No. JPMJSA2102).

REFERENCES

[1] S. Hawken, B. Avazpour, M. S. Harris, A. Marzban, and P. G. Munro, "Urban megaprojects and water justice in Southeast Asia: Between global economies and community transitions," *Cities*, vol. 113, Jun. 2021, doi: 10.1016/j.cities.2020.103068.

[2] T. B. Erickson, J. Brooks, E. J. Nilles, P. N. Pham, and P. Vinck, "Environmental health effects attributed to toxic and infectious agents following hurricanes, cyclones, flash floods and major hydrometeorological events," *Journal of Toxicology and Environmental Health - Part B: Critical Reviews*, vol. 22, no. 5–6. Taylor and Francis Inc., pp. 157–171, Aug. 18, 2019. doi: 10.1080/10937404.2019.1654422.

[3] A. Muthalagu, Y. Lian, R. M. Ravindran, and A. Qureshi, "Impacts of Floods on the Indoor Air Microbial Burden," *Aerosol Air Qual. Res.*, vol. 24, no. 1, Jan. 2024, doi: 10.4209/aaqr.230191.

[4] A. R. K. Gollakota *et al.*, "Bioaerosols: Characterization, pathways, sampling strategies, and challenges to geo-environment and health," *Gondwana Res.*, vol. 99, pp. 178–203, 2021, doi: 10.1016/j.gr.2021.07.003.

[5] O. Ginn *et al.*, "Detection and Quantification of Enteric Pathogens in Aerosols near Open Wastewater Canals in Cities with Poor Sanitation," *Environ. Sci. Technol.*, vol. 55, no. 21, pp. 14758–14771, Nov. 2021, doi: 10.1021/acs.est.1c05060.

[6] D. D. Y. Nga, V. H. Nhung, N. T. Nhan, and T. T. Hien, "Study on the concentration, composition, and recovery rate of bacterial bioaerosols after rainfall in Ho Chi Minh City," *Environ. Monit. Assess.*, vol. 196, no. 3, 2024, doi: 10.1007/s10661-024-12442-3.

[7] K. H. Kim, E. Kabir, and S. A. Jahan, "Airborne bioaerosols and their impact on human health," *Journal of Environmental Sciences (China)*, vol. 67. Chinese Academy of Sciences, pp. 23–35, May 01, 2018. doi: 10.1016/j.jes.2017.08.027.

[8] Y. Rao, H. Li, M. Chen, Q. Fu, G. Zhuang, and K. Huang, "Characterization of airborne microbial aerosols during a long-range transported dust event in eastern China: Bacterial community, influencing factors, and potential health effects," *Aerosol Air Qual. Res.*, vol. 20, no. 12, pp. 2834–2845, 2020, doi: 10.4209/aaqr.2020.01.0030.

[9] S. Romano, "Bioaerosols: Composition, Meteorological Impact, and Transport," *Atmosphere (Basel).*, vol. 14, no. 3, 2023, doi: 10.3390/atmos14030590.

[10] E. Can-Güven, "The current status and future needs of global bioaerosol research: a bibliometric analysis," *Int. J. Environ. Sci. Technol.*, vol. 19, no. 8, pp. 7857–7868, 2022, doi: 10.1007/s13762-021-03683-7.

[11] M. Zhang, J. Zuo, X. Yu, X. Shi, L. Chen, and Z. Li, "Quantification of multi-antibiotic resistant opportunistic pathogenic bacteria in bioaerosols in and around a pharmaceutical wastewater treatment plant," *J. Environ. Sci. (China)*, vol. 72, pp. 53–63, 2018, doi: 10.1016/j.jes.2017.12.011.

[12] P. Zheng *et al.*, "Nanopore sequencing technology and its applications," *MedComm*, vol. 4, no. 4, pp. 1–34, 2023, doi: 10.1002/mco2.316.

[13] UNG, "(PDF) Status of biological water quality of main rivers connected to Tonle Sap lake, Cambodia," no. May 2021, pp. 0–3, 2018, [Online]. Available: https://www.researchgate.net/publication/351984741_Status_of_biological_water_quality_of_main_rivers_connected_to_Tonle_Sap_lake_Cambodia

[14] K. Phan and K. W. Kim, "Health risk assessment of trace elements in the Tonle Sap Great Lake and the Tonle Sap River in Cambodia during the rainy season," *J. Water Health*, vol. 21, no. 5, pp. 547–559, 2023, doi: 10.2166/wh.2023.222.

[15] C. C. Lim *et al.*, "Harmful algal bloom aerosols and human health," *eBioMedicine*, vol. 93, p. 104604, 2023, doi: 10.1016/j.ebiom.2023.104604.

[16] R. Kurmayer, L. Deng, and E. Entfellner, "Role of toxic and bioactive secondary metabolites in colonization and bloom formation by filamentous cyanobacteria Planktothrix," *Harmful Algae*, vol. 54, pp. 69–86, 2016, doi: 10.1016/j.hal.2016.01.004.

[17] A. J. Prussin, D. O. Schwake, and L. C. Marr, "Ten questions concerning the aerosolization and transmission of Legionella in the built environment," *Build. Environ.*, vol. 123, pp. 684–695, 2017, doi: 10.1016/j.buildenv.2017.06.024.

Water and Environment (WAE)

3 Sub-parallel Sessions

"WAE-2: Advanced Water Quality and Pollution Assessment"

Chair: Dr. MUON Ratha; Co-chair: Dr. SANG Davin

No.	Topic
1	Assessment of Water Quality along Bassac River and Its Tributaries in Agricultural Communities *Authored by: Monychot Tepy CHANTO, Sylvain MASSUEL 3, Mallorie HIDE 2,4,5, Pavly GOV, Chanchakriya SAM, Kimyeun OEUR, Sovatey HAK, Sokleaph CHENG, Vuthy CHORK, Chanthol PENG, Anne-Laure BAÑULS*
2	Chemical Characterization of Acid Mine Drainage from Artisanal Mine Site in Chong Plash Commune, Keo Seima District, Mondulkiri Province, Cambodia *Authored by: Srey Noch CHEA, Chandoeun ENG, Muoy Yi HENG, Laymey SRENG, Chorda PEN, Kannan HAS, Bunlim HEAN, Pharong VA, Somphirom LENG*
3	Impact of City Activities on Macroplastic Accumulation in Two Tonle Sap Lake Tributaries *Authored by: Por Y CHHENG, Vouchlay THENG, Kong CHHOUN, Pinnara KET, Davin SANG, Rina HEU, Leang Gek MENH, Solida SIM*
4	Monitoring and Evaluation of Macroplastic Pollution in Siem Reap River *Authored by: Chanmonida HENG, Rina HEU , Leanggek MENH, Solita CHAN, Solida SIM, Sreynim MOL, Sarapich Vatey MAO, Vouchlay THENG, Pinnara KET, Davin SANG, Kong CHHUON*
5	Photoinduced Reactive Intermediates Formation from Microplastics and Leachate Dissolved Organic Matter *Authored by: Vinhteang KAING, Chihiro YOSHIMURA*
6	Water Quality Assessment of Groundwater in Mondolkiri Province, Cambodia: Implications for Agriculture *Authored by: Solita CHAN, Sarapich Vatey MAO, Sreynim MOL, Sreyvin THA, Techleang NAEUNG, Leanggek MENH, Rina HEU*
7	Assessing Macroplastic Pollution in Kampong Khleang River: Urban vs. Rural Pollution and Seasonal Variability *Authored by: Leanggek Menh, Rina HEU, Pinnara KET, Davin SANG, Vouchlay THENG, Vilde Kloster SNEKKEVIK, Rachel HURLEY, Solita CHAN, Solida SIM, Chanmonida Heng, Kong CHHUON*
8	Monitoring Macroplastic Accumulation in the Sangker River: Distribution Patterns and Pollution Hotspots *Authored by: Phoungsy PHEAP, Davin SANG, Rina HEU, LeangGek MENH, Solida SIM, Vouchlay THENG , Kong CHHUON, Pinnara KET, Vilde Kloster SNEKKEVIK, Rachel HURLEY 5*
9	Pesticide Contamination in Groundwater of Agriculturally Intensive Regions of Cambodia: A Case Study of Koh Thum and Kanghot *Authored by: Vuthy RUOS, Khy Eam EANG, Chanvorleak PHAT, Sereyvath YOEUN, Mathilde SESTER, Paul BAUDRON*
10	Development of Coconut Filter to Remove Microplastics at Sea Salt Farm Kampot Province, Cambodia *Authored by: Sonarith Sam, Vouchlay THENG, Chanthol Peng, Seyha DOEURN, Oudam HENG, Sovathana PHUONG*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessment of Water Quality along Bassac River and Its Tributaries in Agricultural Communities

Monychot Tepy Chanto [1,2*], Sylvain Massuel [3], Mallorie Hide [2,4,5], Pavly Gov [5], Chanchakriya Sam [5], Kimyeun Oeurn [5,6], Sovatey Hak [5], Sokleaph Cheng [5], Vuthy Chork [1], Chanthol Peng [1*], Anne-Laure Bañuls [2,4*]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *MIVEGEC, University of Montpellier-IRD-CNRS, Montpellier, France*
[3] *G-Eau (IRD), University of Montpellier, Montpellier, France*
[4] *LMI Drug Resistance in Southeast Asia, Phnom Penh-Montpellier, France*
[5] *Medical Biology Laboratory, Institut Pasteur du Cambodge, Phnom Penh, Cambodia*
[6] *Department of Life Sciences, University of Science and Technology of Hanoi (USTH), Hanoi, Vietnam*

[*] *Corresponding authors: tepy.itc@gmail.com; anne-laure.banuls@ird.fr*

Abstract: *The Bassac River and its tributaries play a crucial role in supporting agricultural activities in Kandal province, Cambodia. This region is characterized by intensive agricultural activities, which are crucial for local livelihoods but also pose potential risks to water quality. Practices such as the application of fertilizers and pesticides, along with other anthropogenic inputs, can significantly alter the hydrochemistry of the river system. Monitoring of physicochemical and major ions is therefore essential for assessing the health of the aquatic environment and its suitability for agricultural use. This study aims to assess the concentrations of major ions in the Bassac River and its tributary channels (Prek) (Prek) during the dry season, focusing on their implications for irrigation water quality in agricultural communities. Surface water samples were collected from 8 sites, where pH, electrical conductivity (EC), and temperature were measured in situ. Major ions (Na^+, NH_4^+, K^+, Mg^{2+}, Ca^{2+}, PO_4^{3-}, F^-, Cl^-, NO_2^-, Br^-, NO_3^-, and SO_4^{2-}) and alkalinity were analyzed in the laboratory, while specific irrigation hazards were also calculated. The result of the study shows that sodium adsorption ratio (SAR) and %Na suggested 'Excellent' quality with low sodium hazard. These results are further supported by USSL and Wilcox diagrams, which mostly classified Prek waters as C2S1 ('Excellent to Good'), suitable for irrigation with moderate leaching. Inter-Prek sites showed slightly higher salinity (C3S1) but were still suitable for moderately salt-tolerant crops. The nutrient pollution index (NPI) indicated minimal nutrient pollution during the dry season, but this may change during the wet season or fertilizer application periods. However, borderline magnesium hazard (MH) values and localized salinity highlight the system's sensitivity and the need for careful management. Overall, the Bassac system's water appears generally suitable for irrigation during the dry season, especially water from the tributaries.*

Keywords: Bassac River; Ion Chromatography; Major Ions; Water Quality; Hydrochemistry

1. INTRODUCTION

Inland waterways, particularly rivers and their tributaries, are vital ecosystems providing essential resources for human societies globally [1]. They serve as critical sources for drinking water, fishery, irrigation for agriculture, transportation routes, sanitation, and recreation [2, 3]. In Southeast Asia, agriculture is one of the key economic sectors therefore, the health of these river systems is directly linked to food security and economic stability. Water pollution by human activity is a global concern. The escalating degradation of freshwater ecosystems, driven by anthropogenic activities, causes ecological damage and poses significant threats to human health, consequently diminishing the availability of these resources for essential irrigation, domestic, fishing, and industrial applications [4].

The Bassac River, a major distributary of the Mekong River, along with its intricate network of tributaries (known locally as 'Preks'), forms the lifeblood for agricultural communities in the Kandal province, Cambodia. This region is characterized by intensive agricultural activities, which are crucial for local livelihoods but also pose potential risks to

[1*] Corresponding author: Monychot Tepy Chanto
E-mail: tepy.itc@gmail.com; Tel: +855-67711207

[2*] Corresponding author: Anne-Laure Bañuls
E-mai: anne-laure.banuls@ird.fr; Tel: +33-467416180

water quality [5]. Practices such as the application of fertilizers and pesticides, along with other anthropogenic inputs, can significantly alter the hydrochemistry of the river system. The region is prone to seasonal flooding and the relation between flood events and water quality has never been studied. High-water periods generally flush and dilute water, while low-water periods tend to concentrate dissolved elements.

Changes in water chemistry, particularly the concentrations of major dissolved ions (e.g., sodium, potassium, calcium, magnesium, ammonium, chloride, nitrite, nitrate, phosphate, sulfate), can directly affect the suitability of water for irrigation, potentially impacting soil structure, crop productivity [6, 7]. Monitoring the concentration of these major ions is therefore essential for assessing the health of the aquatic environment and its fitness for agricultural use. Electrical conductivity (EC), pH, and specific ion ratios are also key indicators used to evaluate potential hazards like salinity and sodium accumulation, which can degrade soil quality and hinder plant growth [8, 9]. Furthermore, elevated levels of certain ions, such as nitrate and phosphate, can indicate nutrient pollution. This study focuses on assessing the concentrations of major ions and related water quality parameters in the Bassac River and its key tributaries within the agricultural landscape of Kandal province during the dry season.

2. METHODOLOGY

2.1 Sampling sites and sample collection

During a dry season, 8 surface water samples of Bassac River and its tributary channels (Prek) were collected from Kandal province, Cambodia. A study area map with sampling sites and locations is shown in the Figures 1. Two sites at S'ang district were chosen as upstream input into the hydrological system (Upstream-Bassac represented the water from Bassac River, and Upstream-PO as the water from Prek Ambel). Four sampling sites are in Sampeou Poun municipality, including P-01, P-02, and P-03 are water from Prek, In-01, and In-02, which are water from the inter-prek (agriculture field), and Downstream sampling site is water from Bassac river as the output of the system.

All samples were immediately filtered through 0.22 µm membrane (Millipore) using Swinnex system into 50ml Falcon tube. The samples for measuring cations were acidified with nitric acid to pH<2. All the containers were sealed and kept refrigerated at about 4 °C until further analysis.

2.2 Analytical procedures

Basic water quality parameters were measured in situ. Water electrical conductivity (EC), pH, and water temperature (T) were measured using portable meters (Hana Combo HI8130). The total concentrations of soluble salts in irrigation water could be expressed for classification as low (EC ≤ 250 µS/cm), medium (EC = 250–750 µS/cm), high (EC = 750–2250 µS/cm), and very high (EC = 2250–5000 µS/cm) salinity.

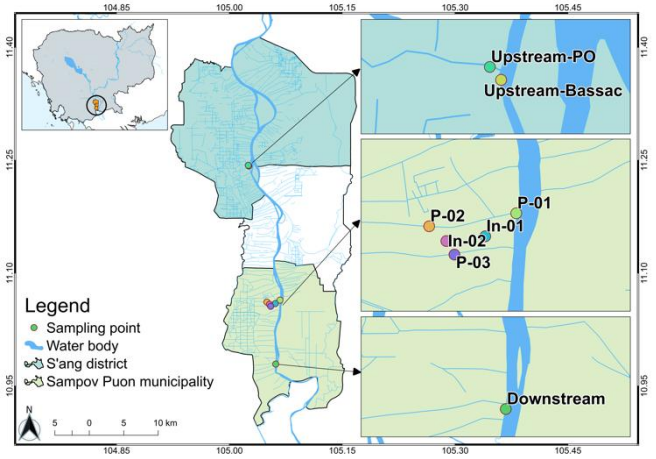


Fig. 1. Location of sampling sites in the study area.

Major ion components, including cations (Na^+, NH_4^+, K^+, Mg^{2+}, and Ca^{2+}) and anions (PO_4^{3-}, F^-, Cl^-, NO_2^-, Br^-, NO_3^-, and SO_4^{2-}), were measured by ion chromatography (IC, Shimadzu Corporation, Japan). The results are expressed in mg/L but were converted to mEq/L for use in subsequent calculations related to water quality.

Total alkalinity (expressed as calcium carbonate) was measured following the standard titration method with 0.01M sulfuric acid (H_2SO_4). The amount of alkalinity is calculated by Equation 1 [5].

$$Alkalinity = \frac{N_{acid} \times V_{acid} \times 50000}{V_{Sample}} \quad \text{(Eq. 1)}$$

where N_{acid} is the normality of H_2SO_4, V_{acid} is the volume of H_2SO_4 at pH = 4.5, V_{Sample} is the volume of the sample, and 50000g is the equivalence of $CaCO_3$.

2.3 Ion balance error

To evaluate the accuracy of the analytical results among the major anions and cations, the Percentages Ion balance error (IBE) was calculated as in Equation 2 [10].

$$\text{IBE} = \frac{\Sigma cations - \Sigma anions}{\Sigma cations + \Sigma anions} \times 100 \quad \text{(Eq. 2)}$$

The IBE values of the samples in the present study were below 10%, indicating that the results are reliable for hadrochemical assessment.

2.4 Surface water quality for irrigation

The irrigation water quality was assessed with reference to water quality indices (WQIs) such as sodium adsorption ratio (SAR), residual sodium carbonate (RSC), magnesium hazard (MH), and percentage sodium (%Na), with respect to the measured water parameters [8, 10, 11]. All ions concentration used are equivalent concentrations (mEq/L).

SAR assesses sodium hazard by considering calcium and magnesium concentrations, with high SAR values increasing soil Na^+ and causing sodium to replace Ca^{2+} and Mg^{2+}, making soil hard and compact, reducing permeability, and hindering crop growth. An excess amount of Mg^{2+} over Ca^{2+} also decreases the quality of irrigation water. It also decreases crop yields. High MH can reduce yields by raising soil alkalinity, despite Mg^{2+} being an essential nutrient [5, 6]. The calculations of these indices are shown in Equations (2, 3).

$$SAR = \frac{Na^+}{\sqrt{(Ca^{2+}+Mg^{2+})/_2}} \quad \text{(Eq. 2)}$$

$$MH = \left(\frac{Mg^{2+}}{Ca^{2+}+Mg^{2+}}\right) \times 100 \quad \text{(Eq. 3)}$$

%Na is a key index for classifying irrigation water quality because sodium reduces soil permeability, and irrigation water with high %Na causes sodium to accumulate in the soil, restricting plants' ability to absorb water and nutrients, ultimately lowering crop yields [5, 9]. The calculations of %Na following Equation 4.

$$\%Na = \left(\frac{Na^+ + K^+}{Na^+ + Ca^{2+} + Mg^{2+} + K^+}\right) \times 100 \quad \text{(Eq. 4)}$$

To further evaluate the suitability of the water for irrigation, the US Salinity Laboratory diagram (USSL) and Wilcox diagram are used, where EC is plotted against SAR and %Na, respectively [7, 8].

RSC evaluates the effects of carbonate and bicarbonate; high RSC values, due to carbonate minerals increasing Na^+ as Na_2CO_3, can degrade soil physical properties [5, 6]. Equation 5 shows a calculation of RSC.

$$RSC = Alkalinity - (Ca^{2+} + Mg^{2+}) \quad \text{(Eq. 5)}$$

The ranges of water suitability classes of the above parameters are shown in Table 1.

Table 1. Quality ranges classification of WQIs for irrigation.

Indices	Ranges	Water quality
SAR	<10 mEq/L	Excellent
	10-18 mEq/L	Good
	18-26 mEq/L	Doubtful
	>26 mEq/L	Unsuitable
MH	<50%	Suitable
	>50%	Unsuitable
%Na	<20%	Excellent
	20–40%	Good
	40–60%	Permissible
	60–80%	Doubtful
	>80%	Unsuitable
RSC	<1.25 mEq/L	Good
	1.25–2.5 mEq/L	Doubtful
	>2.5 mEq/L	Unsuitable
	>5 mEq/L	Harmful
NPI	<1 mEq/L	No pollution
	1-3 mEq/L	Moderate polluted
	3-6 mEq/L	Considerable polluted
	>6 mEq/L	Very high polluted

2.5 Nutrient pollution

The nutrient pollution index (NPI) is used to assess pollution levels in surface water (especially in slow-moving water bodies), particularly from nutrients like phosphate and nitrate, which primarily enter rivers through agricultural activities, synthetic fertilizers, fish farms, domestic sewage, surface runoff, and floodwaters [12, 13]. The below formula (6) is applied to calculate the NPI (Table 1):

$$NPI = \frac{C_N}{MAC_N} + \frac{C_P}{MAC_P} \quad \text{(Eq. 6)}$$

where $C_{N/P}$ represents the nitrate and phosphate concentration, $MAC_{N/P}$ indicates the World Health Organization's maximum allowable limits of 50 mg/L for nitrate and 5 mg/L for phosphate in surface water.

3. RESULTS AND DISCUSSION

3.1 Water quality assessment

The results of the hydrogeochemical analysis for water samples are given in Table 2. Water sample temperatures ranged from 29 to 31 °C, consistent with a tropical climate. The pH values range from mildly acidic (6.57) to mildly alkaline (7.56), all falling within the FAO permissible range of 6.5 to 8.4 [14]. All sites have EC values <700 µS/cm, indicating adequate for the FAO irrigation water standard (<700 µS/cm), except sites In-01 and In-02, which have EC values up to 790 µS/cm and 1290 µS/cm accordingly. However, the EC values in the Bassac River and Prek measured in this study are higher than those reported by Tweet et al. [15]. This difference may be attributed to the variation in sampling seasons, as their study was conducted during the rainy season (June and September), whereas the present study was carried out in the dry season (February).

The alkalinity, expressed as calcium carbonate (CaCO3), is generally classified as soft (0–60 mg/L), moderately hard (61–120 mg/L), hard (121–180 mg/L), and very hard (>180 mg/L) [16]. From the alkalinity results, the water of the study area is moderately hard (Upstream PO, P-01, P-03, Downstream Bassac, In-02) and hard water (Upstream Bassac, P-02, In-01).

The concentration order of the main cations is $Ca^{2+} > Na^+ > Mg^{2+} > K^+ > NH_4^+$, and for the main anions is $Cl^- > SO_4^{2-}$

> NO_3^- > F^- > Br^- > NO_2^- > PO_4^{3-}. The majority of cations are Ca^{2+} and Na^+, while anions are Cl^- and SO_4^{2-}.

In summary, water taken from agricultural fields (Inter-prek sites; In-01 and In-02) has higher EC and elevated concentrations of ions compared to water from Prek. This is primarily due to the impact of agricultural practices, including fertilizer application and soil amendments, which introduce and mobilize salts and nutrients in the soil and water.

3.2 Water quality indices

The results of the Water quality indices are given in Table 2. The SAR and %Na values represent the sodium hazard to soil aggregates by irrigation. Most samples can be defined as the excellent quality (SAR < 10; Na% < 30%), only the samples In-01 and In-02, with Na% of 55.03% and 50.07%, are classified as permissible.

These results are further supported by the Wilcox diagram (Figure 2) and the USSL diagram (Figure 3), which were generated using EC, N%, and SAR values, respectively. From Wilcox diagram, all the Prek waters are in the area of the 'Excellent to Good' zone, while the Inter-Prek waters (In-01 and In-02) are in the "Good to Permissible" zone in the Wilcox diagram. These imply that the Prek waters at present are suitable for agricultural activities and will not lead to soil problems. However, with the enhancement of human activities, the monitoring of water quality should be a long-term and serious mission [17].

Table 2. Hydrogeochemical values for water samples. Ion concentrations, NPI, SAR, and RSC are given in mEq/L, %NA and MH (%), alkalinity (mg/L), EC (µS/cm), and Temperature (°C).

Parameters	Upstream Bassac	Upstream PO	P-01	P-02	In-01	In-02	P-03	Downstream
pH	6.57	7.15	7.34	7.56	6.78	6.66	6.85	7.56
EC	370	390	330	360	790	1290	390	330
Temperature	29.2	29.3	31.4	29.9	30.2	32.6	29.5	29.9
Alkalinity	166	118	120	132	108	102	140	112
F^-	0.016	0.016	0.017	0.016	0.027	-	0.017	0.017
Cl^-	1.275	1.268	1.111	1.356	4.259	5.788	1.226	1.115
NO2$^-$	0.004	0.001	-	0.004	0.002	-	0.001	-
NO3$^-$	0.041	0.061	0.045	0.047	0.003	-	0.006	0.050
Br^-	0.002	0.003	0.003	0.003	0.004	0.007	0.003	0.002
SO_4^{2-}	0.365	0.327	0.306	0.346	0.771	5.249	0.709	0.312
PO_4^{3-}	-	-	-	-	-	0.110	-	-
Na^+	0.863	0.679	0.660	1.058	4.458	6.049	1.167	0.678
K^+	-	0.159	-	-	0.108	0.811	-	-
Mg^{2+}	1.333	1.188	1.145	1.181	1.650	2.903	1.247	1.152
Ca^{2+}	1.357	1.124	1.124	1.222	2.082	3.938	1.655	1.160
NH_4^+	-	-	-	-	-	-	-	-
NPI	0.05	0.076	0.055	0.058	0.004	0.022	0.007	0.062
%Na	24.279	26.608	22.52	30.573	55.025	50.068	28.687	22.665
MH	49.546	51.384	50.449	49.146	44.217	42.438	42.979	49.843
SAR	0.744	0.632	0.619	0.965	3.264	3.271	0.969	0.63
RSC	0.63	0.05	0.13	0.24	-1.57	-4.80	-0.10	-0.07

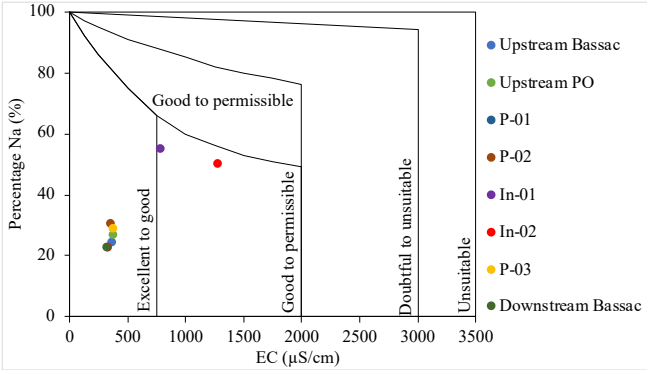


Fig. 2. Wilcox diagram representing the suitability of water for irrigation.

In USSL diagram, EC represents salinity hazard, and SAR represents alkalinity hazard. Similar to the Wilxox diagram, all Prek waters fell into the C2S1 category, indicating medium salinity and low alkalinity, whereas Inter-Prek waters are in C3S1 category, presenting high salinity and low alkalinity. Ions such as Cl^- and SO_4^{2-} in water also have an effect on salinity. The findings suggest that Prek waters appear suitable for irrigation, provided a moderate amount of leaching occurs to manage salinity. The risk of exchangeable sodium reaching dangerous levels is low, and moderately salt-tolerant plants can be cultivated without specific salinity control measures [18].

In this study, 50% of the samples have RSC values less than 1 mEg/L, while the other 50% have negative RSC. These negative RSC values indicate a dominance of calcium and magnesium ions over carbonate and bicarbonate ions. Based

on these RSC findings, all water samples were classified as being of good quality and suitable for irrigation purposes.

Another index used in the evaluation of irrigation water quality is magnesium hazard. The study shows that MH value ranges from 42% to 51%, suggesting that the waters are almost falling into the unsuitable ranges for irrigation.

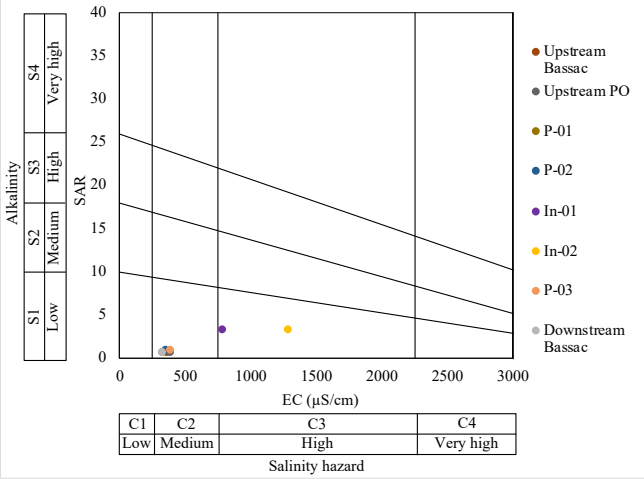


Fig. 3. USSL diagram representing the suitability of water for irrigation.

3.3 Nutrient pollution index (NPI)

The Nutrient Pollution Index (NPI) remained below 1 even in agricultural field sites (Inter-prek; In-01 and In-02). This suggests minimal nutrient pollution from fertilization, likely because sampling occurred post-harvest.

4. CONCLUSIONS

This study assessed dry-season water quality in the Bassac River system in Kandal province, focusing on major ions and irrigation suitability. While pH and temperature were typical, electrical conductivity varied, indicating moderate to high salinity in some areas, and water hardness ranged from moderately hard to hard. Ca^{2+} and Cl^{-} were the dominant ions. Based on irrigation indices like Na%, RSC, and SAR, the water at all sites was generally suitable for irrigation. SAR and %Na indicated 'Excellent' quality for most sites, with low sodium hazard. USSL and Wilcox diagrams confirmed this, classifying most waters as suitable for irrigation. However, MH values bordered the 'unsuitable' threshold in some cases, warranting attention. The NPI was low, suggesting minimal nutrient pollution during this post-harvest sampling period.

Future research should prioritize monitoring across seasons, investigating salinity/magnesium sources, and assessing long-term trends and potential health impacts to support sustainable water resource management in the region.

This study on water quality is part of a broader research effort that also includes ongoing studies on antimicrobial resistance and pathogenic bacteria. All these activities are being conducted within the framework of the PEACH and AFRICAM projects.

ACKNOWLEDGMENTS

The authors are grateful to AFD for the financial support in the framework of PREACTS-AfriCam in Cambodia, and to the University of Montpellier for PEACH project (ExpoSum). We also thank the Platform for Aquatic Ecosystem Research, Institut Pasteur du Cambodge, and IRD for their support with sample collection and laboratory utilities. MTC is supported by French Government Scholarships co-funded by the Ministry of Education Youth and Sports.

REFERENCES

[1] L. Chabib *et al.*, "Assessment of stream quality and health risks in Indonesian river systems: A social analysis and water quality index approach," *Case studies in chemical and environmental engineering,* vol. 11, p. 101200, 2025, doi: 10.1016/j.cscee.2025.101200.

[2] A. U. Haq and S. Muhammad, "Spatial distribution of drinking and irrigation water quality indices of Ghizer River Basin, northern Pakistan," *Environmental science and pollution research international,* vol. 30, no. 8, pp. 20020-20030, 2023, doi: 10.1007/s11356-022-23478-8.

[3] R. Das Kangabam and M. Govindaraju, "Anthropogenic activity-induced water quality degradation in the Loktak lake, a Ramsar site in the Indo-Burma biodiversity hotspot," *Environmental technology,* vol. 40, no. 17, pp. 2232-2241, 2019, doi: 10.1080/09593330.2017.1378267.

[4] V. Kumar, A. Sharma, R. Kumar, R. Bhardwaj, A. Kumar Thukral, and J. Rodrigo-Comino, "Assessment of heavy-metal pollution in three different Indian water bodies by combination of multivariate analysis and water pollution indices," *Human and ecological risk assessment,* vol. 26, no. 1, pp. 1-16, 2020, doi: 10.1080/10807039.2018.1497946.

[5] Ş. Çankaya, M. Varol, and A. Bekleyen, "Hydrochemistry, water quality and health risk assessment of streams in Bismil plain, an important agricultural area in southeast Türkiye," *Environmental pollution (1987),* vol. 331, no. Pt 1, pp. 121874-121874, 2023, doi: 10.1016/j.envpol.2023.121874.

[6] B. Panneerselvam, K. Muniraj, M. Thomas, N. Ravichandran, and B. Bidorn, "Identifying influencing groundwater parameter on human health associate with irrigation indices using the Automatic Linear Model

(ALM) in a semi-arid region in India," *Environmental research,* vol. 202, pp. 111778-111778, 2021, doi: 10.1016/j.envres.2021.111778.

[7] J. Liu and G. Han, "Distributions and Source Identification of the Major Ions in Zhujiang River, Southwest China: Examining the Relationships Between Human Perturbations, Chemical Weathering, Water Quality and Health Risk," *Exposure and health,* vol. 12, no. 4, pp. 849-862, 2020, doi: 10.1007/s12403-020-00343-y.

[8] J. C. Hide, "Diagnosis and Improvement of Saline and Alkali Soils . U.S. Salinity Laboratory Staff; L. A. Richards, Ed. U.S. Dept. of Agriculture, Washington, D.C., rev. ed., 1954. vii + 160 pp. Illus. $2. (Order from Supt. of Documents, GPO, Washington 25, D.C.)," *Science (American Association for the Advancement of Science),* vol. 120, no. 3124, pp. 800-800, 1954, doi: 10.1126/science.120.3124.800.a.

[9] S. Joseph, M. R. Sijimol, J. Thomas, and A. M. Sheela, "Hydrogeochemical characterization and analysis for irrigation applicability of groundwater in the shallow coastal aquifers: a multivariate statistical approach," *International journal of river basin management,* vol. 20, no. 3, pp. 363-374, 2022, doi: 10.1080/15715124.2020.1837143.

[10] S. Ghosh, B. Sarkar, A. Islam, P. K. Shit, A. Quesada-Román, and H. A. R. Gazi, "Surface Water and Groundwater Suitability for Irrigation Based on Hydrochemical Analysis in the Lower Mayurakshi River Basin, India," *Geosciences (Basel),* vol. 12, no. 11, p. 415, 2022, doi: 10.3390/geosciences12110415.

[11] S. Elsayed, H. Hussein, F. S. Moghanm, K. M. Khedher, E. M. Eid, and M. Gad, "Application of Irrigation Water Quality Indices and Multivariate Statistical Techniques for Surface Water Quality Assessments in the Northern Nile Delta, Egypt," *Water (Basel),* vol. 12, no. 12, p. 3300, 2020, doi: 10.3390/w12123300.

[12] F. Ustaoğlu, B. Taş, Y. Tepe, and H. Topaldemir, "Comprehensive assessment of water quality and associated health risk by using physicochemical quality indices and multivariate analysis in Terme River, Turkey," *Environmental science and pollution research international,* vol. 28, no. 44, pp. 62736-62754, 2021, doi: 10.1007/s11356-021-15135-3.

[13] B. O. Isiuku and C. E. Enyoh, "Pollution and health risks assessment of nitrate and phosphate concentrations in water bodies in South Eastern, Nigeria," *Environmental advances,* vol. 2, p. 100018, 2020, doi: 10.1016/j.envadv.2020.100018.

[14] R. S. a. W. Ayers, D.W. "Water Quality for Agriculture, FAO Irrigation and Drainage Paper." Food and Agriculture Organization of the United Nations. (accessed.

[15] S. Tweed *et al.*, "Seasonal influences on groundwater arsenic concentrations in the irrigated region of the Cambodian Mekong Delta," *The Science of the total environment,* vol. 728, pp. 138598-138598, 2020, doi: 10.1016/j.scitotenv.2020.138598.

[16] C. E. Boyd, C. S. Tucker, and B. Somridhivej, "Alkalinity and Hardness: Critical but Elusive Concepts in Aquaculture," *Journal of the World Aquaculture Society,* vol. 47, no. 1, pp. 6-41, 2016, doi: 10.1111/jwas.12241.

[17] P. Li, X. He, Y. Li, and G. Xiang, "Occurrence and Health Implication of Fluoride in Groundwater of Loess Aquifer in the Chinese Loess Plateau: A Case Study of Tongchuan, Northwest China," *Exposure and health,* vol. 11, no. 2, pp. 95-107, 2019, doi: 10.1007/s12403-018-0278-x.

[18] M. El-Sayed and W. M. Salem, "Hydrochemical assessments of surface Nile water and ground water in an industry area – South West Cairo," *Egyptian journal of petroleum,* vol. 24, no. 3, pp. 277-288, 2015, doi: 10.1016/j.ejpe.2015.07.014.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Chemical Characterization of Acid Mine Drainage from Artisanal Mine Site in Chong Plash Commune, Keo Seima District, Mondulkiri Province, Cambodia

Srey Noch CHEA [1*], Chandoeun ENG [1], Muoy Yi HENG [1], Laymey SRENG [1], Chorda PEN [2], Kannan HAS [1], Bunlim HEAN [1], Pharong VA [1], Somphirom LENG [1]

[1] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Ministry of Mines and Energy, HWFC+MX, Preah Moha Ksatreiyani Kossamak Ave (106), Phnom Penh, Cambodia*

[*] *Corresponding author: cheasreynoch288092002@gmail.com*

Abstract: *Mine waste can seriously harm the environment because it lets out large amounts of heavy metals and acid. This happens when sulfide minerals like pyrite oxidize during mining, creating sulfuric acid and heavy metals that affect nearby ecosystems. The issue is exacerbated by releasing additional contaminations from the surrounding rock as the result of Acid Mine Drainage (AMD) formation. This research aims to determine the concentration of heavy metals and some toxic elements and evaluate the environmental impact on water quality and soil quality. At each sampling point, samples were taken from Chong Plash commune, Keo Seima district, and Mondulkiri province, where wastewater and sediment samples. On-site measurement of water samples identified physicochemical parameters that determine pH, electrical conductivity, redox potential, dissolved oxygen, and total dissolved solids, and both samples neither carried out the amounts of chemical parameters such as iron, aluminum, copper, zinc, lead, cadmium, chromium, and cobalt using MP-AES analysis and XRF analysis. As a result of physicochemical parameters, the pH value of water samples was neutralized, while the electrical conductivity and total dissolved solids were higher to indicate high ionic content due to AMD processes and consistent with AMD-impacted waters where the concentration of metals is high. In addition, the moderate DO and TDS indicate strong oxidation of sulfide minerals. There are CP-W4 to CP-W8 sampling points that were most contaminated by AMD products due to the high concentration of Fe, Al, and Ni compared to the maximum allowable concentration based on WHO and MIME guidelines. Additionally, the concentration of metals in sediment samples was increased in CP-S6, CP-S8 and CP-S10 samples. The concentration metals, especially of As, Fe, and Cu, illustrates that the area is heavily affected by AMD from mining-related activities since the oxidation of sulfide minerals exposed during mining activities.*

Keywords: Acid mine drainage, Chong Plash Commune, Heavy metal, MP-AES, Wastewater, Physicochemical Parameter

1. INTRODUCTION

Small-scale and artisanal mining operations have increased recently in Cambodia. Most of those are doing precious metal mining like gold of those are including Mondulkiri, Ratanakiri, Kratie, Kampong Cham, Preah Vihear, and etc. The mining industry is known to cause important environmental pollution, while mining waste heaps are recognized as sources of long-term contamination of water and, in general, of the environment [1]. Mine Waste is a significant risk of environmental contamination, as releases substantial quantities of heavy metals and acids. Drainage from mining by-products, such as tailing and waste rocks often contains high metal concentrations and acidic pH levels due to the chemical and hydrological weathering processes which act on the waste materials. The family artisanal mining without certain management of artisanal tailing and waste rock, the surrounding ground and surface waters will contaminate the leachate [2].

The formation process of acid mine drainage is the oxidation process of sulfide minerals, which is completed by chemical oxidation and biological oxidation [3]. When the sulfide minerals were exposed to the environment's atmospheric (oxygen) with water, caused by the oxidation of sulfide minerals, and produced the sulfuric acid (H2SO4). The most common sulfide minerals such as pyrite (FeS2), chalcopyrite (CuFeS2), arsenopyrite (FeAsS), and pyrrhotite (Fe(1-x)S).

[*] Corresponding author: Srey Noch CHEA
E-mail: cheasreynoch28092002@gmail.com; Tel: +855-975994670

With the participation of water and oxygen, sulfide ore as pyrite was exposed to an oxidative environment in this process, is oxidized to produce H+ ion (acidity), sulfate ion, and metal ion [4], as shown in the reaction Eq. 1. Ferrous iron (Fe2+) further oxidizes to ferric iron (Fe3+) as seen in (Eq. 2). At low pH, this reaction is catalyzed by the presence of bacteria (mainly by Thiobacillus ferroxidase) increasing the rate of reaction by 20 to 1000 times [5]. This process is the key to the formation of AMD.

$$2FeS_2 + 7O_2 + 2H2O \rightarrow 2Fe^{2+} + 4H^+ + 4SO_4^{2-} \quad \text{(Eq. 1)}$$

$$4Fe^{2+} + O_2 + 4H+ \rightarrow 4Fe^{3+} + 2H2O \quad \text{(Eq. 2)}$$

At pH >4, Fe^{3+} precipitates as ferric hydroxide ($Fe(OH)_3$) (Eq. 3), which releases more acid into the environment.

$$Fe^{3+} + 3H2O \rightarrow Fe(OH)_3 + 3H^+ \quad \text{(Eq. 3)}$$

Additionally, AMD is identified by chemical parameters. The typical parameters of AMD are low pH, high acidity, high conductivity, and high concentration of heavy metals (especially iron and aluminum).

Artisanal mine wastes have no right management and are recognized as a source of long-term contamination of water and the environment. The improper management of waste rock and wastewater from artisanal mining has resulted in increased concentrations of heavy metals around the mine site, which is an alarming concern. It is the reason that contamination of the environment is also possible. This research is located at Chong Plash commune, Mondulkiri Province, Cambodia. This study aims to (1) assess the concentration of heavy metals and toxic elements and (2) evaluate the environmental impact of heavy metal.

2. METHODOLOGY

2.1 Site description

The region's geology of this area is characterized by Triassic sandstone interbedded with mudstone, and micro-breccias with some intrusion of granodiorite and surrounded by hornfels. The study area consists of the tailing impoundment contained in a valley and a small lake. The physical-chemical properties of Acid Mine Drainage (AMD) were collected in rainy season. The sampling sites in this research in Chong Plash commune, Mondulkiri province, and eastern Cambodia as shown in Fig. 1. The distance from Phnom Penh and Mondulkiri province is approximately 366 km and take about 7 to 9 hours.

2.2 Sampling and in-situ measurement

The sampling collections were performed in the rainy season of August 2024. The water samples were collected from the processing plant, and some were from tailing. Water sampling points are named as follows: CP-W1 to CP-W10, detailed in Fig. 2 and Fig. 3.

The water samples were collected in acid-rinsed polyethylene bottles and stored refrigerated during transport to the laboratory. The physicochemical parameters, such as pH, electrical conductivity, oxidation-reduction potential, dissolved oxygen, and temperature, were measured in the field using the Hanna HI98194 Multiparameter Meter. The distance from each sampling point was obtained from the actual site. All wastewater samples were taken in three polyethylene bottles (volume 50 ml) and kept in a cooler at 4 °C. Each sample was filtered in the field under vacuum through 0.45 µm, 0.1 µm, and non-filter. One bottle which was filtered with 0.1 µm, was added pure nitric acid (HNO_3) 1% of the volume of the water sample (to preserve for cation).

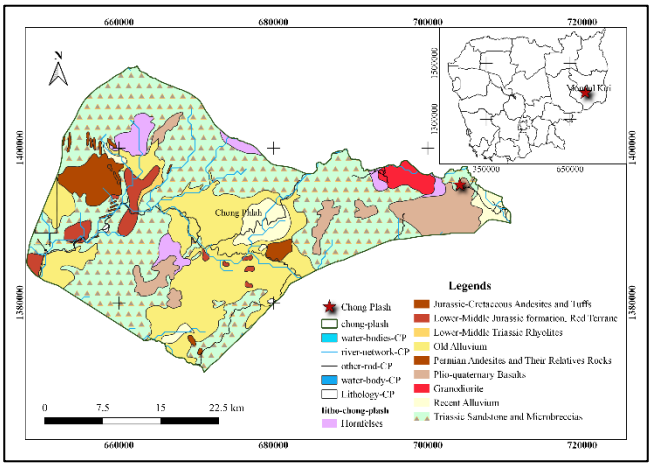


Fig. 1. Map showing the location of study area

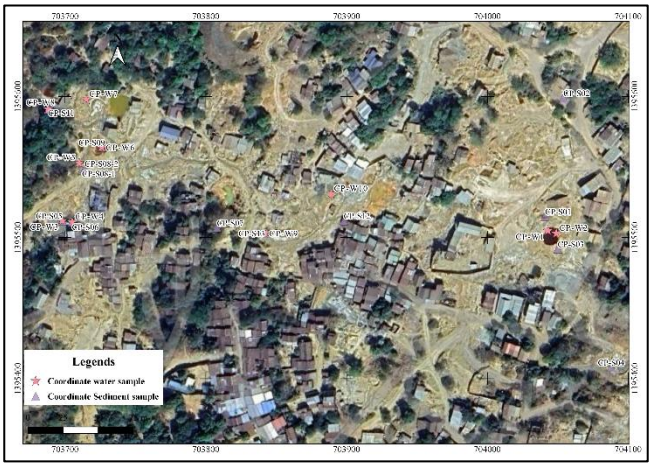


Fig. 2. Mapping of water and sediment sample collection

2.3 Analyte methods

The chemical analysis of water samples was carried out

by microwave plasma-atomic emission spectrometry (MP-AES) at the Laboratory of Nano-structure and Chemical Analysis (ITC). The standard samples were diluted from 1000 ppm to 0.01 ppm. The detection limits were as follows: Zn; 4 $\mu g.L^{-1}$, Fe; 3 $\mu g.L^{-1}$, Cd; 1.2 $\mu g.L^{-1}$, Cu; 0.4 $\mu g.L^{-1}$, Al; 0.5 $\mu g.L^{-1}$, Pb; 2.5 $\mu g.L^{-1}$, Ni; 2 $\mu g.L^{-1}$, Cr; 1.2 $\mu g.L^{-1}$ and Co; 6 $\mu g.L^{-1}$. The major chemical compositions of the sediment samples were conducted by X-ray fluorescence (XRF) in the laboratory of the faculty of Geo-resources and Geotechnical Engineering, ITC, using glass bead samples and synthetic calibration. The point sampling was ground as a powder and compacted with a ring amount of 15 g.



Fig. 3. Photography field illustrating of some sampling points: A) CP-W1 and CP-W2 from small lake; B) CP-W3 and CP-W4 collect from ore extraction plant; C) CP-W5 collect from tributary; D) CP-W6 select from a lake which principiated; E) CP-W7 take from a lake

3. RESULTS AND DISCUSSION

3.1 Physiochemical parameter

Table 1. shown the statistics of physicochemical parameters that were measured on-site, including pH, electrical conductivity, dissolved oxygen (DO), and total dissolved solids (TDS), and conducted in the rainy season of August. The parameters were compared with MIME Cambodia standards and WHO standards for water.

The pH of the wastewater ranged from 6.69 to 6.89, with an average value of 6.78. The tributary pH at each sampling point was above the WHO guideline level and MIME standard (Fig. 4) due to the rise in the water level. In determining the water quality, pH level is a significant consideration as it has an impact on the variety of biological and chemical processes in surface water. With rainwater, all sampling points also become no AMD formation at this site. Although sulfide minerals such as pyrite and chalcopyrite are present in this area to generate AMD, its formation is probably neutralized since rainfall tends to dilute AMD-contaminated water, reducing the concentration of metals and TDS. The conductivity values range from 1.23 mS/cm to 4.80 mS/cm, while the TDS values range between 825.89 mg/L and 3218.46 mg/L. The electrical conductivity value is higher at the CP-W4 (4.80 mS/cm) sampling point due to the high quantity of total dissolved solids (TDS=3218.46 mg/L) in the water. According to WHO [6], the permissible limit for EC (Fig. 6) is 400 µS/cm. A high TDS value (3218.46 mg/L) showed a general level of pollution and one of the parameters that is characteristic of acid mine drainage at the CP-W4 sample. The prescribed limit of TDS (Fig. 5) by MIME standard [7] is 800 mg/L, significantly, the TDS value is above maximum permissible limits in most samples, consistent with AMD-impacted waters where sulfate and metal concentrations are high. Even with a neutral pH level, high TDS and EC identify toxicity to aquatic life and unsuitability for drinking water.

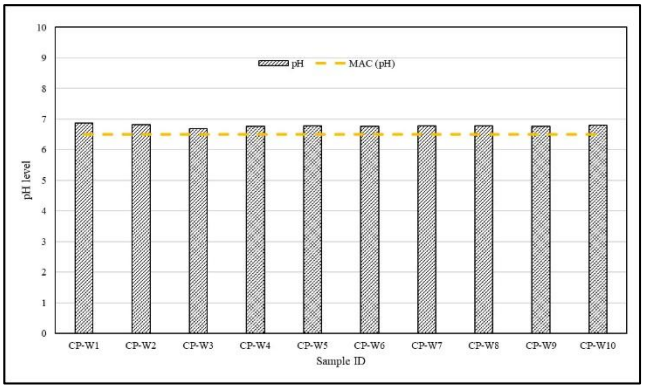


Fig. 4. A diagram of pH level of water sample from in-situ measurement

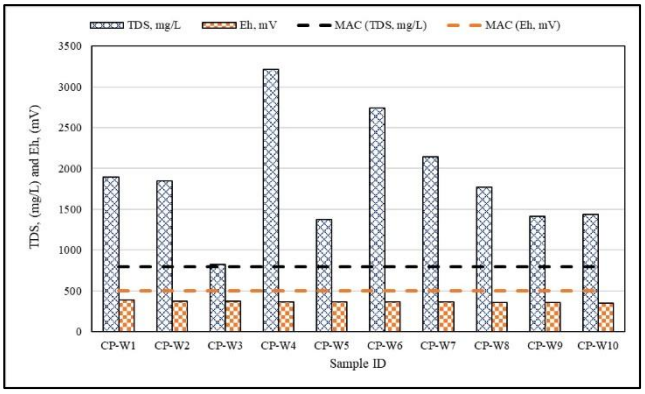


Fig. 5. A diagram of total dissolved solids and redox potential of water samples

The dissolved oxygen ranges from 6.55mg/L to 10.68 mg/L (Fig. 6), and Eh values range between 352.39 and 385.25 mV (Fig. 5). High DO and Eh values indicate strongly oxidizing conditions that can generate sulfuric acid and mobilization of metals. All water samples with high dissolved oxygen values (>5 mg/L) indicated that the quality of surface water is high and supports aquatic life. The WHO standard permissible level for DO [8] is 5 mg/L.

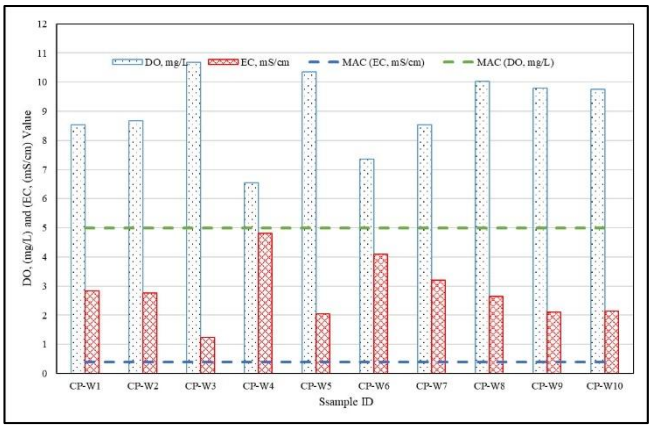


Fig. 6. Diagram of DO and EC values of water samples from in-situ measurement

3.2 Chemical parameters of wastewater

A significant part of the study was the analysis of metals, especially heavy metals as they present general and real environmental problems [9]. Table 2 illustrate the summary of statical of concentration of heavy metals (mg/l) in water sample that collected from the abandoned and active of artisanal mining. Aluminum is detected in water samples from mine tailing with the mean concentration 11.95 mg/L, the lowest value of 0.03 mg/L at CP-W10 and highest value of 62.54 mg/L at CP-W6. Fe has a mean concentration of 262.70 mg/L with lowest value of 2.6 mg/L from CP-W2 and highest value of 1335.24 mg/L from CP-W4. The highest Fe concentration was extremely elevated in nearly all sites, especially, CP-W4 to CP-W8 (Fig. 7), while the concentration of Al exceeded the limit in 6 out of 10 samples, with CP-W6 (Fig. 8). High Al levels are common in AMD due to acid-promoted dissolution, and it is toxic to aquatic life, either. Concentration of metals including Zn, Cu, and Ni show in Fig. 8 Zinc has mean concentration of 0.61 mg/L with a lowest value of 0.14 mg/L at CP-W1 and highest value of 2.44 mg/L at CP-W7 sampling point. However, it wasn't detected in water sample from CP-W2, CP-W4, CP-W9 and CP-W10. Zn concentrations are within WHO and MIME limit but are still concerning at CP-W6 to CP-W8 sampling points. Cu has a mean concentration of 0.54 mg/L in water. Highest concentration value of Cu is 1.86 mg/L from CP-W6 that it is above permissible limit by WHO (2004) and below MIME guideline but it's still toxic to organisms. Ni has a mean concentration of 0.03 mg/L. The higher concentration level of Ni (0.13 mg/L) and (0.07 mg/L) from CP-W6 and CP-W4; respectively is below the permissible limit than standard. The value of Cd, Pb, Co and Cr are below detection limit in water samples, however occasional traces of Cr and Cd were noted due to these metals pose acute toxicity risks event at low concentrations. Despite pH values being within guideline ranges (6.69-6.87), metal concentration remains high that mobilized by rainy season runoff, especially in CP-W4 to CP-W8 samples. This indicate that AMD is having an active effect on the environment, during a season when a decrease in contaminants may be expected.

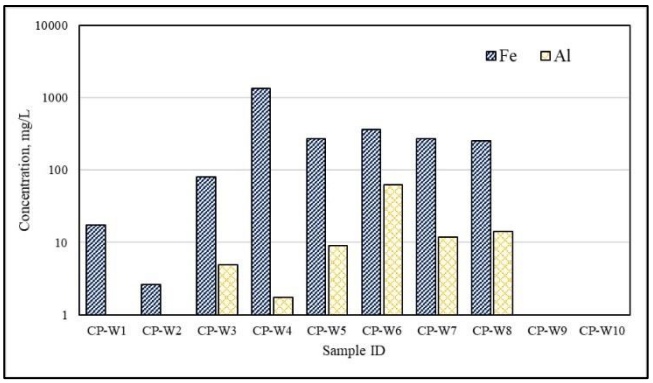


Fig. 7. Diagram of Fe and Al concentation of aqua samples

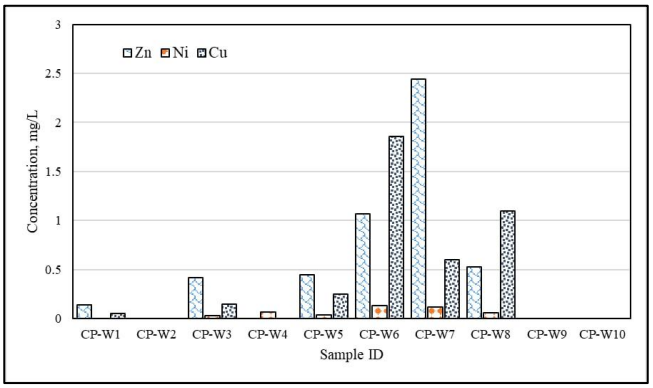


Fig. 8. Concentration of zinc, copper and nickel values of water samples

3.3 Chemical parameters of sediment samples

From Table 3, significant in heavy metal concentration was found in the sediment from all fourteen-sampling point. Detail of statistical analysis of heavy metals in sediment sample. The concentration of heavy metal in sediments samples were higher than those in water samples because sediment is thought to be an important for heavy metals absorption [10]. At each sampling point of heavy metals have mean concentration of Al (75347.70 mg/Kg), Fe (79459.64 mg/Kg), Cu (474.57 mg/Kg), As (18071.05 mg/Kg), Mn (473.14 mg/kg), Pb (83.69 mg/Kg), and Zn (44.56 mg/Kg).

Table 1. Statistical summary of physiochemical parameters; analyzed 10 samples of wastewater. Avg = average; Min = minimum; Max = maximum

Parameter	Temperature, °C	pH	EC, mS/cm	Eh, mV	TDS, mg/L	DO, mg/L
Ave.	35.18	6.78	2.79	365.33	1867.59	9.03
Min.	33.96	6.69	1.23	352.39	825.89	6.55
Max	36.35	6.87	4.80	385.25	3218.46	10.68

Table 2. Statistical summary of chemical parameters with maximum allowable concentration from WHO (2004) and MIME (2004), analyzed 10 samples of wastewater. Avg = average; Min = minimum; Max = maximum, ND = non detect below 0.01 mg per L

Parameter	Zn	Cd	Fe	Cu	Al	Pb	Ni	Cr	Co
					mg/L				
Ave.	0.61	ND	262.70	0.54	11.94	ND	0.03	0.00	0.02
Min	0.14	ND	2.6	0.05	0.03	ND	0.03	0.01	0.01
Max.	2.44	ND	1335.24	1.86	62.54	ND	0.13	0.02	0.08
MAC[a]	3	0.003	0.3	1	0.2	0.01	0.02	0.05	-
MAC[b]	0.5-3	0.003	0.3-1	2	0.05-0.2	0.01	0.07	0.05	-

Note: MAC[a] = maximum allowable concentration from WHO, 2004; MAC[b] = maximum allowable concentration from MIME, 2004.

Table 3. Statistical summary of chemical parameters with maximum permissible limit from WHO (1996), analyzed 14 samples of sediment. Avg = average; Min = minimum; Max = maximum, MPL = Maximum permissible limit; a: set by World Health Organization (WHO)

Sediment samples	Parameters, mg/Kg							
	Al	Fe	Cu	As	Mn	Pb	Ni	Zn
Max.	108481.96	145885.01	1088.03	85994.06	3074.61	200.46	57.37	138.18
Ave.	75347.70	79459.64	474.57	18071.05	473.14	83.69	39.35	44.56
Min.	37471.96	44161.31	178.94	3003.81	54.99	36.98	24.36	15.26
MPL	-	-	36a	20a	-	85a	35a	50a

All sampling point, concentration of arsenic (As) was higher above the maximum permissible limit WHO (1996) standard. The high concentration of arsenic (As) is most likely mining-related, it originates from oxidation of arsenic-bearing sulfide minerals in mine wastes, tailing or exposed rock surface. The concentration values obtain for Ni at artisanal mining, revealed a range of 24.36 mg/Kg to 57.37 mg/Kg. These values didn't exceed for the WHO acceptable limit in soil. Highest value of Ni is in CP-S8 and CP-S10 sampling points. Iron (Fe) was observed to be the most in CP-S8 and CP-S10 sampling point with highest value of 145885.01 mg/Kg and 124703.47 mg/Kg. The values observed for Fe at mine site were within the normal contaminated soil (16443.81 mg/kg) recommended by Cambodia [10]. Lead (Pb) was investigated in sediment sample with value range 36.98 mg/kg to 200.46 mg/kg. Highest value of Pb (200.46 mg/kg) in CP-S10 sampling point that it is above the maximum permissible limit set by WHO (1996) standard. The concentration value of Aluminum and Manganese don't exceed the permissible limit by WHO (1996). The high levels of sediment accumulation show that contamination was occurring for a long time and pose risks of metal remobilization. The wet season produced AMD effects due to the water flow risk, which promoted sulfide oxidation and downstream transport of dissolved metals.

4. CONCLUSIONS

The AMD formation waters have high EC, Eh, TDS, near-neutral pH and moderate DO, are indicated the toxicity to aquatic life and unsustainability of drinking water or irrigation. The values of EC and TDS in particular significantly exceed the WHO and MIME guideline limit, indicating a high load of dissolved ions and contaminants. However, rainfall diluted AMD-contaminated water, reducing concentrations of metals. The increases in Fe, Al, and Ni imply that water is rapidly mobilizing metals into the environment through active interaction with pyritic or sulfide-rich materials. The potential dilution, rainy season runoff, is mobilizing a large number of metals, particularly at CP-W4 to CP-W8 sampling points. The result of sediment in this study (CP-S6, CP-S8 and CP-S10) are heavily contaminated with toxic heavy metals, specifically As, Fe, and Cu that this contamination has serious implication for soil quality. Thus, the results provide strong evidence that the water and sediment in this study are heavily impacted by AMD and that the source of arsenic and other heavy metals is predominantly mining-related rather than geogenic. The concentration of heavy metals clearly indicates that the study area is contaminated by acid mine drainage, primarily due to dissolved metals and sulfate ions from sulfide-bearing oxidation and mining activities. Future work will be collected more water and sediment samples and applied active or passive AMD treatment system to remove dissolved metals from artisanal mine and neutralize acidity.

ACKNOWLEDGMENTS

I would like to pay my highest respect to the Ministry of Mines and Energy for funding to lead this project and the Faculty of Geo-resources and Geotechnical Engineering for providing the laboratory to conduct this research. Additionally, I would like to express all my lectures, friends, and juniors for always helping, guiding, and encouraging me.

REFERENCES

[1] O. C. Modoi, C. Roba, Z. Török, and A. Ozunu, "Environmental risks due to heavy metal pollution of water resulted from mining wastes in NW Romania," *Environ. Eng. Manag. J.*, vol. 13, no. 9, pp. 2325–2336, 2014, doi: 10.30638/eemj.2014.260.

[2] S. Heng *et al.*, "Low-Cost Adsorbent in Treatment of Acid Mine Drainage in Cambodia : Chong," vol. 11, no. 1, pp. 37–45, 2023.

[3] Y. Jiao, C. Zhang, P. Su, Y. Tang, Z. Huang, and T. Ma, "A review of acid mine drainage: Formation mechanism, treatment technology, typical engineering cases and resource utilization," *Process Saf. Environ. Prot.*, vol. 170, no. November 2022, pp. 1240–1260, 2023, doi: 10.1016/j.psep.2022.12.083.

[4] J. Jurković, T. Muhić-Šarac, and M. Kolar, "Chemical characterisation of acid mine drainage from an abandoned gold mine site," *Chem. List.*, vol. 108, no. December, pp. s165–s170, 2014.

[5] T. Genty and M. Benzaazoua, "Neutralization of acid mine drainage in anoxic limestone drains : a laboratory study," no. June 2014.

[6] WHO, "Guidelines for Drinking-water Quality SECOND ADDENDUM TO THIRD EDITION WHO Library Cataloguing-in-Publication Data," *World Heal. Organ.*, vol. 1, pp. 1–103, 2008, [Online]. Available: http://www.who.int/water_sanitation_health/dwq/secondaddendum20081119.pdf

[7] MIME, "Kingdom of Cambodia Drinking Water Quality Standards," *Drink. Water Qual. Stand.*, pp. 1–20, 2004.

[8] WHO, "Guidelines for drinking water quality, vol 2, 2nd ed.," *World Heal. Organiztion, Geneva*, vol. 2, pp. 132–388, 1996.

[9] J. E. Gray and R. G. Eppinger, "Distribution of Cu, Co, As, and Fe in mine waste, sediment, soil, and water in and around mineral deposits and mines of the Idaho Cobalt Belt, USA," *Appl. Geochemistry*, vol. 27, no. 6, pp. 1053–1062, 2012, doi: 10.1016/j.apgeochem.2012.02.001.

[10] K. Phan, "Assessing heavy metal contamination in urban surface soil of Phnom Penh," *Bulle tin Cambod dian Chem. al Soc.*, vol. 8, no. 1–2, pp. 31–35, 2018.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Impact of City Activities on Macroplastic Accumulation in Two Tonle Sap Lake Tributaries

CHHENG Por Y [1], THENG Vouchlay [1,2*], CHHOUN Kong[1], KET Pinnara[1], SANG Davin[1], HEU Rina[1], MENH Leang Gek[1], SIM Solida[1]

[1] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[2] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[*]*Corresponding author: vouchlay.theng@itc.edu.kh*

Abstract: *Macroplastic pollution in freshwater systems poses growing environmental challenges, contributing to habitat degradation and affecting aquatic organisms. This study aimed to compare the macroplastic accumulation in two Tonle Sap Lake tributaries, including the Sangker River, an agricultural area, and the Siem Reap River, a tourism area. Monitoring was conducted at three locations in each river, upstream, midstream, and downstream of the city, to assess the impact of the city on macroplastic pollution. The sampling was conducted in December 2024 and February 2025 using three methods: visual observation, net sampling, and riverbank surveys. The results showed that net sampling in both rivers did not capture any floating macroplastics, and visual observations confirmed that macroplastics were rarely present on the water surface. In Sangker River, midstream locations showed the highest macroplastic accumulation in December 2024, recording 12.78g/m^2, followed by downstream (9.67g/m^2) and upstream (3.75g/m^2), indicating the impact of the city on macroplastic pollution. The highest fraction of macroplastic is Cup (1.41g/m^2), plastic bage (1.24g/m^2), and Plastic bottle (0.94g/m^2). During February 2025, the highest macroplastic was also midstream, 13.56g/m^2, followed by downstream (4.57g/m^2) and upstream (11.3g/m^2). In contrast, the Siem Reap River exhibited lower macroplastic in both December and February, indicating the proper macroplastic management in Siem Reap city. During December, upstream locations showed the highest accumulation, recording 6.63g/m^2 or 95, followed by midstream (1.92g/m^2) and downstream (4.3g/m2). During February, the highest macroplastic was downstream 6.85g/m^2, followed by midstream (0.45g/m^2) and upstream (0.71g/m^2), as the water flow during this period was lower than that in December, which potentially contributed to greater deposition along riverbanks rather than transporting into Tonle Sap Lake. This study provides critical baseline data for freshwater macroplastic pollution in Cambodia and supports the need for targeted riverbank clean-up efforts, community education, and effective waste management policies to mitigate plastic leakage into freshwater systems.*

Keywords: Macroplastic Pollution, River Monitoring, Visual Sampling, Net Sampling, Riverbank Survey, Tonle Sap Lake

1. INTRODUCTION

Plastic pollution, particularly the accumulation of macroplastics (plastic debris >5 mm), has emerged as a major threat to freshwater ecosystems globally, impacting water quality, habitat structure, and biodiversity [1]. Macroplastics are not only highly visible but also dangerous, as they can physically obstruct waterways, cause entanglement or ingestion in aquatic organisms, and serve as long-term sources of microplastics through degradation processes [2]. Their presence disrupts ecosystem functions and services and can adversely affect fisheries, agriculture, and public health, leading to substantial socio-economic consequences for communities dependent on freshwater resources [3].

Urbanization and agricultural activities are key drivers of macroplastic pollution in river systems. Urban areas contribute directly through mismanaged solid waste, runoff from paved surfaces, and littering, while agricultural regions add plastics indirectly via plasticulture residues, irrigation systems, and inadequate rural waste management [4]. Furthermore, river hydrology specifically flow velocity, discharge variability, and seasonal flooding strongly

[1]* Corresponding author: THENG Vouchlay
E-mail: vouchlay.theng@itc.edu.kh*; Tel: +855-81 352 727*

influences macroplastic transport, deposition, and resuspension [5]. High-flow conditions can rapidly redistribute plastic debris downstream, while low-flow or stagnant conditions promote plastic accumulation along banks and floodplains [6].

The Tonle Sap Lake and its tributaries in Cambodia are increasingly recognized as critical zones for freshwater plastic contamination. These rivers, including the Sangker and Siem Reap Rivers, are subject to plastic inputs from urban expansion, tourism, and intensifying agriculture [7]. Despite growing awareness, systematic studies on macroplastic transport, accumulation, and spatial distribution patterns in Tonle Sap's tributaries remain very limited. Existing research often focuses on microplastics or marine environments, leaving a gap in understanding macroplastic dynamics in Cambodian inland waters [8].

Therefore, this study aims to compare the macroplastic accumulation in two key tributaries of the Tonle Sap Lake: Sangker River, which drains a predominantly agricultural landscape, and Siem Reap River, which flows through a major urban and tourism hub. Seasonal surveys were conducted at upstream, midstream, and downstream locations using visual observation, net sampling, and riverbank survey methods. The study seeks to (i) quantify macroplastic densities, (ii) identify major plastic types, and (iii) compare pollution patterns.

2. METHODOLOGY

2.1 Study Area

This study was conducted in two major tributaries of Tonle Sap Lake: Sangker River in Battambang Province and Siem Reap River in Siem Reap Province, both of which pass through provincial capitals before discharging into the lake. The Sangker River flows through Battambang city, home to approximately 200,000 residents, with a surrounding landscape characterized by extensive rice farming, horticulture, and agro-industrial activities. The region's economy is largely agriculture-based, contributing significantly to Cambodia's rice production. In contrast, the Siem Reap River passes through Siem Reap city, which, despite having a smaller population of around 140,000, attracts over 2 million international tourists annually due to its proximity to the Angkor Wat UNESCO World Heritage site. The local economy is heavily reliant on tourism, hospitality, and urban commerce, resulting in different patterns of waste generation and land-use pressure. These contrasting socio-economic and land-use profiles provide a valuable framework for examining how urbanization, tourism, and agriculture influence macroplastic pollution in riverine systems feeding into Tonle Sap Lake.

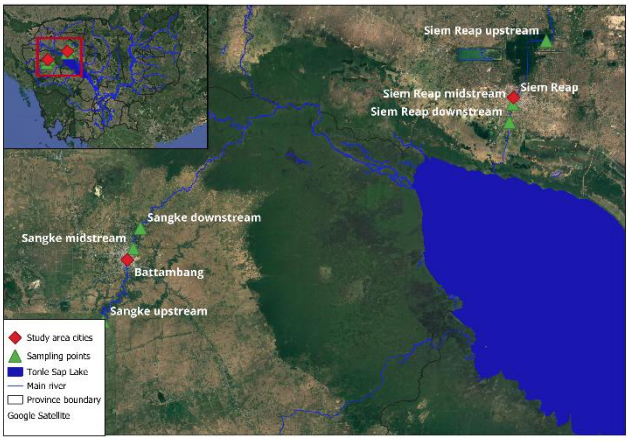


Fig. 1. Map of the study area showing two tributaries of the Tonle Sap Lake, Sangker River (Battambang Province) and Siem Reap River (Siem Reap Province).

Sampling was carried out at three locations per river: upstream, midstream (within the city), and downstream. These locations were selected to evaluate the impact of urban areas on macroplastic accumulation. Monitoring was conducted during two periods: December 2024 (dry season) and February 2025 (dry season). River discharge data was obtained from the Ministry of Water Resources and Meteorology (MOWRAM), Cambodia. The discharge records for the Sangker and Siem Reap rivers were collected for the year 2018 from official hydrological monitoring stations operated by MOWRAM.

2.2 Data Collection Methods

Monitoring was conducted during two periods: December 2024 (dry season) and February 2025 (early flood recession), to capture seasonal variation in microplastic distribution. All samples were collected from 10am to 12pm. This study applied a multi-method approach to assess macroplastic pollution in two tributaries of Tonle Sap Lake Sangker River (Battambang) and Siem Reap River (Siem Reap) using three complementary techniques: Visual observation, Net sampling, and Riverbank survey. Each method was applied at upstream, midstream, and downstream locations in both rivers during two seasonal periods.

Visual Observation: Visual observation surveys were conducted from accessible vantage points such as riverbanks and bridges, following standardized protocols for surface macroplastic monitoring in rivers. At each location, trained observers visually scanned the entire water surface width in front of the station for a fixed period of 15 minutes, as recommended by previous studies on rapid assessment of floating debris [5]. All visible macroplastic items (defined as >5 mm in size) were recorded, with observers noting item type, estimated quantity, and flow direction during the observation period. To reduce observational bias, surveys were performed during daylight under consistent weather

conditions, and where possible, by the same observers across locations and survey dates. Results were reported as total number of items per 15-minute interval, which allows for comparison of floating macroplastic density across sites and times, and aligns with river monitoring practices in other regional studies [9].

Net Sampling: To assess the potential presence of floating and near-surface macroplastics, manta-style net sampling was performed using a rectangular net with a 50 × 50 cm mouth opening and 500 µm mesh size, adapted from microplastic river monitoring protocols [10]. Nets were deployed manually from bridges or riverbanks and towed for 15 minutes in triple redundants, either by anchoring in the current or pulling alongside the river edge, depending on local flow conditions. Collected contents were rinsed into labeled containers and visually inspected on-site for macroplastics (>5 mm). The total mass of macroplastics (g) per tow was recorded. Due to mesh size limitations and river depth, this method primarily targeted items suspended just below the water surface.

Riverbank Survey: Macroplastic accumulation along riverbanks was assessed using a standardized 100 m² quadrat (typically 10 m × 10 m), placed parallel to the shoreline at upstream, midstream, and downstream locations [16,17]. All visible macroplastic debris within each quadrat was collected manually, classified by item type, and weighed using a portable digital scale to determine the total dry mass per site. The results were expressed as macroplastic density (g/m²) by dividing the total weight by the surveyed area. Collected macroplastics were categorized into clearly defined types based on prior studies and field relevance, including: plastic bags (empty and filled), plastic bottles, cups and cup lids, food packaging, sachets, carton/tetrapaks, cutlery, straws, bottle caps/lids, foam fragments, and miscellaneous hard plastics [17]. This method enabled both quantitative analysis and material classification to assess sources and potential land-use drivers of macroplastic pollution.

The average macroplastic mass per square meter was calculated as:

$$M = \frac{\sum_{i=1}^{n} mi}{A}$$

Where M is the macroplastic density (g/m²), m_i is the mass of item i, and A is the total surveyed area (m^2).

3. RESULTS AND DISCUSSION

3.*1 Impacts of city on Macroplastic Accumulation in Siem Reap River*

The results of macroplastic pollution along the Siem Reap River revealed spatial differences in macroplastic presence and variation through the impact of urban and hydrological influences on plastic accumulation (Table 1). On average, macroplastics at riverbank and in the river in December were higher than those in February.

Visual observation detected a clear longitudinal pattern in floating macroplastic density (Table 1). Upstream of the city, no floating macroplastics were observed during either survey period. In contrast, midstream (within the urban center), plastic items averaged 7.5 ± 9.2 items per 15-minute transect, while downstream sites recorded 1.7 ± 2.6 items. This pattern reflects direct inputs from Siem Reap city, where daily human activities, tourism, and littering contribute to midstream surface plastic loads. The decline downstream might contribute to the retaining of microplastic along the river bank, while others are transported further toward Tonle Sap Lake. Notably, the absence of visible plastics upstream confirms minimal local inputs and serves as a control against them. However, visual observation captures only floating items and cannot detect submerged or bank-stranded plastics.

Net sampling using a 50 × 50 cm manta net yielded 0 g of macroplastics at all reaches and on both dates. While this might suggest an absence of macroplastics in the water column as it flowed into Tonle Sap Lake during peak flow.

On average, riverbank surveys provided a cumulative measure of macroplastic accumulation over time. The highest deposition occurred downstream of the city (255 ± 265 g per transect; 4.6 ± 5.5 g m^{-2}), followed by upstream (141 ± 210 g; 3.6 ± 5.6 g m^{-2}), and the lowest was found midstream (22.5 ± 3.5 g; 0.45 ± 0.07 g m^{-2}). These results may seem counterintuitive, given the high midstream floating counts, but can be explained by two key factors: (1) Midstream areas are often actively cleaned (due to tourism and city management), reducing long-term accumulation despite high inputs. (2) Downstream zones act as retention areas, where reduced flow velocity and riverbank geometry favor plastic deposition. Moreover, riverbank surveys allowed classification of macroplastic types. The most common were single-use consumer items: bottle/container lids (21.5%), empty plastic bags (14.6%), food wrappers (9.8%), cups (6.6%), and plastic bottles (5.6%) (Figure 2). These reflect the dominant waste types generated in urban environments and confirm the link between city activities and plastic pollution.

Location	Siem Reap River		Sangker River	
	December	February	December	February
Discharge (m^3/s)	1.35	1.01	4.51	6.42
Visual Observation (items/15 min)				
Upstream	0	0	6.67	0.33
Midstream	8.33	6.67	1	0
Downstream	3.33	0	0.33	4.67
Net Sampling (g)				
Upstream	0	0	0	0
Midstream	0	0	0	0
Downstream	0	0	0	0
Riverbank Surveys (g/m^2)				
Upstream	6.63	0.53	3.75	11.3
Midstream	1.92	0.45	12.78	13.56
Downstream	4.3	6.85	10.11	3.51

Table 1. Presents the macroplastic density observed in the Siem Reap River and the Sangker River (Battambang) across three distinct methods: Visual observation, Net sampling, and Riverbank survey, and categorized by river sections (upstream, midstream, and downstream).

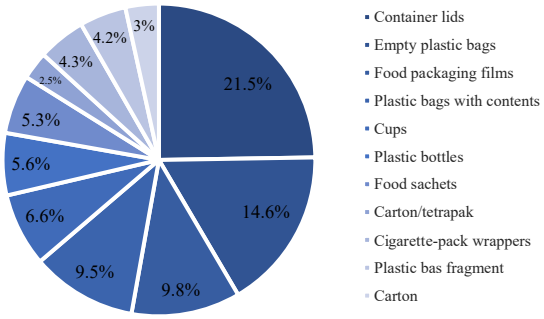


Fig. 2. Macroplastic composition in Siem Reap River

Taken together, these three methods reveal that Siem Reap city is the dominant source of macroplastic inputs to the river. Floating plastics are the most visible midstream, but due to city cleaning efforts at the land and hydrological transport, long-term accumulation is highest downstream.

3.2 Impact of city on Macroplastic Accumulation Sangker Rivers

Visual surveys at Sangker River revealed that on average, floating microplastic in December was higher than that in February. The highest number of floating plastic items was at upstream (3.5 ± 4.5 items per 15-minute transect), followed by downstream (2.5 ± 3.6), and the midstream (0.5 ± 0.8) on average (Table 1). This distribution suggests that rural and peri-urban areas contribute more to visible surface litter, possibly due to uncollected domestic waste and informal dumping near farms and villages. In contrast, floating plastic presence was lower midstream, where the river passes through Battambang city. This might be caused by the low floating aquatic plants in the river which is the obstacle of the plastic transport.

Net sampling using a 50 × 50 cm manta net recorded 0 g of macroplastics across all locations as in Siem Reap River.

The results of the riverbank survey revealed that on average, the riverbank microplastic was similar between December and February. The highest loads of macroplastic accumulation were observed at midstream sites within the urban core of Battambang (678 ± 716 g per transect; 13.56 ± 14.32 g m^{-2}), followed by upstream areas (565 ± 470 g per transect; 11.30 ± 9.40 g m^{-2}), and downstream banks (228 ± 153 g per transect; 4.57 ± 3.06 g m^{-2}). These results indicate that while floating macroplastics were infrequent, accumulation along the riverbanks was highest in urban zones, likely due to continuous inputs from nearby streets, markets, and residential areas. The lower macroplastic densities observed downstream may reflect a combination of hydrological transport, plastic retention, or redistribution into adjacent floodplains. In terms of composition, the survey showed that more than two-thirds of the total macroplastic mass was comprised of four dominant categories: empty plastic bags (23.24%), cups (12.88%), plastic bottles (10.20%), and plastic bags with contents (8.32%) (Figure 3). Additional notable contributors included cigarette box wrappers (6.48%), food packaging (6.20%), and plastic bag fragments (5.43%), with the remaining mass spread across carton/tetrapaks, cup lids, bottle caps, and miscellaneous rigid plastics. These results indicate a strong presence of single-use consumer items, consistent with both urban household waste and agricultural plastic leakage. This composition underscores the need to prioritize plastic bag bans, drink container deposit systems, and waste management upgrades in both rural and urban areas to mitigate further accumulation.

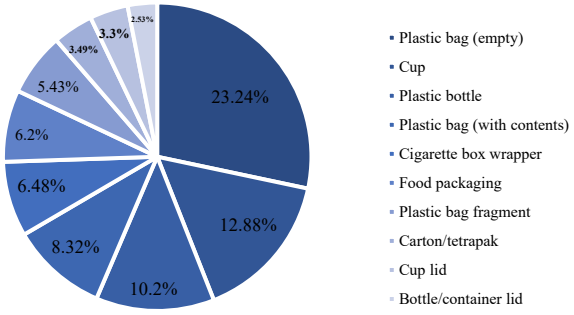


Fig. 3. Macroplastic composition in Sangker River

3.3 Comparison of Macroplastic Accumulation in the Sangker and Siem Reap Rivers

In December 2024, midstream macroplastic density in the Sangker River reached 12.78 g/m², considerably higher than the 1.92 g/m² observed in the Siem Reap River. Upstream and downstream densities followed similar patterns, with Sangker reporting 3.75 and 10.11 g/m², compared to Siem Reap's 6.63 and 4.3 g/m².

In February 2025, midstream macroplastic density in Sangker reached 13.56 g/m², compared to just 0.45 g/m² in Siem Reap. Similarly, upstream and downstream densities were also elevated in Sangker (11.3 and 4.57 g/m²) versus Siem Reap (0.71 and 6.85 g/m², respectively). These differences might reflect contrasting urban dynamics: Battambang, a city of over 200,000 people, lacks comprehensive waste collection, especially in peri-urban zones [11], while Siem Reap though smaller (~140,000 residents) receives over 2 million tourists annually [12] and has implemented stricter waste management to protect its tourism economy [13]. In addition, hydrological factors might also contribute: the Sangker River's low-gradient and wide floodplain promotes macroplastic retention in riparian zones [14], whereas the steeper, faster-flowing Siem Reap River more readily transports plastics downstream or into Tonle Sap Lake, especially during seasonal flow peaks [5].

The dominant macroplastic types in both rivers were single-use consumer plastics, including cups (1.41 g/m²), bags (1.24 g/m²), and bottles (0.94 g/m²), consistent with other studies in Southeast Asia [15]. These findings highlight the need for localized interventions targeted waste infrastructure upgrades in Battambang and continued enforcement in Siem Reap to reduce macroplastic leakage into freshwater systems.

4. CONCLUSIONS

This study provides a comprehensive assessment of macroplastic pollution in two Tonle Sap Lake tributaries the Sangker River (Battambang) and the Siem Reap River (Siem Reap) based on Visual observation, Net sampling, and Riverbank survey methods conducted during December 2024 and February 2025.

The results showed that Sangker River exhibited consistently higher macroplastic densities across all reaches, particularly midstream, where riverbank accumulation peaked at 13.56 g/m². In contrast, the Siem Reap River, despite its urban setting, showed much lower values, with midstream densities as low as 0.45 g/m², likely due to stronger waste management and greater hydrological flushing capacity. The dominant macroplastic items in both rivers were single-use plastics, especially plastic bags (up to 22.3% of total mass), cups (12.5%), and plastic bottles (11.1%), indicating widespread leakage from everyday consumption. Visual observation confirmed midstream floating debris peaks in both rivers 7.5 items per 15-minute count in Siem Reap and 3.5 items in Sangker River.

These findings offered the first detailed baseline data on macroplastic pollution in Tonle Sap Basin and emphasized the need for localized interventions, including: improved urban waste collection, riverbank cleanup campaigns, seasonal monitoring, and public awareness programs. Targeted efforts in high-risk areas such as midstream Sangker and downstream Siem Reap are essential to reduce macroplastic leakage into Tonle Sap Lake and to support national and international water quality goals, including SDG 6 (Clean Water and Sanitation) and SDG 14 (Life Below Water).

ACKNOWLEDGMENTS

This research was supported by the ASEAN-Norwegian Cooperation Project on Local Capacity Building for Reducing Plastic Pollution in the ASEAN Region (ASEANO Phase 2).

REFERENCES

[1] Wagner, M., & Lambert, S. (2018). Freshwater Microplastics: Emerging Environmental Contaminants? Springer, Cham.

[2] Geyer, R., Jambeck, J. R., & Law, K. L. (2017). Production, use, and fate of all plastics ever made. Science Advances, 3(7), e1700782.

[3] Windsor, F. M., Durance, I., Horton, A. A., Thompson, R. C., & Ormerod, S. J. (2019). A Global Change Biology, 25(4), 1207–1221.

[4] Dris, R., Gasperi, J., Saad, M., Mirande, C., & Tassin, B. (2016). Synthetic fibers in atmospheric fallout: A source of microplastics in the environment? Marine Pollution Bulletin, 104(1-2), 290–293.

[5] van Emmerik, T., & Schwarz, A. (2020). Plastic debris in rivers. Wiley Interdisciplinary Reviews: Water, 7(1), e1398.

[6] Lechthaler, S., & Waldschläger, K. (2020). Modelling the transport of macroplastics in rivers: Sources, pathways, and sinks. Science of the Total Environment, 746, 141097.

[7] Tuvikene, A., & Tuvikene, L. (2020). Plastic debris in aquatic ecosystems of Southeast Asia: Sources and solutions. Environmental Science and Pollution Research, 27(29), 36513–36527.

[8] Meijer, L. J. J., van Emmerik, T., van der Ent, R., Schmidt, C., & Lebreton, L. (2021). More than 1000 rivers account for 80% of global riverine plastic emissions into the ocean. Science Advances, 7(18), eaaz5803.

[9] Calcar, C. J., & van Emmerik, T. (2019). Abundance of plastic debris across European and

Asian rivers. Environmental Research Letters, 14(12), 124051.

[10] Lusher, A. L., et al. (2017). Sampling, isolating and identifying microplastics ingested by fish and invertebrates. In Microplastic Contamination in Aquatic Environments (Elsevier).

[11] GIZ Cambodia & MoE (2022). Waste Management Baseline Survey for Battambang Province.

[12] Ministry of Tourism Cambodia (2023). Tourism Statistics Report 2022–2023.

[13] UNDP Cambodia (2021). Sustainable Tourism and Waste Reduction in Siem Reap.

[14] Lechthaler & Waldschläger (2020). Modelling macroplastic transport in lowland rivers. Sci. Total Environ., 746, 141097.

[15] Tuvikene, A. et al. (2020). Plastic pollution in Southeast Asian rivers. Environ. Sci. Pollut. Res., 27(29), 36513–36527.

[16] UNEP (United Nations Environment Programme). (2009). *Guidelines on Survey and Monitoring of Marine Litter*. Nairobi, Kenya: UNEP.

[17] van Emmerik, T., van Calcar, C., van der Hal, N., Kikker, R., & Loozen, M. (2019). Riverbank macroplastic monitoring using citizen science—application and lessons learned from the Rhine and Maas rivers. *Environmental Monitoring and Assessment, 191*(5).

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Monitoring and Evaluation of Macroplastic Pollution in Siem Reap River, Siem Reap Province, Cambodia

Chanmonida Heng [1]*, Rina Heu [1,2], Leanggek Menh [1], Solita Chan [1], Solida Sim[1], Sreynim Mol [1], Sarapich Vatey Mao [1], Vouchlay Theng [1], Pinnara Ket [1], Davin Sang [1], Kong Chhuon [1]

[1] *Faculty of Hydrology and Water Resource Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: chanmonida1@gmail.com*

Abstract: *Plastic pollution in freshwater ecosystems has emerged as a significant environmental challenge, with macroplastic waste adversely impacting water quality, aquatic biodiversity, and overall ecosystem health—particularly in rapidly developing regions such as Siem Reap, Cambodia. This study seeks to quantify the flux and accumulation of macroplastic in various sections of the Siem Reap River and to evaluate seasonal variations in macroplastic density during both the rainy and dry seasons. Field surveys were conducted at three crucial sites—upstream, midstream, and downstream—where macroplastic samples were collected using visual observation, riverbank survey and net sampling methods Plastic litter density (PLD) peaked at 76.9 g/m^2 at downstream Kampong Kleang in December, while the highest item count reached 428 items at the same location in February. The collected macroplastics were categorized and analyzed to assess the extent of pollution. To further evaluate macroplastic contamination, the results indicate significant seasonal differences, with increased macroplastic flux during the rainy season attributed to heightened surface runoff and hydrological transport. Furthermore, the urban sections of the river demonstrated a higher accumulation of macroplastics compared to upstream areas, underscoring the impact of human activities on pollution levels. This research offers valuable baseline data for understanding the dynamics of macroplastic pollution in the Siem Reap River. The findings advocate for enhanced waste management policies and effective environmental conservation strategies. Future studies should utilize GIS-based mapping for spatial analysis and investigate viable mitigation measures, including policy interventions and community-focused initiatives.*

Keywords: Freshwater ecosystems; Macroplastic pollution; Seasonal variations; Siem Reap River; Waste management policies.

1. INTRODUCTION

1.1 Background

Plastic pollution is a significant environmental issue, especially in freshwater ecosystems. Macroplastics, defined as plastic items larger than 5 mm, pose a threat to rivers by harming aquatic life and degrading water quality. In Cambodia, rapid urban growth and poor waste management are worsening this problem, with rivers like the Siem Reap and Kampong Khleang transporting plastic waste to larger bodies of water such as Tonle Sap Lake. The Siem Reap River, a tourist area, experiences high levels of pollution from single-use plastics, while the Kampong Khleang River, surrounded by floating villages, is affected by daily activities like fishing and waste disposal. Improper waste management and seasonal flooding contribute further to plastic pollution. Despite increasing awareness, research on macroplastics in Cambodian rivers is limited. This study will focus on macroplastic pollution in the Siem Reap and Kampong Khleang rivers, using field surveys and sampling to identify pollution hotspots and sources. The findings will guide policymakers, environmental organizations, and local communities in improving plastic waste management and protecting freshwater ecosystems in Cambodia.

1.2 Problem Statement

[1]* Corresponding author: Chanmonida Heng
E-mail: chanmonida1@gmail.com; Tel: +855-17 436 565

The increasing macroplastic pollution in Cambodia's freshwater ecosystems, particularly in the Siem Reap River and Kampong Kleang, threatens aquatic biodiversity, water quality, and local communities. These rivers transport substantial plastic waste into Tonle Sap Lake, but comprehensive data on the volume of macroplastics is lacking. Seasonal changes in water depth can influence the movement of this debris, potentially increasing pollution concentrations at certain times. Contributing factors include inadequate waste disposal, low public awareness, and continuous plastic introduction. This study aims to quantify macroplastic influx into Tonle Sap and explore how seasonal water depth variations affect plastic transport. The findings will inform policy and waste management strategies to reduce pollution and protect Cambodia's freshwater ecosystems.

1.3 Objectives

This research will evaluate the level, distribution, and movement of macroplastic pollution in the Siem Reap River, which includes the urban area of Siem Reap and the rural region of Kampong Kleang. The river is important for both the environment and the local economy because it feeds into Tonle Sap. We will use standardized monitoring methods to measure the amount of macroplastic present and how it accumulates. We will also estimate how much plastic is transported from the Siem Reap River into Tonle Sap Lake. To accomplish these objectives, the study will quantify the flux and accumulation of macroplastic across different sections of the Siem Reap River and Kampong Kleang River, comparing the levels of pollution in urban and rural areas. It will also evaluate the seasonal variations in macroplastic density during both the rainy and dry seasons.

2. METHODOLOGY

2.1 Study Area

This study examined macroplastic pollution in the Siem Reap River in Siem Reap Province, Cambodia, focusing on its role in transporting plastic waste to Tonle Sap Lake. Sampling took place at five sites: three in the central river (upstream, midstream, downstream) to assess urban and rural impacts, and two in Kampong Kleang (upstream and downstream) to analyze transport in a less urbanized area. The study sites were chosen for their environmental significance and proximity to human activities. The findings shed light on plastic pollution levels in the Siem Reap River and its effects on Tonle Sap Lake.

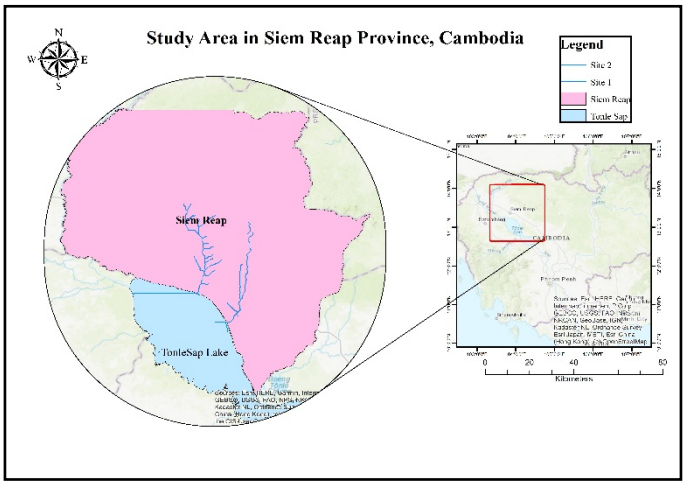


Fig 2.1.1 Map of Siem Reap Province and selected river

2.2 Data Collection Methods

This study adopted a multi-method approach to collect quantitative and qualitative data on macroplastic pollution in the Siem Reap River and Kampong Kleang River. Field data were gathered during three separate periods—December 2024, February 2025, and April 2025—to capture seasonal dynamics within the dry season and transitional rainfall period. Five sampling sites were selected based on their environmental relevance, accessibility, and proximity to human activity: three within the Siem Reap River (upstream, midstream, downstream) and two along the Kampong Kleang River (upstream and downstream).

To ensure comprehensive assessment, the study employed four main data collection methods, each tailored to capture different aspects of macroplastic pollution: visual observation, net sampling, riverbank survey, and plastic composition analysis. These methods allowed for cross-validation and a more complete understanding of macroplastic presence, transport, and accumulation.[3]

2.2.1 Visual Observation

Visual observation was used to record floating macroplastic items in the river channel. This method involved manually counting visible plastic items larger than 2.5 cm as they passed a defined cross-section of the river. Observations were conducted from a vantage point—such as a bridge or elevated riverbank—that allowed a clear view of the river's surface. Each observation session lasted 15 minutes, during which all floating plastic items within the observable section were counted.[3] Where possible, the river was divided into multiple segments (or "sweeps") to ensure full coverage. For example, if the full width of the river could not be seen from

a single vantage point, the observer rotated their focus across 2–3 segments, spending 5 minutes per segment. The plastic flux (items per second) was later calculated based on observed counts and timing. While this method offers rapid assessment and minimal equipment, it is subject to limitations such as observer bias, low-light conditions, and obstructed visibility due to floating vegetation or weather.[3]

2.2.2 Net Sampling

Net sampling provided quantitative data on the physical presence and mass of floating macroplastic. A fine mesh sampling net (approx. 5 mm mesh size) was placed across the surface of the river, perpendicular to the water flow. The net remained in position for 15 minutes per sampling session to allow for sufficient collection of debris.[3]

Once retrieved, all collected items were rinsed, separated from natural materials (like leaves or twigs), and air-dried. The plastic items were then counted to determine the total number of plastic pieces, weighed using a digital scale to assess total mass (grams), and sorted by category (e.g., plastic bag, bottle, styrofoam, etc.).

2.2.3 Riverbank Survey

Macroplastics were collected on the riverbank following the method of Rech *et al*.(2014)[4] with modifications. The surveys were conducted to assess the accumulation of macroplastic debris deposited along the banks of the river. At each site, a 50-meter-long transect was established parallel to the water's edge, with a width of approximately 2 meters inland from the riverbank. All visible macroplastic items within this defined area were collected manually using gloves and trash bags, counted and recorded by type, and photographed for documentation. This method is particularly useful during the dry season, when lower water levels expose more of the riverbank. It allowed the research team to identify accumulation hotspots, determine the impact of nearby land use (e.g., tourism, housing, fishing), and understand how river flow affects macroplastic deposition.[3]

2.2.4 Plastic Composition Analysis

To better understand the nature of macroplastic pollution, all collected items were categorized by type according to a standardized framework. The categories included bottles, such as water bottles and detergent containers; food wrappers, including items like chips, instant noodles, and candy; Styrofoam products, such as foam food boxes and cup holders; cigarette boxes, which encompassed both hard and soft plastic packaging; and a miscellaneous category for other plastic fragments and unidentified plastics.Composition analysis was done **monthly** for each site, allowing for the comparison of trends over time and between locations. The relative proportions of each category were calculated to create composition profiles, which were visualized using bar and pie charts. These profiles provided insights into the **dominant sources of pollution** and guided interpretation of land-based waste behaviors.

2.3 Data analysis

The density (PLD) and composition (PLC) of the collected macroplastics were determined using the computations based on the study of Abreo *et al.* (2019):

- The **Plastic Litter Density (PLD)** was calculated using the equation:

$$PLD = \frac{Number\ of\ items}{Total\ area\ sampled\ (L \times W)\ \ (m^2)} = Item/m^2 \qquad \text{(Eq. 1)}$$

The number of macroplastic items collected was divided by the total sampling area (50 m^2) to determine the density of plastic pollution in the study area.

- The **Plastic Litter Composition (PLC)** was determined using:

$$PLC = \frac{Number\ of\ items\ per\ category}{Total\ number\ of\ items\ in\ all\ category} \times 100 \qquad \text{(Eq. 2)}$$

This method categorized the collected macroplastics into different types (eg, plastic bags, bottles, foam containers) and calculated their percentage in relation to the total collected items.

3. RESULTS AND DISCUSSION

3.1 Macroplastic Litter Flux

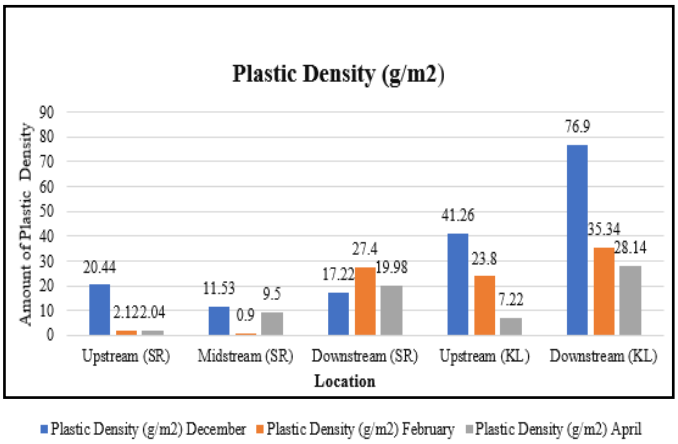


Fig. 3.1.1 Comparison of Plastic Density Across Sampling Sites and Seasons.

The plastic density observed at the five sampling sites across December, February, and April revealed both spatial and temporal variations: Downstream Kampong Kleang (KL) consistently recorded the highest plastic density, peaking at 76.9 g/m² in December, followed by 35.34 g/m² in February, and 28.14 g/m² in April. This indicates it is a major accumulation hotspot, likely influenced by local waste discharge and limited waste management practices. Upstream KL and Siem Reap River sites showed more moderate levels, with plastic density generally decreasing from December to April. This decline in plastic density over time can be partly attributed to seasonal hydrological changes. December and February represent the peak dry season, when river flow is lower and plastic tends to accumulate. In contrast, April marks the beginning of the rainy season, when increased flow may transport plastics downstream or disperse them further, reducing visible accumulation at sampling points.

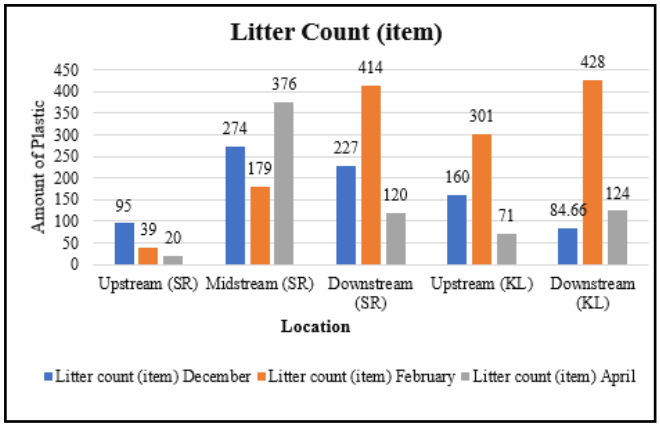


Fig. 3.1.2 Number of Macroplastic Items Collected at Sampling Sites (Dec-Apr)

The number of plastic litter items collected followed seasonal and spatial trends. February recorded the highest litter counts across several sites, with 414 items at Downstream SR and 428 at Downstream KL. In contrast, April counts were significantly lower at nearly all sites, except for Midstream SR, which had a relatively high count of 376 items despite lower plastic density. This indicates that while the total mass of plastic (g/m²) may decline, fragmented or lightweight plastics, such as bags and wrappers, can persist in large numbers during early rains. Although all data were collected before or at the beginning of the rainy season, a clear seasonal trend emerges, showing that both plastic density and item count were generally higher in the dry season, particularly in December and February, when plastic accumulates due to low flow and human activities. The notable reduction in April suggests that the onset of the rainy season begins to flush and disperse the litter accumulated plastics, especially at upstream and midstream sites.

This seasonal variation aligns with findings from regional studies, where macroplastic levels peak during dry periods and decline during wet seasons due to enhanced water flow and runoff dynamics. In summary, macroplastic pollution in both Siem Reap and Kampong Kleang rivers show clear temporal and spatial patterns, with downstream areas and dry months experiencing the highest levels of plastic accumulation.

These differences between item count and density indicate that plastic type and size influence seasonal behavior. For example, heavier or sedimented plastics may remain in place or become buried, while lightweight items are more easily transported with the onset of rain.

3.2 Macroplastic in Riverbank

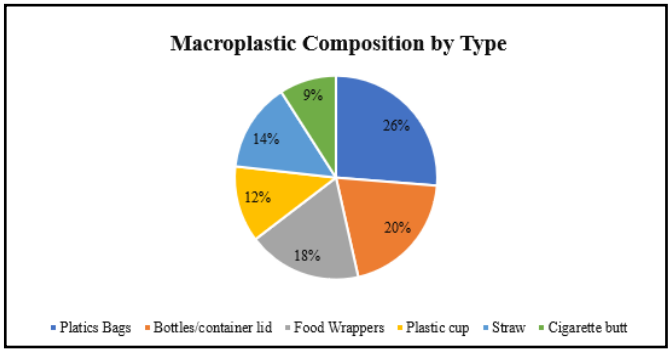


Fig. 3.2.1 Macroplastic Composition by Type Across Sampling Time

The results show that **plastic bags** were the most prevalent category, accounting for **26%** of the total collected items. This indicates the widespread use and improper disposal of single-use plastic bags, particularly in both urban and rural settings along the rivers. Their lightweight nature and high availability make them one of the most common plastic pollutants in freshwater environments. The second most common type was bottles and container lids, comprising 20% of the total. These are typically discarded beverage containers and household packaging, which are commonly used and often found in areas with high foot traffic, such as tourist zones and local markets. **Food wrappers** represented 18% of macroplastics. These items include flexible packaging used for snacks, instant noodles, and other processed foods. Their prevalence suggests that consumption patterns—particularly of convenience foods—are contributing significantly to plastic pollution. **Plastic cups** and straws contributed 12% and 14%, respectively. These items are often used in food stalls, restaurants, and floating villages, and their high occurrence reinforces the impact of local food-related waste on riverine environments. Finally, **cigarette butts**, although relatively smaller in size, made up 9% of the macroplastic waste. These are typically made of cellulose acetate, a form of plastic that does not biodegrade easily. Their presence indicates improper disposal of smoking-related waste, particularly along the riverbanks where people congregate.

3.3 Discussion

The study examined macroplastic pollution in the Siem Reap and Kampong Kleang rivers, primarily linked to human activities from nearby residential communities. The most common form of debris was household waste, especially single-use plastic food packaging. These findings align with Willis et al. (2017)[5], who stated that plastic litter is often sourced near its origin, especially in urban and residential areas. The midstream and downstream sections of the Siem Reap River had the highest amounts of macroplastic debris due to dense populations and local activities, while the upstream section had the least due to lower urbanization.This pattern is consistent with Rech et al. (2014)[4], who observed that plastic concentrations tend to decrease with increasing distance from river mouths and densely populated zones. Similarly, the Kampong Kleang River showed more debris downstream compared to upstream, influenced by upstream transport and factors like fishing activities and riverbank structures that disrupt flow, as noted in the study by van Emmerik et al (2019)[6] Hydrological and meteorological factors, such as rainfall and tidal influences, significantly affected macroplastic flux[2]. Overall, geographical location, residential density, river morphology, weather, and tidal effects play crucial roles in macroplastic distribution, highlighting the need for targeted waste management strategies and community interventions to reduce plastic pollution in freshwater systems.

4. CONCLUSIONS

This study investigated macroplastic pollution in the Siem Reap River and Kampong Kleang River, with a focus on understanding its spatial distribution, seasonal variation, and composition. The research was conducted through three field campaigns during the dry season and early transition to the rainy season (December 2024, February 2025, and April 2025) using a combination of visual observation, net sampling, riverbank surveys, and composition analysis. The results revealed that plastic density and litter count were highest in downstream areas, particularly at Kampong Kleang, indicating significant accumulation zones where macroplastics are trapped before entering Tonle Sap Lake. This pattern reflects both the influence of hydrological flow and the impact of human activity—especially in areas with limited waste management and high dependency on plastic-based materials. Seasonal analysis showed that plastic pollution was generally higher during the dry season, with the highest densities recorded in December. By April, the onset of the rainy season had begun to disperse or flush accumulated waste downstream. This trend highlights the importance of considering seasonal hydrodynamics when assessing riverine plastic pollution. In terms of plastic composition, the most common macroplastic items were plastic bags (26%), bottles and container lids (20%), and food wrappers (18%). These items were consistent across sites and sampling periods, emphasizing the dominance of single-use consumer plastics in the pollution profile. The presence of cigarette boxes, straws, and cups further supports the connection between waste generation and daily community behavior. Overall, this research provides important baseline data for the monitoring of macroplastic pollution in Cambodian freshwater ecosystems. It underscores the need for targeted interventions, including: Improved solid waste collection and infrastructure in riverside communities, public awareness campaigns to reduce single-use plastic consumption, and policy frameworks that address both urban and rural sources of plastic pollution.

.**ACKNOWLEDGMENTS**

This work was supported by ASEAN-Norwegain Cooperation Project on Local Capacity Buidling for Reducing Plastic Pollution in the ASEAN Region (ASEANO Phase 2).

REFERENCES

[1] T. Van Emmerik and A. Schwarz, “Plastic debris in rivers,” *WIREs Water*, vol. 7, no. 1, p. e1398, Jan. 2020, doi: 10.1002/wat2.1398.

[2] M. Liro, A. Zielonka, and P. Mikuś, “First attempt to measure macroplastic fragmentation in rivers,” *Environ. Int.*, vol. 191, p. 108935, Sep. 2024, doi: 10.1016/j.envint.2024.108935.

[3] R. Hurley *et al.*, “Measuring riverine macroplastic: Methods, harmonisation, and quality control,” *Water Res.*, vol. 235, p. 119902, May 2023, doi: 10.1016/j.watres.2023.119902.

[4] S. Rech, V. Macaya-Caquilpán, J. F. Pantoja, M. M. Rivadeneira, D. Jofre Madariaga, and M. Thiel, “Rivers as a source of marine litter – A study from the SE Pacific,” *Mar. Pollut. Bull.*, vol. 82, no. 1–2, pp. 66–75, May 2014, doi: 10.1016/j.marpolbul.2014.03.019.

[5] K. Willis, C. Maureaud, C. Wilcox, and B. D. Hardesty, “How successful are waste abatement campaigns and government policies at reducing plastic waste into the marine environment?,” *Mar. Policy*, vol. 96, pp. 243–249, Oct. 2018, doi: 10.1016/j.marpol.2017.11.037.

[6] T. Van Emmerik and A. Schwarz, “Plastic debris in rivers,” *WIREs Water*, vol. 7, no. 1, p. e1398, Jan. 2020, doi: 10.1002/wat2.1398.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Photoinduced Reactive Intermediates Formation from Microplastics and Leachate Dissolved Organic Matter

Vinhteang Kaing[1,2*], Chihiro Yoshimura[2]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Civil and Environmental Engineering, Institute of Science Tokyo, 2-12-1, Meguro-Ku, Tokyo, 152-8552, Japan*

[*] *Corresponding author: vinhteang.kaing@itc.edu.kh*

Abstract: *In water environment, microplastics (MPs) and their released dissolved organic matter (DOM) produce reactive intermediates (RIs) under solar irradiation. These RIs react with plastics and pollutants, leading to their degradation and subsequent environmental processes. Therefore, this study quantified RIs photoproduction from MPs and their leachate DOM and investigated the relationship between the MP type and the production of RIs. Four types of MPs, aliphatic polymers (polypropylene (PP), and polymethyl methacrylate (PMMA)), and aromatic polymers (polystyrene (PS), and polyethylene terephthalate (PET)), were prepared for a size range from 300-1000 µm and used for photochemical experiment. The photoproduction of excited triplet-state of DOM or polymer ($^3DOM^*$, $^3RH^*$), singlet oxygen ($^1O_2$), hydroxyl radical (•OH), and superoxide ion ($O_2^{\bullet-}$) was examined under simulated sunlight using chemical probes. The results showed the average quantum yield of RIs (Φ_{RIs}) from MP-DOM were 0.30×10^{-2} for $\Phi_{^3DOM^*}$, 0.29×10^{-2} for $\Phi_{^1O_2}$, 0.48×10^{-5} for $\Phi_{O_2^{\bullet-}}$, and 0.12×10^{-3} for $\Phi_{\bullet OH}$. Based on statistical analysis, Φ_{RIs} from MP-DOM among different RIs and MPs followed the order $\Phi_{^3DOM^*} \approx \Phi_{^1O_2} > \Phi_{\bullet OH} > \Phi_{O_2^{\bullet-}}$ and $\Phi_{PMMA} > \Phi_{PP} \approx \Phi_{PS} \approx \Phi_{PET}$, respectively. These results indicated that $^3DOM^*$, $^1O_2$, and •OH are the dominant species in RIs production from MP-DOM, while $O_2^{\bullet-}$ plays a minimal role. Our study showed that the RI production from the DOM of aliphatic polymers is comparable to that from aromatic polymers. The calculated value of $\Phi_{^3RH^*}$, $\Phi_{^1O_2}$, $\Phi_{O_2^{\bullet-}}$, and $\Phi_{\bullet OH}$ from the MP surface were $(3.47 \pm 2.72)\times10^{-4}$, $(23.39 \pm 17.57)\times10^{-4}$, $(0.16 \pm 0.04)\times10^{-4}$, and $(0.22 \pm 0.21)\times10^{-4}$, respectively. Statistical analysis suggested that the chemical structure of MPs did not significantly influence RI production, while $\Phi_{^1O_2}$ from MP surface showed the highest value among RIs. By comparing the RI production rates of leachate DOM and MP surfaces in MP suspension, MPs exhibited a higher production rate of all RI types, exceeding 70%. These results suggested that MPs play a significant role in the RIs production and related photochemical reactions in water environment. This study highlights the potential production of RIs from leachate DOM and MPs, which may enhance the photodegradation of plastics and other co-existing pollutants in aquatic environments, thereby contributing to a better understanding of fate and remediation of plastics in water environment.*

Keywords: Microplastics; DOM; Photoinduced Reactive Intermediates; Water Environment

1. INTRODUCTION

The environmental fate of microplastics (MPs) is currently a global concern. During the degradation in water environment, MPs release low molecular weight compounds as dissolved organic matter (DOM). Under sunlight, such DOM and MPs possibly produce reactive intermediates (RIs), including excited triplet-state of DOM (^{3}DOM*), excited triplet-state of MP (^{3}RH*), singlet oxygen ($^1O_2$), and hydroxyl radical (•OH), and superoxide ion ($O_2^{\bullet-}$). These RIs play a critically important role [1], [2]. This is because RIs promotes the degradation of plastics and subsequent environmental processes [3], [4]. For instance, generated RIs can react with plastics surface resulting in accelerates the aging and fragmentation of plastics [1]. Such RIs reacts with plastics and increases oxidative functional groups on the surface, which commonly increased its negative surface charge and stability in water further accelerating the aggregation of plastics [2], [5]. Additionally, they significantly decrease the persistence of coexisting contaminants (e.g., pesticides and pharmaceuticals) in water [6], [7], [8], while they also exhibit ecotoxicity to organisms, causing oxidative stress and cell damage [9], [10]. The photodegradation and RIs production of plastics in the aquatic environment are controlled by physicochemical properties of plastics (e.g. polymer structure, additives, physical properties, and crystallinity) and

[*] Corresponding author: Vinhteang Kaing
E-mail: vinhteang.kaing@itc.edu.kh

environmental conditions (e.g., water quality, water content, and biofilm) [4], [11]. Therefore, this study quantified RIs photoproduction from MPs and their leachate DOM and investigated the relationship between MP type and the production of RIs. The study specifically tested the hypothesis that the photo-reactivities of MP-derived DOM (MP-DOM) and MPs depends on the chemical structures of the MPs.

2. METHODOLOGY

1.1 Microplastics preparation

For production MPs, plastic pellets were ground in a mixer mill (MM301, Retsch, Japan) with two chrome steel grinding jars and balls. The ground plastic particles were sieved through meshes of 1000 µm and 300 µm to obtain MPs, representing the dominant microplastic size in rivers and lakes as 51% of microplastic particles ranges from 355 – 999 µm [12]. This sieving minimized the effect of particle size on the photoproduction of RIs in the following photochemical experiments. The obtained MPs were then placed into 50 mL centrifuge tubes, filled with Milli-Q water. The suspension was sonicated for 60 minutes and soaked in water for 24 hours to ensure thorough dispersion of the MP particles. Then MPs were sieved through a mesh of 300 µm again to ensure particles smaller than 300 µm being removed. Subsequently, MPs were further rinsed with ethanol and Milli-Q water three times respectively to eliminate the potential influence of residual substances from the synthesis process or storage medium (e.g., surfactants and amine catalyst). Then, the microplastic samples were freeze-dried for 48 hours and stored in brown vials in a desiccator. The light absorption spectra of MPs were determined using a UV–vis spectrophotometer (Shimadzu UV-2600, Japan) using diffuse reflectance spectroscopy technique with a wavelength range from 200 to 800 nm.

1.2 Photochemical experiment

Photochemical experiments were conducted to determine the ^{3}RH*, ^{3}DOM*, $^1O_2$, •OH, and $O_2^{\bullet-}$ produced from the MP suspension and leachate DOM using the chemical probe technique. All experiment was carried out with triplicate under simulated sunlight in the photoreactor (HELIOS.Xe, Koike Precision Instruments, Japan) equipped with a 300W Xenon short arc lamp (Ushio, Japan). The average spectral lamp irradiance was 1.71×10^{-5} Einstein L^{-1} s^{-1} (wavelength from 280 – 400 nm).

Photoproduction of RIs from MP suspension and leachate DOM was determined using chemical probes, including 2,4,6-trimethylphenol (TMP, 0.1 mM), furfuryl alcohol (FFA, 0.1 mM), nitrobenzene (NB, 0.5 mM), and 2,3-bis-(2-methoxy-4-nitro-5-sulfophenyl)-2h-tetrazolium-5-carboxanilide (XTT, 0.1 mM), respectively [13], [14]. For RIs production from MP suspension, we added 0.1 g of MPs and 10 mL of chemical probe solution (0.1 mM of TMP, 0.1 mM of FFA, 0.5 mM of NB, and 0.5 mM of XTT) to each quartz tube. At the intervals of 0, 3, 6, 9, 12, and 24 hours of irradiation, 0.55 mL of reaction solution was collected, filtered through a mixed cellulose ester filter (pore 0.45 µm), and injected into brown vial. The remaining concentration of TMP, FFA, and NB was analyzed by high performance liquid chromatography (HPLC; Prominence UFLC, Shimadzu, Japan) equipped with a Kinetex C18 100A reversed-phase column (5 µm, 4.6 × 250 mm; Phenomenex Co. USA). For photoproduction experiment of $O_2^{\bullet-}$, the concentration of XTT formazan was measured using a UV-Visible spectrophotometer at 470 nm. The measured pseudo-first-order reaction rate constants of TMP, FFA, and NB were used to calculate the steady-state concentration of ^{3}RH*, $^1O_2$, and •OH. For RIs production from DOM experiment, the isolated DOM solutions from 6-hour irradiated MP suspensions were subjected for this experiment, in which we obtained 0.15 mL aliquots of 5 mL reaction solution at 0, 3, 6, 9, and 12 hours. Then, probe concentrations were determined by HPLC to calculate their pseudo-first-order rate constants and steady-state concentrations.

The quantum yield (Φ) of RI photoproduction was then calculated as the RI production rate divided by adsorbed photon flux. The RI production rate of MP surface was obtained by subtracting the RI production rate of DOM from RI production rate of MP suspension. The adsorbed photons of MP were calculated from the light absorbance of MP and the absolute spectral irradiance of lamp at each tube position, which was determined through chemical actinometry experiment using p-nitroanisole/pyridine as the actinometer. The total irradiance used in the calculation of adsorbed photons was corrected to cover the wavelength range of 280–600 nm.

The light absorption spectra of DOC solution were obtained using a UV–vis spectrophotometer (Shimadzu UV-2600, Japan) for wavelength ranging from 200 to 800 nm. Additionally, fluorescence excitation-emission matrices (EEMs) of the leachate DOM solution were acquired by florescence spectrofluorometer (Shimadzu RF5300 PC, Japan) in a 1 cm quartz cuvette (excitation 230 – 455 nm with 5 nm interval, emission 290 – 702 nm with 4 nm interval). Based on the obtained florescence and light absorption spectra, we determined fluorescence index (FI), humification index (HIX), biological index (BIX), major peaks (*b*, *t*, *a*, *m*, and *c*), absorption coefficients at 254 and 300 nm (a_{254} and a_{300}), the specific UV absorbance at 254 nm and 280 nm ($SUVA_{254}$ and $SUVA_{280}$), the E2:E3, the spectral slope coefficients ($S_{275–295}$, $S_{350–400}$, and $S_{300–600}$), and the spectral slope ratio of $S_{275–295}$ to $S_{350–400}$ (S_R) using the staRdom package in R program ver. 4.3.2.

3. RESULTS AND DISCUSSION

3.1 *Photoproduction of RIs from leachate DOM*

The calculated steady-state concentration of RIs produced from DOM (cases without MPs in solution) ranged from $9.26 \times 10^{-17} - 8.56 \times 10^{-16}$ M for ^{3}DOM*, $0.00 - 1.39 \times 10^{-15}$ M for $^1O_2$, $2.84 \times 10^{-18} - 1.68 \times 10^{-17}$ M for $O_2^{\bullet-}$, and $1.27 \times 10^{-15} - 1.55 \times 10^{-15}$ M for •OH. There is no significant difference in the steady-state concentration of RIs between MP-DOM types (two-way ANOVA, $p = 0.66$); however, a significant difference was observed among RIs types (two-way ANOVA, $p < 0.001$). According to Duncan's multiple comparison test at significant level 0.05, the order of their steady-state concentrations was $[\bullet OH]_{ss} > [^1O_2]_{ss} > [^3DOM^*]_{ss} > [O_2^{\bullet-}]_{ss}$.

PET-DOM had the highest specific UV absorbance at 254 nm (SUVA254) value (1.00 L mgC-1 m-1) likely due to the release of aromatic compounds.

The calculated production rates of RIs from leachate DOM of the 6-hour irradiated MP suspension ranged from $4.63 \times 10^{-11} - 4.28 \times 10^{-10}$ M s^{-1} for ^{3}DOM*, $0.00 - 3.83 \times 10^{-10}$ M s^{-1} for $^1O_2$, $1.42 \times 10^{-12} - 8.40 \times 10^{-12}$ M s^{-1} for $O_2^{\bullet-}$, and $4.79 \times 10^{-11} - 5.79 \times 10^{-11}$ M s^{-1} for •OH. A significant difference in RIs production rate was observed among RIs types (two-way ANOVA, $p < 0.01$), but no significant difference was found between MP-DOMs (two-way ANOVA, $p > 0.05$). The production rates of $^1O_2$ ($(2.00 \pm 1.81) \times 10^{-10}$ M s^{-1}) were similar to those of ^{3}DOM* ($(1.86 \pm 1.68) \times 10^{-10}$ M s^{-1} and their production rates were higher than other RIs. The production of $O_2^{\bullet-}$ from MP-DOMs exhibited the lowest value at $(3.68 \pm 3.25) \times 10^{-12}$ M s^{-1}, while •OH production rate was $(4.84 \pm 0.51) \times 10^{-11}$ M s^{-1}.

Similar to the steady-state concentrations, the quantum yields of RIs varied across four MP-DOM types. PMMA-DOM exhibited the highest $\Phi_{^3DOM^*}$ (0.50×10^{-2}), followed by PS-DOM (0.23×10^{-2}), with an overall average of 0.26×10^{-2}. For $\Phi_{^1O2}$, PMMA-DOM showed the highest yield (0.53×10^{-2}), while PET-DOM produced none, resulting in an average yield of 0.29×10^{-2} across the MP-DOM types. The $\Phi_{\bullet OH}$ ranged from 0.08×10^{-2} (PET-DOM) to 0.18×10^{-2} (PMMA-DOM), with an average of 0.12×10^{-2}. The lowest values of RIs among RI types were observed for $\Phi_{O2\bullet-}$, ranging from 0.07×10^{-4} (PET-DOM) to 0.81×10^{-4} (PS-DOM), averaging 0.48×10^{-4}. These results highlight the variability in ROS production across MP-DOM types. Statistical analysis revealed a significant difference in Φ_{RIs} among RI types (two-way ANOVA, $p < 0.01$) and MP-DOMs (two-way ANOVA, $p < 0.05$) (Fig. 1). The Ducan's multiple comparison test (at significant level of $\alpha = 0.01$) further from confirmed differences Φ_{RIs} among RIs following the $\Phi_{^3DOM^*} \approx \Phi_{^1O2} > \Phi_{\bullet OH} > \Phi_{O2\bullet-}$, where $^1O_2$ and ^{3}DOM* showed no significant difference in quantum yield. Additionally, among MP-DOMs, the quantum yield followed the order $\Phi_{PMMA} > \Phi_{PP} \approx \Phi_{PS} > \Phi_{PET}$ (at $\alpha = 0.05$). These results indicated that ^{3}DOM*, $^1O_2$, and •OH are the dominant species in RIs production from MP-DOM, while $O_2^{\bullet-}$ plays a minimal role. Our study showed that the RI production from the DOM of aromatic polymers is not significantly higher than that of aliphatic polymers.

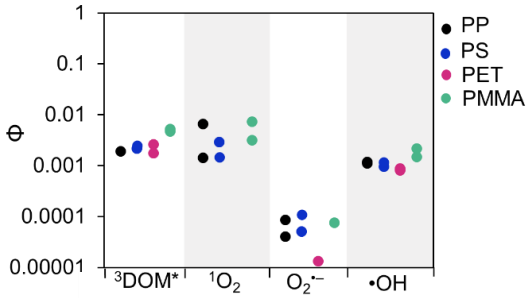


Fig. 1: Quantum yield of RIs production rate from leachate DOM ($n = 2$; missing values equal zero).

3.2 *Relation between RIs and properties of leachate DOM*

This study determined the production of RIs from leachate DOM from MP under irradiation. We further investigated the relationship between RIs production from DOM and their properties to explain which factors determine RIs production, the potential predictor of Φ_{RIs}. As the results showed a significant correlation between Φ_{RIs} and DOM optical and fluorescence properties ($p < 0.05$) (Fig. 2). $\Phi_{^3DOM^*}$ was strongly positively correlated with HIX and a_{254}, suggesting that highly aromatic or humic-like DOM favors $\Phi_{^3DOM^*}$. $\Phi_{^1O2}$ was weak positive correlation with $S_{274\text{-}295}$ and c, meaning more protein-like, less aromatic DOM might promote singlet oxygen production. $\Phi_{O2\bullet-}$ has a strong negative correlation with FI, and positive with E2:E3, suggesting that more microbial-like or low molecular weight DOM reduces superoxide production.

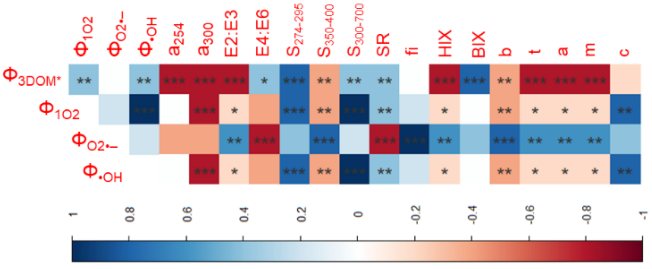


Fig. 2: Spearman's correlation matrix of quantum yield of RIs and optical properties of DOM. The color represents the value of the correlation coefficient: blue means positive correlation, red means negative correlation, white means near zero correlation, stronger colors mean stronger correlations. Significant level: $p < 0.001$ ~ "***", $p < 0.01$ ~ "**", $p < 0.05$ ~ "*", $p > 0.05$ ~ " ") (n=4).

For •OH, $\Phi_{\bullet OH}$ has a strong positive correlation with $S_{300\text{-}700}$ and a_{300}, indicating that broader absorbance and optical density favor hydroxyl radical formation. Fluorescence Indices (FI, BIX), which indicate microbial-like DOM,

negatively correlate with most RIs productions, particularly Φ_{3DOM*}. PARAFAC components such as *a* and *m* (humic-like) had positive correlations with RIs production (especially Φ_{3DOM*}), while *b* and *t* (protein-like) had negative correlations with Φ_{3DOM*}. These finding indicated that aromatic MP-MDOM generally promote Φ_{3DOM*} and Φ_{1O2} but suppress $\Phi_{O2\bullet-}$ and $\Phi_{\bullet OH}$. Broad absorbance and larger molecular size are favorable for producing highly reactive species like •OH.

3.3 Photoproduction of RIs from MPs surface

The calculated production rates of RIs from MP surface ranged from $7.67 \times 10^{-10} - 49.19 \times 10^{-10}$ M s^{-1} for $^1O_2$, $3.88 \times 10^{-12} - 36.41 \times 10^{-12}$ M s^{-1} for •OH, and $2.41 \times 10^{-10} - 3.56 \times 10^{-10}$ M s^{-1} for ^{3}RH*, and $7.11 \times 10^{-12} - 20.77 \times 10^{-12}$ M for $O_2^{\bullet-}$. Our study could observe the production of $O_2^{\bullet-}$ from MPs using XTT probes. However, some studies (e.g., Pan et al., [15]) could not detect $O_2^{\bullet-}$ production MPs when using the 5,5-dimethylpyrroline-n-oxide (DMPO) trapping agent due to complex effects of generated unstable DMPO-OOH. By comparing the production rate of RIs between leachate DOM and MP surface in the suspension, MPs exhibited a higher concentration of all RIs types, exceeding 70%, indicating the significant role of MP surfaces in radical production (Fig. 3). These results suggest that in the environment, MPs play crucial roles in the production of radicals and related reactions.

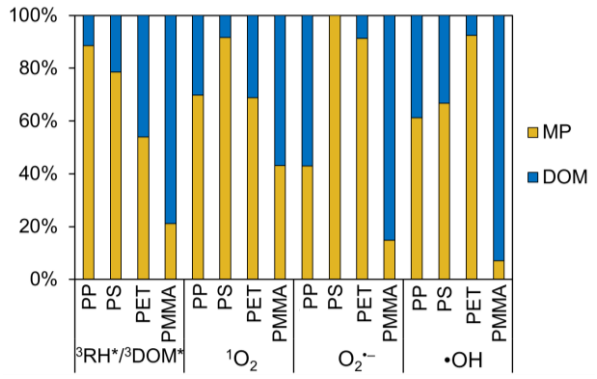


Fig. 3: Contribution of RIs production produced by MP surface and leachate DOM to the total production rate from suspension.

The calculated value of Φ_{3RH*}, Φ_{1O2}, $\Phi_{O2\bullet-}$, and $\Phi_{\bullet OH}$ were $(3.47 \pm 2.72) \times 10^{-4}$, $(23.39 \pm 17.57) \times 10^{-4}$, $(0.16 \pm 0.04) \times 10^{-4}$, and $(0.22 \pm 0.21) \times 10^{-4}$, respectively (Fig. 4). Statistical analysis revealed a significant difference in quantum yields among the RIs (two-way ANOVA, $p < 0.001$), whereas no significant difference was observed among MP types (two-way ANOVA, $p > 0.05$). These results indicate that the photoproduction of RIs from MPs is not determined by MP type. Duncan's multiple comparison test (at $\alpha = 0.001$) further demonstrated that quantum yield of $^1O_2$, which exhibited the highest value, was significantly difference from those of other RIs. Conversely, the quantum yield of ^{3}RH*, $O_2^{\bullet-}$, and •OH did not show significant differences among themselves. These finding highlight $^1O_2$ as the dominant RIs produced by MP, underscoring its role as the primary reactive species driving photo-transformation of MPs.

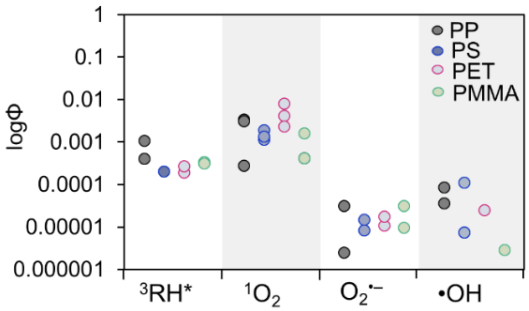


Fig. 4: Quantum yield of RIs production from MPs in expressed in proportion to mass (a) ($n = 3$ for $^1O_2$ and •OH, and $n = 2$ for ^{3}RH* and $O_2^{\bullet-}$).

4. CONCLUSIONS

This study quantified RIs photoproduction from MPs and their leachate DOM. The results found that ^{3}DOM*, $^1O_2$, and •OH are the dominant species in RIs production from MP-DOM, while $O_2^{\bullet-}$ plays a minimal role. The results of this study indicate that the photo-reactivities of MP-DOM are not significantly influenced by the chemical structure of MP, suggesting that other factors play a more dominant role in determining photochemical behavior of MP-DOM. Despite the higher aromatic content, the production of RIs from DOM of aromatic polymers is not necessarily higher than that from aliphatic polymers. Moreover, the study highlights that the optical properties of leachate DOM can serves as key indicators for explaining Φ_{RIs}, enabling the estimation of Φ_{RIs} through these properties. This research lays the groundwork for the development of comprehensive models to investigate the indirect photolysis processes of MP-DOM in natural aquatic environments. For the RI production from MP surface, the study revealed that $^1O_2$ is the dominant RIs produced by MP, indicating that $^1O_2$ is the major driver of MP photoreaction and mediated by indirect photolysis in the environment. Additionally, the photoproduction of RIs from MPs was not determined by MP type. This study underscores the potential RIs production from leachate DOM and MPs in enhancing the photodegradation of plastics and co-existing pollutants, offering insights into the fate and remediation of plastics in water environments.

This study focused on the leachate DOM and their associated RIs, and RIs production from pristine MPs under irradiation. However, in natural environments, plastic debris can persist for centuries to millennia, leading to alterations in surface properties, such as functional groups and physicochemical characteristics [16]. These modifications may affect the leaching behavior and photolytic reactivity of

MPs, which were not considered in this study. For the future research it is necessary to investigate the long-term leaching behavior and the environmental impacts of MP-DOM on photochemical processes. Furthermore, studies on the evolution of surface and chemical properties, along with RI production from MPs undergoing environmental aging, are essential to fully understand the environmental fate of MPs under sunlight environment.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 23KK0191 and the Nexus, Young Talents Exchange and Capacity Development Program, funded by Japan Science and Technology Agency (JST).

REFERENCES

[1] J. Duan, Y. Li, J. Gao, R. Cao, E. Shang, and W. Zhang, "ROS-mediated photoaging pathways of nano- and micro-plastic particles under UV irradiation," *Water Res.*, vol. 216, Jun. 2022, doi: 10.1016/J.WATRES.2022.118320.

[2] K. Zhu *et al.*, "Long-term phototransformation of microplastics under simulated sunlight irradiation in aquatic environments: Roles of reactive oxygen species," *Water Res.*, vol. 173, p. 115564, Apr. 2020, doi: 10.1016/J.WATRES.2020.115564.

[3] P. Liu *et al.*, "New Insights into the Aging Behavior of Microplastics Accelerated by Advanced Oxidation Processes," *Environ. Sci. Technol.*, vol. 53, no. 7, pp. 3579–3588, Apr. 2019, doi: 10.1021/ACS.EST.9B00493/SUPPL_FILE/ES9B00493_SI_001.PDF.

[4] V. Kaing, Z. Guo, T. Sok, D. Kodikara, F. Breider, and C. Yoshimura, "Photodegradation of biodegradable plastics in aquatic environments: Current understanding and challenges," *Sci. Total Environ.*, vol. 911, p. 168539, Feb. 2024, doi: 10.1016/j.scitotenv.2023.168539.

[5] Y. Liu, Y. Hu, C. Yang, C. Chen, W. Huang, and Z. Dang, "Aggregation kinetics of UV irradiated nanoplastics in aquatic environments," *Water Res.*, vol. 163, Oct. 2019, doi: 10.1016/j.watres.2019.114870.

[6] R. Ding, Z. Ouyang, L. Bai, X. Zuo, C. Xiao, and X. Guo, "What are the drivers of tetracycline photolysis induced by polystyrene microplastic?," *Chem. Eng. J.*, vol. 435, May 2022, doi: 10.1016/j.cej.2022.134827.

[7] A. M. de Souza, A. L. Santos, D. S. Araújo, R. R. de B. Magalhães, and T. L. Rocha, "Micro(nano)plastics as a vector of pharmaceuticals in aquatic ecosystem: Historical review and future trends," *J. Hazard. Mater. Adv.*, vol. 6, p. 100068, May 2022, doi: 10.1016/j.hazadv.2022.100068.

[8] X. Zhang *et al.*, "Effects and mechanisms of aged polystyrene microplastics on the photodegradation of sulfamethoxazole in water under simulated sunlight," *J. Hazard. Mater.*, vol. 433, p. 128813, Jul. 2022, doi: 10.1016/J.JHAZMAT.2022.128813.

[9] C. Chen *et al.*, "Impacts of microplastics on organotins' photodegradation in aquatic environments," *Environ. Pollut.*, vol. 267, Dec. 2020, doi: 10.1016/j.envpol.2020.115686.

[10] A. Das, "The emerging role of microplastics in systemic toxicity: Involvement of reactive oxygen species (ROS)," *Sci. Total Environ.*, vol. 895, p. 165076, Oct. 2023, doi: 10.1016/j.scitotenv.2023.165076.

[11] J. Duan *et al.*, "Weathering of microplastics and interaction with other coexisting constituents in terrestrial and aquatic environments," *Water Res.*, vol. 196, May 2021, doi: 10.1016/j.watres.2021.117011.

[12] P. L. Lenaker, A. K. Baldwin, S. R. Corsi, S. A. Mason, P. C. Reneau, and J. W. Scott, "Vertical Distribution of Microplastics in the Water Column and Surficial Sediment from the Milwaukee River Basin to Lake Michigan," *Environ. Sci. Technol.*, vol. 53, no. 21, pp. 12227–12237, Nov. 2019, doi: 10.1021/acs.est.9b03850.

[13] M. O'Connor, S. R. Helal, D. E. Latch, and W. A. Arnold, "Quantifying photo-production of triplet excited states and singlet oxygen from effluent organic matter," *Water Res.*, vol. 156, pp. 23–33, Jun. 2019, doi: 10.1016/J.WATRES.2019.03.002.

[14] X. Wu *et al.*, "Photo aging of polypropylene microplastics in estuary water and coastal seawater: Important role of chlorine ion," *Water Res.*, vol. 202, Sep. 2021, doi: 10.1016/j.watres.2021.117396.

[15] J. Pan, S. Zhang, X. Qiu, L. Ding, X. Liang, and X. Guo, "Molecular Weights of Dissolved Organic Matter Significantly Affect Photoaging of Microplastics," *Environ. Sci. Technol.*, Jul. 2024, doi: 10.1021/acs.est.4c04608.

[16] X. F. Wei, M. Bohlén, C. Lindblad, M. Hedenqvist, and A. Hakonen, "Microplastics generated from a biodegradable plastic in freshwater and seawater," *Water Res.*, vol. 198, p. 117123, Jun. 2021, doi: 10.1016/J.WATRES.2021.117123.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Water Quality Assessment of Groundwater in Mondolkiri Province, Cambodia: Implications for Agriculture

Solita Chan [1]*, Sarapich Vatey Mao [1], Sreynim Mol [1], Sreyvin Tha [1], Techleang Naeung [1], Leanggek Menh [1], Rina Heu [1]

[1] *Faculty of Hydrology and Water Resource Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.*

[*]*Corresponding author: chonsolitae20201709@gmail.com*

Abstract: *Groundwater is one of essential resources in Cambodia, supporting ecosystems, drinking water, and farming. However, in some areas, the water quality is affected by high levels of chemicals and ions, which can harm crops and their growth. This study focuses on the water quality of groundwater in the Pech Chreada and Kaoh Nhaek districts of Mondolkiri Province. Twenty water samples were collected and tested for important factors like pH (using a HORIBA pH110 meter), total hardness, total phosphate, total nitrogen, iron, and alkalinity (using an MD600 photometer). The results showed large differences in water quality. Total hardness ranged from low (9 mg/L) to high (330 mg/L), which could affect how plants absorb nutrients. There were also changes in the amounts of nitrogen and iron, which could impact crop growth and irrigation systems. High alkalinity levels could cause micronutrient shortages and make pesticides less effective. Although the pH levels were mostly acceptable for irrigation, they did fluctuate, which could hurt plant health. Phosphate levels were generally low, but occasional spikes raised concerns about possible eutrophication, which can lead to excessive algae growth and oxygen depletion in water bodies. These results emphasize the need to monitor water quality to support sustainable farming and protect the environment. Poor water quality can weaken crops, reduce yields, and increase farming costs. If water conditions are not properly managed, long-term soil health could also decline, further affecting agricultural productivity. Proper water management, including regular testing and treatment, is crucial for preventing damage to crops and keeping ecosystems in balance for future generations.*

Keywords: Agricultural Impact; Environmental Impact; Groundwater; Sustainable Water Management; Water Quality

1. INTRODUCTION

Water supports important ecosystem services, functions as a non-substitutable input to crop and animal production, and is essential for sustaining economic growth, health and resilience of a country. However, water is also a finite resource and is becoming increasingly scarce in many parts of the world with more countries facing water stress due to climate change. Water pollution from domestic, industrial, and/or agricultural sources can severely affect the availability of the resources for various uses, as well as human and environmental health. Agriculture, as the largest single user of freshwater resources, is also a significant source of chemical and organic pollution to surface water and groundwater resources, causing human illnesses, loss of biodiversity, contamination of marine ecosystems from land-based activities, closure of drinking water sources due to nutrients and toxic algal blooms, and global contamination by persistent organic pesticides.

Growing urbanization is increasing the demand for water and producing more wastewater. Different sectors such as agriculture, industry and households will face stronger competition for scarce water resources. Achieving sustainable urban development, including food and water security, requires sustainable production and consumption patterns by incorporating water valuation into integrated water resources management and transforming food systems. Globally, about 330 km³ of urban wastewater, 660 km³ of industrial wastewater (including cooling water) and an estimated 1260 km³ of agricultural drainage effluent are annually discharged untreated into the environment [1] (Mateo- Sagasta, Zadeh & Turral, 2018), affecting about 29 million hectares of irrigated farm land [2] (Thebo et al., 2017).

2. Study area

This study was conducted in Bu Sra, Srae Ampum Commune, Pech Chreada District, and Srae Sangkum, Sok Sant Commune, Kaoh Nhaek District, in Mondolkiri Province. Water samples were collected from both surface water and groundwater in Mondolkiri.

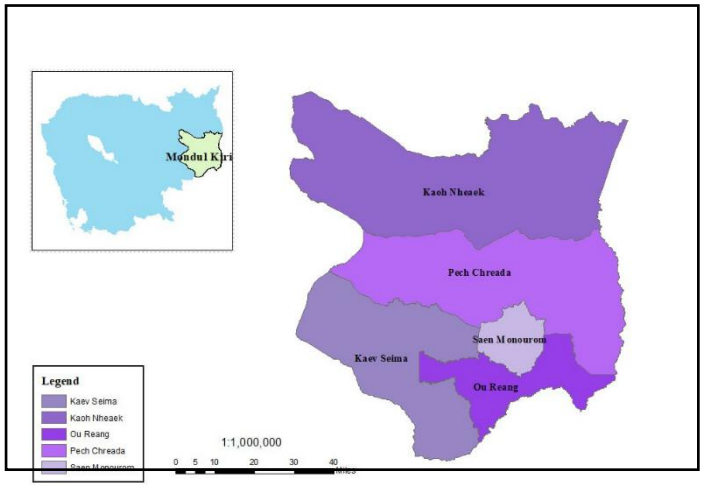


Figure 1 Map for collection sample

3. Methodology

3.1. Hardness analysis

The Hardness (Caco3) was analyzed by photometer MD 600. It has a range of Hardness from 2 – 50 mg/l Caco3 by using reagent HARDCHECK. First, use to clean vials (24 mm) and add 10 ml of water sample. Place the vial in the sample chamber making sure that the Marks are aligned. Press the zero key. Remove the vial from the water sample. Add one HARDCHECK P Tablet straight from the foil to the water sample and crush the tablet using a clean stirring rod. Close the vial tightly with the cap and swirl until the tablet is dissolved. Place the vial in the sample chamber making sure that the Marks are aligned. Press the Test key. Wait for a reaction period of 5 minutes. After the reaction period is finished the measurement starts automatically. The result is shown on the display as total hardness.

3.2. Total Nitrogen analysis

The Total Nitrogen (TN) was analyzed by Photometer MD600. It has a range of nitrogen from 0.5 – 25 mg/l N by using regent one Vario TN Persulfate Rgt Powder pack and one Vario TN regent A powder pack and one Vario TN regent B powder pack. First open two TN Hydroxide LR digestion vials and add the contents of one Vario TN Persulfate Rgt Powder pack. Add 2 ml deionized water to the prepared vial (this is the blank). Add 2 ml of the water sample to the other prepared vial (this is the sample). Close the vials with the cap and shake to mix the contents (at least 30 seconds). Heat the vial for 30 minutes in the preheated reactor at a temperature of 100ºC. After 30 minutes remove the vial from the reactor and allow the vial to cool to room temperature. Open the cooled digestion vial and add the contents of one Vario TN regent A powder pack to each vial. Close the vial with the caps and shake to mix the contents (at least 15 seconds). Press (Inter) key, wait for a rection period of 3 minutes. After the reaction period is finished proceed as follows. Open the digestion vials and add the contents of one Vario TN

regent B powder pack to each vial. Close the vials with the caps and shake to mix the contents (at least 15 seconds). Press [Enter] key. Wait for a reaction period of 2 minutes. After the reaction period is finished proceeding as follows. Open two TN Acid LR/HR (Reagent C) vials and add 2 ml of the digested, treated blank to one vial (this is the blank). Add 2 ml of the digested, treated water sample to the other TN Acid LR/HR vial (this is the sample). Close the vials with the caps and swirl the vials gently several times to mix the content. Place the vial (the blank) in the sample chamber making sure that the marks are aligned. Press ZERO key. Wait for a reaction period of 5 minutes. After the reaction period is finished the measurement starts automatically. Remove the vial from the sample chamber. Place the vial (the sample, Note 10) in the sample. chamber making sure that the marks are aligned. Press TEST key.The result is shown in the display in mg/l Nitrogen.

3.3. Iron analysis

Photometer MD600 analyzed the Fe. It has a range of Fe from 0.02 – 3 mg/l Fe by using reagent Vario Ferro F10 powder pack. First, fill a clean vial (24 mm) with 10 ml of the water sample, and close it tightly with the cap. Place the vial in the sample chamber making sure that the Marks are aligned. Press the zero key. Remove the vial from the sample chamber. Add the contents of one Vario Ferro F10 Powder Pack straight from the foil to the water sample. Close the vial tightly with the cap and swirl several times to mix the contents. Place the vial in the sample chamber making sure that the Marks are aligned. Press the TEST key and wait for a reaction period of 3 minutes. After the reaction period is finished the measurement starts automatically. The result is Shown in the display in mg/l Fe.

3.4. Total Phosphate analysis

Photometer MD600 analyzed the total phosphate (TN). It has a range of phosphate from 0.02 – 1.1 mg/l PO4 by using regent Vario Potassium Persulfate F10 Powder Pack and Vario Phosphate Rgt F10 Powder. First, open the white cap of one digestion tube PO4-P acid reagent and add 5 ml of the water sample. Add the contents of one Vario Potassium Persulfate F10 Powder Pack straight from the foil to the vial. Close he vial with the cap and invert several times to mix the contents. Heat the vial for 30 minutes in the preheated reactor

at a temperature of 100C. After 30 minutes remove the vial from the reactor and allow the vial to cool to room temperature. Open the cooled digestion vial and add 2 ml 1.54 N Sodium Hydroxide solution to the vial. Close the vial with the cap and invert gently several times to mix the contents. Place the vial in the sample chamber making sure that the Marks are aligned. Press zero key. Remove the vial from the sample chamber. Add the contents of one Vario Phosphate Rgt F10 powder pack straight from the foil to the vial. Close the vial tightly with the cap and swirl several times to mix the contents (approx. 10-15 sec). Place the vial in the sample chamber making sure that the Marks are aligned. Press TEST key. Wait for reaction period of 2 minutes. After the reaction period is finished the measurement starts automatically. The result is Shown in the display in mg/l PO4.

.

3.5. Alkalinity analysis

The process for measuring alkalinity in water involves adding acid to a water sample and tracking the pH change until the sample reaches a specific pH endpoint. This process is called titration, using sulfuric acid (H_2SO_4) with a concentration of 0.02 N.

First, pour 100 mL of the pure water sample into a beaker.Measure and record the initial pH of the water before titrating with sulfuric acid.Add sulfuric acid gradually to the water sample while continuing to monitor the pH.Continue adding the acid until the pH reaches the specific endpoint, typically around 4.5.After reaching the endpoint, calculate the alkalinity based on the volume of acid required to reach the endpoint.

3.6. pH analysis

pH is a measure of how acidic or basic water is. The range goes from 0 to 14, with 7 being neutral. We measure pH using the pH equipment (Portable pH Meter Model PH 110 HORIBA). Calibrate the equipment by placing the probe into a beaker containing distilled water. Wait until the equipment is fully calibrated. After calibration, pour the sample into a clean beaker and measure its pH. When the reading on the screen stabilizes, that is the result.

4. Results and Discussion

4.1. Hardness analysis

Total hardness is the sum of the calcium and magnesium concentrations, both expressed as calcium carbonate, in milligrams per liter (mg/L). You can determine your water's hardness based on these concentrations of calcium carbonate: Below 75 mg/L - is generally considered soft. 76 to 150 mg/L - moderately hard. The sampling of KCAC- 01A, KCAC-01B, and KCAC- 01C have the highest concentration range from 240 mg/l to 330 mg/l (shown in figure 11). KV-01B, PCAC-01B, KV-01A, PV-01C, PV-01A have a concentration range from 50 mg/l to 120 mg/l and the sampling of K Downstream-A, K Downstream-B, KV-01C, KV-02A, KV-02B have a range from 41 mg/l to 48 mg/l. The lowest concentration of Hardness has a range from 9 mg/l to 36 mg/l represent by PCAC-01A, PCOF-01B, PCAF-01A, PCOF-01C, PCA-C01C and PV-01B. The high hardness of water is more difficult for plants to absorb and break down than soft water. It tends to bind up soil nutrients, making it more difficult for plants to absorb what they need. To try and compensate, farmers will tend to increase their rate of irrigation, which leads to other problems.

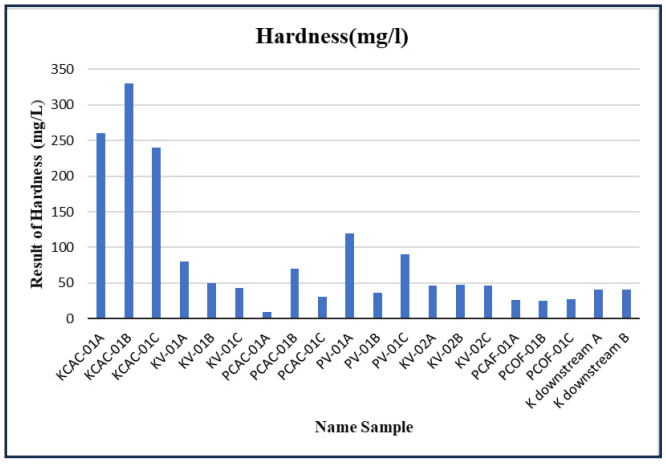


Figure 2: Result of Hardness analysis

4.2. Total Nitrogen

Total nitrogen (TN) is the sum of all nitrogen forms in a substance, including: Ammonia nitrogen (NH3), Organic nitrogen (nitrogen in amino acids and proteins), Nitrite (NO2), and Nitrate (NO3). TN can be natural or man-made, and can enter the body through food, drinking water, or air pollution. The water sample of KCAC-01B, PV-01C, KV-02C have a highest concentration range from 22 mg/l to 49 mg/l**).** KV-01C, PCAC-01B, PV-01B, KV-02B, and K downstream-B have mediums range from 5.7 mg/l to 16.6 mg/l. The lowest concentration of water sample is sampling KCAC-01A, KCAC-01C, KV-01A, KV-01B, PCAC-01C, PV-01A, KV-01A, PCAF-01A, PCOF-01B, PCOF-01C, and PCAC-01A have a range from ≤0.5 mg/l to 0.7 mg/l. Too much nitrogen in water can overstimulate crop growth, which can delay maturity, reduce quality, and upset production. It can also cause an explosion of foliar growth at the expense of flower formation, fruit set, and root growth. Total Nitrogen in water also can cause algae to grow faster than ecosystems can handle, which can harm water quality, food resources, and habitats. Additionally, high concentration of Total nitrogen can be influenced by multiple factors including fertilizers, manure, and runoff from agricultural fields.

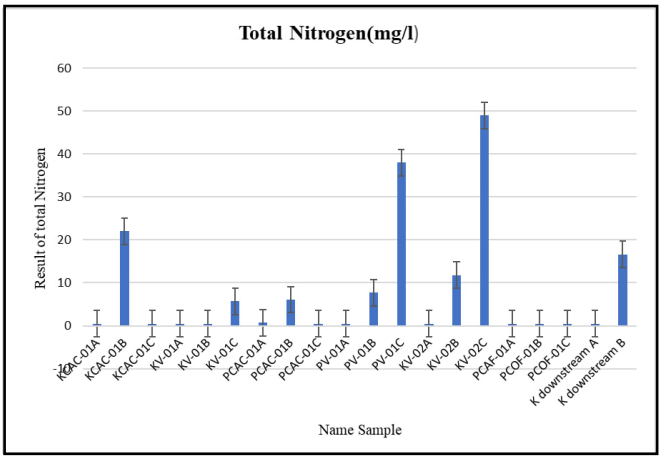


Figure 3: Result of Total Nitrogen analysis

4.3.Iron analysis

The lowest concentration of iron water sample has a range from ≤0.02 mg/l to 0.7 mg/l represent by sampling KCAC-01A, KCAC-01B, KCAC-01C, PV-01A, PCAC-01C, PCAC-01B, KV-01A, and PCAC-01A. K downstream-B, PCAC-01A, KV-02A, KV-02C, PV-01C, KV-02B, KV-01C, K downstream-A has medium concentration range from 0.89 mg/l to 1.2 mg/l. The sampling of PCOF-01C, PCOF-01B, PCAF-01A, and PV-01B has a highest concentration range from 1.42 mg/l to 4.1 mg/l. High levels of iron may also allow the growth of iron bacteria, which produce a slime that can clog filters and plumbing. Although these bacteria can make water taste or smell bad, there do not harm your health. The water with iron levels may cause clogging of drip irrigation emitters and also lead to iron rust stains, and discoloration on foliage plants in overhead irrigation applications. Iron natural occurs in rocks and soil. It can dissolve into groundwater through the weathering of iron-bearing minerals.

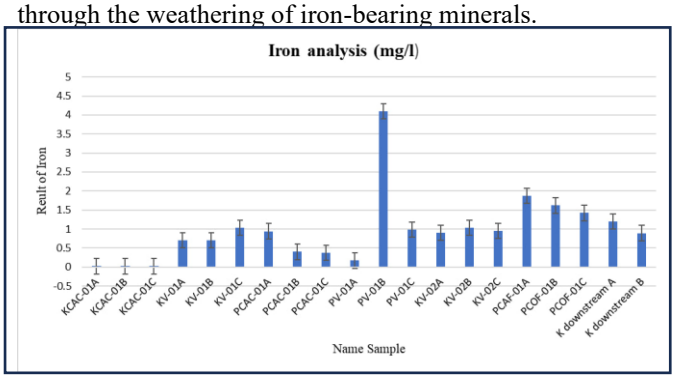


Figure 4: Result of Iron analysis

4.4.Alkalinity

The sampling of KCAC-01C, KCAC-01A, KCAC-01B has the highest value range from 395 mg/l to 400 mg/l. K downstream-B, KV-02B, PCAC-01B, KV-01A, KV-01B, PCAC-01C, KV-01C, PV-01C has medium of value range from 61 mg/l to 102 mg/l, and the lowest value of water sample has a range from 13 mg/l to 60 mg/l, represented by PCAC-01A, PCAF-01A, PCOF-01C, PCOF-01B, PV-01B, K downstream-A and KV-02C. High alkalinity water can increase the pH of the growing medium, which can lead to micronutrient deficiencies, toxicity, and poor plant quality. And, it can reduce the effectiveness of some pesticides, floral preservatives, and growth regulators. Additionally, high concentration of alkalinity can be influenced by multiple factors including Agricultural lime, weathering rates, and the dissolution of carbon dioxide from the atmosphere can also affect water alkalinity.

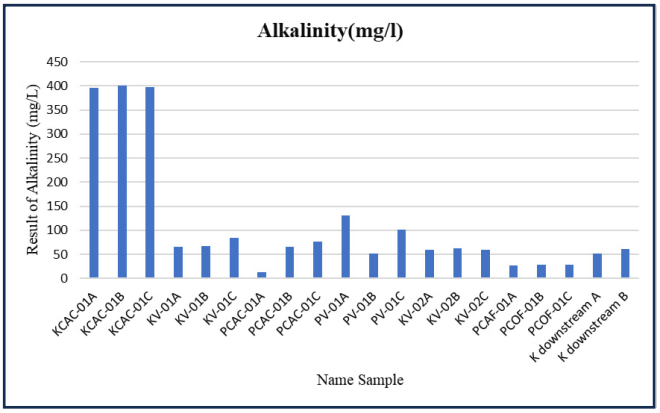


Figure 5: Result of Alkalinity analysis

4.5.pH analysis

pH is a measure of how acidic or basic water is, and is reported on a scale from 0 to 14. Water has a neutral pH of 7, which indicates that it is neither acidic or basic. pH is one of the most important parameters of water quality. The sampling in the is categorize of pH concentration. Higher pH levels from 7.31 to 7.30 are show at KCAC-01B, KCAC-01C and KCAC_01A with the largest point. PCOF-01B, PV-01B, KV-02C, KV-02B, KV-02A, K downstream B, PCAC-01B, KV-01B, KV-01A,KV-01C, PCAF-01A, PV-01A, PV=01C has a concentration range from 6.53 to 6.93. The lowest concentration is range from 6.12 to 6.46 are represented by PCAC-01A, PCAC-01C, PCOF-01c and K downstream A. The ideal pH range for irrigation water is 6.5–8.4. Water with a pH outside this range can cause nutritional imbalances or contain toxic ions. And also Water with an adverse pH can slowly change the pH of soil. However, soil is buffered and resists change, so it may be easier to correct the soil pH than the water's pH. The pH affects a variety of chemical reaction that occur in aquatic environments, including complexations, oxidation- reduction reaction and acid-base reaction (Ndur et al., 2020).

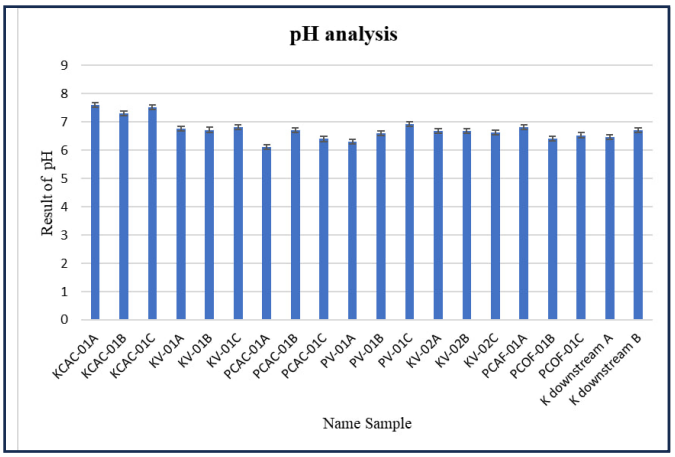


Figure 6: Result of pH analysis

4.6.Total Phosphate

The Sampling of KCAC-01A, KCAC-01B, KV-01C, PCAC-01A, PCAC-01C, PCAF-01A, PCAF-01B, PV-01B, K downstream-B, KCAC-01C and KV-02B have a lowest concentration range from ≤0.02 mg/l to 0.15 mg/l. KV-01A, KV-02C, PV-01A, PCOF-01C, KV-01B, PV-01C have concentration from 0.16 mg/l to 0-27 mg/l. The highest concentration of sampling ranges from 0.32mg/l to 0.41 mg/l represents KV-02A, PCAC-01B and K downstream A. The high concentration of phosphate in water can make the water has a low dissolved oxygen and also can affect the pH levels in water, and Excess nutrients like phosphorus can lead to cultural eutrophication, which is a serious threat to water quality. Phosphorus is a nutrient that is essential for plants and animals. However, too much phosphorus in water can be caused by human activities like using lawn fertilizers, wastewater treatment facilities, and storm water runoff.

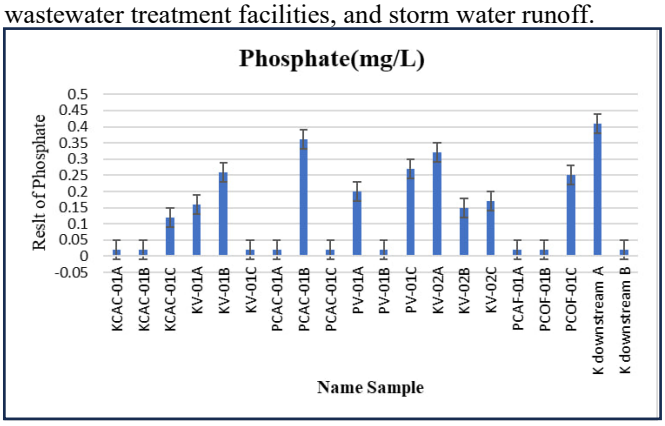


Figure 7: Result of Total phosphate

5. Conclusion

This study aimed to assess the water quality in the Pech Jreada and Kaoh Nhaek Districts of Mondolkiri Province by analyzing water samples using an MD600 photometer. The analysis revealed significant variations in water quality parameters, which have important implications for agricultural and environmental management. Total hardness levels ranged from low (9 mg/L to 36 mg/L) to high (240 mg/L to 330 mg/L), indicating potential challenges for plant nutrient absorption in areas with elevated hardness. Total nitrogen concentrations varied widely, with higher values in some samples, potentially accelerating plant growth but also risking water quality degradation due to nutrient overload. Iron concentrations also ranged from low to high, with some sites experiencing levels that could cause clogging of irrigation systems and plant discoloration.

Alkalinity levels showed considerable variation, with high alkalinity potentially leading to micronutrient deficiencies and reduced pesticide effectiveness. While pH levels generally fell within the acceptable range for irrigation, slight variations could still affect plant health and soil conditions over time. Phosphate concentrations were mostly low but occasionally reached higher levels, which could contribute to eutrophication in water bodies, ultimately degrading water quality and disrupting aquatic ecosystems.

These findings highlight the importance of regularly monitoring key water quality parameters to ensure sustainable agricultural practices. Proper management of irrigation systems, targeted water treatment, and careful nutrient application are necessary to mitigate the potential adverse effects on crop productivity and long-term environmental health in the study area.

Acknowledgement

I would like express thank to Project for Strengthening Engineering Education and Research for Industrial Development in Cambodia (Laboratory-Based Education Project), for financial support and the Water Supply and Treatment Laboratory for the experimental facilities.

References

[1] M. M. M. Syeed, M. S. Hossain, M. R. Karim, M. F. Uddin, M. Hasan, and R. H. Khan, "Surface water quality profiling using the water quality index, pollution index and statistical methods: A critical review," Jun. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.indic.2023.100247.

[2] E. S. W A R R A N T Y and Y. M. A D E I N J A P A N, "Salinity Resistivity Conductivity Ion ORP pH DO TDS Handheld Water Quality Meters 100 Series." [Online]. Available: www.horiba-laqua.com

[3] M. Shariati-Rad and S. Heidari, "Classification and determination of total hardness of water using silver nanoparticles," *Talanta*, vol. 219, Nov. 2020, doi: 10.1016/j.talanta.2020.121297.

[4] S. Karastogianni, S. Girousi, and S. Sotiropoulos, "pH: Principles and Measurement," in *Encyclopedia of Food and Health*, Elsevier Inc., 2015, pp. 333–338. doi: 10.1016/B978-0-12-384947-2.00538-9.

[5] J. K. Trick, M. Stuart, and S. Reeder, "Contaminated Groundwater Sampling and Quality Control of Water Analyses," in *Environmental Geochemistry: Site Characterization, Data Analysis and Case Histories: Second Edition*, Elsevier, 2018, pp. 25–45. doi: 10.1016/B978-0-444-63763-5.00004-5.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessing Macroplastic Pollution in Kampong Khleang River: Urban vs. Rural Pollution and Seasonal Variability

Leanggek Menh[1*], Rina Heu [1,2], Pinnara Ket [1], Davin Sang [1], Vouchlay Theng [1], Vilde Kloster Snekkevik [3]

Rachel Hurley [3], Solita Chan [1], Solida Sim [1], Chanmonida Heng [1], Kong Chhuon [1]

[1] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Norwegian Institute for Water Research (NIVA), Oslo, Norway*

[*]*Corresponding author: menh.leanggek@itc.edu.kh*

Abstract: Rapid population growth and economic development have significantly increased plastic waste generation in Cambodia. A recent World Bank research study identified that the large quantity of plastic litter in Siem Reap can carry plastic debris to the river. Kampong Khleang (KPK) is a permanent community within the floodplain of the lake, with an economy based on fishing and surrounded by flooded forests that can be impacted by plastic pollution. Macroplastic (MP) litter can reduce fish production and abundance by harming the aquatic ecosystem. Therefore, all the issues can cause a serious threat to freshwater biodiversity and ecosystem services. So, this study was conducted to quantify the flux and accumulation of MP in different sections of the KPK river, comparing urban and rural pollution levels, and to assess the seasonal variation of MP pollution by comparing litter accumulation and density during the rainy and dry seasons. Data were collected at two selected sampling sites located along the KPK River. The midstream is located at Spean Betong Phumi Prey Pros (13.148782, 104.217235), and the upstream is located at Tnal Dach bridge on National Road 6A (13.178974, 104.245268). The method of visual observation involves human analysis to observe and record visible MP loads passing through a cross-section of the river channel, and net sampling uses a 2 cm mesh size, lowered 5 cm below the surface of the water, to intercept and collect MP flowing in the river. The last method is a riverbank survey to estimate and collect MP within 50 meters of the riverbank edges. The collected MP was measured by weight and categorized based on its type. The higher proportion of plastic bag fragments and empty plastic bags (35% combined) indicates improper disposal upstream. Downstream, the highest percentage of plastic litter is cigarette butts at 15%, indicating frequent human activity in the area. The result of the density shows that the downstream is higher than the upstream. In the dry season, litter accumulation is higher downstream, reaching nearly 900 units compared to about 400 units upstream. After field observations and data analysis, illustrate a high accumulation of MP such as plastic bottles, bags, cigarette butts, and food packaging. Overall, MP pollution is consistently higher downstream than upstream in the KPK River. These pollutants can threaten local biodiversity and compromise the safety of food sources and the sustainability of community life.

Keywords: Ecosystem, Fish production, Kompong Khleang River, Macroplastic, Pollution

1. INTRODUCTION

Kampong Khleang commune is located in Sotrnikum District on the northern lake edge, about 35 km east of Siem Reap town, is a permanent community within the flood plain of the lake with an economy based on fishing and surrounded by flooded forest. It is significantly larger, with nearly 3.5 times the population of Kompong Pluk Commune, making it the largest community on the Lake. Mangrove forests are dominant in this floating commune. Dried fish are the specialty of this commune [1].

Over recent decades, a combination of rapid population growth and economic development has led to an exponential increase in the volume of solid waste, including plastic waste, in Cambodia. According to a UNDP report 2018 [2]. The rapidly growing volume of waste has led to several increasingly urgent problems. A recent World Bank research (2020)3 identified the large quantity of plastic litter in Phnom Penh Canals, Siem Reap Canal, and Sihanoukville Beach, which potentially carries plastic debris to the Mekong River and eventually to the oceans. Major plastic litter, found in these areas, includes plastic wrappers, shopping bags, bottles, and bottle caps [3].

The nano, micro, and macro forms of plastics are reported to be linked with microbiological, chemical, and particle toxicities, and other hazardous consequences in aquatic

[1*] Corresponding author: Leanggek Menh
E-mail: menh_leanggek@itc.edu.kh; Tel: +855-70 226 621

ecosystems. Fish are considered a mostly affected biotic component in response to diverse forms of environmental stress in the aquatic ecosystem [4]. Plastic litter may cause a notable reduction in total fish production or abundance in an aquatic habitat due to its adverse effect on the existing ecosystem [5]. Therefore, all the aforementioned issues ultimately can cause a serious threat to freshwater biodiversity and ecosystem services.

Most studies on plastic pollution in rivers focus on measuring how plastic moves through the water. Following a review, 39 studies from around the world were found that measured large plastic waste, known as macroplastics (MP), in rivers. These studies used five main methods, which can be grouped into two types: observation-based and interception-based. Observation-based methods involve watching the river and recording visible plastic without actually collecting it. Interception-based methods, on the other hand, involve catching plastic using tools or barriers so that it can be collected, counted, and categorized [4]. To address these knowledge gaps, this study was conducted to quantify the flux and accumulation of MP in different sections of the Kampong Kleang River (KPK), comparing urban and rural pollution levels and assessing the seasonal variation of MP pollution by comparing litter accumulation and density during dry seasons.

2. METHODOLOGY

2.1 Study Area

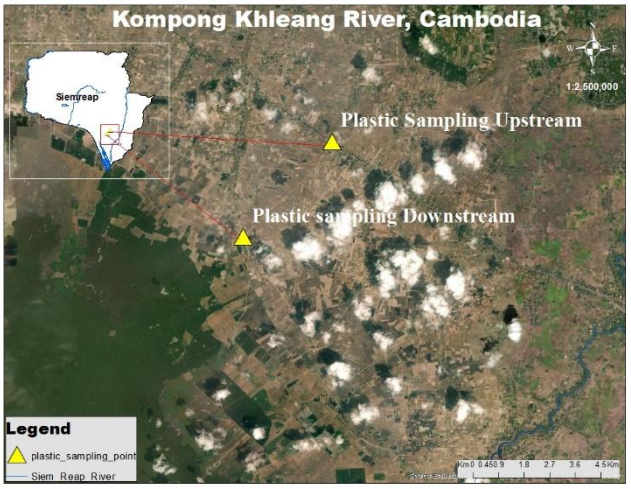


Fig 1. 1 Study area in Kampong Khleang River

Two sampling sites were selected in the KPK Commune in this study. It was divided into 2 sites, downstream is on Spean betong Phumi Prey Pros with coordinates 13.148782, 104.217235, and Upstream is on Tnal dach bridge, National 6A road with coordinates 104.245268,13.178974. The sampling method is divided into 3 sessions. The first session starts from 10 am to 11 am. The second session starts from 11:00 am to 12:00 pm. The third session starts from 12:00 pm to 1:00 pm. The method of visual observation involves human analysis to observe and record visible MP loads passing through a cross-section of the river channel. It conducted plastic flow monitoring by standing on a bridge overlooking the river and observing for a continuous period of 15 minutes. During this time, we counted all visible plastic items floating downstream and categorized them based on type, such as plastic bottles, bags, food packaging, and other single-use plastics. Net sampling uses a 2 cm mesh size, lowered 5 cm below the water's surface, to intercept and collect MP flowing in the river. The last method is a riverbank survey to estimate and collect MP within 50 meters of the riverbank edges. The materials and equipment used for field data collection included a handcrafted manta net (50x50 cm), a high-precision balance for measuring collected samples, protective gloves, and face masks. These materials were carefully selected to meet the methodological requirements of the study and ensure both the safety of the researchers and the integrity of the samples, thereby maintaining high standards of accuracy and safety throughout the fieldwork process.

- Net

Nets comprise 2 cm mesh screens that are lowered to a water depth of 5 cm and used to intercept MP that is flowing in the river. They are typically deployed from a fixed point, such as a bridge, but they may also be used from a boat [6] or from the riverbank or bed [7]. Nets can be utilized to measure floating, sub-surface, or even benthic transport of MP using buoys and weights to position them in the desired location within the water column [8]. They can also be deployed in fleets (several nets connected to the same frame) to sample different parts of the water column simultaneously [9]. Nets are deployed for predefined measurement times or until the net becomes full or clogged with debris. Once the measurement period is over, the nets are retrieved from the water. Their contents are then inspected to isolate, count, categorize, and weigh the MP component. The use of nets represents the most commonly utilized physical interception-based technique for measuring MP flows.

- Visual Observation

Visual observation is the most commonly utilized technique for measuring MP flows in rivers. Briefly, human analysts observe and record the visible MP load passing a cross-section of the river channel or a portion of it during a predefined measurement period. A prerequisite is the availability of a vantage point for viewing the river. This is typically a bridge that passes over the river, it can also be performed from the riverbank for narrower streams [11,12]. Visual observation is, therefore, limited to locations where such a vantage point exists and can be safely accessed [12]. In wider channels, it will not be possible to observe the full

width of the river from a single point, and in this case, it is necessary to define a cross-section portion or observation track that represents a proportion of the river width [13]. Van Emmerik et al. (2018) [14] recommend dividing the channel cross-section into several portions and observing each of these for a measurement period that allows all defined portions to be observed in under one hour, deemed to be a time lag where stream conditions will be relatively stationary. Where the full river width cannot be monitored within an appropriate time window, the results from the measurements are.

- Macroplastic Collection on Riverbank

Macroplastics (MP) were collected on the riverbank following the method of Rech *et al.* (2014) [4] with modifications. The area was divided into three sections: the downstream and upstream (Table 1). The sampling sites were heterogeneous depending on their accessibility. Each section had three quadrats (10 m x 5 m) with a 50-m distance in between quadrats. MP found in the quadrats was collected and placed in a container. The MP was then weighed and characterized. Collection of MP from the same site was done 3 times per six months from 17 Dec 2024–04 April 2025. The collected MP litter was classified based on its uses such as food packaging, toiletries, plastic bag, plastic bottle, disposable cups, plastic caps, medical waste, cigarette butts, plastic fragments, cigarette cartons, diaper/sanitary pads, ropes, and others [15].

2.2 Data analysis

The Plastic litter Density (PLD) of collected macroplastic was used by the method of Abreo et al. (2019) [16]:

$$\text{PLD} = \frac{\text{Number of items}}{\text{Total area sampled (L×W)}} \quad \text{(Eq 1)}$$

PLD: Plastic Litter Density (Item/m^2)

L: Length of area (m)

W: Width of area (m)

The Plastic litter Density (PLC) of collected macroplastic was;

$$PLC = \frac{Number\ of\ items\ per\ caterggory}{Total\ number\ of\ items\ in\ all\ catergory} * 100 \quad \text{(Eq 2)}$$

3. RESULTS AND DISCUSSION

3.1 The result of the flux and accumulation of the riverbank macroplastic survey in urban and rural pollution

- Plastic litter Composition in Upstream (urban pollution)

The pie chart displays the distribution of various types of plastic waste found in the upstream area. The most common type is plastic bag fragments (20%), followed by empty plastic bags (15%), plastic bottles (12%), and cup lids (10%). Other notable items include bottle/container lids and cups (both at 9%), bottle labels (7%), and smaller proportions of food containers**,** straws**,** cigarette box wrappers**,** plastic cables, and plastic bags with contents, each comprising between 2% to 5% of the total litter. The composition of plastic litter upstream reveals a presence of single-use plastic items**,** particularly those related to packaging and food/beverage consumption. The high proportion of plastic bag fragments and empty plastic bags (35% combined) indicates improper disposal or degradation of bags over time, possibly due to wind or water movement fragmenting them after they've been discarded. Their lightweight nature also makes them more likely to be transported from nearby roads or fields into the water. Plastic bottles and their components (bottle/container lids and labels) account for a combined 28%, suggesting frequent usage and disposal of bottled drinks, possibly by travelers or passersby near the national road**.** The presence of cup lids, cups, and food containers (24% collectively) also points toward roadside food consumption or takeout culture contributing to littering. Items like plastic cables and cigarette box wrappers**,** although smaller in percentage, indicate other miscellaneous household or personal waste being disposed of improperly.

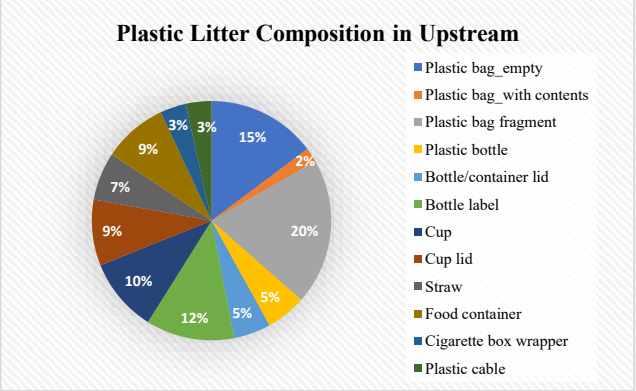


Fig 3. 1 The result of plastic litter Composition in the upstream

- *Plastic litter Composition in the Downstream (rural pollution)*

This is followed by cigarette butts at 15%, indicating frequent human activity in the area, possibly from nearby households or passersby. Plastic bottles account for 12%, and sachet food wrappers make up 10%, both commonly associated with daily consumer habits. Other categories include bottle or container lids and plastic bag fragments, each at 7%, and cup lids and plastic bags with contents, contributing 6% and 5% respectively. Although empty plastic

bags account for only 3%, their presence remains due to their tendency to be carried by wind or water, eventually contributing to environmental degradation. Less common, but still present, are straws (2%), cigarette box wrappers (2%), lighters (1%), bottle labels (2%), and carton/tetrapaks (2%). This composition reflects the influence of nearby residential and agricultural activities, where household waste and convenience products are often discarded improperly. The presence of trees and open fields likely contributes to the accumulation of wind-blown or water-transported litter. The high percentage of cigarette butts also suggests informal gatherings or a lack of designated disposal areas.

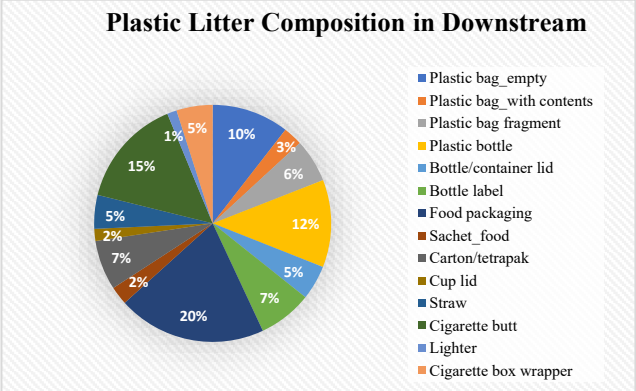


Fig 3. 2 The result of plastic litter Composition in the Downstream

- Plastic Litter Density

The bar graph illustrates the plastic litter density (items/m²) recorded upstream and downstream at three time points: December 2024, February 2025, and April 2025. In December, there were two sites, with the downstream area showing a higher density (9.96 items/m²) compared to the upstream (1.60 items/m²), suggesting substantial accumulation or input of litter along the waterway. By February, both upstream and downstream densities increased to 6.02 and 7.88 items/m² respectively, reducing the gap between the two, which may indicate increased litter input or runoff during this period. However, in April, plastic litter densities dropped markedly at both sites, 1.42 items/m² upstream and 2.48 items/m² downstream, indicating a possible seasonal decline in litter or the impact of cleanup efforts. Throughout the period, downstream consistently exhibited higher litter densities, highlighting the need for targeted waste management interventions downstream, as well as continuous monitoring of upstream sources, especially given the notable increase observed in February. Overall, despite being rural, the downstream area's close interaction with daily human activities and agricultural practices makes it more vulnerable to plastic litter accumulation, whereas the upstream's urban but more physically separated and possibly better-managed environment contributes to its lower litter density.

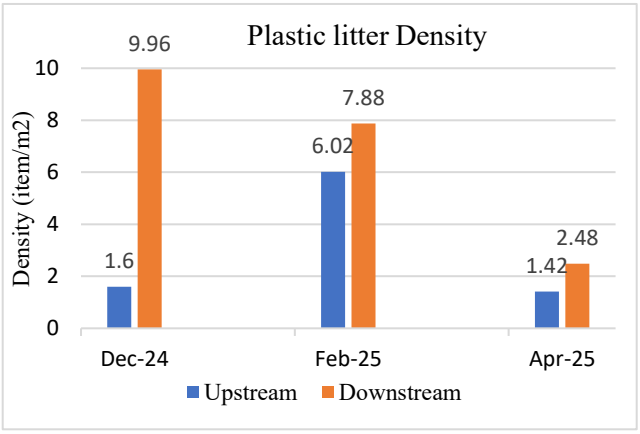


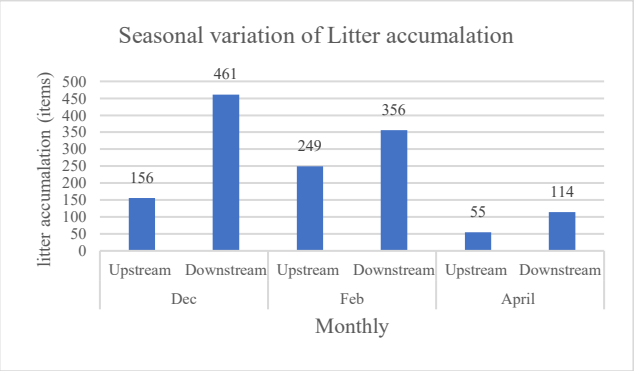
Fig 3. 3 The result of plastic litter Density in the upstream and the downstream

3.2 Seasonal variation of macro-plastic pollution by comparing litter accumulation and density during the dry season

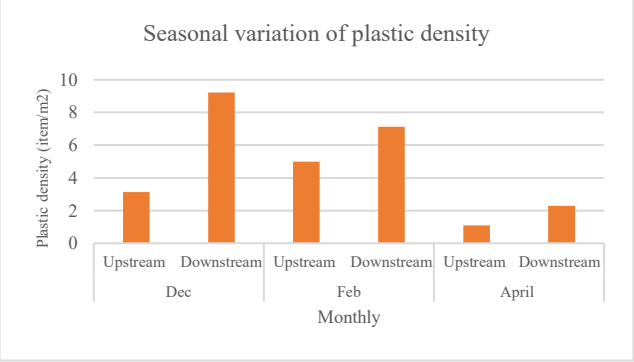
In the upstream area, litter accumulation increased from 156 items in December to 249 items in February, then decreased to 55 items in April. The initial rise in February could be due to wind-blown litter or dumping from passersby, as there are no nearby sources of regular human or agricultural activity. The sharp decrease in April aligns with the progression of the dry season when minimal runoff and human influence lead to reduced litter input. The downstream area, which is surrounded by houses, paddy fields, and a pond, showed a different pattern. Litter accumulation was highest in December at 461 items and dropped to 114 in April. The high levels of litter in December are likely due to daily human activities, agricultural runoff, and potential dumping near water bodies. Even during the dry season, when rainfall is minimal, downstream areas continue to receive litter through direct human activities or slow-moving surface flow from nearby land uses. The presence of the pond might also act as a temporary sink that traps and releases litter over time.

In the dry season, the upstream area displayed a fluctuating trend from December to April. In December, plastic density upstream was approximately 3.2 items/m²**.** This increased to about 5.1 items/m² in February, before dropping sharply to around 1.1 items/m² in April. These changes occurred despite the absence of nearby human settlements, paddy fields, or ponds in the upstream area. The increase in February may be attributed to wind-blown plastic or occasional waste from visitors or upstream activities, even though no consistent local sources of plastic are present. The decrease by April aligns with the peak of the dry season, when

limited rainfall and surface runoff reduce the transport of plastic waste into the stream. In the downstream area, plastic density was consistently higher due to the presence of nearby houses, paddy fields, and a pond—all potential sources of plastic waste. In December, the density was highest at 9.2 items/m², indicating heavy input from human and agricultural activities. This decreased to 7.2 items/m² in February, possibly due to slightly reduced runoff or collection efforts, but remained high because of continuous local waste input. By April, plastic density further declined to 2.6 items/m², likely influenced by the peak of the dry season, when the absence of rain limits surface runoff and waste transport. However, despite the seasonal drop, downstream plastic density in April was still more than double that of the upstream area, highlighting the persistent pressure from nearby human activities.



(a) Seasonal variation of litter accumulation of riverbank plastic pollution



(b) Seasonal variation of plastic density of riverbank plastic pollution

Fig 3. 4 (a) (b) The result of Seasonal variation of riverbank macroplastic pollution in Kampong Khleang River.

4. CONCLUSIONS

This study indicated the growing concern about macroplastic pollution in the Kampong Khleang River, located within the Siem Reap Province, Cambodia. Field observations and data collection in this area revealed a significant accumulation of macroplastics, particularly single-use items such as plastic bottles, bags, and food packaging. The findings show that downstream areas consistently exhibited higher waste densities than upstream, especially in December and February. This pattern suggests that waste is either accumulating as it moves downstream or being introduced along the way. Although upstream areas recorded a slightly higher density in April, overall, downstream regions are more severely impacted by plastic pollution. These pollutants not only threaten local biodiversity but also compromise the safety of food sources and the sustainability of community life.

ACKNOWLEDGMENTS

This work was supported by the ASEAN-Norwegian Cooperation Project on Local Capacity Building for Reducing Plastic Pollution in the ASEAN Region (ASEANO Phase 2).

REFERENCES

[1] M. Solieng, “World Geographical Dictionary On Cambodia Prepared by,” no. May 2013, pp. 1–33, 2013.

[2] UNDP, “UNITED NATIONS DEVELOPMENT PROGRAMME ANNUAL REPORT 2018,” 2018.

[3] UNDP, “Project Title : Combatting Marine Plastic Litter in Cambodia Implementing Partner : National Council for Sustainable Development Start Date January 2021 End Date : July 2023 L PAC Meeting Date : 10 December 2020 Brief Description The project aims to prevent,” no. December 2020, 2023.

[4] R. Hurley *et al.*, “Measuring riverine macroplastic: Methods, harmonization, and quality control,” *Water Res.*, vol. 235, p. 119902, 2023, doi: https://doi.org/10.1016/j.watres.2023.119902.

[5] S. Kirschke, T. H. M. Van Emmerik, S. Nath, and C. Schmidt, “Barriers to plastic monitoring in freshwaters in the Global South,” vol. 146, no. May, pp. 162–170, 2023.

[6] M. C. W. Haberstroh, Charlotte J, Mauricio E. Arias, Zhewen Yin, “Effects of hydrodynamics on the cross-sectional distribution and transport of plastic in an urban coastal river,” pp. 0–2, 2020, doi: 10.1002/wer.1386.

[7] C. Munari *et al.*, “Temporal variation of floatable

plastic particles in the largest Italian river, the Po," *Mar. Pollut. Bull.*, vol. 171, no. August, p. 112805, 2021, doi: 10.1016/j.marpolbul.2021.112805.

[8] D. Morritt, P. V Stefanoudis, D. Pearce, O. A. Crimmen, and P. F. Clark, "Plastic in the Thames : A river runs through it," *Mar. Pollut. Bull.*, vol. 78, no. 1–2, pp. 196–200, 2014, doi: 10.1016/j.marpolbul.2013.10.035.

[9] C. J. Van Calcar and T. H. M. Van Emmerik, "Abundance of plastic debris across European and Asian rivers Abundance of plastic debris across European and Asian rivers," 2019.

[10] R. Crosti, A. Arcangeli, I. Campana, M. Paraboschi, and D. González, "' Down to the river ': amount, composition, and economic sector of litter entering the marine compartment, through the Tiber River in the Western Mediterranean Sea," *Rend. Lincei. Sci. Fis. e Nat.*, no. 0123456789, pp. 1–8, 2018, doi: 10.1007/s12210-018-0747-y.

[11] P. Tasseron, H. Zinsmeister, L. Rambonnet, A. F. Hiemstra, D. Siepman, and T. van Emmerik, "Plastic hotspot mapping in urban water systems," *Geosci.*, vol. 10, no. 9, pp. 1–11, 2020, doi: 10.3390/geosciences10090342.

[12] T. Van Emmerik, P. Vriend, and E. C. Peereboom, "Roadmap for Long-Term Macroplastic Monitoring in Rivers," vol. 9, no. February, pp. 1–8, 2022, doi: 10.3389/fenvs.2021.802245.

[13] D. González-fernández and G. Hanke, "Toward a Harmonized Approach for Monitoring of Riverine Floating Macro Litter Inputs to the Marine Environment," vol. 4, no. March, pp. 1–7, 2017, doi: 10.3389/fmars.2017.00086.

[14] T. Van Emmerik *et al.*, "A Methodology to Characterize Riverine Macroplastic Emission Into the Ocean," vol. 5, no. October, pp. 1–11, 2018, doi: 10.3389/fmars.2018.00372.

[15] J. Martins and P. Sobral, "Plastic marine debris on the Portuguese coastline : A matter of size ?," *Mar. Pollut. Bull.*, vol. 62, no. 12, pp. 2649–2653, 2011, doi: 10.1016/j.marpolbul.2011.09.028.

[16] N. A. S. Abreo, "Anthropogenic Marine Debris (AMD) in Mangrove Forests of Pujada Bay, Davao Oriental, Philippines," pp. 38–53, 2020, doi: 10.21463/jmic.2020.09.1.03.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Monitoring Macroplastic Accumulation in the Sangker River: Distribution Patterns and Pollution Hotspots

Phoungsy PHEAP [1*], Davin SANG [1,2], Rina HEU [1,3], LeangGek MENH [1], Solida SIM [1], Vouchlay THENG [1,4], Kong CHHUON [1], Pinnara KET [1], Vilde Kloster SNEKKEVIK [5], Rachel HURLEY [5]

[1] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *University-Industry Linkage, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[4] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[5] *Norwegian Institute for Water Research (NIVA), oslo, Norway*

[*]*Corresponding author: pheapphoungsy@gmail.com*

Abstract: *Macroplastic pollution is an increasing environmental concern in freshwater ecosystems, contributing to habitat degradation and posing risks to aquatic life. This study presents a comprehensive monitoring approach to assess macroplastic contamination in the Sangker River at three locations: upstream, midstream, and downstream. Three methods—visual sampling, net sampling, and riverbank surveys—were employed to evaluate spatial distribution patterns, plastic types, and pollution hotspots. Monitoring was conducted over two periods, in December 2024 and February 2025. The results indicate that visual sampling and net sampling did not detect floating or suspended plastics during the observation periods, suggesting that macroplastics were not significantly present in the water column. However, the riverbank survey provided substantial evidence of macroplastic pollution, with notable variations across locations and monitoring periods. In December 2024, the highest accumulation was observed downstream, with 121 plastic bags and 450 cigarette butts, whereas upstream recorded 40 single-use plastic cups. By February 2025, plastic pollution remained prevalent, with 48 plastic bags and 47 food packaging items recorded upstream, while downstream recorded 35 food packaging items. These findings indicate that macroplastic pollution primarily accumulates along riverbanks rather than floating in the water. This study highlights the need for targeted mitigation efforts to address macroplastic accumulation in riverine environments. The findings contribute to the growing body of research on freshwater plastic pollution and provide valuable data to support policy decisions for effective waste management and pollution control strategies.*

Keywords: Macroplastic Pollution, River Monitoring, Visual Sampling, Net Sampling, Riverbank Survey, Freshwater Contamination

1. INTRODUCTION

Plastic pollution in aquatic environments presents significant risks to ecosystem health and human livelihoods [1,2]. Rivers, as major conduits, are critical in transporting plastic waste from land to marine environments, making freshwater systems essential areas for pollution studies [3]. Macroplastics, defined as plastic items larger than 5 mm, can harm aquatic organisms through ingestion and entanglement, degrading habitats and disrupting ecosystems [4].

In Cambodia, plastic waste generation is rapidly increasing due to urbanization, population growth, and inadequate waste management infrastructure [5]. The Sangker River, which flows through Battambang Province—a major urban and tourism hub—plays a pivotal role in transporting plastic waste to Tonle Sap Lake. Human activities, including improper disposal, fishing, and waste discharge from floating villages, significantly contribute to macroplastic pollution. Seasonal flooding further exacerbates this issue by mobilizing land-based plastics into the river [6].

While research on plastic pollution in Southeast Asia is growing, there remains a significant gap in studies targeting macroplastic accumulation in Cambodian rivers. Understanding spatial patterns, accumulation hotspots, and

[1]Corresponding author: Phoungsy PHEAP
E-mail: pheapphoungsy@gmail.com; Tel: +855-96 686 9813

pollution sources is vital for developing targeted mitigation strategies [7]. This study aims to quantify the influx of macroplastics into Tonle Sap Lake via the Sangke River using three methods: upstream, midstream, and downstream in the Sangker River, employing visual sampling, net sampling, and riverbank surveys conducted at upstream, midstream, and downstream locations during three monitoring periods (December 2024, February 2025 and April 2025). It will investigate how seasonal variations in water depth affect the transport and accumulation of plastic waste, as well as identify the main sources of pollution. The findings will provide valuable insights to inform policy development and improve waste management strategies, with the aim of reducing macroplastic pollution and protecting Cambodia's freshwater ecosystems.

2. METHODOLOGY

2.1 Study area

The Sangker River, located in Battambang Province, Cambodia, flows into Tonle Sap Lake. It is characterized by a mix of urban, peri-urban, and rural landscapes. Three sampling sites were selected to represent different sections of the river: upstream (less urbanized area), midstream (densely populated and active area), and downstream (near the river's confluence with Tonle Sap Lake).

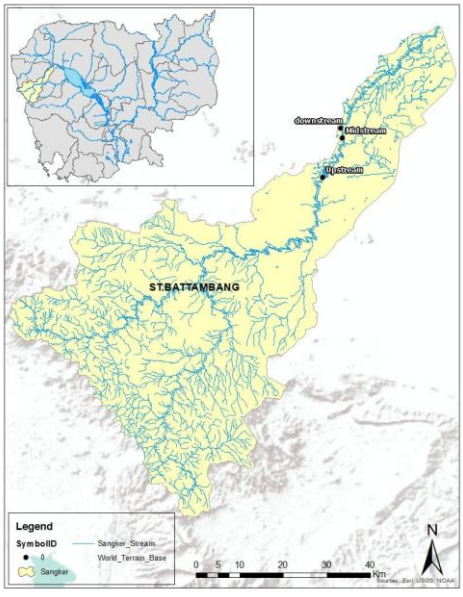


Fig. 1. *Data collection in Sangker River.*

2.2 Method selection and Harmonization Approach

The monitoring methods used in this study were selected based on the Global Plastic Pollution Observatory (GPPO) Methods, Harmonisation, and Quality Control framework [8]. This framework promotes standardized, repeatable, and scalable techniques for plastic monitoring across river systems to enhance data comparability and quality control.

To ensure consistency with international protocols, we employed a combination of three complementary methods:

- Visual surface observation for floating plastics
- Net-based sampling for the physical collection of floating macroplastics
- Riverbank surveys for assessing stranded macroplastic accumulation

Each method targets different compartments of the riverine plastic transport system (surface water and bank accumulation) while adhering to standardized protocols.

2.3 Sampling Protocols

2.3.1 Visual observation of floating Macroplastics

Visual observations were conducted from elevated fixed points (bridges) according to GPPO protocols. Observers scanned the river surface within a defined width for 15-minute intervals. Only macroplastic items larger than 2.5 cm were recorded, consistent with the minimum size thresholds recommended by van Emmerik et al. (2023). Observations were repeated three times per site on each monitoring day.

2.3.2 Net sampling of floating Macroplastics

A 0.5 m × 0.5 m rectangular net with a 2 mm mesh size was deployed in the river current for 15-minute intervals to collect floating macroplastic debris. The net was suspended from bridges to avoid interference from riverbanks, as recommended by standardized methods [8]. Flow velocity was measured using a portable flow meter to calculate the sampled water volume. Sampling was replicated three times at each site.

2.3.3 Riverbank survey of floating Macroplastics

Riverbank plastic accumulation was assessed using 5 m × 10 m survey plots at each site. Surveys followed GPPO harmonized guidelines, which recommend standardizing plot sizes, categorizing debris types based on a harmonized typology, and recording both item counts and weights [8]. Each site was surveyed twice, during two different seasonal conditions (December 2024, February 2025, and April 2025).

2.4 Quality control measures

To ensure consistency and minimize observer bias, the following quality control measures were implemented:

- Observers underwent training sessions based on GPPO visual identification manuals before field surveys.
- Standardized data sheets were used for recording all observations and collections.
- Duplicate surveys were conducted for 10% of riverbank plots to assess inter-observer reliability.
- All collected plastic debris was photographed, labeled, and stored for later verification.

2.5 Data Analysis

Plastic litter density (PLD) and plastic litter composition (PLC) were calculated using standardized equations proposed in GPPO methods [8]. Spatial and seasonal variations in macroplastic density and composition were analyzed descriptively across sites and monitoring periods.

$$\text{PLD} = \frac{\text{Number of items}}{\text{Total area sampled (L×W)}} \quad \text{(Eq 1)}$$

PLD: Plastic Litter Density (Item/m^2)

L: Length of area (m)

W: Width of area (m)

The Plastic litter Density (PLC) of collected macroplastic were;

$$PLC = \frac{Number\ of\ items\ per\ caterggory}{Total\ number\ of\ items\ in\ all\ catergory} * 100 \quad \text{(Eq 2)}$$

The number of plastic items in a river can be determined through a visual counting method, often conducted from bridges, where observers count the floating litter within a defined area. This method is used to estimate the amount of macroplastic in rivers and can be combined with other sampling methods like grab sampling on riverbanks and in the water. Anything man-made such as fishing gear, plastic bags, beverage bottles, balloons, food wrappers, and even vessels can become marine debris through dumping, improper waste management, litter that is blown or washed out to sea through storm drains, and extreme natural events which can transport both small and large items into the ocean.

3. RESULTS AND DISCUSSION

3.1 Environmental conditions and impact on plastic distribution

During the study periods (December 2024 and February 2025), the Sangker River experienced moderate seasonal fluctuations in water levels and flow, with the highest flow observed in the downstream region. In December, following the dry season, water levels were higher, especially downstream, which may have carried more plastic waste downstream. In contrast, February 2025 saw more stable flow conditions. The weather was relatively dry during both monitoring periods, with temperatures ranging from 24°C to 30°C. These environmental conditions likely contributed to higher accumulation of plastics on riverbanks, particularly downstream, where water levels fluctuated more significantly.

Table 1. *Environmental conditions*

Date	Location	Flow (m^3/s)	Weather	Distance Bridge to surface (m)	Obs. Height (m)
Dec 15, 2024	Upstream	0.04725	Sunny	10	12
	Midstream	0.07617	Cloudy	4.9	7.5
	Downstream	0.057	Sunny	10.7	11
Feb 18, 2025	Upstream	0.03623	Sunny	10	12
	Midstream	0.05256	Sunny	4.9	7.5
	Downstream	0.64906	Sunny	9.7	11
Apr 03, 2025	Upstream	0.2031	Sunny	10.1	12
	Midstream	0.199	Sunny	5.2	7.5
	Downstream	0.1324	Sunny	9.53	11

3.2 Macroplastic Density and Spatial Distribution

Macroplastic densities along the Sangker River varied spatially and temporally across the three sampling locations (upstream, midstream, and downstream) and three monitoring periods (December 2024, February 2025, and April 2025). The highest density of macroplastics was recorded upstream during February 2025, with 45.2 items/m^2, while the midstream site exhibited the lowest density, with 3.36 items/m^2 (Fig.2). In April 2025, macroplastic densities slightly decreased at all locations but remained highest downstream.

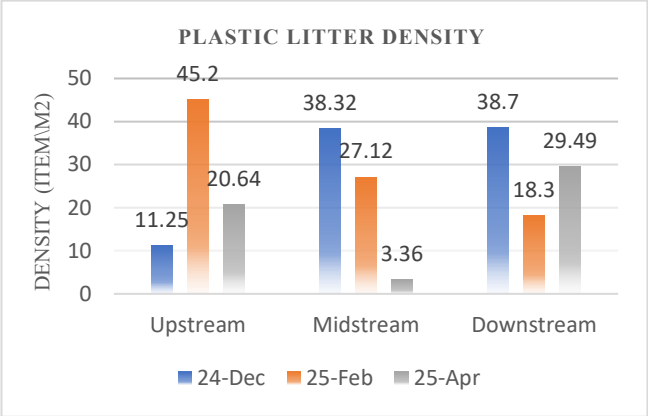


Fig. 2. *Macroplastic density (items/m²) at each sampling location and period*

These results suggest that downstream areas accumulate more macroplastic debris, likely due to hydrodynamic processes that transport waste from urbanized and populated midstream regions toward the river month.

3.3 Composition of collected Macroplastics

The composition analysis revealed that plastic bags, food packaging, and cigarette butts were the dominant types of macroplastics collected during both monitoring periods (Fig. 3). In December 2024, plastic bags represented (5%) of the collected items, cup (45%) and cup lip (9%), followed by food packaging (1%), plastic straw (6%), pastic bottle (7%) cigarette butts –plastic_bag fargment –cigarette box (0%) and other (27%) of the collection itmes.

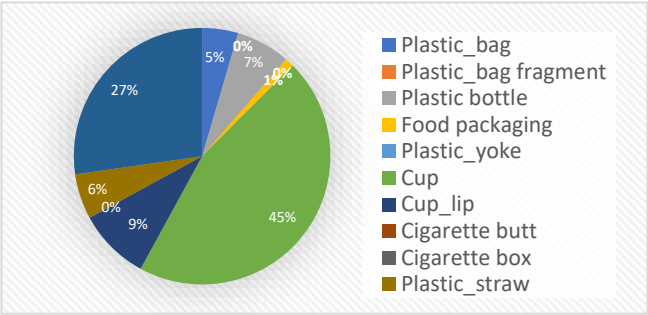


Fig. 3. *Composition of macroplastics by category (%) Dec-2024.*

(Fig.4) In February 2025, plastic bag respresented (20%) of the collection items, plastic bottle (7%), cigarette but (10%), cup plastic-straw and other (8%), follow by food packaging and plasticc_bag fragment (15%)of the collection itmes.

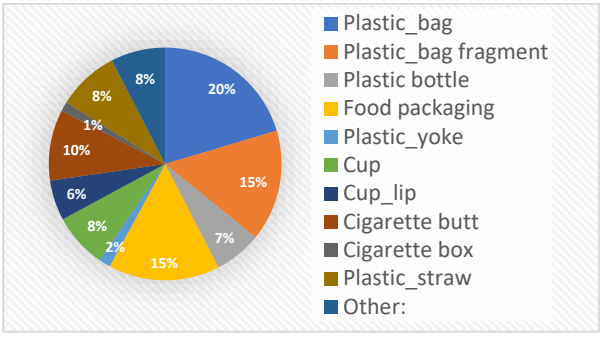


Fig. 4. *Composition of macroplastics by category (%) Feb-2025.*

(Fig.5). In April 2025, plastic straw and plastic bottle (12%) of the collection items, plastic_bag (18%) and plastic_bag fargment (2%), cup (23%), cup lip (7%), followed by food pakaging (19%), plastic yoke and other (3%) of the collection itmes.

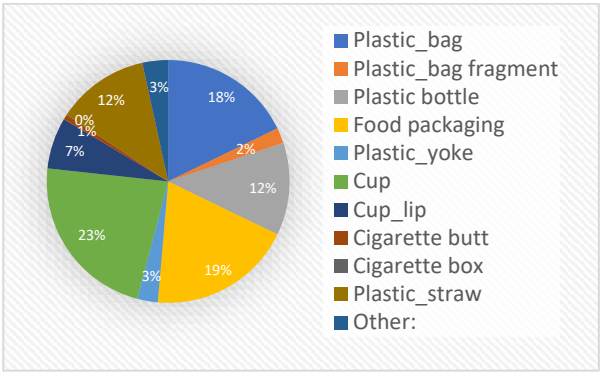


Fig. 5. *Composition of macroplastics by category (%) April-2025.*

The predominance of plastic bags and single-use items highlights the urgent need for targeted waste management policies focusing on these plastic types in Battambang province.

3.4 Net Sampling and Visual survey results

During the net sampling campaigns at all three locations, no macroplastics were detected either visually or through net retrievals. Volumes filtered during net sampling ranged from 3.97 to 19.2 m³ per deployment (Table 2). Similarly, visual sampling from bridges, conducted three times daily at each site, did not detect any floating macroplastic objects larger than 2.5 cm.

Table 2. *Net sampling Results showing volume filtered and plastics detected*

Location	**Sampling Date**	**Volume Filtered (m³)**	**Plastics Detected**

Upstream	Dec 2024	9.0	0
Midstream	Dec 2024	8.5	0
Downstream	Dec 2024	10.0	1
Upstream	Feb 2025	10.8	0
Midstream	Feb 2025	15.0	0
Downstream	Feb 2025	19.2	0
Upstream	April 2025	6.0	0
Midstream	April 2025	5.7	0
Downstream	April 2025	3.97	1

The absence of floating plastics aligns with findings from previous studies [9], which highlight that in many river systems, macroplastics tend to accumulate along riverbanks and floodplains rather than remaining suspended in the water column. Additionally, riverine hydrology and flow conditions during the dry season may promote sedimentation and entrapment of plastics in vegetation or riverbanks.

3.5 Temporal variation in Macroplastic pollution

A comparison between the three monitoring periods showed a slight decrease in macroplastic densities from December 2024, February 2025 to April 2025 (Fig.2). This reduction may be attributed to seasonal effects, including lower river flows and human activities. February 2025 coincides with higher tourism and festive activities in Battambang, which could contribute to increased waste generation.

Moreover, the characteristics of plastic debris (predominantly lightweight single-use plastics) suggest that seasonal flooding and high flows during the dry season may mobilize and redistribute macroplastics further downstream or on to floodplains.

3.6 Pollution hotspots identification

Based on riverbank surveys, the downstream site consistently showed higher macroplastic accumulation compared to midstream and upstream sites. This finding indicates that the lower stretch of the Sangke River acts as a pollution hotspot. The proximity of the downstream site to denser urban settlements and floating villages likely contributes to higher levels of mismanaged waste entering the river system. Identification of these hotspots is crucial for prioritizing mitigation efforts, including community engagement, enhanced waste collection services, and localized cleanup initiatives.

The Sengker River contributes significantly to the transport of macroplastics to Tonle Sap Lake, primarily through hydrological connectivity, runoff during rainy seasons, and human activities that lead to improper waste disposal. This transport poses serious risks to the lake's ecosystem, water quality, and local economies. Effective strategies are needed to mitigate these impacts and protect both the river and lake environments for future generations.

4. CONCLUSIONS

This study provides valuable insights into the spatial distribution in macroplastic pollution in the Sangker River. Our findings show that downstream sections of the river accumulate the highest plastic debris, mainly due to hydrodynamic transport and seasonal flooding. The results underscore the need for a multi-pronged approach to address plastic pollution, which includes improving waste management systems, targeting high-pollution hotspots, and promoting local awareness campaigns. Future research should continue to monitor macroplastic accumulation in this river system and expand the monitoring network to other rivers in Cambodia to assess broader trends in plastic pollution. Macroplastics significantly disrupt ecosystems, harm wildlife, pose potential health risks to humans, and impact economic activities. Addressing the issue of macroplastic pollution is crucial for protecting environmental integrity and public health.

ACKNOWLEDGMENTS

This work was supported by ASEAN-Norwegian Cooperation Project on Local Capacity Building for Reducing Plastic Pollution in the ASEAN Region (ASEANO Phase 2).

REFERENCES

[1] J. R. Jambeck, et al., "Plastic waste inputs from land into the ocean," *Science*, vol. 347, no. 6223, pp. 768-771, 2015.

[2] L. C. M. Lebreton, et al., "River plastic emissions to the world's oceans," *Nature Communications*, vol. 8, p. 15611, 2017.

[3] T. Van Emmerik and A. Schwarz, "Plastic pollution in rivers: A global perspective," *Environmental Science & Technology*, vol. 54, no. 23, pp. 14227-14240, 2020.

[4] F. M. Windsor, et al., "Impacts of marine plastic pollution on freshwater systems," *Science Advances*, vol. 5, no. 7, p. eaaw6279, 2019.

[5] World Bank, "Cambodia: Addressing the plastic waste challenge," World Bank Report, 2022. [Online]. Available: https://www.worldbank.org/en/news/feature/2022/cambodia-plastic-waste

[6] A. Sato, et al., "Seasonal variations in riverine plastic pollution: A case study of Southeast Asia," *Environmental Pollution*, vol. 268, p. 115500, 2021.

[7] R. A. Pettipas, et al., "Assessing plastic pollution in rivers: Methods and standards," *Science of the Total Environment*, vol. 554-555, pp. 155-166, 2016.

[8] T. Van Emmerik, et al., "Harmonized methods for monitoring macroplastics in rivers: A guide to standardized data collection," *Global Plastic Pollution Observatory (GPPO)*, 2023.

[9] A. González, et al., "Floating plastics in freshwater environments: A global review," *Environmental Pollution*, vol. 276, p. 116602, 2023.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Pesticide Contamination in Groundwater of Agriculturally Intensive Regions of Cambodia: A Case Study of Koh Thum and Kanghot

Vuthy Ruos [1], Khy Eam Eang [2*], Chanvorleak Phat [2], Sereyvath Yoeun [2], Mathilde Sester [3, 4], Paul Baudron [5]

[1] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *CIRAD, Aïda, Univ Montpellier, Montpellier, France*
[4] *CIRAD, UPR Aïda, Institut Technologique du Cambodge, Phnom Penh, Cambodia*
[5] *UMR G-EAU, French Institute for Sustainable Development, University of Montpellier, France*

[*] *Corresponding author: khyeam@itc.edu.kh*

Abstract: *Pesticides are any substances or mixture of substances used for preventing, eliminating, repelling, reducing, and controlling weeds, insects, and other organisms in a wide variety of agriculture and non-agriculture. The use of pesticides in intensive agricultural areas of Cambodia poses significant risks of pollution of groundwater resources, highlighting the need for rigorous monitoring to protect public health and preserve ecosystems. This study investigates pesticide contamination in groundwater from Koh Thum District, Kandal Province, and Kanghot area, Battambang Province; groundwater is somehow used for domestic purpose and agriculture in both regions. Ten groundwater samples from Koh Thum and ten from Kanghot were collected from tube wells and push points, representing a variety of local hydrological conditions. The samples were processed using solid-phase extraction (SPE) and analyzed by gas chromatography-mass spectrometry (GC-MS) equipped with an automated identification and quantification system (451-compound database) at the PAER Laboratory, Institute of Technology of Cambodia (ITC). Results revealed detectable residues of atrazine (herbicide), parathion and chlorfenapyr (insecticides), mefenoxam and azoxystrobin (fungicides) in the groundwater samples from Koh thum while mefenoxam (fungicide), and chlorfenapyr and spirodiclofen (insecticides) were detected in the groundwater samples from Kanghot. The presence of these pesticides highlights potential contamination from agricultural runoff, raising concerns about chronic exposure risks for communities. This study could remark the urgency for systematic groundwater quality monitoring, stricter regulatory enforcement, and sustainable farming practices to mitigate environmental and health impacts. The findings contribute to regional pesticide management strategies and emphasize the need for targeted interventions in Cambodia's agriculturally intensive areas.*

Keywords: GC-MS; Intensive Agriculture; Pesticides; SPE

1. INTRODUCTION

Pesticides are defined as chemical substances that are mainly used in agriculture in order to protect, destroy, eliminate or mitigate any pests or other diseases that cause damage to crops [1]. Pesticides are classified in many ways by their toxicity, what pests they target, how they work, and their chemical structure. People are most commonly exposed to pesticides through the food they eat, but these chemicals can also enter the body through the skin, mouth, eyes, or by inhaling [2]. In Cambodia's agricultural sector, pesticide use is on the rise, particularly in the production of rice and vegetables. More than 3.2 million liters of pesticides, containing more than 100 different types of compounds, are used in agriculture in Cambodia every year [3].Water contaminated with pesticides is known to cause a range of other illnesses like asthma, dermatitis, as well as vomiting, coma, and even death. The hazard of pesticide was classified into extremely hazardous (Ia), highly hazardous (Ib), moderately hazardous (II), slightly hazardous (III), and unlikely to present acute hazard (U). Pesticides with organic constituents typically have more complex chemical structures, such as carbamate insecticides, organophosphorus insecticides, and chlorohydrocarbon insecticides. Insecticides, fungicides, herbicides, and rodenticides are among the functional groups that are categorized according to the specific pest that they are designed to control [4].

[1]* Corresponding author: Khy Eam Eang
E-mail: khyeam@itc.edu.kh; Tel: +855-12 311 110

Pesticides are mobile because they drift to new regions in the form of aerosols, through surface runoff, leaching, and infiltration processes, to reach waterways and cause contamination [5], as illustrated in **Fig. 1**. Therefore, the objective of this study is identifying the presence of pesticide in the groundwaters in both regions.

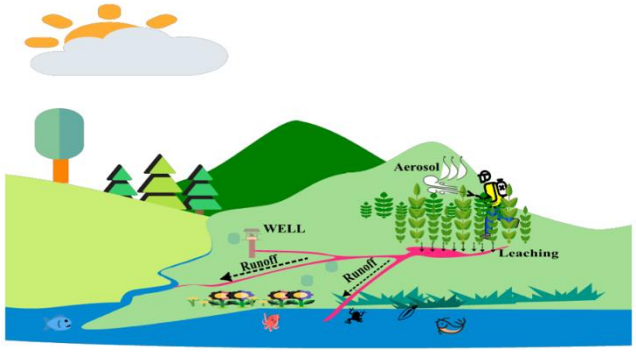


Fig. 1. Ways in which pesticides spread into environment

2. METHODOLOGY

2.1 Site description and sample collection

The study of pesticide residue analysis from Koh Thum in Kandal Province and Kanghot in Battambang Province (**Fig. 2**), was conducted on November 23-24, 2024. Twenty groundwater samples were collected in duplicate and stored in 1 L polypropylene bottles. All samples were well labelled with information such as sampling site, type of sample, and sampling date. The samples were filtrated on-site through 2 µm glass microfiber filters (Whatman, GF/CTM) in order to remove suspended solids and debris, stored in iced box with temperature of approximately 4 ºC and transported to PAER lab of the ITC.



Fig. 2. Mapping of sampling sites in (a) Koh Thom and (b) Kanghot

2.2 Chemical reagents and consumables

The sodium phosphate buffer solution (pH=7.0, 1 mol/L), dichloromethane (ACS BASIC), acetone, n-hexane, and acetonitrile were HPLC grade purchasing from RCI Labscane Limited. Pure water (PW) was extracted from Puric-α ultra-pure water Technology (Organo, Japan). Sodium sulfate anhydrous was purchased from Sigma-aldrich, Inc. InertSep (PLS-3) and InertSep Slim-J (AC2) were purchased from GL Sciences, Japan.

2.3 Sample preparation

The solid-phase extraction (SPE) was performed after complete filtration of the water samples (fiberglass filter, 0.45 µm). The SPE first started from the sorbent PLS3 cartridge attached with AC cartridge. They were activated by conditioning with adding 5 mL of dichloromethane, 5 mL of acetone, and 5 mL of purified water, two times subsequently. The purpose of the conditioning step was to ensure that the cartridge was wet and clean. At that time, 1 mL of sodium phosphate buffer solution (1 mol/L, pH 7.0) was added to each liter of filtrated water sample to adjust pH, then it was shaken to homogenize the solution and left to settle for 1h. Then the samples were passed through the cartridges by a vacuum pump at a flow rate of approximately 15 mL/min. The set of cartridges were dried under evaporation with a nitrogen gas stream for 40 min. Next, the set of dried cartridges were disassembled, and wash the PLS3 cartridge with 2 mL acetone and 5 mL dichloromethane. Also, add 5 mL acetone to wash the set of PLS3 attached to the AC cartridge by eluting and switched on the pump at the end. After that, the mixed eluted solvent in the glass tube was concentrated with a nitrogen gas stream for approximately 1h until 1 mL of concentration volume. Then, 10 mL of n-hexane was added to the 1 mL concentrated solution in the previous step and dehydrated by eluted passing through the sodium sulphate (Na_2SO_4) as the filter. The solution obtained was concentrated again to concentration volume 1 mL using nitrogen gas stream, and finally transferred into the vial; stored at -20 ºC subjected to GC-MS analysis as show in **Fig. 3.**

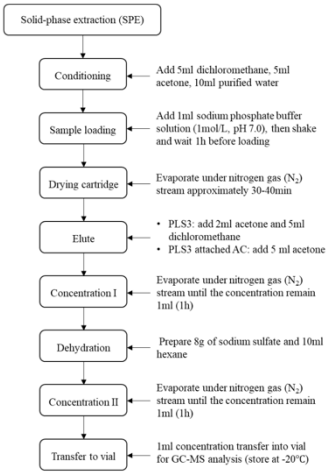


Fig. 3. Experimental flow chart of solid phase extraction

2.4 GC-MS analysis

Gas chromatography and mass spectrometry (GC-MS) equipment, model TQ8040 series (Shimadzu, Japan), was used to screen pesticide chemicals. The MS database contained information on 451 pesticide compounds. With a length of 30 meters, a thickness of 0.25 meters, and a diameter of 0.25 millimeters, the DB-5ms was employed as the column. Samples were injected into the GC-MS system using a 1 microliter volume in spitless injection mode. The temperature of the column oven was first set at 4°C for two minutes before being raised to 310°C for five minutes at a rate of 8°C/min. The flow rates for the GC column were 1.23mL/min and 50mL/min, respectively, for the ultra-pure helium that served as the carrier gas. The contact was 300°C and the ion source was 200°C in temperature. The samples were analyzed in scanned mode in the mass rang of 33-600 m/z. The identification of pesticide residue was performed by the data treatment system, which calculates the monoisotopic mass, predicts the structural formula of compounds and compares them using the MS database.

Table 1. GC-MS Analytical Conditions for Pesticide Screening

Parameter	Condition
Instrument	Shimadzu TQ8040 GC-MS
Column	DB-5ms (30 m × 0.25 mm ID × 0.25 µm film thickness)
Injection Volume	1 µL
Carrier Gas	Helium (Ultra-pure)
Flow Rate	1.23 mL/min (GC column), 50 mL/min (Purge flow)
Oven Temperature Program	- Initial: 40°C (hold 2 min) - Ramp: 8°C/min to 310°C (hold 5 min)
MS Interface Temp.	300°C
Ion Source Temp.	200°C
Mass Scan Range	33–600 m/z (Scan mode)
Database	451 pesticide compounds (Monoisotopic mass, structural formula comparison)

2.5 Database software for simultaneous analysis

The GC-MS works with Automated Identification and Quantification System with a Database (AIQS-DB). This database (AIQS-DB) was designed by professor Kiwao Kadokami. This database contains the mass spectra, retention times, and calibration curves for about 1,000 substances including 450 compounds of pesticide, 194 compounds of CH, 150 compounds of CHO, 113 compounds of CHN(O), 14 compounds of PPCPs, 12 compounds of CHS(NO), 8 compounds of CHP(NOS); permitted simultaneous identification and quantification of about 1,000 plus substances without the use of chemical standards.

2.6 Calibration for analysis

A 10 μl of IS (Internal Standard) was added into each samples solution before conducted the GC-MS measuring. In addition, the external standard solutions were prepared 0.025ppm, 0.05ppm,0.1ppm, 0.125ppm, 0.5ppm and 1ppm which diluted from stock standard solution 10 ppm in order to identified the concentration of 40 target pesticide compounds while other pesticide compounds were detected in a semi-quantitative analysis.

3. RESULTS AND DISCUSSION

3.1 Physical-chemical parameters

The water quality results in **Table 2** from Koh Thom and Kanghot show significant variations in electrical conductivity (EC), pH, temperature, depth, and water level, which may reflect differences in groundwater contamination, geology, and agricultural influence. In Koh Thom, extremely high electrical conductivity (EC) values were recorded, particularly in shallow wells W01_1 (9210 µS/cm); W01_2(6200 µS/cm), suggesting severe salinity intrusion or contamination from agricultural runoff containing fertilizers and pesticides [6]. Moderately elevated EC levels in other Koh Thom wells (653–2320 µS/cm) indicate varying degrees of aquifer pollution. In contrast, Kanghot exhibited lower EC values (722–1371 µS/cm), reflecting generally better water quality, though the elevated reading in W09 (1371 µS/cm) points to localized contamination. The pH data shows near-neutral to alkaline conditions (Koh Thom: 7.15–7.97; Kanghot: 7.5–8.4), with Kanghot's higher alkalinity potentially slowing pesticide degradation and increasing persistence in groundwater. The temperature remained consistently warm (29–31°C) across all samples, which may accelerate both pesticide breakdown and leaching processes. Depth profiles reveal that shallow wells (<3 m) in Koh Thom have the highest EC, indicating greater vulnerability to surface contamination, while deeper wells (>25 m) in both regions generally show better water quality.

3.2 Pesticide residues detected in Koh Thum and Kanghot

Groundwater samples collected from the study areas revealed the presence of multiple pesticide residues. In this study area, 6 pesticides were detected and were classified into different classes based on target organisms and toxicity hazardous according to world health organization in **Table 3**. Specifically, the following active compounds were detected: Atrazine (herbicide), Azoxystrobin (fungicide), Chlorfenapyr (insecticide), Parathion (insecticide), Spirodiclofen

Table 2. Chemical and physical parameter of Groundwater

Location	Sample code	E.C(µS /cm)	pH	Temperature (℃)	Depth (m)	Water level (m)
Koh Thom	W01_1	9210	7.15	31.8		1.7
	W01_2	6200	7.39	29.7	60	20
	W02	1244	7.97	30.7		0.16
	W03	653	7.9	29.6	25	2.5
	W04	2320	7.93	30.8		2.4
Kanghot	W05	797	8.206	30.1	25	17.2
	W06	722	8.4	30	45	14
	W07	779	7.5	30.3	25	12.6
	W08	802	7.9	30.3	40	14.2
	W09	1371	7.7	29.8	67	19.35

(acaricide/insecticide), and Mefenoxam (fungicide). These pesticides represent different chemical classes and target a broad spectrum of pests, including weeds, fungi, mites, and insects. in **Fig. 4**, pesticides were detected which the main majority determined by fungicides (67%), insecticides (25%), and herbicides (8%). Moreover, **Fig. 5** illustrated that most of slightly hazardous (25%), followed by unlikely hazardous (17%). pesticides detected were moderately hazardous (58%), slightly hazardous (25%), followed by unlikely hazardous (17%).

Table 3. Toxicity and functional role classification of detected pesticide

N^0	Pesticide	Type	Toxicity Category
1	Atrazine	Herbicide	Moderately toxic
2	Azoxystrobin	Fungicide	Slightly toxic
3	Chlorfenapyr	Insecticide	Moderately toxic
4	Parathion	Insecticide	Highly toxic
5	Spirodiclofen	Acaricide/Insect	Slightly toxic
6	Mefenoxam	Fungicide	Moderately toxic

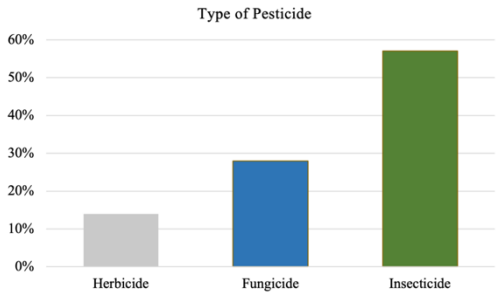


Fig. 4. Types of pesticides detected in Koh Thum and Kanghot

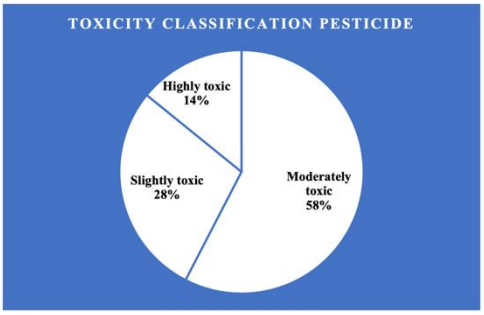


Fig. 5. Toxicity classes of pesticides

3.3. Pesticide analysis in groundwater

The GC-MS spectrum for Semi quantitative analysis of pesticide identification was injecting samples, and compared with the instrumental database and the NIST library version 14. The GC-MS semi quantitative studies provided the first indication of the presence of pesticides in the groundwater. The majority of the 451 compounds in the MS database, the Kandal province's Koh Thum district were beneficial and show pesticide contamination in this location. **Table 4** illustrates presence of atrazine, parathion, chlorfenapyr, mefenoxam, and azoxystrobin in the groundwater samples from Koh thum. chlorfenapyr spirodiclofen, mefenoxam were detected in the groundwater samples from Kanghot, the pesticide screening results from Koh Thom and Kanghot reveal concerning but variable groundwater contamination patterns. In Koh Thom, targeted detection was observed with atrazine present only in W01_2 and the highly toxic parathion exclusively in adjacent shallow well W01_1, demonstrating significant spatial variability even between proximate sampling points. Azoxystrobin was additionally detected in W02, indicating fungicide infiltration. Kanghot exhibited a different contamination profile, with chlorfenapyr found in W05 and W06, spirodiclofen in W05, and mefenoxam in W07, these findings suggest that Kanghot's pesticide profile differs both in terms of the specific compounds present and their distribution pattern compared to Koh Thom. atrazine was detected only in W01_2 (Koh Thom), it is possible localized agricultural runoff. Parathion was found only in

W01_1 showed no atrazine contamination but instead contained parathion, a highly toxic organophosphate insecticide that has been banned or restricted in many countries. This striking difference between adjacent wells suggests highly localized contamination patterns, possibly due to varying agricultural practices, well construction characteristics, or subsurface flow paths. A third detection in Koh Thom occurred at W02, which contained azoxystrobin, a modern fungicide used in crop protection, despite its proximity to W01_2, indicating point-source contamination rather than widespread leaching, Azoxystrobin (a fungicide) was detected in W02. a widely used fungicide, but no other pesticides were detected. but W08, and W09 had no pesticide detections, The detection of parathion, a highly toxic and often restricted insecticide, raises immediate public health concerns, particularly in communities relying on shallow groundwater. The variability between adjacent wells (W01_1 and W01_2) underscores the influence of well depth, localized farming practices, and potential point- source pollution in determining contamination risks[1]. Several wells in both regions (W03, W04, W08, W09) showed no detectable levels of the screened pesticides. This absence could indicate either genuine lack of contamination or pesticide concentrations below detection limits. The presence of multiple pesticides in some wells but not others highlight the patchy nature of groundwater contamination, likely reflecting variations in local agricultural practices, soil characteristics, and hydrogeological conditions [1,6]. The results demonstrate that pesticide contamination is occurring but varies significantly by location, suggesting that mitigation strategies may need to be tailored to specific local conditions and contamination patterns.

4. CONCLUSIONS

The detection of pesticide residues including atrazine, parathoxon, chlorfenapyr, mefenoxam, azoxystrobin, and spirodiclofen in groundwater samples from Koh Thum (Kandal Province) and Kanghot (Battambang Province) underscores the pervasive impact of agricultural practices on water quality in Cambodia. These findings highlight a critical environmental and public health concern, as groundwater serves as a primary water source for local communities. The presence of multiple pesticide classes (herbicides, insecticides, fungicides, and acaricides) suggests diffuse contamination from agricultural runoff, emphasizing the need for improved pesticide management and monitoring. The groundwater quality assessment in Koh Thom and Kanghot reveals significant pesticide contamination, with distinct spatial patterns between the two regions. Koh Thom exhibits higher vulnerability, particularly in shallow wells, where hazardous pesticides such as parathion (W01_1) and atrazine (W01_2) were detected despite their close proximity highlighting localized contamination sources and variable aquifer susceptibility. The presence of azoxystrobin (W02) further confirms fungicide infiltration in Koh Thom's groundwater. In contrast, Kanghot shows a more sporadic pesticide distribution, with chlorfenapyr (W05, W06), spirodiclofen (W05), and mefenoxam (W07) detected, suggesting either reduced agricultural chemical usage or more protective hydrogeological conditions.

Table 4. Pesticide detected in groundwater samples

Pesticide detected	Samples code									
	W01_1	W01_2	W02	W03	W04	W05	W06	W07	W08	W09
Atrazine	-	+	-	-	-	-	-	-	-	-
Azoxystrobin	-	-	+	-	-	-	-	-	-	-
Chlorfenapyr	-	-	-	-	-	+	+	-	-	-
Parathion	+	-	-	-	-	-	-	-	-	-
Spirodiclofen	-	-	-	-	-	+	-	-	-	-
Mefenoxam	-	-	-	-	-	-	-	+	-	-

ACKNOWLEDGMENTS

This extended abstract was produced with the financial support of "Laboratoire Mixte International: Laboratoire d'Excellence en co-Ingénierie des Agrosystems Durables" Project under IRD fund. The authors would like to deeply thank to all project members for their kind support and cooperation in every sampling campaign.

REFERENCES

[1] N. Khan, G. Yaqub, T. Hafeez, and M. Tariq, "Water and Life in Tonle Sap Lake (Chihiro Yoshimura, Rajendra Khanal etc.) (Z-Library).pdf," *J. Chem.*, vol. 2020, pp. 1–7, Jun. 2020, doi: 10.1155/2020/5497952.

[2] S. Kai, S. Yeoun, and C. Phat, "Analysis of Pesticide Residues in Sediment from Chhnok Tru, Kampong Chhnang Using Different Extraction Methods," 2020.

[3] P. Ngin, P. Haglund, S. Proum, and J. Fick, "Pesticide screening of surface water and soil along the Mekong River in Cambodia," *Sci. Total Environ.*, vol. 912, p. 169312, Feb. 2024, doi: 10.1016/j.scitotenv.2023.169312.

[4] *WHO Recommended Classification of Pesticides by Hazard and Guidelines to Classification, 2019 Edition*, 2nd ed. Geneva: World Health Organization, 2020.

[5] T. Mutunga, S. Sinanovic, and C. S. Harrison, "Integrating Wireless Remote Sensing and Sensors for Monitoring Pesticide Pollution in Surface and Groundwater," *Sensors*, vol. 24, no. 10, p. 3191, May 2024, doi: 10.3390/s24103191.

[6] L. Thourn, C. Phat, M. Soung, S. Sieng, S. Heng, and S. Yoeun, "Identification of Pesticide Contamination in Water Sources Surrounding Agrochemical-Free Rice Farming in Battambang Province," 2022.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Development of Coconut Filter to Remove Microplastics at Sea Salt Farm, Kampot Province, Cambodia

Sonarith Sam[1], Vouchlay Theng[1*], Chanthol Peng[2], Seyha Doeurn[2], Oudam Heng[2], Sovathana Phuong[2]

[1] *Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: vouchlay.theng@itc.edu.kh*

Abstract: *The increasing contamination of microplastics (MPs) in marine environments poses significant risks to aquatic ecosystems and human health, particularly in sea salt production, where seawater is the primary raw material. This study aimed to develop a sustainable coconut-based filter to remove MPs from seawater at sea salt farms. Utilizing the natural fibrous structure of coconut coir—an abundant, biodegradable, and renewable material—the filter was designed to efficiently capture MP particles while allowing seawater to pass through for salt production. The experiment was conducted in a batch glass column (10 cm in width, 10 cm in length, and 40 cm in height). Coconut coir was cleaned and dried at 60 °C/24h in an oven before being placed into the column, where gravel 10 cm and sand 4 cm served as supporting layers. The optimal conditions of coconut coir were determined by comparing two coconut coir lengths: short fiber (2–5 cm) and long fiber (6–29 cm). After selecting the coconut coir length, the density of coconut coir varied from 37 g/dm^3 to 148 g/dm^3, and the thickness varied from 3 to 5 cm. Key parameters were measured to evaluate filtration efficiency, including MP removal efficiency, salinity concentration, and effluent flow rate. The results showed that the salinity did not decrease in all experiments. MP removal efficiency of short fiber (90 %) was higher than that of the long fiber (45%). The optimal coconut coir density was obtained at 111.11 g/dm^3, with a thickness of 3 cm, a removal efficiency higher than 90 %, and an effluent flow rate of 0.045 dm^3/s. These findings highlight the potential of coconut coir, particularly short fibers, as an effective and eco-friendly solution for mitigating microplastic contamination in sea salt. This study provides a promising approach to enhancing environmental sustainability and food safety in marine-based industries.*

Keywords: Coconut coir; Coconut filter; Microplastic; Sea salt farm

1. INTRODUCTION[1]

Microplastics (MPs, <5 mm) are pervasive pollutants in marine ecosystems, originating from plastic degradation, synthetic textiles, and industrial waste. Larger plastic items, including textiles and cosmetics, break down into tiny pieces known as microparticles. This process is influenced by various factors, including temperature, physical and chemical interactions, and exposure to UV light. Around 51 trillion plastic particles can be found spread across the surface waters of the globe [1], [2]. It has become a problem for humans. The occurrence and impact of MPs in commercial sea salts that humans consume have not been well studied so far [3]. Microplastic pollution has become a pressing global concern, infiltrating various environmental compartments, including marine ecosystems [4].

Salt is one of the marine products consumed by humans. Based on its origin, salt is categorized into sea, lake, rock, river, and borehole salt. Sea salts are the most common type of salt available and consumed worldwide [5, 6, 7]. Recently, one of the most significant concerns regarding the quality of sea salts has been the contamination of salt products with microplastics (MPs), because Sea salt is made by evaporating seawater and forming salt crystals [8], and it is well known that the marine environment (sea and ocean) is the world's largest receiver of mismanaged plastic waste [9]. Research in various countries indicates that sea salt has been contaminated with MP [10]. Especially in sea salt preparation, where MPs can accumulate in the final result, their presence in seawater

[1*] Corresponding author: Vouchlay Theng
E-mail: vouchlay.theng@itc.edu.kh; Tel: +855-81352727

endangers marine life and human health. Expensive and non-biodegradable, conventional filtration techniques like membrane-based systems call for sustainable substitutes.

Coconut coir, in particular, is noteworthy due to its porous structure, high surface area, and low cost, making it a sustainable material suitable for large-scale use. Coir is a hard, versatile, natural fiber available in large quantities, and is a renewable resource. Coir fiber is obtained by dehusking the coconut and separating it from the pithy material of the fibrous mesocarp by the retting process [11]. The coir fiber is the thickest and most resistant of all commercial natural fibers. A low decomposition rate is the key advantage for making durable products [12]. Coconut coir possesses several advantageous properties that make it well-suited for microplastic removal. Its fibrous nature allows for adsorption over a wide area, enabling effective trapping and retention of microplastics. Additionally, coconut coir is hydrophilic, meaning it has a strong affinity for water, facilitating the capture of microplastics suspended in wastewater. Coir fibers have a rough, multi-layered surface with natural pores that mechanically entangle MPs [13]. Moreover, coconut coir is biodegradable and poses minimal environmental risks compared to synthetic adsorbents commonly used in wastewater treatment. Several studies have investigated the potential of coconut-based materials for general performance in wastewater treatment. For instance, Kasmuri et al., and [14] investigated the utilization of a blend of coconut husk and rice husk as activated carbon for wastewater treatment.

This study aims to evaluate the efficacy and feasibility of using coconut coir in expediting microplastic removal within sea salt production processes. It will meticulously examine the capabilities of coconut coir in removing diverse types and dimensions of microplastics across various operational scenarios, with a specific focus on comparing the microplastic removal efficiency between short (2–5 cm) and long (6–29 cm) coconut coir fibers.

2. METHODOLOGY

2.1 Materials preparation

To ensure the reliability and reproducibility of experimental outcomes, the protocols were adopted for coconut coir preparation with treatment procedures to mitigate batch-to-batch variations in material properties. The pre-treatment process began with washing the coconut coir thoroughly to remove any dirt and debris. This was done using tap water. The washed coconut coir, both short and long length ranges from (2–5 cm) and (6–29 cm) respectively, was then dried in an oven at 60 ◦C for 24 hours to reduce moisture content and prevent mold growth. Once dried, the coconut coir is weighed, then put into the column and compressed with pressure to obtain coconut coir fibers of uniform size and thickness. Lastly, the width and length, and height of the coconut coir are noted down to calculate the density of the coconut coir.

2.2 Experimental set-up

In the experimental setup, batch systems were employed to investigate the efficacy of microplastic removal using coconut coir fibers. Synthetic contaminated with microplastics were treated under natural gravity conditions. The experiment will proceed by placing all the supporting layers in the order gravel 10 cm, sand 4 cm, and the prepared coconut coir fibers of various condition will be placed in batch glass column, (10 cm) in width, (10 cm) in length, and (40 cm) in height (Fig.1).

The columns were then filled with 1 L of synthetic MP seawater (31g NaCl with 1 L of DI water) containing microplastics ranging from 300µm to 1mm, and the system was allowed to equilibrate. The coir fibers were incorporated into the experimental setup, and the effluent rate will be measured. While proceeding with the experiment, the effluent concentration of MPs and salinity were monitored continuously. Experiments were replicated, and statistical analysis techniques were employed to quantify variability and determine the significance of observed differences, ensuring robust data interpretation. These efforts collectively aimed to control potential variable sources and enhance the reliability of results.

Fig.1 below shows an experimental setup. For the lab-scale experiment, glass has been chosen as a container to avoid using plastic material.

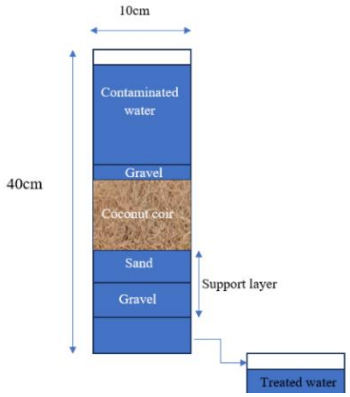


Fig. 1. Experimental setup

2.3 Analytical method

The density of coconut coir is studied to determine microplastic capture efficiency, where different packing densities can create varying filtration efficiencies [15]. Researchers aim to find the optimal density that maximizes

the removal of microplastics while maintaining an acceptable flow rate for the sea salt production process by using Eq.1.

Density

$$\rho = \frac{m}{V} \quad \text{(Eq. 1)}$$

where ρ is the density of coconut coir (g/dm^3 or g/l), m is the mass of coconut coir used(g), and V is the volume of synthetic seawater or real seawater (dm^3 or l).

In addition to density, the flow rate analysis is conducted to determine the filter's performance and practical application, such as filter efficiency capacity, to assess how much seawater can pass through the filter per unit of time while effectively removing microplastics. This determines the filter's throughput and overall processing capacity, with ease of optimization for filter improvements without significantly slowing down the process. This helps in achieving maximum microplastic removal efficiency with a balance between effective filtration (small enough pores to catch microplastics) and maintaining an acceptable flow rate for field application.

$$Q = \frac{V}{t} \quad \text{(Eq. 3)}$$

where Q is the flow rate of discharged effluent (l/s). V is the volume of discharge water (l). t is the discharge duration (s).

Moreover, to find the removal efficiency of coconut coir, a multi-step process of microplastic analysis is involved. Begin with sieving of the treated effluent through a 146 µm pore size sieve to isolate particulate matter. The filtered residues were then rinsed using DI water before oxidation with hydrogen peroxide (H_2O_2) solution. If the residues are rich in organic matter from coconut coir, then the density separation is carried out. Finally, the supernatant is filtered through a 0.45 µm pore size cellulose nitrate filter paper to calculate the removal efficiency by using Eq. 2.

$$\text{Re}\ (\%) = \frac{C_i - C_f}{C_i} \times 100 \quad \text{(Eq. 2)}$$

where, Re is the removal efficiency of microplastics by coconut coir (%). C_i is the initial concentration of microplastic in contaminated water or synthetic seawater (mg/l). C_f is final concentration of microplastic after passing through the coconut coir (mg/l).

3. RESULTS AND DISCUSSION

3.1 Effect of coconut coir length on MP removal

In this study, MPs of polypropylene (PP) were selected as adsorbents due to their high detection frequency and potential toxicological effects in aquatic environments [16].

Table 1. shows the study compared short (2–5 cm) and long (6–29 cm) coconut coir fibers, revealing a significant difference in MP removal efficiency; short fibers achieved 90% removal, while long fibers only reached 45%. This disparity can be attributed to the higher surface area per unit mass, enhancing interactions with MPs, and increasing adsorption capacity for fine particles [13], where long coir had faster flow but lower retention, indicating larger pore spaces. Long coir's faster flow but lower retention aligns with hydraulic studies on natural fiber filters [17]. Short fibers have a higher chance of being trapped within the coir's irregular pore structure. Long coir may bridge gaps between coir fibers, allowing them to pass through more easily, and Short coir has more exposed surface area per unit mass, increasing interactions with coir fibers [18]. The short coir was chosen to perform an optimal density test.

Fig. 2. Show the comparison of the removal efficiency between short and long coconut coir

Table 1. Effect of coconut coir length on MP removal

Coir	Mass (g)	Height (cm)	Density(g/dm^3)	Salinity (g/l)	Re (%)
Short	10	1	111.11	31	54
	30	3	111.11	31	90
Long	10	1	111.11	31	23
	30	3	111.11	31	45

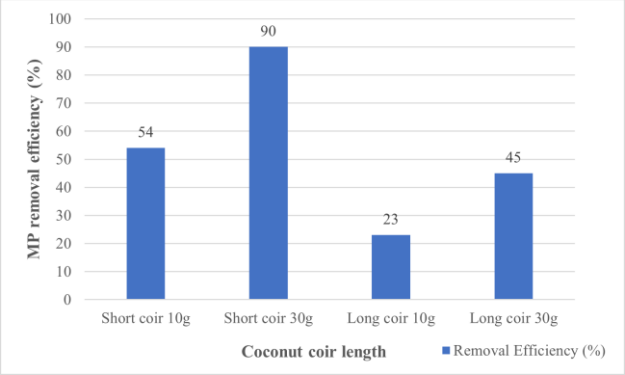


Fig. 2. Compare the removal efficiency between short and long coconut coir

3.2 Optimal density

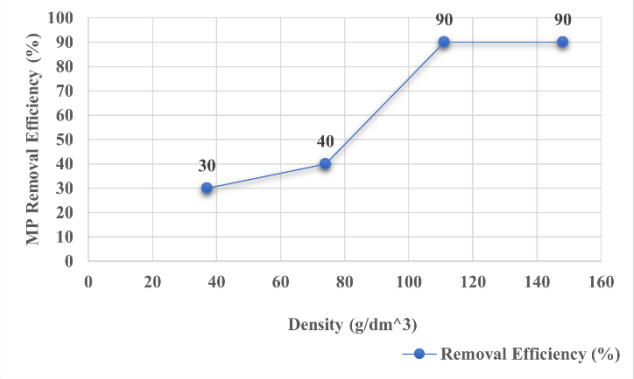
After obtaining the optimal result of coconut coir, Table 1 shows the optimal density that provides the best performance in removing MPs. At 111.11 g/dm^3, up to 90 % of MP removal is achieved with 30 g of coconut core used. At 111.11 g/dm^3, the flow rate is 0.045 L/s, which is an optimal flow rate among all density operation conditions. There remains a limited body of literature examining the removal efficiency of coconut-based filter

media, particularly concerning density variations and their impact on filtration performance [19], [20]. Current research on the density-dependent removal efficiency of coconut-based filters appears to be insufficient, highlighting a gap in the literature that warrants further investigation. Additional studies are needed to explore how variations in filter density influence contaminant removal effectiveness. In this study, at a density of 148.15 g/dm^{3,} bigger than 90% of the removal efficiency was achieved, but the flow rate decreased. Moreover, the process of manually compacting the coconut coir in the column is rather difficult.

Table 2. Optimal density

Mass (g)	Height (cm)	Density(g/dm^3)	Q (L/s)	Re (%)
10	3	37.04	0.07	30
20	3	74.07	0.059	40
30	3	111.11	0.045	90
40	3	148.15	0.032	>90

Fig. 3. Compare the removal efficiency between different densities of short coconut coir



3.3 Optimal thickness

Table 3 shows that Coconut coir removes MPs through physical filtration (trapping in fibrous matrix) and adsorption (hydrophobic interactions, electrostatic forces) [21].This study was conducted by varying the thickness from 1cm to 5cm. Higher than 90% removal efficiency, was achieved similar to Ganesan et al.'s 2023 study, which focused on optimizing column height for microplastic filtration using coconut coir, found that a 5 cm thickness (at 100 g/dm³ density) removed 92% of microplastics (MPs), compared to 75% removal at a 2 cm thickness. This indicates that increasing the depth of the coir filter significantly enhances microplastic removal efficiency. Increasing coir weight (e.g., 30g to 50g) likely improves removal efficiency due to higher density and more contact opportunities with MPs. However, the 3 cm thickness provides sufficient time and removal capacity for MPs to be captured effectively, which is 90%, while considering the flow rate. Moreover, the salinity did not decrease in all experiments.

Table 3. Optimal thickness

Mass (g)	Height (cm)	Density (g/dm^3)	Q (L/s)	Salinity (g/l)	Re (%)
10	1	111.11	0.06	31	30
30	3	111.11	0.045	31	90
50	5	111.11	0.04	31	>90

3.4 Variation of the removal efficiency of coconut coir in different types of MP

After obtaining the optimal condition of coconut coir in Tables 1, 2, and 3, coconut coir with mass 30 g at a density of 111.11 g/dm^3, and a thickness of 3 cm is chosen as the optimal condition while considering the flow rate. Table 4 presents data on the variation of the removal efficiency of coconut coir in removing two types of MP, polypropylene (PP) and polyethylene terephthalate (PET). Three different concentrations of MP were studied, using 1, 5, and 10 mg/l. The removal efficiency of PP increases from 90% to 97%, while the removal efficiency of PET rises from 87% to 95%, respectively. This suggests that coconut coir performs better at higher MP concentrations, possibly due to enhanced adsorption or filtration mechanisms at greater particle loads[22]. PP is removed more efficiently (90–97%) compared to PET (87–95%) across all tested concentrations. The difference may be attributed to surface chemistry, hydrophobicity, and particle size, while the variation in trapped MP concentrations depends on polymer type, size, shape, and environmental conditions [23]. PP is more hydrophobic, allowing stronger interactions with coconut coir fibers [24]. PET, being slightly more polar, may have weaker adsorption. Therefore, this result highlights the potential of coconut coir as a green filtration material for MP removal, with PP showing slightly better removal rates than PET. Compared to the study of Zharkeno et al., there is a notable improvement in removal efficiency. This is related to the difference in size of MP used in the experiment, which ranges from 100 µm to 500 µm, smaller than this study, where the pristine MP used ranges from 300µm to 1mm [19, 22]. Further optimization (e.g., chemical modification of coir) could enhance PET removal efficiency.

Table 4. Variation of the removal efficiency of coconut coir in removing two types of MP

MP	Concentration (mg/l)	Re (%)	STDEV

PP	1	90	0.5
	5	92	1.54
	10	97	0.5
PET	1	87	0.64
	5	91	1
	10	95	0.275

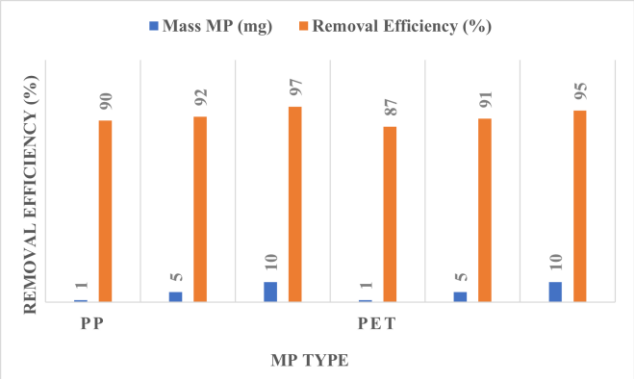


Fig. 3. Removal capacity of coconut coir for different Microplastic types and concentrations

4. CONCLUSIONS

The study has illuminated the significant role of coconut coir in addressing the pervasive challenge of microplastic pollution in seawater. Through meticulous experimentation and analysis, we have uncovered the microplastic removal ability of coconut coir under various conditions. Notably, the use of short coconut coir (2cm to 5cm) led to a substantial increase in both specific surface area and pore volume, with typical values rising from 54% to 90% better than the long coir's removal efficiency. The optimal coconut coir density was obtained at 111.11g/dm^3, with a thickness of 3cm, a removal efficiency is up to 90%, and an effluent flow rate of 0.045dm^3/s. For instance, at microplastic concentrations of 1 mg/L, 5 mg/L, and 10 mg/L, coconut coir displayed removal capacities ranging from 90% to 97% for PP and 87% to 95% for PET. Remarkably, the coconut coir filter achieves a notable result in removing microplastics from synthetic seawater while not decreasing any salinity content, leading to a bright direction for coconut filter implementation at the sea salt farm. Therefore, these findings highlight the potential of coconut coir, particularly short fibers, as an effective and eco-friendly solution for mitigating microplastic contamination in sea salt. Finally, this research provides a promising approach to enhancing environmental sustainability and food safety in marine-based industries.

ACKNOWLEDGMENTS

The author would like to acknowledge the Asian Blue Innovation Challenge (ABIC) for funding this research. Thanks to the Platform for Aquatic Ecosystem Research (PEAR) for providing laboratory facilities.

REFERENCES

[1] J. Gigault *et al.*, "Current opinion: What is a nanoplastic?," *Environ. Pollut.*, vol. 235, pp. 1030–1034, Apr. 2018, doi: 10.1016/j.envpol.2018.01.024.

[2] Y. L. Cheng, J.-G. Kim, H.-B. Kim, J. H. Choi, Y. Fai Tsang, and K. Baek, "Occurrence and removal of microplastics in wastewater treatment plants and drinking water purification facilities: A review," *Chem. Eng. J.*, vol. 410, p. 128381, Apr. 2021, doi: 10.1016/j.cej.2020.128381.

[3] M. Sivagami *et al.*, "Extraction of microplastics from commonly used sea salts in India and their toxicological evaluation," *Chemosphere*, vol. 263, p. 128181, Jan. 2021, doi: 10.1016/j.chemosphere.2020.128181.

[4] K. Ziani *et al.*, "Microplastics: A Real Global Threat for Environment and Food Safety: A State of the Art Review," *Nutrients*, vol. 15, no. 3, p. 617, Jan. 2023, doi: 10.3390/nu15030617.

[5] Q. Zhang *et al.*, "A Review of Microplastics in Table Salt, Drinking Water, and Air: Direct Human Exposure," *Environ. Sci. Technol.*, vol. 54, no. 7, pp. 3740–3751, Apr. 2020, doi: 10.1021/acs.est.9b04535.

[6] A.-A. Ujjaman Nur *et al.*, "Microplastic contamination in processed and unprocessed sea salts from a developing country and potential risk assessment," *Chemosphere*, vol. 308, p. 136395, Dec. 2022, doi: 10.1016/j.chemosphere.2022.136395.

[7] H. Lee, A. Kunz, W. J. Shim, and B. A. Walther, "Microplastic contamination of table salts from Taiwan, including a global review," *Sci. Rep.*, vol. 9, no. 1, p. 10145, Jul. 2019, doi: 10.1038/s41598-019-46417-z.

[8] M. S. Hossain, M. Z. Hossain, and S. R. Chowdhury, "An Analysis of Economic and Environmental Issues Associated with Sea Salt Production in Bangladesh and Thailand Coast".

[9] Z. Yuan, R. Nag, and E. Cummins, "Ranking of potential hazards from microplastics polymers in the marine environment," *J. Hazard. Mater.*, vol. 429, p. 128399, May 2022, doi: 10.1016/j.jhazmat.2022.128399.

[10] “A Review on Extraction Processes of Salts from Different Salt Lakes and their Environmental Impact in Industry,” *Lett. Appl. NanoBioScience*, vol. 11, no. 4, pp. 4016–4039, Oct. 2021, doi: 10.33263/LIANBS114.40164039.

[11] K. G. Satyanarayana, A. G. Kulkarni, and P. K. Rohatgi, “Structure and properties of coir fibres,” *Proc. Indian Acad. Sci. Sect. C Eng. Sci.*, vol. 4, no. 4, pp. 419–436, Dec. 1981, doi: 10.1007/BF02896344.

[12] L. C. Hao, S. M. Sapuan, M. R. Hassan, and R. M. Sheltami, “Natural fiber reinforced vinyl polymer composites,” in *Natural Fibre Reinforced Vinyl Ester and Vinyl Polymer Composites*, Elsevier, 2018, pp. 27–70. doi: 10.1016/B978-0-08-102160-6.00002-0.

[13] N. H. A. Norhasnan *et al.*, “Physicomechanical Properties of Rice Husk/Coco Peat Reinforced Acrylonitrile Butadiene Styrene Blend Composites,” *Polymers*, vol. 13, no. 7, p. 1171, Apr. 2021, doi: 10.3390/polym13071171.

[14] N. Kasmuri, N. F. M. Dzulkifli, N. A. Ismail, N. Zaini, and Z. Yaacob, “An investigation of a mixture of coconut husk and rice husk as activated carbon for treatment of wastewater,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1019, no. 1, p. 012048, Apr. 2022, doi: 10.1088/1755-1315/1019/1/012048.

[15] K. P. Indulekha, P. K. Jayasree, K. Balan, L. Dhanesh, and J. S. Vinod, “Laboratory Evaluation of Permeability Characteristics of Cocologs,” *J. Nat. Fibers*, vol. 18, no. 12, pp. 2127–2138, Dec. 2021, doi: 10.1080/15440478.2020.1723779.

[16] J. Daily and M. J. Hoffman, “Modeling the three-dimensional transport and distribution of multiple microplastic polymer types in Lake Erie,” *Mar. Pollut. Bull.*, vol. 154, p. 111024, May 2020, doi: 10.1016/j.marpolbul.2020.111024.

[17] L. Q. N. Tran, T. N. Minh, C. A. Fuentes, T. T. Chi, A. W. Van Vuure, and I. Verpoest, “Investigation of microstructure and tensile properties of porous natural coir fibre for use in composite materials,” *Ind. Crops Prod.*, vol. 65, pp. 437–445, Mar. 2015, doi: 10.1016/j.indcrop.2014.10.064.

[18] L. Yin *et al.*, “Microplastics retention by reeds in freshwater environment,” *Sci. Total Environ.*, vol. 790, p. 148200, Oct. 2021, doi: 10.1016/j.scitotenv.2021.148200.

[19] L. M. De Oliveira Cruz, B. G. L. A. Gomes, A. L. Tonetti, and I. C. S. Figueiredo, “Using coconut husks in a full-scale decentralized wastewater treatment system: The influence of an anaerobic filter on maintenance and operational conditions of a sand filter,” *Ecol. Eng.*, vol. 127, pp. 454–459, Feb. 2019, doi: 10.1016/j.ecoleng.2018.12.021.

[20] Y. Zharkenov *et al.*, “Utilizing biofilm-enhanced coconut coir for microplastic removal in wastewater,” *Case Stud. Chem. Environ. Eng.*, vol. 9, p. 100726, Jun. 2024, doi: 10.1016/j.cscee.2024.100726.

[21] F. Liu, N. Nord, K. Bester, and J. Vollertsen, “Microplastics Removal from Treated Wastewater by a Biofilter,” *Water*, vol. 12, no. 4, p. 1085, Apr. 2020, doi: 10.3390/w12041085.

[22] “Microplastic retention in green walls for nature-based and decentralized.”

[23] “Carbon-Based Adsorbents for Microplastic Removal.”

[24] M. Ateia, T. Zheng, S. Calace, N. Tharayil, S. Pilla, and T. Karanfil, “Sorption behavior of real microplastics (MPs): Insights for organic micropollutants adsorption on a large set of well-characterized MPs,” *Sci. Total Environ.*, vol. 720, p. 137634, Jun. 2020, doi: 10.1016/j.scitotenv.2020.137634.

Water and Environment (WAE)

3 Sub-parallel Sessions

"WAE-3: Hydrology and Climate: Driving Innovation for Resilient Water Systems"

Chair: Dr. THENG Vouchlay; Co-chair: Mr. SOK Kimhuy

No.	Topic
1	Spatio-temporal Distribution of Electric Conductivity, Stable Isotopes of Water and Water Level along Tonle Sap River *Authored by: Ratha DOUNG, Paul BAUDRON, Vuthy CHORK, Sambo LUN, Sytharith PEN 1, Khy Eam EANG, Kimhuy SOK, Sylvain MASSUEL, Jonathan Van HANJA, Marie BALZAN, Julia MEERMAN, Kong CHHUON*
2	Daily Tidal Dynamics at Chaktomuk Confluence by a 2D Hydrodynamic Model *Authored by: Sytharith PEN, Vuthy CHORK, Sambo LUN, Ratha DOUNG*
3	Dynamic Seasonal Variations in River Mixing Patterns at a Major Confluence-Diffluence Site (Chaktomuk Junction, Cambodia) Revealed by Hydrochemical Tracing *Authored by: Vuthy CHORK, Paul BAUDRON, Ratha DOUNG, Sambo LUN, Sytharith PEN, Khyeam EANG, Kimhuy SOK, Sylvai MASSUEL, Jonathan Van HANJA, Marie BALZAN, Julia MEERMAN, Voitna PRAK, Kong CHHUON*
4	Historical Changes in Extreme Daily Temperature Indices in the Last 3 Decades in Cambodia *Authored by: Ty SOK, Por Y CHHENG, Meng Hour HUOT, Ilan ICH, Penglong KOUN, Rattana CHHIN*
5	Contemporary Changes in Precipitation Patterns and Extreme Rainfall Indices in Cambodia *Authored by: Meng Hour HOUT, Khamkeo HOAN, Ravon SOY, Rattana Chhin, Ty SOK, Layheang SONG*
6	Assessment of Inundation at Prek System of Koh Thum Inducing by Bassac River and Regional Precipitation Pattern *Authored by: Kim Hort LY, Khy Eam EANG, Sytharith PEN, Ratha DOUNG, Vuthy CHORK*
7	Review on Historical Tropical Hurricanes, Typhoons, and Cyclones Frequency, Intensity, and Trajectory *Authored by: Por Y CHHENG, Chanchakriya KHEN, Lyhieng LAY, Menghour HUOT, Phy Sophearum PHY, Ty SOK*
8	Evaluating Water-Saving Irrigation Methods for Rice Cultivation, Case Study in Cambodia *Authored by: Mengheak PHOL, Chetra PHEUN, Chankethya SUN, Vanndy LIM, Veasna TOUCH, Pinnara KET, Chengxiang MA*
9	The Dependency of the Planting Date on Rainfall in Paddy Fields in the Pursat River Basin *Authored by: Kimhor SIENG, Veasna LY, Sithov PHAT, Menghour HUOT, Ty SOK*
10	Hydraulic Propagation of Pesticides at the Watershed Scale: Tonle Sap River Case Study *Authored by: Ech CHIN, Khy Eam EANG, Sombath KEO, Paul BAUDRON*

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Spatio-temporal distribution of electric conductivity, stable isotopes of water and water level along Tonle Sap river

Ratha Doung [1*], Paul Baudron [2], Vuthy Chork [1], Sambo Lun [1], Sytharith Pen [1], Khy Eam Eang [1], Kimhuy Sok [1], Sylvain Massuel [2], Jonathan Van Hanja [2], Marie Balzan [2], Julia Meerman [3], Kong Chhuon [1]

[1] *Coastal and Wetland Environmental Lab, Institute of Technology of Cambodia, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *French National Research Institute for Sustainable Development*
[3] *Avans Hogeschool, the Netherlands*

[*]*Corresponding author: doung.ratha@gmail.com*

Abstract: *The largest freshwater body in South-East Asia, Tonle Sap Lake in Cambodia, is believed to absorb significant volumes of water from the Mekong during the wet season and to release water back to the confluence in dry season. However, in the last decades, rapid anthropogenic and climate changes have significantly impacted the lake's hydrology, drastically modifying the Mekong's flood pulse through the Tonle Sap River and its flow pattern. Transience of water flow and storage/feeding in floodplains require specific attention. This study provides better understanding on this new uncertain regime. Weekly EC and water level data from 12 sites revealed three distinct hydrological phases influencing mixing patterns along the river. Stable isotope results corroborate EC-based mixing patterns, highlighting transitions during the flood pulse. Findings of this research will be beneficial to anticipate water quality changes (and renewal) and to anticipate fish migrations along Tonle Sap River. Further effort to improve findings include to compare the high spatio-temporal resolution of water elevation data from our DGPS with similar data from surface Water Ocean Topography Satellite (SWOT) and to consider the hydro-chemical-dynamic modelling of the Tonle Sap river flow.*

Keywords: Tonle Sap Lake, Electrical conductivity, water isotopes, anticipate fish migration

1. INTRODUCTION

The Tonle Sap Lake in Cambodia is the largest permanent body of fresh water in Southeast Asia. Cambodian floodplains, including the Tonle Sap floodplain, contain the most extensive wetland habitats in the Mekong system [1]. Each monsoon, the Tonle Sap River reverses flow, allowing Mekong waters to enter the lake and support migratory fish vital to Cambodia's fisheries. Yet, current methods to monitor and predict this reversal are limited. The InverSap project addresses this by combining long-term hydrochemical and hydrodynamic data with new, strategically timed measurements. This integrated approach aims to improve predictions of water flow and quality, supporting agencies like the Cambodian Fisheries Administration, Tonle Sap Authority, and Phnom Penh Water Supply Authority in enhancing monitoring and management.

This study aims to answer: (1) How does EC vary across the Tonle Sap River during different flow regimes? (2) What are the spatial patterns of water level and EC in relation to discharge? (3) Can isotopic signals confirm water source mixing during reversal periods?

This research work intended to build Cambodian research capacity by providing tools, equipment, and interdisciplinary skills, and by strengthening existing partnerships. It is based on the idea that hydrochemical signatures can trace water sources and real-time mixing in the Mekong-Tonle Sap system—key for understanding fish migration and water quality shifts. The study focuses on the Tonle Sap River from Phnom Penh to the lake's mouth with

[1*] Corresponding author: Ratha Doung
E-mail: doung.ratha@gmail.com; Tel: +855-77 558 500

the ultimate goal of developing a user-friendly tool to predict river flow dynamics and fish biodiversity.

The specific objective of this research aims to generate a high-resolution, river-scale assessment of hydrological and water quality interactions along the Tonle Sap River by implementing a systematic, weekly monitoring program. The study focuses on 12 strategically selected sampling sites spanning from the Chaktomuk confluence—where the Mekong, Bassac, and Tonle Sap Rivers meet—to the entrance of the Tonle Sap Lake. At each location, measurements of electrical conductivity (EC), Differential Global Positioning System (DGPS) coordinates, and stable water isotopes are collected to capture spatial and temporal variations in water composition.

The core objective is to trace the influence of the Mekong River on the Tonle Sap River's water quality during different hydrological seasons. Due to the distinct difference in EC values between the Mekong River (typically higher) and the Tonle Sap River (lower), EC serves as a reliable and cost-effective tracer for identifying the extent and timing of Mekong River water intrusion into the Tonle Sap system. By analyzing EC gradients along the river, the study seeks to pinpoint the reach of Mekong-sourced water and better understand the mechanisms governing water exchange between the two rivers.

This investigation is crucial for improving our understanding of the Tonle Sap River's seasonal hydrodynamics and its connectivity with the Mekong River, particularly in the context of upstream hydropower development, climate variability, and changes in land use. The findings will support improved water resource management and help anticipate ecological and socio-economic impacts in the Tonle Sap–Mekong system.

2. METHODOLOGY

2.1 Study area

The study area covers the Chaktomuk junction—where the Mekong, Bassac, and Tonle Sap Rivers converge—and extends along the Tonle Sap River toward the Tonle Sap Lake, spanning approximately 120 kilometers from Phnom Penh. A total of 12 monitoring sites were strategically selected along this stretch to conduct in-situ measurements of electrical conductivity (EC), record water levels using Differential GPS (DGPS), and collect water samples for stable isotope analysis. The location of sample site is illustrated in Figure. 1.



Fig. 1. Location of samplings at along Tonle Sap River from Chaktomuk to Tonle Sap Lake

2.2 Data collection

To achieve the objectives of this research, a weekly field campaign has been conducted continuously since May 2024 and is ongoing at the time of writing. During each field campaign, in-situ measurements of electrical conductivity (EC), water level, and benchmark location were systematically performed. EC was measured using a freshly calibrated hand-held sensor (ProfiLine pH/Cond 3320; Figure. 2, while water levels and precise location data were recorded using a Differential GPS device (A3RTK-Pro; Figure. 3. Additionally, water samples for stable isotope analysis were collected from each of the 12 monitoring sites using 200 mL sampling bottles.



Fig. 2. Hand-held measuring device ProfiLine pH/Cond 3320



Fig. 3. DGPS A3RTK-Pro



Fig. 4. Water Isotope Samples

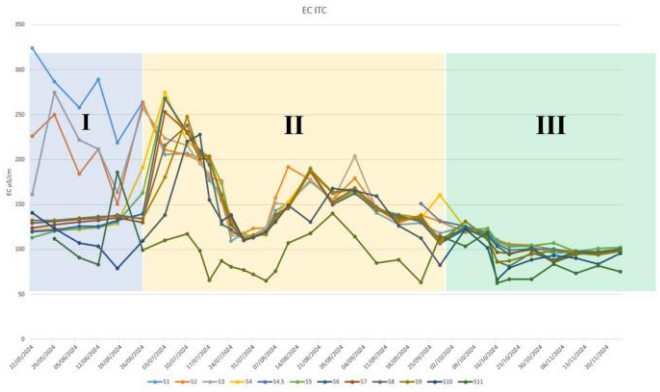


Fig. 5. EC (µc/cm) in particular location throughout the year

2.3 Data Analysis

Electrical Conductivity (EC) and water level data were analyzed to assess spatial and temporal variations in water characteristics along the Tonle Sap River. Measurements were collected at 12 key locations spanning from Chaktomuk in Phnom Penh to the mouth of the Tonle Sap Lake. Data collection was synchronized using Differential GPS (DGPS) to ensure high positional accuracy and consistency across sites.

For EC data, both instantaneous and averaged values were calculated to evaluate salinity gradients and potential mixing zones between Mekong and Tonle Sap waters. Seasonal trends were identified by comparing measurements across different hydrological periods, particularly during the monsoon-driven flow reversal.

Water level measurements obtained via DGPS were used to map elevation profiles along the river course. These data helped track water flow direction, identify inversion points, and validate hydrodynamic conditions. Cross-sectional comparisons among the 12 sites were made to analyze river slope, backwater effects, and the timing of water level rise or fall during seasonal transitions.

Statistical analysis, including correlation and regression, was performed to explore relationships between EC values, water levels, and geographical position. All data were processed using ArcGIS tools to visualize spatial patterns and identify critical zones of hydrochemical change.

3. RESULTS AND DISCUSSION

3.1. EC value of Tonle Sap River

Based on nearly one year of continuous field campaigns, the collected electrical conductivity (EC) data were analyzed and are presented in Figure 5. Over the course of the monitoring period, the spatial and temporal variations in EC values revealed three distinct patterns, which were categorized into three zones: Zone I, Zone II, and Zone III.

Zone I: corresponds to the dry season and the early stages of the rainy season, during which the reversal flow from the Mekong River into the Tonle Sap Lake had not yet begun. During this period, electrical conductivity (EC) values were relatively high, reflecting limited hydrological connection with the Mekong and the influence of local water sources with naturally higher salinity or lower dilution from rainfall and upstream inflow.

Zone II: represents the period when the hydrological reversal occurs, and the Mekong River starts flowing into the Tonle Sap Lake. EC values during this time remained elevated, indicating the strong influence of Mekong River water with higher conductivity. The mixing of Mekong water with Tonle Sap River water during the reversal phase contributes to the observed EC peaks across many sampling sites.

Zone III: marks the phase when the flow direction returns, and water from the Tonle Sap system begins draining back into the Mekong River. During this period, EC values decreased significantly, reflecting the dilution effect as the water mass shifts, decreased Mekong water influence, and the dominance of lake outflow conditions.

Among all the stations, S11 consistently records relatively lower EC values throughout the monitoring period, suggesting it may be influenced by a fresher or less mineralized water source compared to the others.

3.2. Relationship between EC value with Water Discharge

EC values a long Tone Sap River were taking to correlate the discharge capacity measured at Prek Kdam.

During the wet season for monitoring purposes, the accumulated reverse flow volumes at Prek Kdam (in the Tonle Sap River) have been determined by combining the available data at Kampong Luong (in the Tonle Sap Lake) and Phnom Penh Port (in the Tonle Sap River) [2].

The chart below (Figure. 6) combines electrical conductivity (EC) data from multiple stations along the Tonle Sap River (TSR) with discharge measurements taken at Prek Kdam (located between stations S4 and S5). This comparison offers valuable insight into how river flow dynamics influence water quality across time.

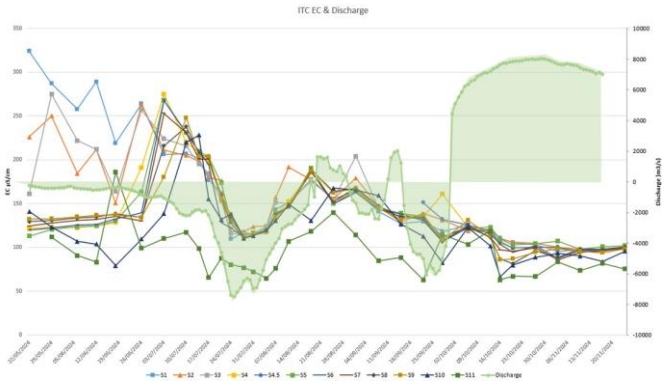


Fig. 6. Relationship between EC value along Tonle Sap River and Discharge Rate Measured at Prek Kdam station

Discharge Drop Signals Complete Flow Reversal. After the distinct rise and synchronization of EC values seen earlier in the timeline, the discharge curve at Prek Kdam shows a sharp downward trend. This sudden decline marks the end of the propagation phase and indicates that the reversal of flow is fully underway. The Tonle Sap River, which previously received inflow from the Mekong, now begins to drain back toward the Mekong River system.

Positive Discharge Indicates Temporary Flow Toward Chaktomuk. During this reversal period, discharge values temporarily turn positive. This change in sign reveals that the flow direction is momentarily toward Chaktomuk, rather than feeding into the Tonle Sap. Chaktomuk is a major junction point where the Mekong splits into multiple channels, including the Tonle Sap and Bassac Rivers. This momentary shift reflects the dynamic hydrology of the region, where seasonal backflows regulate the floodplain and lake system.

This stage also coincides with more variable EC readings, as the system undergoes reorganization and mixing due to the reversed flow and varying water sources.

The spatial distribution of electrical conductivity (EC) values for six selected field campaigns—Campaign 1 (22 May 2024), Campaign 6 (26 June 2024), Campaign 13 (28 July 2024), Campaign 18 (18 August 2024), Campaign 27 (16 October 2024), and Campaign 33 (24 November 2024)—is illustrated in Figures 8 through 13, respectively. These campaigns were chosen to represent key hydrological phases, including the dry season, the onset and peak of the flow reversal, and the return flow toward the Mekong River, providing insight into the temporal and spatial variations of EC along the Tonle Sap River.

3.3. Relationship between Water Elevation and Discharge

The water level at each station was measured using DGPS during each survey campaign. Water level fluctuations in the Tonle Sap River (TSR) were observed to vary over time. These fluctuations are primarily influenced by the bidirectional flow between the Mekong River and the Tonle Sap Lake, which changes with the seasons (rainy and dry). By comparing the discharge data provided by the Mekong River Commission (MRC), measured at the Prek Kdam station, the relationship between the water level in the TSR and the discharge rate is illustrated in Figure 7 below.

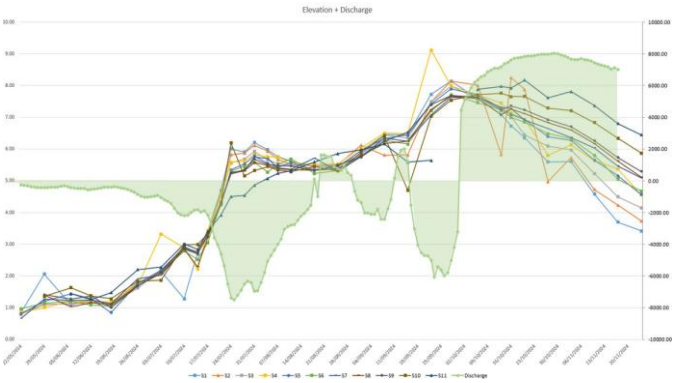


Fig. 7. Relationship between Water Elevation along Tonle Sap River and Discharge Rate Measured at Prek Kdam station

At the beginning of the field campaign in the third week of June 2024, the water elevation was relatively low, around 0.8m above mean sea level. This period marked the onset of the flow reversal, characterized by a low discharge rate of approximately 500 m³/s. As the discharge from the Mekong River into the Tonle Sap Lake increased, the water level rose significantly, reaching a peak elevation of 8.2m above mean sea level. The maximum discharge rate during this period reached around 7,500 m³/s. The flow direction began to shift back to normal in early October 2024, at which point the water level in the Tonle Sap River gradually declined from its peak, eventually reaching around 0.5m above mean sea level.

The data indicates a strong correlation between the discharge rate of the Mekong River and the water level in the Tonle Sap River. During the reversal period beginning in June 2024, increased inflow from the Mekong significantly raised water levels, peaking in response to the maximum discharge. Conversely, as the flow shifted back to normal in October, water levels in the Tonle Sap River gradually declined. This highlights the seasonal and dynamic hydrological interaction between the Mekong River and the Tonle Sap system.

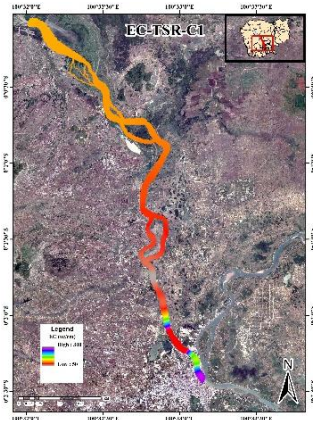


Fig. 8. EC distribution on 22 May 2024

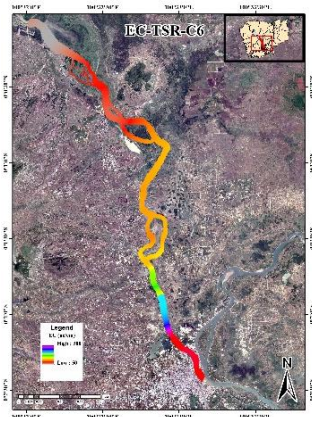


Fig. 9. EC distribution on 26 June 2024

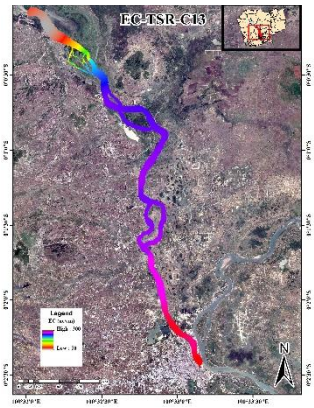


Fig. 10. EC distribution on 28 July 2024

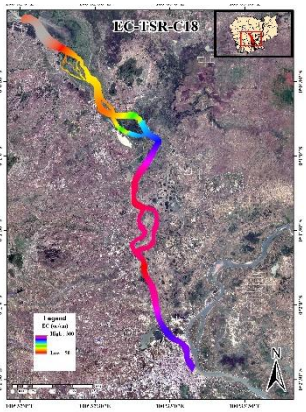


Fig. 11. EC distribution on 18 August 2024

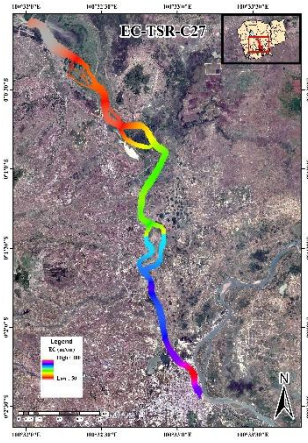


Fig. 12. EC distribution on 16 October 2024



Fig. 13. EC distribution on 24 November 2024

Figures 8 to 13 illustrate the spatial variation of electrical conductivity (EC) along the Tonle Sap River across six selected campaigns.

In the early stages, such as Campaign 1 (22 May 2024) and Campaign 6 (26 June 2024), EC values are generally higher, particularly in the downstream section near the Chaktomuk junction. This pattern reflects the dry season and early rainy season conditions, where limited freshwater inflow results in more concentrated EC levels. By Campaign 13 (28 July 2024) and Campaign 18 (18 August 2024), during the onset and peak of the flow reversal, EC values remain elevated along much of the river, especially in the middle and upstream sections. This suggests strong influence from the Mekong River's higher-conductivity water entering the Tonle Sap system. In contrast, Campaign 27 (16 October 2024) and Campaign 33 (24 November 2024) show a noticeable decline in EC values throughout the river, particularly in the downstream areas. These patterns correspond to the post-reversal period when water begins flowing back from the Tonle Sap Lake toward the Mekong River, resulting in lower EC due to greater dilution and reduced Mekong influence.

4. CONCLUSIONS

This research shows that electrical conductivity (EC) is an effective and sensitive indicator for understanding hydrodynamic interactions between the Tonle Sap River (TSR) and the Mekong River. Observed EC variations confirmed that both station location and water type (mainstream, side stream, or floodplain) strongly influence values, while elevation patterns reflect both local and broader environmental factors. Ultimately, EC proved to be a responsive metric, revealing flow direction and side stream influence on the main channel. Elevation dramatically changed spatially as revealed by our high-resolution and discharge comparisons showed that flow inversion occurs in distinct phases, not as a single event.

Combining large- and small-scale data improved overall reliability, offering a clearer picture of system behavior. While occasional equipment failures occurred, satellite-based sampling locations helped maintain data continuity.

Future research will integrate SWOT satellite altimetry and stable isotope mapping to enhance predictive models of Tonle Sap's flood dynamics, supporting sustainable management under climate and infrastructure pressures.

These findings could help support decision-making through enhanced understanding of the dynamics of flow inversions and storage of floodwater in the floodplains. Therefore, these findings will help better anticipate both flow and quality changes along the river and enhance water budget calculations in the great lake.

ACKNOWLEDGMENTS

This research was conducted as part of a work package within the main project titled *Anticipating the Inversions of the Tonle Sap River (InverSAP)*, supported by the French

Ministry for Europe and Foreign Affairs through the French Embassy in Cambodia. The authors gratefully acknowledge this financial support.

REFERENCES

[1] M. Kummu, J. Sarkkula, J. Koponen, and J. Nikula, "Ecosystem management of the Tonle Sap Lake: An integrated modelling approach," *Int. J. Water Resour. Dev.,* vol. 22, no. 3, pp. 497–519, Sep. 2006.

[2] Mekong River Commission, "Article 6B for planning purposes," *Procedures for the Maintenance of Flows on the Mainstream (PMFM),* [Online]. Available: https://pmfm.mrcmekong.org/planning/6b/. [Accessed: Apr. 25, 2025].

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Daily Tidal Dynamics at Chaktomuk Confluence by a 2D Hydrodynamic Model

Sytharith Pen[1*], Vuthy Chork[1], Sambo Lun[2], Ratha Doung[1]

[1]*Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2]*Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: sytharithpen@itc.edu.kh*

Abstract: *Chaktomuk confluence possesses a peculair complex hydraulic characteristic. It is characterized by the four rivers namely Mekong upstream, Mekong down stream, Tonle Sap, and Bassac. In addition to the unique inverse flow between the Mekong river into the Tonle Sap river in the wet season, the confluence is affected by the downstream tidal effect in the Bassac river by extent from the Mekong Delta. The results showed that the effect is significant in dry season. In the dry season, the fluctuation of the tidal effect is clearly observable in the sub-daily period where water in the Tonle Sap flows in two directions. The tidal effect is observable far upstream into the Tonle Sap. The changing in flow direction affects the raw water pumping station from which the water is pumped to the water supply treatment plant. Each river is inherent with different water quality parameters and water quality. This affects the operation and management of the water supply associated with the amount of chemical use for water treatment. As a result, the dynamic of the flow at the confluence was determined to sub-daily level. The result is beneficial for the water supply authority to manage and plan their operation. In addition, the potential threats on aquatic ecology from the tidal effects are identified.*

Keywords: Chaktomuk Confluence, Climate Change, Inverse flow

1. INTRODUCTION

Chaktomuk junction comprises of an extraordinary characteristic of the Tonle Sap River connect to the Mekong River and joins at Chaktomuk Junction and splits into the Bassac River and lower Mekong River.

It is well established that the water level rises remarkably in the Mekong during rainy season and flow into the Tonle Sap filling the Tonle Sap Lake and its flood plain. In the dry season, however, the reverse flow from the Tone Sap is observed due to the higher water elevation in the Tonle Sap compared to the Mekong In this regard, the Tonle Sap Lake plays the role of both as a reservoir to store and supply the water in the system. A critical perspective to consider in studying the Chaktomuk junction is the strong relationship between flood flow in the Mekong River and the volume of entering the Tonle Sap River and the Great Lake. Regulation of the Mekong River floods would lead to a decrease in water supply to the Lake water. This would happen even if the regulation did not reduce the annual flow but only caused redistribution throughout the year, such as an increase of flow during dry periods and a reduction of flood peaks [1].

It is well established that the reverse flow from the Chaktomuk to the Tonle Sap varies seasonally [2,3,4]. However, the reverse flow takes place even sub-daily time step. The flow direction in the Tonle Sap river fluctuates under the effect of tidal of the downstream of Bassac river.

2. METHODOLOGY

2.1 Study Area

In this analysis, the Chaktomuk confluence is incorporated into the model. The system contains the Mekong upstream, the Mekong downstream, and the Bassac river with distance from the Chaktomuk junction approximately 3 km, 5 km, and 3 km, respectively. As of the river section of the Tonle Sap, the upstream boundary is set to be at Prek Kdam with the approximate distance of 33 km from the junction.

[1*] Corresponding author: *Sytharith PEN*
E-mail: sytharithpen@itc.edu.kh; Tel: +855-16 320 332

The selection of the study area based on the availability of the boundary condition data. In addition, the effect of the tidal effect is expected to be far upstream to the Tonle Sap. The study area is shown in Fig. 1



Fig. 1.: Study Area for the study.

2.2 Data Collection

2.2.1 Bathymetry Data

The bathymetry data is obtained from the Ministry of Publick Works and Transport. The data is obtained by eco-sounder. In the boundary of the channel, the deepest elevation is observed at the elevation of -27 msl as shown in Fig. 2. The bank boundary is improved by the topographical survey using DGPS set.

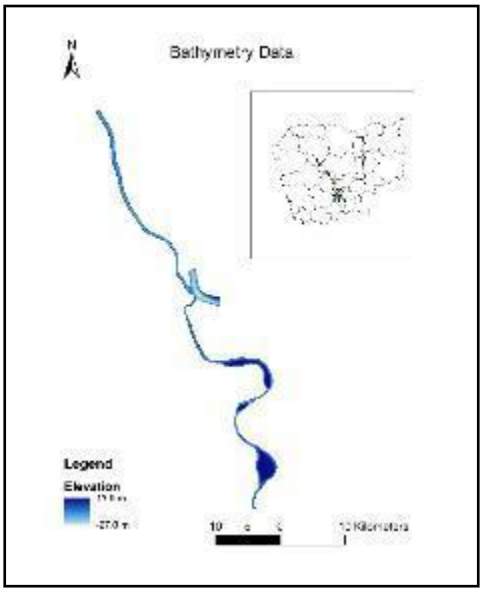


Fig. 2.: Bathymetry data of the river system

2.2.1 Water Elevation Data

The water level at Prek Kdam and Chaktomuk are using as boundary condition and calibration respectively. The water level at the two stations is shown in Fig. 3.

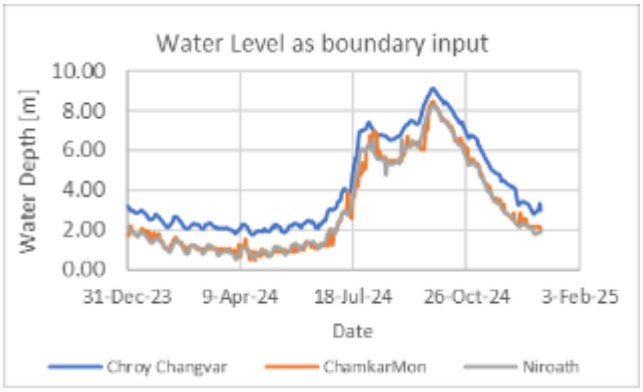


Fig. 3.: Water level data as boundary input

The water level at Chroy Changvar, Chamkar Mon, and Niroath water supply stations are input as boundary condition. The depth is to be converted into mean sea level system to incorporate into the hydraulic model.

2.3 Model approach

For the study, HEC-RAS model [5] was used as modeling approach for the analysis of the hydraulic characteristic of the junction. The model was found to perform well with mode where stage hydrograph is used as boundary condition. Additionally, it is more convenient to edit model geometry with RAS mapper that is built in in the model than to edit with other software. It also permits the users to run multiples calculation plans which make the model easier for result comparison. Furthermore, HEC-RAS represents a power tool that can produce reliable findings when it comes to Manning's n value estimation, [6,7].

2.4 Statistical analysis

For the model assessment, root mean square error (RMSE) is applied for model result and observation result.

$$RMSE = \sqrt{\frac{\sum_i^n (h_i - h_o)^2}{n}} \quad \text{(Eq. 1)}$$

h_i: simulation water level

h_o: observed water level

n: number of observed data

3. RESULTS AND DISCUSSION

3.1 Model performance

The model is calibrated with a peak flow event in the rainy season at Chaktomuk station. The RMSE is evaluated and the model show satisfactory model performance with RMSE = 0.45.

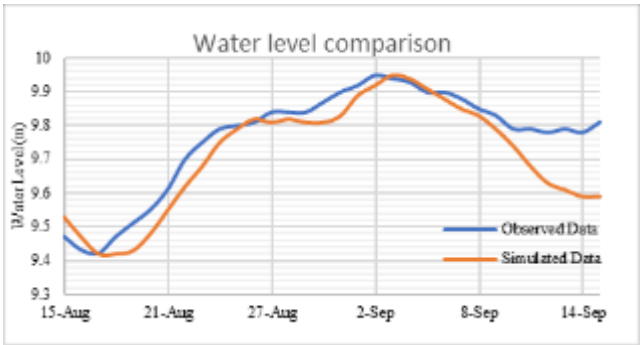


Fig 4.: Comparison of simulated and observed water level at chaktomuk station

3.2 Spatial Velocity Distribution

The model was run to cover the whole year of 2024 with two distinct season variation. In the low flow period, the velocity at Bassac junction is relatively low compare to the Tonle Sap river due to the expansion of the cross section and the mixing of the rivers. Fig. 5 depicts the velocity in the range of approximately 0.5 *m/s* at the junction, while the velocity can reach to the range of 1.0 ~2.0 *m/s* in the high flow period see Fig. 6. The results are consistent with the observation.

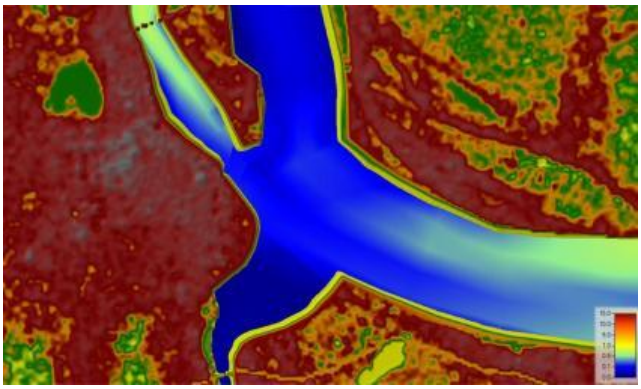


Fig. 5.: Flow velocity at Chaktomuk junction in low flow period

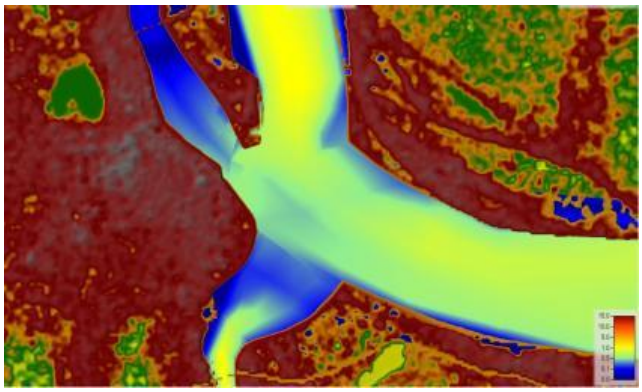


Fig. 6.: flow velocity in high flow period

The understanding of flow velocity in the junction is beneficial for the migration of the fish [8]. The success rate of upstream migration is inverse proportional to the flow velocity. Moreover, the low flow period signifies the normal flow condition from the Tonle Sap bringing more sedimentation. The amount of the sediment is clearly shown in normal flow period as shown in Fig. 7. The contract of sediment is clearly illustrated by the velocity profile in the model. It is observed that the velocity from both Mekong and Tonle river run parallelly. Mixing of the two is observed far downstream of the lower Mekong.

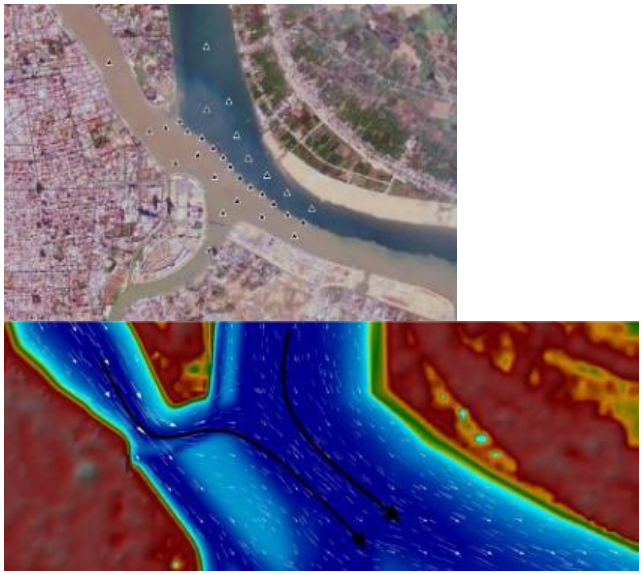


Fig. 7.: Sediment contrast at Chaktomuk junction on site and model result

3.3 Temporal Velocity Distribution

The analysis focuses on the temporal variation of the flow direction. The flow in the Tonle Sap is shown by the model to have two directional flows depending on the boundary condition of the upstream and downstream. The model identifies the different between the upstream water level and downstream water level ascribed to the changing direction of the flow from the reverse flow and normal in the Tonle Sap as shown in Fig. 8.

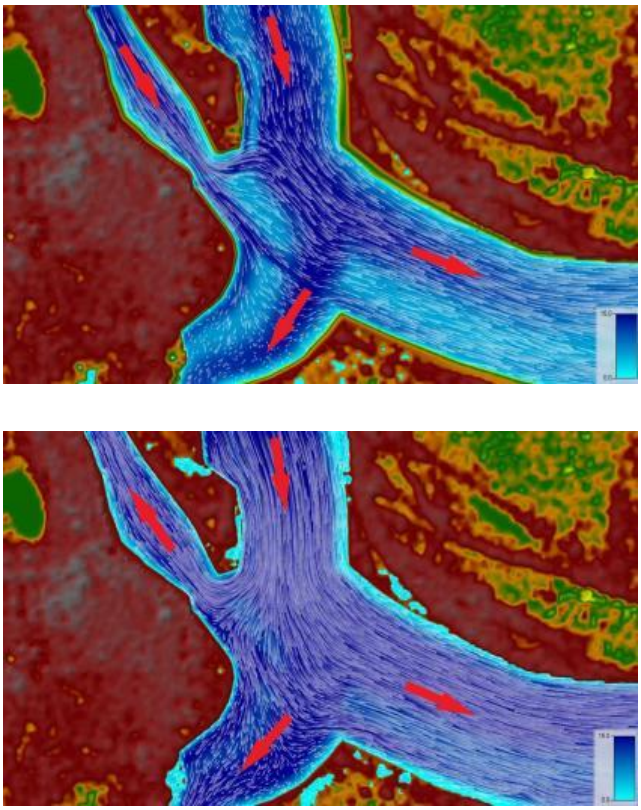


Fig. 8.: Temporal variation of velocity direction

It is important to point out that the flow in the Mekong upstream and the Tonle Sap has different characteristic of the quality parameter, i.e. electric conductivity (EC). According to the field investigation, the EC value in the upper Mekong is generally higher compared to that in the Tonle Sap. There are three raw water stations located in the Tonle Sap, downstream Mekong, and Bassac river. The stations pump the water directly from those rivers. The temporal variation of the velocity direction affects the water quality parameter EC of the water source. Therefore, the result supports the operation of the pumping station.

On the sub-daily basis, the water in the Bassac river can push the water upstream in the Tonle Sap under the tidal effect of the downstream boundary. During this event, high EC is expected in the Tonle Sap. From the analysis, it supports the operator to understand and regulate the pumping schedule to secure high quality station water which results in reduce the production cost.

On top of all that, the situation of the tidal effect becomes more significant in the future due to multiple reasons. Firstly, the increase of the mean sea level which amplifies the magnitude of the tide. Secondly, it is critical in the climate change and anthropogenic change activity in the upstream particularly during the low flow period. The extend of the tidal effect is foreseen to become stronger in the future scenario.

4. CONCLUSIONS

The analysis pointed out the spatial and temporal variation of the flow direction in the Chacktomuk Junction. The magnitude of the flow velocity can affect the migration of the fish. The model, in addition, explain the sediment contrast at the Chaktomuk junction. It is shown that the velocity from the Mekong and the Tonle Sap run parallelly without mixing until further downstream.

The water study pumping stations are located and pump water directly from the river. Therefore, the changing of the flow direction temporally explain and predict the water quality parameter such as EC. The understanding on the temporal variation of EC provides support to the public water supply authority to schedule pumping operation which is economically beneficial to the cost management. The analysis anticipates to incorporate the effect of climate change in particularly critical low flow condition which exacerbates the current situation.

ACKNOWLEDGMENTS

The authors would like to express great thanks for technical and financial support from IRD through INVERSAP Project and by extent to Coastal and Wetland Environmental Research Laboratory.

REFERENCES

[1] Mekong River Commission (2001) Chaktomuk Junction Environmental, Hydrological and Morphological study.

[2] Kummu, M., Tes, S., Yin, S., Adamson, P., Józsa, J., Koponen, J., Richey, J. and Sarkkula, J. (2014), Water balance analysis for the Tonle Sap Lake–floodplain system. Hydrol. Process., 28: 1722-1733. https://doi.org/10.1002/hyp.9718

[3] Fujii, H., Garsdal, H., Ward, P., Ishii, M., Morishita, K., & Boivin, T. (2003). Hydrological roles of the

Cambodian floodplain of the Mekong River. International Journal of River Basin Management, 1(3), 253–266. https://doi.org/10.1080/15715124.2003.9635211

[4] Koponen, J. ; Kummu, M. ; Sarkkula, J. / Modelling Tonle Sap Lake Environmental Change. In: Verhandlungen des internationalen Verein Limnologie. 2005 ; Vol. 29, No. 2. pp. 1083-1086.

[5] Gary W. Brunner, J.C.W. Brent, C. Wolfe (2018). HEC_RAS Application Guide.

[6] Ardıçlıoğlu, M., & Kuriqi, A. (2019). Calibration of channel roughness in intermittent rivers using HEC-RAS model: case of Sarimsakli creek, Turkey. *SN Applied Sciences*, *1*(9), 1–9.

[7] Alzahrani, A. S (2017). Application of two-dimensinal hydraulic modeling in riverine systems using HEC-RAS. University of Dayton.

[8] Hu, J.; Wang, P.; Chen, H.; Liu, Z.; Chen, S. Effects of the Diversity of Flow Velocity on the Upstream Migration Behavior of Grass Carp in the Reaches of Spur Dikes. *Water* 2023, *15*,3091. https://doi.org/10.3390/w15173091

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Dynamic Seasonal Variations in River Mixing Patterns at a Major Confluence-Diffluence Site (Chaktomuk Junction, Cambodia) revealed by Hydrochemical Tracing

Vuthy Chork[1,2,*], Paul Baudron[3], Ratha Doung[1,2], Sambo Lun[2], Sytharith Pen[1,2], Khyeam Eang[1,2], Kimhuy Sok[1], Sylvain Massuel[3], Jonathan Van Hanja[3], Marie Balzan[3], Julia Meerman[4], Voitna Prak[5], Kong Chhuon[2]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *G-EAU, IRD, AgroParisTech, Institut Agro, Cirad, INRAE, University of Montpellier, 361 rue J.F. Breton, BP 5095, 34196 Montpellier Cedex 5, France*

[4] *Environmental Science for Sustainability, Ecosystems and Technology (ESSET), Avans Hogeschool, Lovendijkstraat 63, Breda, Netherlands*

[5] *Department of Natural Resources, General Secretariat of the Tonle Sap Authority, NR 2, Phnom Penh, Cambodia*

[*] *Corresponding author: chork.vuthy@itc.edu.kh*

Abstract: *Mekong river and Tonle Sap are the two main water source in Cambodia while highly transient flow pattern (where water meet) is at Chaktomuk where the confluence-diffluence with varying configurations along time, changing from year to year. The study is about understanding the spatio-temporal dynamics of river water flow and mixing at the Chaktomuk junction of Phnom Penh capital of Cambodia by using hydrochemical tracers. Its unique confluence-diffluence, where four main river branches intersect in varying flow configuration, including shift from normal to reverse flow from Tonle Sap during the wet and dry seasons, makes it a critical area for water management. This study focuses on the four main river branches (Tonle Sap, Bassac, Mekong Upstream, Mekong downstream) using electrical Conductivity (EC) to track the evolution of water origin and mixings through time. Five sampling locations on the four main rivers provides a three-year time-series of EC, completed by high resolution spatial campaigns with continuous surveys and vertical profiling of the stratification of water mixing, providing a 3D image of EC of water dynamics. The results from EC offer novel insights on the dynamics of water mixings, including unexpected sub-fluvial flow and vertical stratification. Understanding these mixing dynamics is crucial for accurately assessing water balance, anticipating quality changes and shifts in ecological patterns, and informing on water supply intake management.*

Keywords: Sub-fluvial flow, Hydrochemical, Chaktomuk, Electrical Conductivity (EC), Spatio-temporal

1. INTRODUCTION

Water quantity and quality are closely related issues that will be made worse by climate change, population increase, urbanization, and water pollution. The changing hydrological cycle may cause a decrease in groundwater recharge, an increase in evaporation, and a diversion from the annual precipitation pattern [1], [2], [3]. Chaktomuk junction (confluence area) of Phnom Penh, Cambodia is a unique confluence-difluence site where two water sources, namely Tonle Sap and Mekong, mix together, and where a new river starts, namely the Bassac river. Mekong River is one of the major rivers of the world. Depending on the reference used, the river stretches somewhere between 4,000 to 4,800 km from its headwaters in the Tibetan Plateau to the South China Sea, making it the 9th to 16th longest river in the world [4]. Tonle Sap River is the outlet of the largest freshwater lake in SouthEast Asia, the Tonle Sap Lake. Its flow reversal during rainy season is a unique phenomenon during which Mekong water reaches the TS Lake. However, water quantity and water quality are frequently viewed as distinct factors, and their intricate interactions are frequently disregarded and

[1]* Corresponding author: Vuthy Chork
E-mail: chork.vuthy@itc.edu.kh; Tel: +855-96 46 45 239

poorly described, underscoring the necessity of integrated water resource study [1], [5], [6]. Electrical conductivity (EC) has been shown to be a good way to track the source of water and its constituents, particularly the main cations and anions, as well as to estimate journey time in surface water. Since EC measurement is simple and affordable, data may be gathered in real time and with a high temporal resolution.[1], [7], [8], [9], [10]. Recent works in Berlin demonstrated the potential of EC as a cost-effective tracer of reverse flow in complex river systems [11].

This study aims at understanding the spatio-temporal dynamics of river water flow and mixing by monitoring time series of EC from the four main branch of the river around the Chaktomuk confluence.

2. METHODOLOGY

2.1 Sampling location

The sampling location is located at confluence area (Chaktomuk junction), Phnom Penh Capital showed in **Fig. 1,** where two water sources (Mekong and Tonle Sap) meet and flow downstream and change from year to year. The interface of Mekong and Tonle Sap is then visible from the surface, as illustrated by color contrasts in **Fig. 1**. This interface always fluctuates in dry season, while during the reverse flow in wet season, Mekong and rainfall are the only water sources to the confluence. Five locations around confluence area shown in were selected to measure the EC and water level on a weekly basis.

Table 1. Information of sampling location

No	Sampling Name	Location
1	S1	Tonle Sap
2	S2	Bassac
3	S3	Mekong Downstream
4	S4	Mekong Upstream
5	S5	Norea



Fig. 1. Study area

2.2 Vertical EC profile

Beside the weekly activity of EC measurement, two campaigns of vertical EC profile were conducted in dry season to see the situation with depth of the mixing interface between Mekong and Tonle Sap water by using CastAway CTD. The measurement was conducted in 2 campaigns on 16 and 19 January 2025. The idea is to collect EC data from the surface until the bottom of the confluence site with interval depth measurement 0.3m.

3. RESULTS AND DISCUSSION

3.1 Spatio-temporal results

The result from weekly measurement of EC from 2023 to 2025 showed the spatio-temporal evolution of EC from the four main branch rivers, considering EC as a tracer of the identity of the water. The results are divided between two water flow patterns: 1) normal flow when the Tonle Sap River flows downwards, i.e., towards the confluence, and 2) reverse flow, when the Tonle Sap River flows upwards, i.e., towards Tonle Sap Lake. During normal flow, EC from the Mekong River (around 300 µS/cm) is always higher than EC from the Tonle Sap River. Intermediate values correspond to mixings

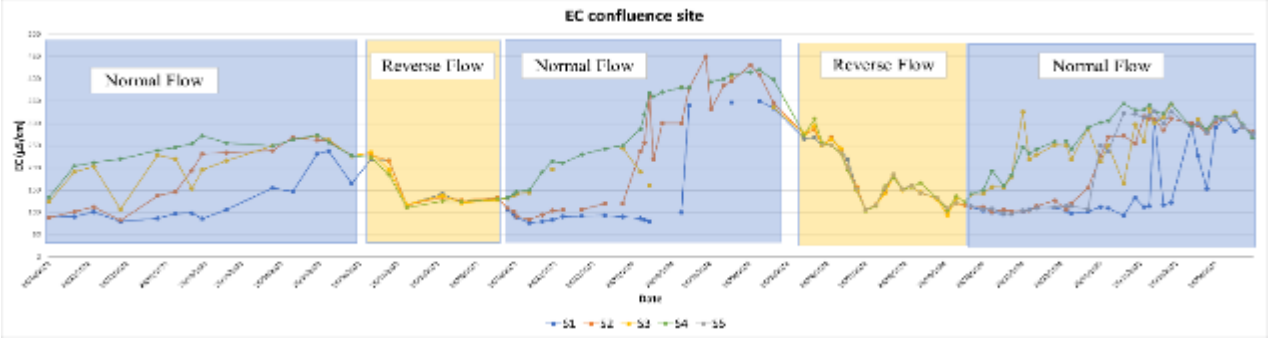


Fig. 2. EC result at confluence area 2023-2025

in the other branches (Bassac River and Mekong River downstream). During reverse flow in the wet season, EC is totally the same all around the confluence (around 100 µS/cm) as Mekong is the only input of water and flow towards all other 3 branches. This low EC is related to the collection of water from rainfall along the Mekong watershed, while higher values in the dry season may be explained by a combination of many factors to be represented, including groundwater inputs along the river. Bassac and Mekong downstream are the two main river that receive varying impact from both Mekong and Tonle Sap. In dry season, EC differentiation (Mekong river vs Tonle Sap river) and mixing between both (Bassac river and Mekong downstream) start from the end of October until the end of May. In Mekong downstream, the EC value in S3 (left bank) is generally similar to EC from Mekong upstream, while at location S5 (right bank), the EC fluctuates between water from Tonle Sap and Mekong water, evidencing varying mixing rates. The EC value in Bassac river has the same as Tonle Sap from the end of October until the end of January. After that, EC evidence shows mixing water from Mekong and Tonle Sap starting to flow to Bassac River, which makes EC increase. In March to April, Mekong-like EC values in Bassac evidence that it receives only water from Mekong due to the water flow from Tonle Sap being lower than water flow from the Mekong.

3.2 Vertical EC profile result

The result from the five locations only shows the situation of EC on the surface. According to the result from **Fig. 2** at the end of normal flow in January 2024, EC values can be very similar in S1 and S5, evidencing a continuous downstream flow from Tonle Sap river, while S2 and S4, i.e. Mekong river and Bassac river, feature similar values. So, there might be something happening below the confluence area. To investigate this hypothesis, two campaigns were conducted in similar conditions (16 Jan 2025, 29 Jan 2025) with CastAway measurements to investigate the mixing area and the situation of Tonle Sap and Mekong water flow below the surface of the confluence area. In the 1st campaign of the measurement,

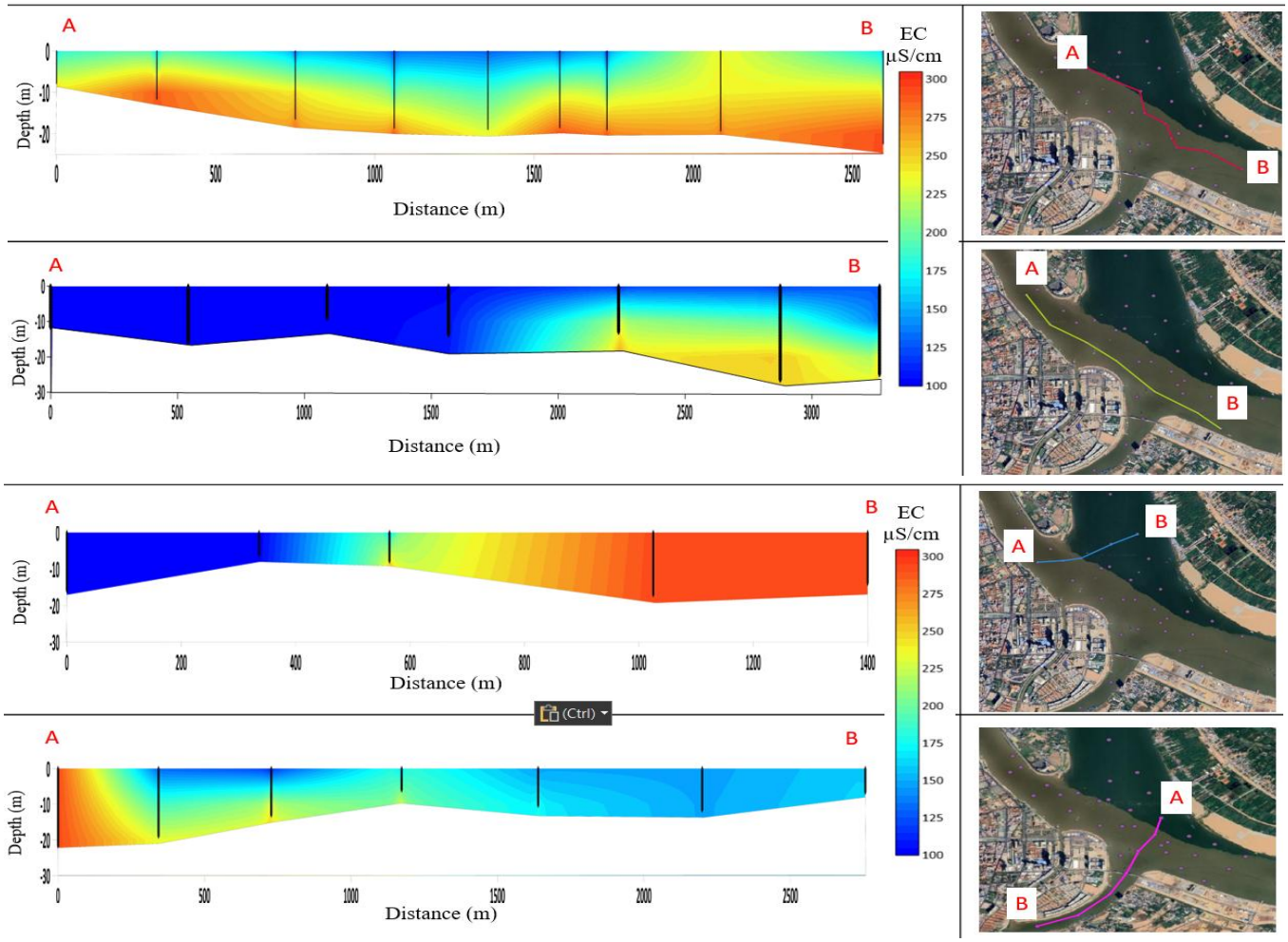


Fig. 3. CastAway result (16/Jan/2025)

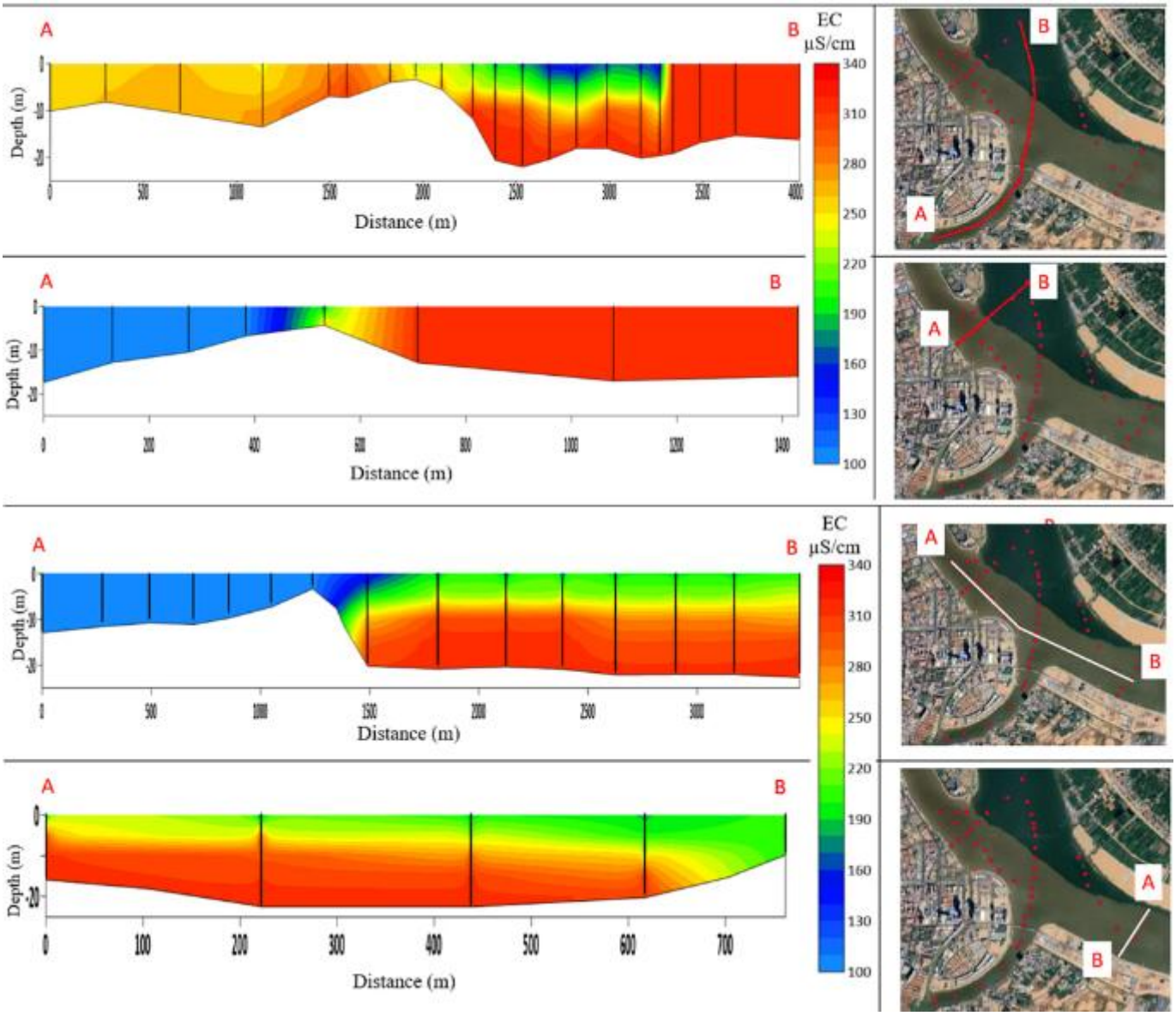


Fig. 4. CastAway result (29/Jan/2025)

cross section 1 is the section along the interface between Tonle Sap and Mekong that shows the mixing EC on the surface (≈ 150 µS/cm) and Mekong water flow below (≤ 270 µS/cm). Section 2 focuses on Tonle Sap to Mekong downstream (Norea site) to show there is less significant contribution of Mekong water below Tonle Sap river, only a mixing. Cross section 3 focuses on Tonle Sap and Mekong upstream also clearly shows the water sources from Tonle Sap (100 µS/cm) and the interface area. The cross section 4 focuses on the section from Mekong upstream to Bassac, the results are showing the water mixing at the interface and slightly mixing flow to Bassac river (150 to 175 µS/cm). The 2nd campaign of the Castaway measurement shows 4 cross sections in **Fig. 4**. The cross section 1 from Mekong upstream to Bassac river. The result show clearly that water from Mekong flow flows below the confluence site and generates highly mixed water flow to Bassac river. The section 2 still clearly represents the Tonle Sap river, Mekong water, and the interface area. Section 3 from Tonle sap to Mekong downstream, expresses high value of EC flow below that represented for Mekong water. Section 4, which is the cross section at Mekong downstream, has no Tonle Sap water above, only mixing water and Mekong water. This is to express that the Mekong water source is starting to reach and have more flow in the confluence area due to the low flow from Tonle Sap.

4. CONCLUSIONS

Inconclusion, EC times series provided timing for anticipating change in flow pattern of Chaktomuk inversion and also showed the highly variable flow pattern revealed by spatial EC distribution. The vertical profile stratification is mandatory in understanding river flow and mixing. The Chaktomuk confluence exhibit dynamic seasonal change influence by discharge from Mekong and Tonle Sap. The reversal and mixing are driven by both hydrological and stratification effect. However, stable isotope will be needed to analyze to clarify and confirm the result from hydrochemistry, and also hydrological modelling need to perform to clarify the river flow in Chatomuk area.

ACKNOWLEDGMENTS

The author is thankful to Coastal and Wetland Environmental Research Lab team members, to the FEF-R Inversap Project funded by the French Ministry of Europe and Foreign Affairs through the French Embassy in Cambodia and IRD for their support for fund support and sample collection.

REFERENCES

[1] C. J. Reith, J. Lewandowski, A. Putschew, T. Goldhammer, J. Filter, and S. Spahr, "Electrical Conductivity as a Tracer for Seasonal Reverse Flow and Transport of Trace Organic Contaminants in River Spree," *Hydrol Process*, vol. 39, no. 2, Feb. 2025, doi: 10.1002/hyp.70084.

[2] G. Konapala, A. K. Mishra, Y. Wada, and M. E. Mann, "Climate change will affect global water availability through compounding changes in seasonal precipitation and evaporation," *Nat Commun*, vol. 11, no. 1, Dec. 2020, doi: 10.1038/s41467-020-16757-w.

[3] P. Däll and J. Zhang, "Impact of climate change on freshwater ecosystems: A global-scale analysis of ecologically relevant river flow alterations," *Hydrol Earth Syst Sci*, vol. 14, no. 5, pp. 783–799, 2010, doi: 10.5194/hess-14-783-2010.

[4] K. N. Irvine, I. G. Droppo, M. Sampson, and J. Sarkkula, "Water quality monitoring and non-destructive suspended sediment size analysis in the Chaktomuk Junction-Tonle Sap Lake Corridor, Cambodia."

[5] Z. Zeng, J. Liu, and H. H. G. Savenije, "A simple approach to assess water scarcity integrating water quantity and quality," *Ecol Indic*, vol. 34, pp. 441–449, 2013, doi: 10.1016/j.ecolind.2013.06.012.

[6] L. A. Candido, G. A. G. Coêlho, M. M. G. A. de Moraes, and L. Florêncio, "Review of Decision Support Systems and Allocation Models for Integrated Water Resources Management Focusing on Joint Water Quantity-Quality," *J Water Resour Plan Manag*, vol. 148, no. 2, Feb. 2022, doi: 10.1061/(asce)wr.1943-5452.0001496.

[7] M. Williams *et al.*, "The use of multiple tracers for tracking wastewater discharges in freshwater systems," *Environ Monit Assess*, vol. 185, no. 11, pp. 9321–9332, 2013, doi: 10.1007/s10661-013-3254-8.

[8] A. Marandi, M. Polikarpus, and A. Jõeleht, "A new approach for describing the relationship between electrical conductivity and major anion concentration in natural waters," *Applied Geochemistry*, vol. 38, pp. 103–109, Nov. 2013, doi: 10.1016/j.apgeochem.2013.09.003.

[9] J. Masse-Dufresne, P. Baudron, F. Barbecot, P. Pasquier, and B. Barbeau, "Optimizing short time-step monitoring and management strategies using environmental tracers at flood-affected bank filtration sites," *Science of the Total Environment*, vol. 750, Jan. 2021, doi: 10.1016/j.scitotenv.2020.141429.

[10] A. D. Kney and D. Brandes, "A graphical screening method for assessing stream water quality using specific conductivity and alkalinity data," *J Environ Manage*, vol. 82, no. 4, pp. 519–528, 2007, doi: 10.1016/j.jenvman.2006.01.014.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Historical Changes in Extreme Daily Temperature Indices in the Last 3 Decades in Cambodia

Ty Sok [1,2 *], Por Y Chheng[2], Meng Hour Huot[1,2], Ilan Ich[3], Penglong Koun[2], Rattana Chhin[3]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[2] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[3]*Ministry of Water Resources and Meteorology, #47 Preah Norodom Bvd, Phnom Penh, Cambodia*

[*]*Corresponding author: sokty@itc.edu.kh*

Abstract: *This study analyzes historical temperature changes in Cambodia from 1971 to 2020, using ERA5 reanalysis data and incorporating a bias correction for observed data. The study employs extreme daily temperature indices—maximum and minimum temperatures, hot days, warm nights, cool days, and cold nights—to measure deviations in temperature patterns. The country average temperature increases of 0.29°C per decade was observed, but with significant regional heterogeneity. Northeastern provinces experienced the most substantial warming, while the southwest coastal and northwest plains showed the least. Analysis of the extreme daily temperature indices reveals decreases in maximum temperatures across several regions, particularly in the Western Highlands, Coastal Plains, and Central Plains. In contrast, minimum temperatures generally decreased countrywide, with the most significant declines observed in Battambang and Kep. These findings highlight the uneven impact of climate change across Cambodia and underscore the need for region-specific adaptation strategies in agriculture, health, and infrastructure development. Further research is recommended to investigate the underlying causes of these regional variations and to enhance predictive modeling capabilities for climate change impacts in Cambodia.*

Keywords: Extreme Daily Temperature; ERA5, Temperature Indices

1. INTRODUCTION

Cambodia experiences a humid tropical monsoon climate, while higher elevations have subtropical conditions. Both regions undergo one wet season and one dry season each year, with their patterns affected by the interannual variability of the El Niño-Southern Oscillation (ENSO) [1].

Recently, climate change trends observed across the country indicate a rise in the frequency and severity of extreme weather events, higher average temperatures, alterations in the timing and length of seasons, and an ongoing rise in sea levels [2]. Between 1971 and 2020, Cambodia's average mean temperature increased by 0.29°C per decade with the greatest changes observed in the northern provinces and most widespread warming during the winter months. Nationwide, average minimum temperatures increased 0.29°C per decade between 1971–2020, while average maximum temperatures increased 0.32°C per decade over the same period [1].

Climate change has significantly altered temperature patterns globally, with noticeable shifts in the frequency and intensity of extreme temperature events. According to the IPCC's Sixth Assessment Report (2021) highlights those anthropogenic activities have been the primary drivers of these changes, leading to increased occurrences of hot days, warm nights, cool days, and cold nights. These indices are critical for assessing the impacts of climate change on various sectors, including agriculture, health, and infrastructure [3].

Therefore, the objective of this study is to assess historical temperature variations across Cambodia from 1971 to 2020 by using ERA5 reanalysis data, which has been bias-corrected against observed station records. The study employs extreme daily temperature indices—maximum and minimum temperatures, hot days, warm nights, cool days, and cold nights—to measure deviations in temperature patterns.

2. METHODOLOGY

[1]Corresponding author: Ty Sok
E-mail: sokty@itc.edu.kh; Tel: +855-11 980 698

2.1 Trends and changes in temperature

The temperature data presented are historical and derived from ERA5 (ECMWF Reanalysis v5) dataset from the European Centre for Medium-Range Weather Forecasts (ECMWF) (1950–2023) [4]. The long-term mean of the model was removed to obtain the anomalous field (by dividing the data by the daily long-term mean). The anomalous field of the bias-corrected model data was then interpolated to a fine scale that matches that of the observation. The interpolated anomalous field was subsequently scaled by the long-term mean of the observation to align the model's spatial feature with the spatial pattern of the observation. The observed data was obtained from the Ministry of Water Resources and Meteorology (MOWRAM) from 1990-2023. Utilizing ERA-5 temperature data, the warming situation in Cambodia was evaluated. The dataset covers around 80 years, from 1940 to 2022. The first 30 years of data (1940–1969) were used as the baseline or reference period for calculating the warming level, while the last 20 years (2003–2022) were selected to represent the current climate in Cambodia. The degree of warming in Cambodia can be determined by adjusting these time parameters.

2.2 Extreme daily temperature indices

Extreme daily temperature indices are metrics used to measure deviations in temperature patterns from historical norms. The key indices under review include:

- **Maximum temperature (TXx):** The annual value of daily maximum temperature.
- **Minimum temperature (TXn):** The annual minimum value of daily minimum temperature.
- **Hot Days (TX90p):** Days where the daily maximum temperature exceeds the 90th percentile of a reference period.
- **Warm Nights (TN90p):** Nights where the daily minimum temperature exceeds the 90th percentile of a reference period.
- **Cool Days (TX10p):** Days where the daily maximum temperature falls below the 10th percentile of a reference period.
- **Cold Nights (TN10p):** Nights where the daily minimum temperature falls below the 10th percentile of a reference period.

3. RESULTS AND DISCUSSION

The warming level based on the ERA5 dataset is shown in Figure 1. using the previously described calculation method. The average temperature increase in Cambodia is 1.42 °C when considering the entire country. The map shows that Kampong Speu, Kampong Chhnang, Kampong Thom, Kampong Cham, Tboung Khmum, Kratie, Mondul Kiri, Ratanak Kiri, Stung Treng, Pursat, Battambang, and Banteay Meanchey have the warmest temperatures, which range from 1.40 to 1.60 °C. The coastal regions get the least amount of warming, which is between 0.80 and 1.10 °C. This indicates that Cambodia's temperature is 1.23 °C warmer than the global average.

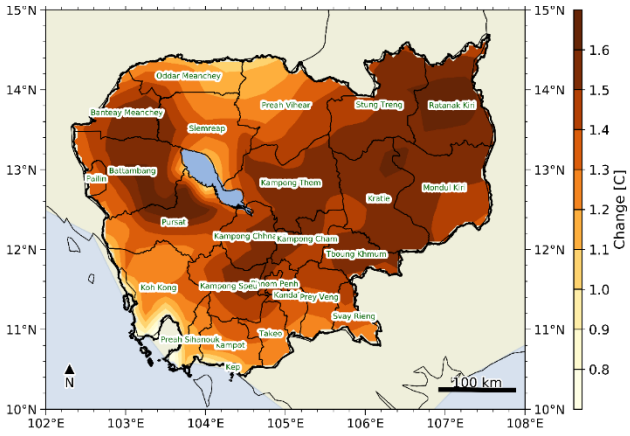


Figure 1: Average Spatial Annual Mean Temperature Trends Nationally per Decade. Warming Status in Cambodia: A Comparative Analysis using ERA5 Data. The reference climate period is 1940-1969, while the current climate status is analyzed for the period 2003-2022.

In the Western Highlands and Coastal Plain, provinces like Kampong Speu, Pursat, and Kep have experienced decreases in maximum temperatures, with Kampong Speu showing the largest drop at -0.25°C per decade (Figure 2). In contrast, Kampot has seen an increase of 0.06°C per decade. Minimum temperatures in this region are consistently decreasing, with Battambang (-0.42°C/decade) and Kep (-0.43°C/decade) showing significant declines.

In the Central Cambodia Plains, maximum temperatures are also decreasing, particularly in provinces as such Kampong Thom (-0.30°C/decade), while Kandal and Svay Rieng are experiencing slight increases of 0.02°C and 0.06°C per decade, respectively. Minimum temperatures are mostly falling, with Banteay Meanchey (-0.48°C/decade) and Otdar Meanchey (-0.59°C/decade) having the largest decreases. Kampong Thom, however, shows a slight increase in minimum temperatures (+0.07°C/decade).

In the Eastern Highlands and Plains, the decrease in maximum temperatures is even more pronounced, especially in provinces such as Mondul Kiri (-0.40°C/decade). Minimum temperatures in the region are also dropping, with Preah Vihear showing the largest decrease (-0.77°C/decade). However, Kratie shows a slight increase in minimum

temperatures (+0.07°C/decade), contrasting with the overall cooling trend.

Most provinces are experiencing decreases in both maximum and minimum temperatures, particularly in the highlands and plains. Some urbanized areas like Kampot, Kandal, and Svay Rieng are seeing slight increases, likely due to factors such as urbanization.

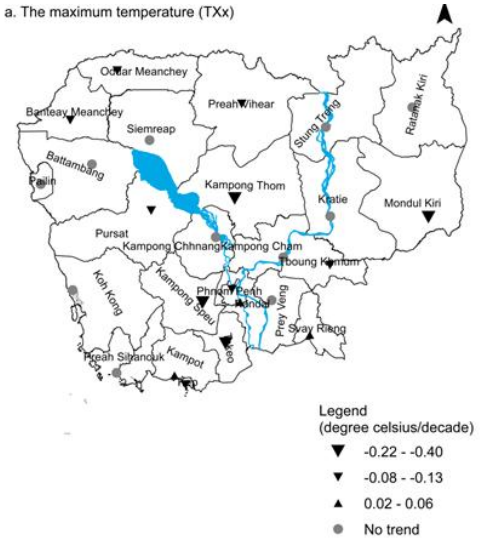


Figure 2: Spatial temperature trends in different regions of Cambodia per decade for annual, focusing on (a). The maximum temperature (TXx).

In the Western Highlands and Coastal Plain, most provinces show minimal changes in extreme temperatures. Notably, Koh Kong shows an increase in hot day maximum (+0.15 days/decade) and a decrease in warm night maximum (-0.25 days/decade), indicating more frequent hot days and fewer warm nights. Kampong Speu province exhibits decline in both hot day maximum (-0.07 days/decade) and cold day minimum (-1.00 days/decade). Kep experiences an increase in cold nights (+0.71 days/decade), while Battambang shows a decrease (-0.67 days/decade).

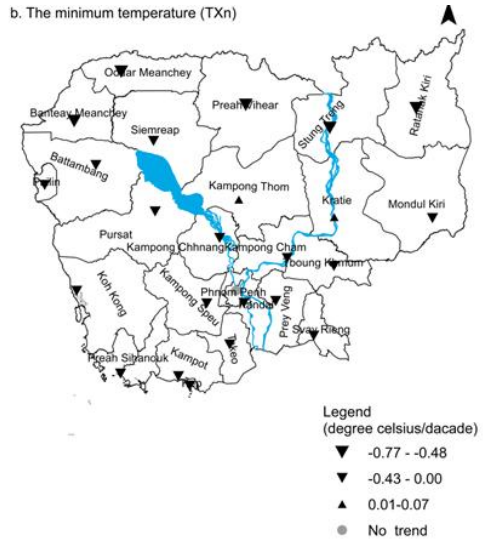


Figure 3: Spatial temperature trends in different regions of Cambodia per decade for annual, focusing on (b). The minimum temperature (TXn)

In the Eastern Highlands and Plains, Stung Treng and Kratie both provinces show decreasing trends in hot day maximum (-0.17 days/decade). Kratie also shows significant decreases in both cold day minimum (-1.15 days/decade) and cold night minimum (-2.27 days/decade). Preah Vihear, on the other hand, shows a minor increase in cold night minimum (+0.50 days/decade).

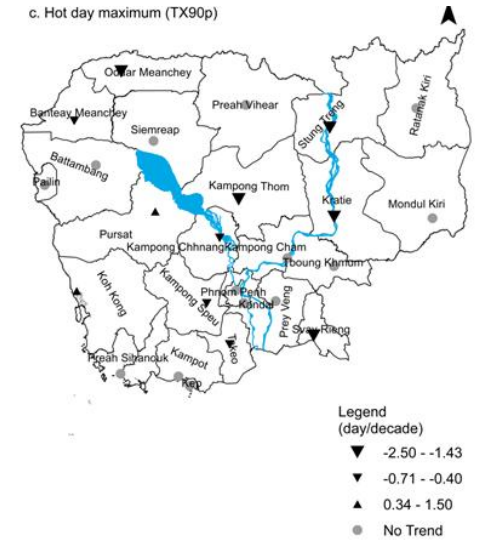
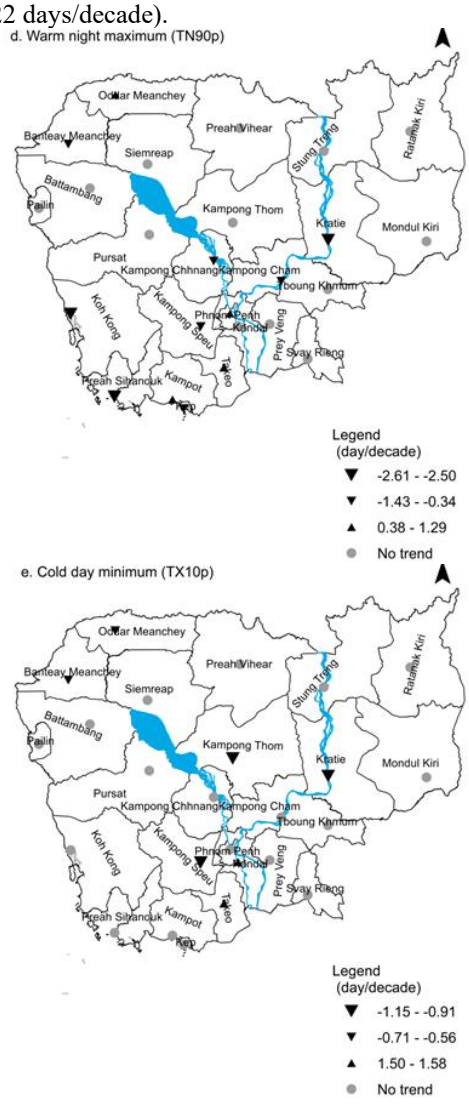


Figure 4: Spatial temperature trends in different regions of Cambodia per decade for annual, focusing on (c). Hot day maximum (TX90p)

In the Central Cambodia Plains, varied temperature trends are observed. Banteay Meanchey province has declines in both hot day maximum (-0.06 days/decade) and warm night maximum (-0.14 days/decade), as well as decreases in cold

day minimum (-0.56 days/decade) and cold night minimum (-0.46 days/decade). Oddar Meanchey exhibits a decrease in hot day maximum (-0.25 days/decade) but an increase in warm night maximum (+0.13 days/decade) and cold night minimum (+0.77 days/decade). Kampong Thom shows a decline in hot day maximum (-0.16 days/decade) and a reduction in cold days (-0.91 days/decade). Phnom Penh has an upward trend in a maximum of warm night (+0.08 days/decade). Interestingly, Takeo displays mixed trends, with a decrease in hot day maximum (-0.04 days/decade) but increases in warm night maximum (+0.04 days/decade), cold day minimum (+1.50 days/decade), and cold night minimum (+2.22 days/decade).



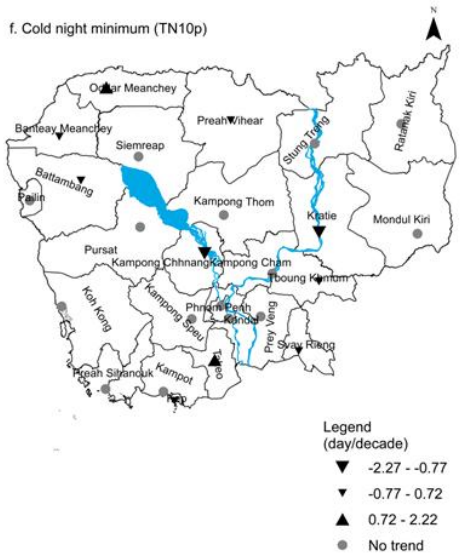


Figure 5: Spatial temperature trends in different regions of Cambodia per decade for annual, focusing on the following indicators: (d). warm night maximum (TN90p), (e). cold day minimum (TX10p), and (f). cold night minimum (TN10p).

4. CONCLUSIONS

This study analyzed the historical changes in extreme daily temperature indices in Cambodia over the past three decades. The results indicate a clear warming trend, with an increase in the frequency of warm days and nights and a decrease in cold extremes. The average temperature increase in Cambodia is 1.42 °C when considering the entire country. The map shows that Kampong Speu, Kampong Chhnang, Kampong Thom, Kampong Cham, Tboung Khmum, Kratie, Mondul Kiri, Ratanak Kiri, Stung Treng, Pursat, Battambang, and Banteay Meanchey have the warmest temperatures, which range from 1.40 to 1.60 °C. The coastal regions get the least amount of warming, which is between 0.80 and 1.10 °C. This indicates that Cambodia's temperature is 1.23 °C warmer than the global average. These changes highlight the growing intensity of heat-related events, which may pose challenges for agriculture, water resources, and public health. The findings underline the need for effective climate adaptation strategies to reduce future risks and enhance resilience to extreme temperature events in Cambodia. Further research is recommended to investigate the underlying causes of these regional variations and to enhance predictive modeling capabilities for climate change impacts in Cambodia.

ACKNOWLEDGMENTS

The authors would like to acknowledge the Ministry of Water Resources and Meteorology for providing dataset.

REFERENCES

[1] S. G. Koh, "Climate risk country profile: Cambodia (2024)," *World Bank,* 2024.

[2] G. Oudry, K. Pak, and C. J. I. O. f. M. Chea, Phnom Penh, "Assessing vulnerabilities and responses to environmental changes in Cambodia," 2016.

[3] J. Lynn and N. J. C. C. Peeva, "Communications in the IPCC's Sixth Assessment Report cycle," vol. 169, no. 1, p. 18, 2021.

[4] H. Hersbach *et al.*, "ERA5 hourly data on single levels from 1940 to present, Copernicus Climate Change Service (C3S) Climate Data Store (CDS)[data set]," ed, 2023.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Contemporary Changes in Precipitation Patterns and Extreme Rainfall Indices in Cambodia

Meng Hour Hout [1,2*], Khamkeo HOAN [1], Ravon Soy [1], Rattana Chhin[3], Ty Sok [1,2], Song Layheang[1,2*]

[1] *Faculty of Hydrology and Water Resources, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Ministry of Water Resources and Meteorology, #47 Preah Norodom Bvd, Phnom Penh, Cambodia*

[*] *Corresponding author: houtmenghour78@gmail.com*

Abstract: *This study investigates the recent shifts in precipitation patterns and the escalation of extreme rainfall events in Cambodia from 1980 to 2023, a region significantly affected by climate variability. This research study utilizes historical climate data and ERA5 reanalysis data incorporated bias correction for observed data, enhanced by statistical analyses and climate modeling, to identify and quantify changes in precipitation patterns and the frequency of extreme rainfall events across Cambodia's region. Statistical tools including the Mann-Kendall Test and Sen's Slope Estimator were employed to analyze the data collected from each province rainfall stations, ensuring a comprehensive evaluation of temporal and spatial precipitation variations. The results indicate a pronounced increase in extreme rainfall events across Cambodia, especially in the western highlands and coastal areas, while eastern regions have seen a decrease in overall precipitation. These findings are critical for developing adaptive strategies to manage water resources effectively and mitigate the risks associated with extreme rainfall, essential for sustaining agricultural productivity and reducing flood and drought risks. Future research will focus on refining predictive models for better accuracy in forecasting and planning, which is vital for policy-makers and stakeholders in the region.*

Keywords: Cambodia; Climate Change; Extreme Rainfall; Precipitation Patterns

1. INTRODUCTION

Cambodia's climate is governed by a tropical monsoon system, with distinct wet and dry seasons shaped by seasonal wind patterns and interannual variability. However, in recent decades, the reliability of these seasonal cycles has diminished due to the increasing influence of global climate change [1]. Variations in rainfall distribution, timing, and intensity—especially the rise in extreme weather events—have raised growing concerns for critical sectors such as agriculture, water resources, and disaster management [2].

Recent studies have indicated an upward trend in both the frequency and magnitude of extreme rainfall events in Southeast Asia, including Cambodia [3]. These changes contribute to more frequent flash floods in low-lying and urbanized areas, while irregular dry spells increase the risk of drought and water scarcity. Regions such as the Tonle Sap Basin and Mekong floodplain have been particularly susceptible to these changes, where agriculture and rural livelihoods remain climate-sensitive [4]. Furthermore, El Niño–Southern Oscillation (ENSO) events have amplified these shifts, with El Niño typically suppressing rainfall and La Niña increasing flood risk through enhanced monsoon activity. This interannual variability adds complexity to Cambodia's climate system, making long-term planning and risk reduction more challenging.

In this context, a comprehensive understanding of long-term rainfall trends and extreme precipitation indices is essential. This study analyzes historical climate observations and ERA5 reanalysis data from 1980 to 2023, applying the Mann-Kendall test and Sen's Slope estimator to detect spatial and temporal changes in Cambodia's rainfall regime. By identifying regions most affected by intensifying rainfall extremes and drying trends, this research provides critical insights to inform climate-resilient planning and adaptation policy at both national and sub-national levels.

[1*] Corresponding author: Hout Meng Hour
E-mail: houtmenghour78@gmail.com; Tel: +855-17 721 153

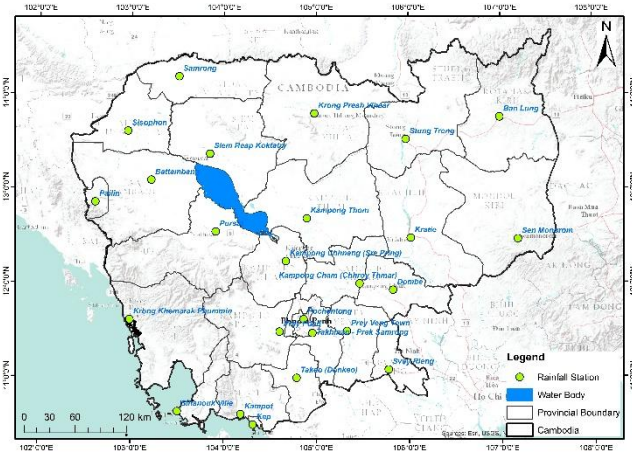


Figure 1 Observed Rainfall Station Location for each Provincial in Cambodia

2. METHODOLOGY

2.1 Data Sources

This research utilizes observed station-level data from the Ministry of Water Resources and Meteorology (MOWRAM) and bias-corrected ERA5 reanalysis data (1979–2023) [5]. Spatial coverage includes 25 stations across four climatic regions: Coastal, Mekong Delta, Tonle Sap, and Upper Mekong.

2.1 Extreme Rainfall Indices

Extreme rainfall indices are measures used to quantify and analyze heavy precipitation events. These indices can help in understanding the intensity, duration, and frequency of extreme rainfall events, which is vital for flood forecasting, water management, and assessing the impacts of climate change. Extreme rainfall indices included in this report are: RX1day, R50, R100, CWD, and CDD (Table 1).

2.2 Mann-Kendall (MK) Test and Sen's Slope

The Mann-Kendall (MK) Test and Sen's Slope help identify and quantify trends in data. The MK Test determines whether a statistically significant trend exists, while Sen's Slope measures the rate of change over time. Key results from the MK Test include the Z statistic and the P-value. The Z statistic indicates whether there is an increasing or decreasing monotonic trend, while the P-value shows if the trend is statistically significant. In this study, a P-value less than 0.05, at a 95% confidence level, suggests that the trend is likely real and not due to random variation. Sen's Slope estimates the magnitude of the trend: a positive value indicates an increasing trend, while a negative value indicates a decreasing trend.

Table 1 Definitions of extreme rainfall indices used in this study

Index	Index name	Index definitions	Units
RX1day	Max 1-day precipitation	Annual maximum 1-day precipitation	mm
R50	Number of very heavy precipitation days	Annual count of days when daily precipitation >=50mm	day
R100	Number of very heavy precipitation days	Annual count of days when daily precipitation >=100mm	day
CDD	Consecutive dry days	Maximum number of consecutive dry days	day
CWD	Consecutive wet days	Maximum number of consecutive wet days	day

3. RESULTS AND DISCUSSION

3.1 Trends and changes in rainfall

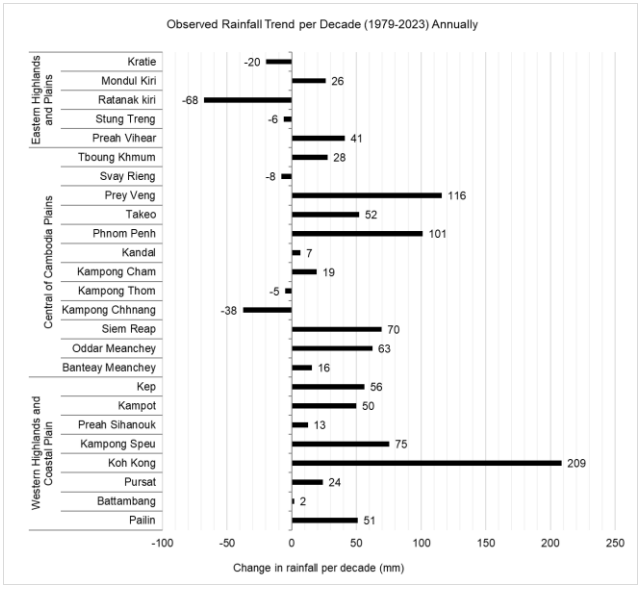


Figure 2 Observed rainfall trend per decade (1979-2023) annually across various province and Cambodian regions. This bar chart represents the magnitude of rainfall changes per decade, highlighting the variability in decadal from 1979 to 2023.

In the Western Highlands and Coastal plain, Koh Kong province observed the most significant increase in rainfall per decade (+209 mm/decade), followed by Kampong Speu (+75

mm/decade). Other significant increases in the Central of Cambodia Plains included Prey Veng (+116 mm/decade), Phnom Penh (+101 mm/decade), Siem Reap (+70 mm/decade), Oddar Meanchey (+63 mm/decade), and Takeo (+52 mm/decade) per decade. Other significant increases in the Central of Cambodia Plains included Prey Veng (+116 mm/decade), Siem Reap (+70 mm/decade), Otdar Meanchey (+63 mm/decade), and Takeo (+54 mm) per decade. In contrast, the Eastern Highlands and Plains showed mixed trends; Ratanak Kiri experienced a decrease per decade (−68 mm/decade), whereas Mondul Kiri recorded an increase (+26 mm/decade).

3.1 Extreme Rainfall Indices

Between 1990 and 2023, Cambodia experienced a general trend of increasing maximum daily precipitation, more days with very heavy precipitation, shorter dry spells, and longer wet spells. In the western highlands and coastal plains, there is a mixed trend in maximum 1-day precipitation (Figure 3). Provinces like Koh Kong, Kep, and Kampong Speu saw a large increase of more than 8 mm/decade, whereas areas along the coast, such as Kampot and Preah Sihanouk, experienced a decrease. The central plains show a slight trend toward both decreasing and increasing precipitation, with a notable decrease of more than 4 mm/decade in Kampong Chhnang. In the eastern highlands and plains, there is a general trend of increasing precipitation, especially in the northeastern provinces like Ratanak Kiri and Preah Vihear. Stung Treng is an exception, showing a decreasing trend.

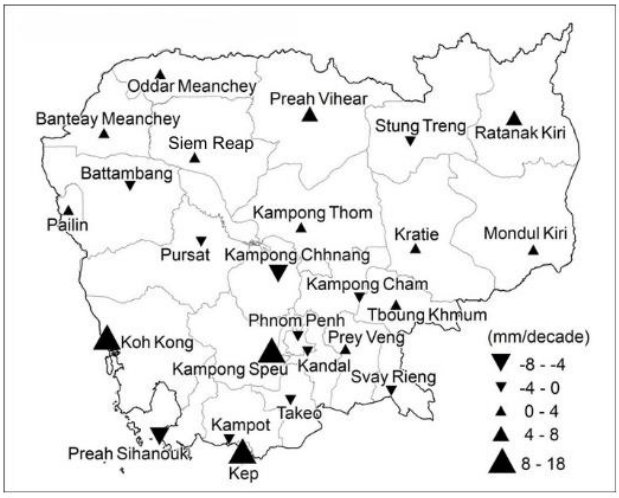


Figure 3: Change trend (day per decade) of at the station of each province in Cambodia between 1990 and 2023. Normal triangles indicate an increasing trend of Maximum 1-day Precipitation, while upside-down triangles indicate a decreasing trend. The size of each triangle indicates the magnitude of an increasing or a decreasing trend. Grey circles indicate no trend.

Only coastal provinces like Koh Kong, Preah Sihanouk, and Kampong Speu in the western highlands and coastal plains show an increase in the number of days with precipitation greater than 50 mm (Figure 4a). In contrast, the other provinces did not experience any trends. The central plains show no trend except in Kampong Chhnang, Prey Veng, and Oddar Meanchey. There is only a notable increase in Mondul Kiri and Preah Vihear in the eastern highlands and plains. Most regions show no significant trend in days with precipitation greater than 100 mm over the past three decades (Figure 4b). However, Koh Kong stands out in the coastal plain with a marked increase.

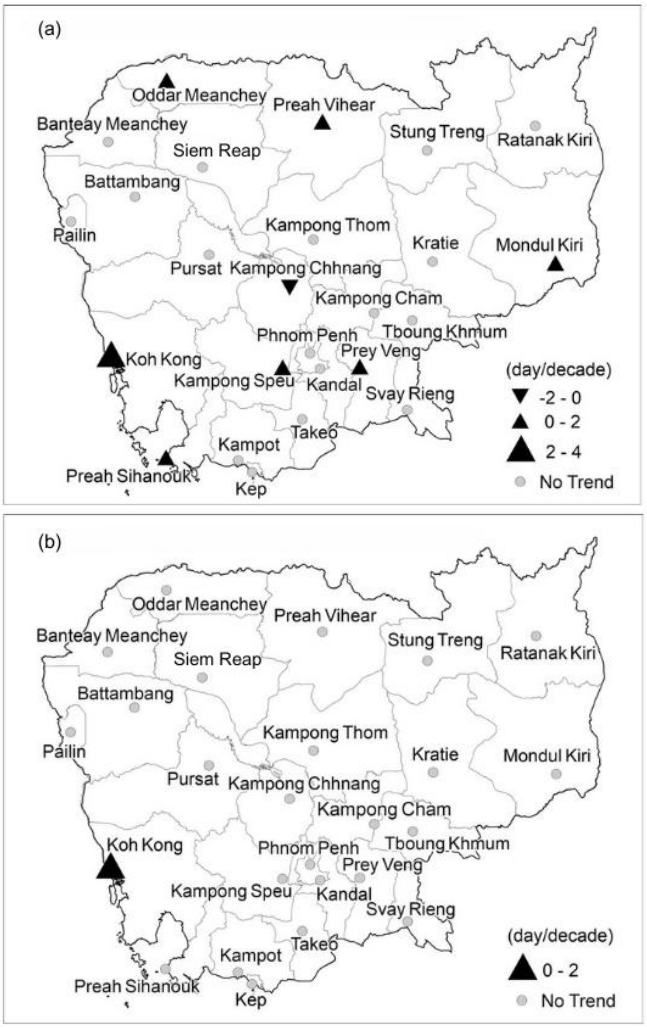


Figure 4: Change trend (day per decade) of (a) Number of days with precipitation larger than 50 mm and (b) Number of days with precipitation larger than 100 mm at the station of each province in Cambodia between 1990 and 2023. Normal triangles indicate an increasing trend of Maximum 1-day Precipitation, while upside-down triangles indicate a decreasing trend. The size of each triangle indicates the magnitude of an increasing or a decreasing trend. Grey circles indicate no trend.

Except for Kratie, all provinces show a decreasing trend in consecutive dry days. The central plains experienced a

significant decrease of more than 8 days/decade (Figure 5a). In the western highlands and coastal plains, there is a notable increase in consecutive wet days, particularly in Koh Kong, Preah Sihanouk, and Kep (Figure 5b). In the central plains, only Takeo experienced a significantly increasing trend, whereas the same trend occurred in Mondul Kiri in the eastern highlands and plains.

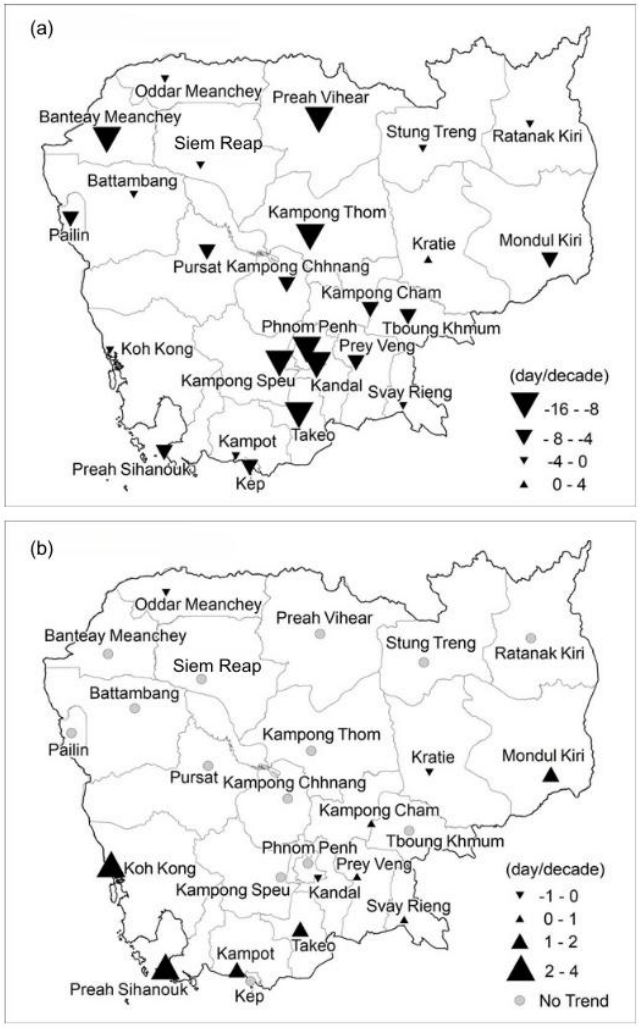


Figure 5: Change trend (day per decade) of (a) Number of consecutive dry days and (b) Number of consecutive wet days at the station of each province in Cambodia between 1990 and 2023.

Cambodia is highly vulnerable to climate-related hazards, including river floods, droughts, extreme heat, wildfires, and sea level rise. From 1987 to 2007, major floods and droughts significantly impacted Cambodia's provinces and are expected to become more frequent in the future. Flooding by midcentury could cause billions of dollars in damage without proper risk planning, increasing the risks of landslides and other hazards. Moreover, provinces with productive agricultural activities face the greatest climate hazard risks from droughts, which have affected millions and challenge water resource management. Additionally, extreme heat, wildfires, and sea level rise pose threats to biodiversity and human health.

4. CONCLUSION

This study provides clear evidence of shifting precipitation regimes in Cambodia over the past four decades, marked by an increase in the frequency and intensity of extreme rainfall events. Spatial analysis indicates that the western highlands and coastal provinces are experiencing more intense rainfall, while parts of the eastern highlands are trending toward drier conditions.

These findings are critical for developing adaptive strategies to manage water resources effectively and mitigate the risks associated with extreme rainfall, essential for sustaining agricultural productivity and reducing flood and drought risks. Future research will focus on refining predictive models for better accuracy in forecasting and planning, which is vital for policy-makers and stakeholders in the region.

ACKNOWLEDGMENTS

Authors would like to acknowledge Ministry of Water Resources and Meteorology for dataset.

REFERENCES

[1] IPCC. (2021). Sixth Assessment Report: Climate Change 2021 – The Physical Science Basis.

[2] ADB & World Bank Group. (2021). Climate Risk Country Profile: Cambodia.

[3] Hijioka, Y. et al. (2014). Asia. In: Climate Change 2014: Impacts, Adaptation, and Vulnerability.

[4] Mekong River Commission (MRC). (2018). State of the Basin Report.

[5] Duong, S., Song, L., & Chhin, R. (2023). Precipitation Projection in Cambodia Using Statistically Downscaled CMIP6 Models. Climate, 11(12), 245.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Assessment of Inundation at Prek System of Koh Thum Inducing by Bassac River and Regional Precipitation Pattern

Kim Hort Ly[1], Khy Eam Eang[2*], Sytharith Pen[2], Ratha Doung[2], Vuthy Chork[2]

[1]*Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2]*Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: khyeam@itc.edu.kh*

Abstract: *Flooding is a constant devastating threat to human settlements, infrastructure and crops especially those low-lying places like Koh Thum in Cambodia. This study aims to define areas that are gravely affected and also delineate threatening areas on map with hydrological modeling, and identify the most severely affected areas and outline high-risk zones on a map using hydrological modeling. Hydrological and hydraulic modeling are used to evaluate characteristics of flooding such as inflow from major tributaries, namely Prek Chann, Prek Touch and Prek Thum and downstream exits. The modeling framework is input with key parameters, including rainfall runoff, terrain elevation, and land use classification, to determine the depth and extent of flooding within the study area. The area has a high susceptibility to flooding, particularly in the eastern section, where multiple channels deliver water, leading to significant pooling. Studies indicate that areas with lower elevations and closer proximity to water channels experience greater flooding depths, with some sections being significantly affected. Enhanced drainage systems and reinforcements of embankments and use of sustainable land use planning stand out among these measures. The study also provides invaluable insights into assessing flood risks comprehensively by integrating hydrological models together with spatial analysis. Future research efforts must concentrate on refining accuracy of models by incorporating dynamic flood data in real time in addition to advanced remote sensing techniques and climate change projections. Also, involving the local community in participatory assessments of flood risk significantly enhances effectiveness of policies relating to flood management. This research makes important contributions to understanding the dynamics of flooding in Koh Thum and highlights the key role of a data-driven approach in reducing flood risk.*

Keywords: Bassac River, Flood Risk Assessment, Hydraulic Modeling, Inundation Patterns, Koh Thum

1. INTRODUCTION

Flooding in Cambodia results from a mix of natural and human-driven factors. The natural influences include the Mekong River, Tonle Sap Lake, and seasonal monsoons, which cause regular flooding of nearby plains. Tonle Sap Lake uniquely expands due to reverse flow from the Mekong during the rainy season (May to October), intensifying floods in low-lying areas like Koh Thum District. Human activities such as deforestation, urban growth, and dam construction have disrupted natural water flow and drainage, increasing flood risks. Climate change further worsens the situation by causing more intense rainfall and rising sea levels [1].

Despite increasing flood challenges, floodplains in tropical deltas like the Mekong remain essential for agriculture, fisheries, and livelihoods, relying on regular flood cycles. However, upstream changes have disrupted flood patterns, threatening food security and ecosystems. Traditional monitoring tools like MODIS and Landsat are limited by low resolution and cloud cover. To improve monitoring, the Water Level–Flood Link (WAFL) method was developed, combining high-resolution Sentinel-1 and Sentinel-2 satellite data with long-term water level records. Applied in the Cambodian Upper Mekong Delta, WAFL has successfully reconstructed over 30 years of flood patterns with high accuracy [2]. This makes it a powerful tool for understanding local flood dynamics in data-scarce settings and for informing risk reduction strategies and adaptive management, particularly in vulnerable regions like Koh Thum where traditional monitoring may fall short.

Koh Thum District in Kandal Province, Cambodia, is highly vulnerable to seasonal flooding due to its proximity to

[*] Corresponding author: Khy Eam Eang
E-mail: khyeam@itc.edu.kh; Tel: +855-12 311 110

the Mekong River. While these floods play a vital role in supporting agriculture and maintaining ecological balance, they also inflict substantial damage on homes, crops, and infrastructure, severely affecting livelihoods and food security in the region. The absence of a reliable flood dynamics models largely due to limited hydrological data, inadequate modeling tools, and insufficient topographic information has resulted in poor flood forecasting and a lack of effective early warning systems. These challenges are further exacerbated by the impacts of climate change and rapid urbanization, which have disrupted natural drainage patterns and altered rainfall regimes, leading to more frequent and intense flooding events, ecological degradation, and soil erosion [3].

2. STUDTY AREA AND METHODOLOGY

2.1 Study Area

The study area is located in Koh Thom District, Kandal Province, Cambodia, between the Bassac River and a network of smaller waterways known as the Prek system, including Prek Chann, Prek Touch, and Prek Thum. This low-lying floodplain region plays a vital role in local agriculture, with rice farming being the dominant land use, supported by seasonal irrigation from both the Bassac and Prek channels. The area's hydrology is highly dynamic, with floodwaters during the rainy season expanding the connectivity between water bodies, enriching soil fertility, and supporting fish habitats. Local livelihoods depend heavily on water resources for farming and fishing, while the region also supports diverse ecosystems typical of the Tonle Sap-Mekong floodplain. However, environmental challenges such as seasonal flooding, sediment loss, water quality degradation, and unregulated land use are increasingly affecting the sustainability of the area. Basic infrastructure, including rural roads and bridges, helps maintain connectivity, although access can be limited during heavy rains. The area is also rich in Khmer cultural traditions, with rivers and Preks playing important economic and spiritual roles in community life.

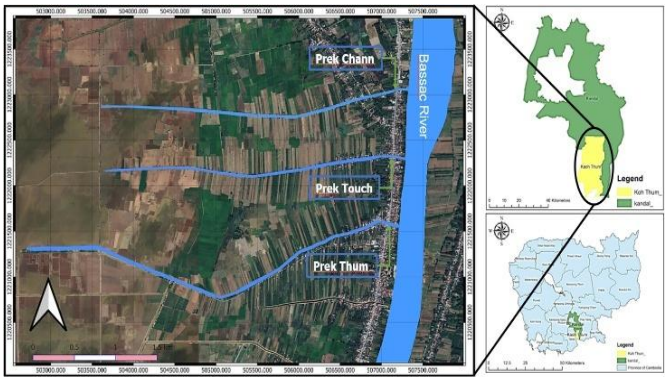


Fig. 1. Study Area at Prek and Bassac River

2.2 Methodology

A two-dimensional (2D) hydraulic simulation was conducted using the Hydrologic Engineering Center's River Analysis System (HEC-RAS) to analyze floodplain behavior in the Bassac River and Prek System, including Prek Chann, Prek Touch and Prek Thum, located in Koh Thum district. The study utilized hydrodynamic modeling integrated with digital terrain data, water level inputs, and Manning's roughness coefficients to simulate flow conditions and flood dynamics. Field data were collected using DGPS technology to obtain accurate elevation, depth, and location measurements across the study area. The data input period covered the wet season, from 01 May to 30 November 2022, to capture critical high-flow conditions relevant for floodplain assessment.

For development of a high-resolution Digital Elevation Model (DEM), accurate geospatial data is being collected using Differential Global Positioning System (DGPS) technology. A DEM is a digital representation of the Earth's surface topography, and its accuracy is crucial for various geospatial analyses. DGPS, a highly precise positioning method, enhances the accuracy of conventional GPS data by correcting signal errors through reference stations. These data points will form the foundational input for generating a detailed and reliable DEM, which is essential for applications such as hydrological modeling, flood risk assessment, and infrastructure planning. Additionally, water level data is being gathered from a monitoring station installed along the Bassac River at Koh Thum. This station is equipped with automated sensors that continuously record variations in river water levels over time. The observational data collected from this site provides valuable insights into the river's behavior, supports the calibration and validation of hydrological models, and plays a key role in informing flood management and water resource planning strategies in the region. Together, DGPS topographic data and water level measurements together enhance understanding of local hydrology, aiding sustainable development and disaster risk management.

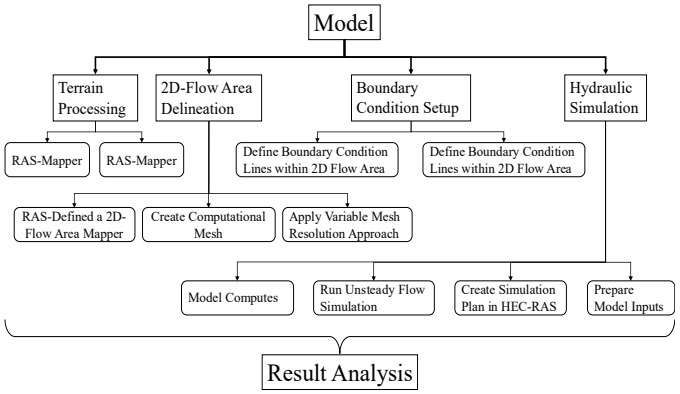


Fig. 2. Overall analysis workflow

2.2.1 Terrain Processing

A high-resolution Digital Elevation Model (DEM) was imported into RAS Mapper which in turn we used to develop the topographic surface required for accurate hydraulic modeling. We generally use the DEM which is in GeoTIFF format to produce a terrain layer that which in detail presents elevation data in the project's coordinate system [4]. This terrain file we found to be very fine scale in its representation of land features which in turn enabled very precise simulation of overland flow and the Bassac River's channel morphology. By use of this very detailed elevation data, we were able to very precisely define cross sections, flow paths and floodplain extents which in turn enhanced the reliability of flood propagation, water surface elevation and flow velocity simulations.

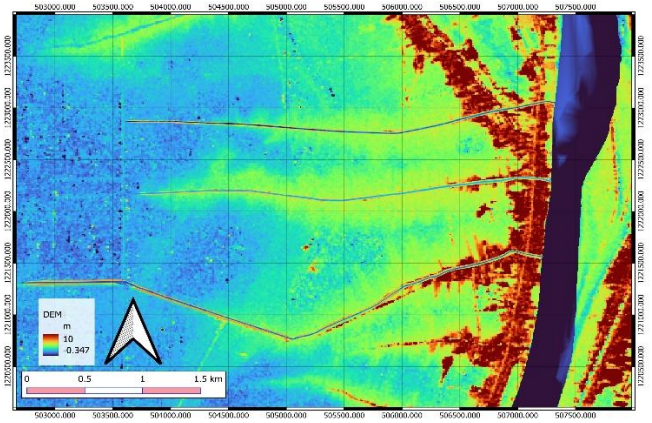


Fig. 3. DEM of Bassac River and Prek

2.2.2 2D-Flow Area Delineation

We defined a 2D-Flow Area in RAS Mapper which includes the region of interest that also has the Bassac River and its floodplain. In this case we divided the area into a computational mesh that is our spatial frame work for which the hydraulic equations that determine flow dynamics are solved. We had a very careful configuration of the mesh in which we focused on the accurate capture of key flow processes at hand also at the same time we did not sacrifice the computational efficiency. Also, we used a variable mesh resolution approach in which we put in finer cells in the most complex hydraulic areas like around the river channel, the banks, and structures to get more detail and do it precisely; also, we used coarser cells in the less dynamic floodplain areas in which the action is not as much in order to reduce the computer resources required. That which is present is a model which is able to very well represent the movement of the water, the flow interactions, and the floodplain connectivity throughout the whole domain [5].

2.2.3 Boundary Condition Setup

Boundary condition lines were established within the 2D Flow Area to define the hydraulic behavior at both the upstream and downstream extents of the model domain. At the upstream boundary, inflow conditions were represented using either hydrographs time-varying flow data or constant flow values, depending on the nature of the simulation. These inputs simulate the volume and timing of water entering the system during typical flood events or specific design scenarios [6]. At the downstream boundary of the model, outflow conditions were defined using either a stage hydrograph estimated based on the downstream channel slope and a rating curve that establishes a relationship between the water surface elevation and the corresponding discharge. These boundary conditions are essential for representing realistic hydraulic behavior at the model's outlet and limits. By accurately simulating the water surface and flow rates at these boundaries, the model can effectively replicate how water enters, moves through, and ultimately exits the system across a range of hydrologic scenarios, including both typical and extreme flow conditions.

2.2.4 Hydraulic Simulation

An unsteady flow simulation was carried out using the defined terrain, geometry, and boundary conditions to model the dynamic behavior of water across the study area. This type of simulation captures time-varying flow responses, making it suitable for representing flood events or other transient hydraulic conditions. The model was executed over a specified simulation period, with carefully selected computational and output time steps to ensure numerical stability and accurate resolution of rapidly changing flow patterns. A simulation plan was created within HEC-RAS, integrating the 2D Flow Area, boundary condition lines, terrain data, and flow inputs into a unified setup. This plan serves as the executable framework that governs the unsteady flow analysis, enabling the software to calculate water surface elevations, velocities, and flow distributions throughout the domain over time [4].

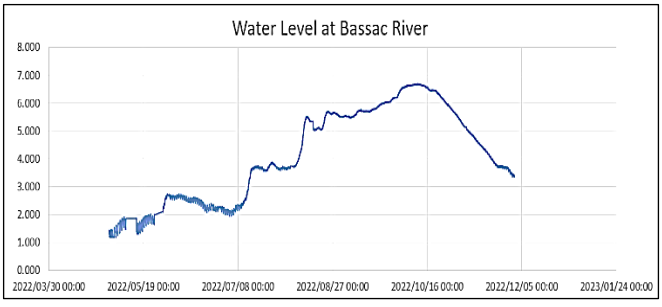


Fig. 4. Water Level (May-Nov)

2.2.5 Flood Risk Analysis

Post-simulation, maximum values of water surface elevation (WSE), depth, and velocity were analyzed within RAS Mapper. These outputs were used to identify flood-prone zones, assess flow patterns, and evaluate risk areas. Visual inspection and spatial overlays facilitated interpretation of hydraulic responses within the modeled area.

3. RESULTS AND DISCUSSION

3.1 Result

The time frame from May 1 to November 30, 2022, was chosen for the flood simulation as it corresponds to the peak of Cambodia's monsoon season, when the region experiences the highest rainfall, river discharge, and flood activity. Within this period, September 22, 2022, was specifically selected because it marked the peak flood event, exhibiting the highest recorded values of water surface elevation, flow velocity, flood depth, and flood extent. This date represents the most severe flooding conditions observed during the study, making it ideal for analyzing worst-case scenarios. Simulating flood dynamics during this critical period allowed for accurate assessment of overbank flooding and inundation patterns, using parameters such as depth, velocity, and water surface elevation. These indicators provided valuable insights for flood risk assessment and informed mitigation planning in the vulnerable Koh Thum District.

3.1.1 Depth

The dynamic interaction between the main river and the Prek system, a network of smaller interconnected channels and reservoirs. This relationship suggests active water exchange influenced by both natural and anthropogenic factors. Key observation points include the prominent main channels connecting the river to the Prek system, which often exhibit bidirectional flow. Typically, the river predominantly feeds the Prek system.

During the high discharge period (May 1 to November 30), characteristic of the monsoon season, overbank flooding occurs. This is when the river overflows its banks, spilling into the adjacent plains and into the Prek system. This process not only recharges the groundwater but also expands the flooded areas, facilitating the exchange of water, nutrients, sediments, and freshwater. Such interactions are essential for supporting the Prek system's ecosystem and play a crucial role in human activities such as agriculture, fisheries, and transportation.

While the simulation does not explicitly include regulation structures like weirs or sluice gates, their presence could significantly influence water flow patterns, especially for irrigation or flood control purposes. The region's generally flat topography aids in water spread, though minor variations in elevation can still impact flow direction.

The flood inundation map reveals a clear depth gradient, with the deepest flooding near the river and shallower depths toward the west, influenced by terrain and limited drainage. Areas near the eastern riverbank, especially low-lying urban zones and road crossings, are most flood-prone, with depths reaching up to 15 meters. Floodwaters extend westward, affecting agricultural land and rural communities. Flow generally follows the natural elevation slope, but drainage systems, canals, and breached infrastructure can intensify flooding in certain areas. These findings are vital for targeted flood risk management and infrastructure planning.

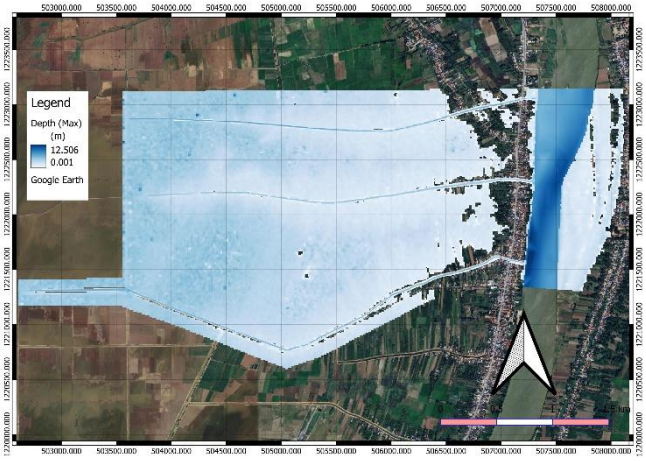


Fig. 5. Water Depth (22 September2022)

3.1.2 Velocity

The velocity analysis derived from the hydraulic model output indicates that the majority of the flooded area experiences relatively low flow velocities, predominantly below 2 m/s. This is visually represented by widespread dark blue shading on the velocity map, which is characteristic of broad floodplains where water disperses and slows down. In contrast, higher velocity zones illustrated in lighter blue to green shades are concentrated along narrow, well-defined pathways. These likely correspond to main river channels, confined flow paths, or areas of constricted flow near urban infrastructure.

The high velocities, approaching 1–1.5+ m/s are minimal or entirely absent in the simulation results. This suggests a limited risk of severe erosion or structural damage associated with fast-moving water, with such conditions likely confined to localized areas, if present at all.

The flow velocity distribution reveals predominantly slow-moving floodwaters, with faster currents exceeding 2 m/s confined mainly to primary river channels and engineered structures. Central and western floodplain areas exhibit stagnant or slow flow, characteristic of natural floodplain

conditions. Flow is driven by terrain gradients, moving from higher to lower elevations, with drainage paths showing reduced velocities. In hydrodynamic analysis, distinguishing between average, instantaneous, vector, and scalar velocities is essential, as each offers valuable insights into flood dynamics and helps identify potential hazard zones.

Accurate velocity analysis in hydraulic modeling relies on several interrelated factors, including displacement, time, flow rate, water depth, channel slope, and cross-sectional area. Understanding the spatial and temporal distribution of flow velocities is crucial for flood risk assessment, infrastructure design, and effective floodplain management.



Fig. 6. Velocity of Water (22 September2022)

3.1.3 Water Surface Elevation

The hydrodynamic model output illustrates the spatial distribution of water surface elevations (WSE) during the modeled flood event, using a color scale ranging from green (representing lower elevations) to pink and white (indicating higher elevations). The highest water surface elevations are concentrated within the main river channel on the right side of the map, signifying deeper or upstream flow conditions. In contrast, lower elevations dominate the left and central portions of the inundated area, corresponding to shallow, slow-moving water spread across floodplains and low-lying agricultural land.

The background terrain model reveals a general slope from left to right, which aligns with the overall direction of water flow from higher to lower elevations. This topographic gradient helps explain the observed distribution of water surface elevations and complements the flow velocity patterns noted in the earlier analysis.

The WSE map shows water surface elevations ranging from about 6.49 to 6.87 meters, with a mostly uniform green color indicating a gentle slope typical of flat floodplains. The slight elevation increases from northwest to southeast aligns with the natural flow direction. This minimal variation suggests low energy gradients and slow-moving water, resulting in widespread, stable flooding with low localized erosion risk but extensive floodplain coverage.

Overall, the WSE analysis confirms a typical floodplain hydrodynamic behavior: shallow, slow-moving water over gently sloping terrain, punctuated by deeper, faster-moving flows within the main river channels and confined pathways.

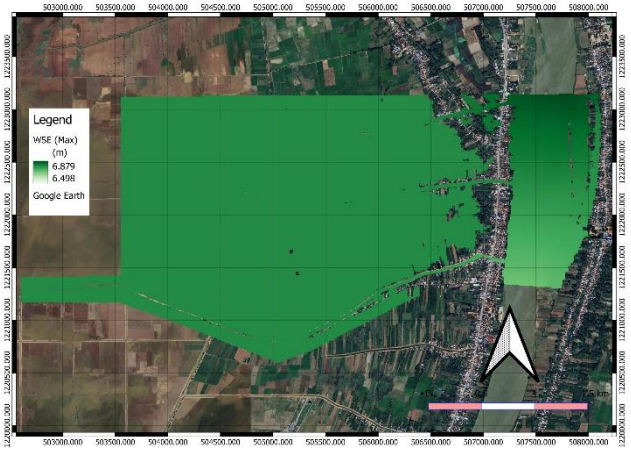


Fig. 7. Water Surface Elevation (22 September2022)

3.2 Discussion

The hydrodynamic simulation results underscore the intricate interaction between the main river and the Prek system a network of smaller interconnected channels and reservoirs highlighting a dynamic water exchange influenced by both natural hydrology and potential anthropogenic modifications. During the monsoon season (May 1 to November 30), elevated discharge from the river leads to significant overbank flooding. This flooding recharges groundwater, disperses water across adjacent plains, and nourishes the Prek system with essential sediments and nutrients, supporting agricultural, ecological, and socio-economic activities. The exchange primarily flows from the river to the Prek system, although localized bidirectional flow occurs along major connecting channels.

The flood inundation map shows deeper flooding near the river, tapering off toward the Prek system and western plains. Flat terrain and poor drainage cause prolonged waterlogging, particularly affecting urban zones along the eastern riverbank and road crossings, while rural areas and farmland to the west are also impacted.

The velocity analysis confirms a predominantly low-energy flood event. Most inundated areas show flow velocities below 2 m/s, consistent with the behavior of broad floodplains where water disperses and slows. Higher velocities, occasionally exceeding 2 m/s, are generally limited to narrow channels, urban drainage systems, or confined flow paths. However, no areas display dangerously high velocities

(above 1.5 m/s), indicating a low risk of severe erosion or structural damage.

The water surface elevation (WSE) map shows a gentle gradient from northwest to southeast (6.49–6.87 m), consistent with natural drainage patterns. Uniform WSE values confirm shallow, slow-moving floodwaters spread across a flat floodplain, with variations linked to channel depth and terrain features.

Overall, the integrated analysis of flood depth, velocity, and WSE illustrates a classic overbank flood event driven by seasonal discharge and terrain-controlled flow paths. While not extreme in energy, the flood poses risks due to its spatial extent, prolonged inundation, and interaction with human infrastructure. Understanding these dynamics is essential not only for developing resilient flood management strategies, infrastructure design, and land-use planning in vulnerable lowland environments, but also for optimizing agricultural practices. In flood-prone regions like Koh Thum, detailed knowledge of flood depth, duration, and flow velocity can directly inform crop planning enabling farmers to select flood-resilient crop varieties, adjust planting calendars, and implement strategic field drainage solutions. This data-driven approach helps minimize crop losses, enhance soil productivity, and ensure more sustainable agricultural outcomes amid increasingly unpredictable flood events.

4. CONCLUSIONS

This study effectively demonstrated the application of a two-dimensional hydraulic model using HEC-RAS to simulate and assess flood dynamics within Koh Thum District a low-lying, flood-prone area situated along the Bassac River in Cambodia's Mekong Delta. The modeling outcomes emphasized the strong hydrological connection between the Bassac River and the Prek system (Prek Chann, Prek Touch, and Prek Thum), driven primarily by seasonal monsoonal rainfall and terrain-induced flow interactions. Overbank flooding during the wet season resulted in significant water exchange, with floodwaters dispersing into secondary channels and across agricultural floodplains, contributing to both ecosystem functions and socioeconomic vulnerability.

Key hydrodynamic indicators including flood depth, water velocity, and surface elevation revealed spatial variations that are closely aligned with topographic gradients and drainage efficiency. Flood depths reached over 1.5 meters near the main river, gradually reducing below 0.5 meters toward peripheral Prek areas, while flow velocities remained generally low (<2 m/s), indicative of typical floodplain behavior but some time higher velocities, occasionally exceeding 2 m/s. The relatively uniform water surface elevation distribution further confirmed the gentle slope and flat nature of the terrain, reinforcing the predictability of flow direction and the risk to low-lying infrastructures.

The study identifies critical flood-prone areas, especially at urban edges and in poorly drained rural zones, offering valuable insights for flood mitigation, land-use planning, and infrastructure design. It emphasizes the importance of using high-resolution terrain data and 2D hydraulic modeling in areas with limited traditional monitoring. The research advocates for integrating engineering-based modeling with community-focused planning. Future work should include real-time hydrological data, advanced remote sensing, and climate change projections, along with participatory assessments to improve early warning systems, resilience, and sustainable development in vulnerable communities.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the support of the Institut de Recherche pour le Développement (IRD) through the Wat-Health project. We also extend our sincere appreciation to colleagues at the Institute of Technology of Cambodia for their valuable support and collaboration.

REFERENCES

[1] Mekong River Commission Secretariat "Annual Mekong Hydrology, Flood and Drought Report 2019: Drought in the Lower Mekong Basin". Vientiane, Lao PDR, Dec. 2021.

[2] C. Orieschnig *et al.*, "A Multi-Method Approach to Flood Mapping: Reconstructing Inundation Changes in the Cambodian Upper Mekong Delta," *J. Hydrol.*, vol. 610, p. 127902, Jul. 2022.

[3] Z. Xu *et al.*, "Morphological Characteristics of Cambodia Mekong Delta and Tonle Sap Lake and Its Response to River-Lake Water Exchange Pattern," *J. Water Resour. Prot.*, vol. 12, no. 04, pp. 275–302, 2020.

[4] K. Vashist and K. K. Singh, "HEC-RAS 2D modeling for flood inundation mapping: a case study of the Krishna River Basin," *Water Pract. Technol.*, vol. 18, no. 4, pp. 831–844, Apr. 2023.

[5] G. Dahal, R. K. Regmi, and S. Adhikari, "Optimization of Model Parameters of HEC-RAS 2D Model on Flood Inundation Mapping: A Case Study of Kankai River Basin".

[6] A. Juárez, A. Adeva-Bustos, K. Alfredsen, and B. O. Dønnum, "Performance of A Two-Dimensional Hydraulic Model for the Evaluation of Stranding Areas and Characterization of Rapid Fluctuations in Hydropeaking Rivers," *Water*, vol. 11, no. 2, p. 201, Jan. 2019.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Review on Historical Tropical Hurricanes, Typhoons, and Cyclones Frequency, Intensity, and Trajectory

CHHENG Por Y[1], KHEN Chanchakriya[1], LAY Lyhieng[1], Huot Menghour[1,2], Phy Sophearum[1], Ty SOK[1,2 *]

[1] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[2] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O.BOX 86, Phnom Penh Cambodia.*

[*] *Corresponding author: sokty@itc.edu.kh*

Abstract: *This review examines the existing literature on the spatial and temporal variability of tropical cyclone (TC) activity and its impacts across Indochina. Studies show varying trends in TC landfall frequency and intensity, with significant regional heterogeneity. While some areas show an increase in TC landfalls, others experience relatively few events. The review highlights the relative scarcity of research focused on Cambodia, despite its vulnerability to TC-related hazards. Cambodia presents a case study in low TC landfall density, contrasting with neighboring countries that experience more frequent and intense events. The literature highlights the importance of considering regional differences in vulnerability assessments and the implementation of targeted adaptation strategies. Further research is needed to resolve inconsistencies in TC data, improve modeling capabilities, and better understand the influence of climate change on regional TC patterns. A more integrated approach, combining climatological and socio-economic data, is needed to inform effective disaster risk reduction and adaptation strategies in Indochina.*

Keywords: Tropical Hurricanes; Typhoons; Cyclones; Intensity; Trajectory

1. INTRODUCTION

Tropical cyclones (TCs) which include hurricanes, typhoons, and cyclones depending on the ocean basin are powerful natural hazards that cause substantial human, environmental, and economic losses. Their development and movement are influenced by various atmospheric and oceanic factors such as sea surface temperature (SST), sea level pressure, and wind circulation. In recent decades, concerns have grown over the potential changes in TC behavior due to climate variability and global warming, particularly with respect to their frequency, intensity, and spatial trajectories.

Recent studies have highlighted an increasing trend in tropical cyclone intensity globally, with some regions experiencing more frequent landfall events, while others show a decrease. Within the Indochina Peninsula including Cambodia, Vietnam, and Laos such trends show strong spatial heterogeneity. From 1979 to 2021, the region experienced over 120 TCs and TDs (tropical depressions), with a modest increase in frequency in the post-2000 era. Observations point to greater atmospheric moisture content, warmer SSTs, and intensified monsoonal flows as key contributors to this shift. However, despite this regional activity, Cambodia has experienced relatively few severe landfalling cyclones and remains understudied in the broader context of tropical cyclone impact and preparedness.

This review seeks to address this gap by synthesizing historical data on TC landfalls, trajectories, and precipitation contributions across Indochina, with a particular focus on Cambodia. The objective is to better understand how regional and seasonal climatic conditions shape TC behavior and to assess Cambodia's relative vulnerability. The scope of this study includes the analysis of long-term climatological data from multiple meteorological agencies and the interpretation of spatial trends in cyclone-induced precipitation and storm paths over a 40-year period.

By focusing on both the physical characteristics of cyclones and their regional manifestations, this study aims to inform future disaster risk reduction strategies and support

[1*] Corresponding author: Ty SOK
E-mail: sokty@itc.edu.kh; Tel: +855-11 980 698

climate adaptation planning in Southeast Asia's inland and coastal areas.

2. Tropical cyclones and Tropical Depressions in Indochina

There were 123 TCs (tropical cyclones) and TDs (tropical depressions) making landfall in Indochina from 1979-2021 (59 from 1979-2000 and 64 from 2001-2021, Figure 1) [1]. The number of TCTD in the latter period is slightly higher than that in the former period, partly due to a noticeable increasing trend in annual precipitation patterns, making TC/TD landfall frequency evident post-2000. Further explanation from the same study lies in analyses for seasonal mean precipitation water, SST (sea surface temperature), sea level pressure, and wind patterns in 2001–2021. Higher precipitation water values are observed in 2001–2021, with notable increases in open oceans, suggesting more water vapor transport to Indochina. Warmer SSTs in the South China Sea and Philippine Sea, particularly in the northern region, also indicate warmer air holding more moisture. Additionally, a low-pressure system from eastern China to northern Indochina during 2001–2021 is linked to a stronger westerly wind pattern from the Indian Ocean, pointing to a robust Asian monsoon channeling increased moisture to Indochina.

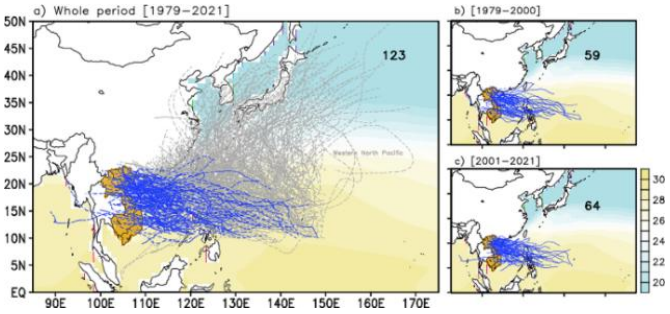


Fig. 1: Trajectories of tropical cyclones (TC) and tropical depressions (TD) from the South China Sea and the Philippine Sea (1979-2021) are shown in (a). Blue lines indicate TCs and TDs that made landfall in Indochina (Cambodia, Laos, and Vietnam), while grey lines show those moving northward or elsewhere. The corner number shows TCs and TDs landing in the mainland. Shading represents the long-term mean sea surface temperature (1979-2021), and the orange polygon outlines Indochina. Panels (b) and (c)display the landfall trajectories in Indochina from May to December for 1979–2000 and 2001–2021, respectively [1]

3. Tropical cyclones and Tropical Depressions in Cambodia

Covering a span of four decades (1979–2021), TC contributions to the annual precipitation are 6% in Cambodia, increasing to 7% when TD were taken into account (Figure 2). Moreover, the contribution is approximately 3% (approximately 50 mm per year) in southern Cambodia since TC and TD from the South China Sea and Philippines Sea seldom reach that deep of the tropics.

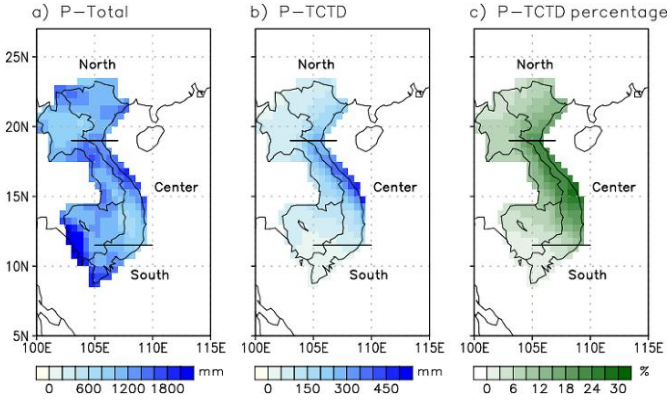


Fig. 2: Spatial distribution of climatological precipitation displaying a) annual precipitation, b) precipitation induced by TCTD (P-TCTD), and c) its contribution to the climatology (P-TCTD percentage) (Ho et al., 2024) [1].

Cambodia has experienced 29 land falling TCs with a relatively low density (1.6 x 10-4) of landfall TCs per square kilometer (Table 1), indicating while tropical cyclones do occur, they are relatively rare when considering the country's overall size. This low density could indicate that Cambodia is not a primary target for TCs compared to other regions in Southeast Asia. However, there were 33 records that were given total intensity measurements of more than one intensity state during landfall, in which 24 and 9 are characterized as tropical depressions and tropical storms. The absence of typhoons and super typhoons suggests that Cambodia is not commonly affected by the most severe cyclonic events, characterized by wind speeds of 119 km/h and 240 km/h, respectively, possibly due to its geographical location and the natural path of storm systems in the region.

Table 1: Number of Land falling TCs, Land falling TCs per a Square Kilometer

Type of Extreme Disaster	Number
Land falling TCs	29
Land falling TCs/ km^2 x 10^{-4}	1.6
Entries	34
Land falling Intensity Values	33
Tropical Depression	24
Tropical Storm	9
Typhoon	0
Super Typhoon	0

Note: Actual Landfalls (Including Multiple Landfalls by a Single TC), Intensity at Landfall Values from the JTWC Best Track Archive, and Number of Landfalls by JTWC Intensity Categories for Cambodia [2].

Inland countries in Area 3 such as Laos and Cambodia are typically impacted by TCs that first make landfall along the Vietnam coastline (Figure 3) [2]. Moreover, Cambodia experiences a large number of tropical depressions as many as Thailand and Laos. To a certain extent, tropical storms, traversing westward from Vietnam, usually hit the northeastern and eastern Cambodia provinces. Based on the same figure, no typhoons and super typhoons have been observed in the country.

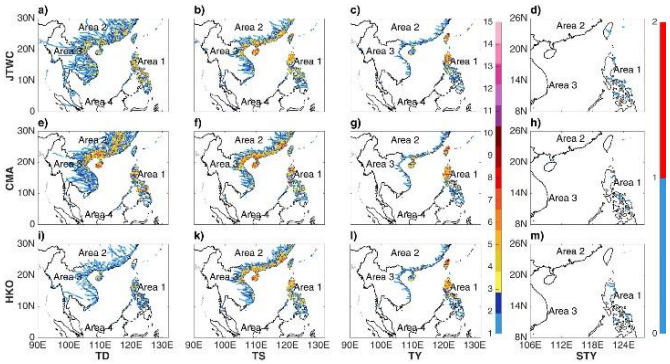


Fig. 3: TC landfall frequency in 0.25° × 0.25° grids sorted by the Joint Typhoon Warning Center (JTWC) intensity classification for: (a) Tropical Depression; (b) Tropical Storm; (c) Typhoon; and (d) Supertyphoon in the JTWC dataset; and (e)-(f)-(g)-(h); (i)-(k)-(l)-(m) as in (a)-(b)-(c)-(d) but for the China Meteorological Administration (CMA), and Hong Kong Observatory (HKO) agency records [2].

The TC Density is scattered in the riparian countries, with a greater density of TCs in the eastern of the lower Mekong River Basin, which essentially includes Cambodia (Figure 4) [3]. Generally, the spatial coverage of the MRB affected by TCR for 1983–2016 at annual scale shows high variability with a mean percentage of 66.6%; while 2002 is an extreme year of the lowest TCR influence (23.9%).

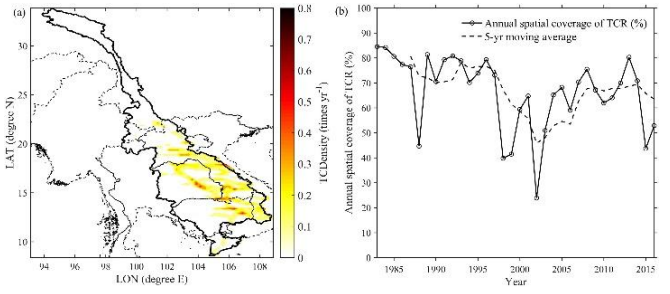


Fig. 4: (a) Spatial patterns of annual mean TC Density (times yr−1) and (b) annual spatial coverage of TC rainfall (TCR in %) across the Mekong River Basin for 1983–2016 [3].

4. CONCLUSIONS

This study provides a comprehensive overview of the historical trends and changes in tropical cyclone (TC) activity across the Indochina region from 1979 to 2021, with a focus on Cambodia. The findings reveal that although the overall number of TCs and tropical depressions (TDs) making landfall in Indochina has slightly increased in recent decades, Cambodia experiences relatively few direct impacts compared to neighboring countries. Specifically, only 29 landfalling cyclones have been recorded in Cambodia over the last four decades, with none reaching typhoon or super typhoon intensity.

Despite this low landfall density, Cambodia still receives a non-negligible amount of its annual rainfall from TCs and TDs up to 7% in certain areas indicating an indirect but significant influence on the country's hydrological cycle. The spatial distribution of TC-related precipitation suggests a stronger impact in northeastern and eastern provinces, particularly from weakened systems traversing westward from Vietnam. Changes in sea surface temperatures, increased moisture content, and evolving monsoonal patterns have contributed to shifts in cyclone trajectories and rainfall distribution across the region.

These findings highlight the need to consider not only direct cyclone strikes but also the broader climatic interactions that affect inland nations. Cambodia's vulnerability, while less apparent in terms of storm intensity, remains considerable due to its exposure to residual rainfall and flooding from storms originating in surrounding regions.

Future research should prioritize high-resolution modeling and enhanced regional data sharing to better understand the dynamics of TC behavior in low-impact zones. An integrated approach that combines meteorological, hydrological, and socio-economic data is essential to develop tailored adaptation and disaster resilience strategies for communities across Indochina.

ACKNOWLEDGMENTS

Authors would like to acknowledge the Hydro-Met and Disaster Management Laboratory for supporting this work, and the Ministry of Water Resources and Meteorology for providing the necessary datasets.

REFERENCES

[1] Ho, T. N. H., Wang, S. Y. S., & Yoon, J. H. (2024). Trends and variations of tropical cyclone precipitation contributions in the Indochina Peninsula. Theoretical and Applied Climatology, 1-13.

[2] Tran, T. L., Ritchie, E. A., & Perkins-Kirkpatrick, S. E. (2022). A 50-year tropical cyclone exposure climatology

in southeast Asia. Journal of Geophysical Research: Atmospheres, 127(4), e2021JD036301.

[3] Chen, A., Ho, C. H., Chen, D., & Azorin-Molina, C. (2019). Tropical cyclone rainfall in the Mekong River Basin for 1983–2016. Atmospheric Research, 226, 66-75.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Evaluating Water-Saving Irrigation Methods for Rice Cultivation, Case Study in Cambodia.

Mengheak Phol[1*], Chetra Pheun[3] Chankethya Sun[3], Vanndy Lim[3], Veasna Touch[3], Pinnara Ket[1], Chengxiang Ma[2]

[1]*Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2]*Gansu Academy for Water Conservancy, China*
[3]*Cambodian Agricultural Research and Development Institute (CARDI)*

[*]*Corresponding author: phol_mengheak@gsc.itc.edu.kh*

Abstract: *Due to water scarcity caused by global climate change and limited water sources, rice cultivation in Cambodia has been worsened by drought and unavailable of water for irrigation. However, regarding Cambodia, there is still a significant gap in scientific research on the feasibility and performance of these water-saving irrigation methods for rice cultivation. This study aims to reduce water use, improve water productivity, and identify suitable irrigation practices for dry-season rice cultivation at the Cambodian Agricultural Research and Development Institute (CARDI) by comparing two water-saving methods with the conventional approach. The experiment was split-plot based on a randomized complete block design with 3 treatments and 3 replications. The treatments were conventional irrigation (CK), Shallow-deep-Shallow irrigation (S), and Shallow and Frequent irrigation (Q). The results of this study showed that Shallow-deep-Shallow (S), Shallow and Frequent (Q) significantly improved water saving by 32% to 53% over CK ($p < 0.05$). Water productivity (WP) ranged from 0.61 to 0.89 kg/m³. However, our results showed that water-saving irrigation method did not significant effect on grain yield, canopy cover, and plant height ($p > 0.05$), compared to CK treatment, and the average yield of rice ranged from 3.3 tons/ha, 3.4 tons/ha, and 3.8 tons/ha in S, Q, and CK. Our findings demonstrate that implementing water-saving irrigation, particularly the Q method, has potential for rice cultivation on sandy loam soils by significantly enhancing water use efficiency without negatively affecting rice yield.*

Keywords: Dry season rice, Water productivity, Water-savings, Yield.

1. INTRODUCTION

Drought conditions in Cambodia significantly restrict water resources during the dry season, thereby introducing uncertainty for agricultural practitioners. Projections indicate an increase in the duration of dry periods, more intense rainfall events, and rising temperatures [1,2]. With the population anticipated to double by 2050 [3], enhancing agricultural production, particularly dry-season rice yields, is imperative. Traditional paddy rice farming practices depend heavily on continuous flooding, requiring 5–10 cm of water, which results in substantial water demand [3,4]. To advance water use efficiency, various methodologies have been developed, including Alternate Wetting and Drying (AWD), aerobic rice, and rain-gathering irrigation [5,6]. Even minimal water savings can yield significant societal and environmental benefits [7]. The AWD technique, developed by the International Rice Research Institute (IRRI), allows for the soil to dry between irrigation events while often maintaining yields comparable to those achieved through continuous flooding [8,9]. However, the effectiveness of this method varies based on soil type, water conditions, and rice variety [5,10]. In Cambodia, the implementation of Alternate Wetting and Drying (AWD) has resulted in modest yield increases of 6–8% and water savings of 5.7–7.3%; however, these benefits have been insufficient to drive widespread adoption [11]. A recent study by Phoeurn et al. (2024) reported that AWD with a 15 cm threshold (AWD15) reduced water use by 10–30% in sandy loam soils; however, no significant improvements were observed in clay soils. These savings were still lower

[1*] Corresponding author: Mengheak Phol
E-mail: phol_mengheak@gsc.itc.edu.kh; Tel: +855-68 719 412

compared to other studied by [3]. Given the limitations associated with AWD, it is critical to explore alternative water-saving techniques that can achieve greater water efficiency without compromising rice yield. This study aims to reduce water use, improve water productivity, and identify suitable irrigation practices for dry-season rice cultivation at the Cambodian Agricultural Research and Development Institute (CARDI) by comparing two water-saving methods with the conventional approach. We hypothesize that the implementation of new water-saving irrigation methods, specifically shallow-deep-shallow (S) irrigation and shallow and frequent (Q) irrigation, will significantly reduce water usage and improve water productivity without negatively impacting rice yields, compared to the conventional continuous flooding method (CK) during the dry season in Cambodia. These methods are expected to offer substantial benefits to farmers, particularly in the context of increasing water scarcity and limited water availability.

2. METHODOLOGY

2.1 Study Area

The experiment was conducted at the Cambodian Agricultural Research and Development Institute (CARDI), Phnom Penh, Cambodia. With latitude and longitude were 11° 28'37 31" N and 104° 48' 29.19" E. CARDI is located in Prateah Lang Commune, Dangkor district, 20 kilometers south of Phnom Penh, in the dry season from 11th January to 19th April. The climate in the study area is a tropical climate with an average monthly minimum and Maximum air temperature of 24 and 34 °c during the planting period. The total average daily precipitation is 0.57 mm, mainly from January to April. the daily temperature and rainfall during the growing period in 2023 are represented in (Fig. 1). The total experiment size is 46.6 m × 16.2 m and was randomized and split into nine blocks with three different treatments and three replications. The rice variety is Sen Kraon-01 which is recommended by CARDI.

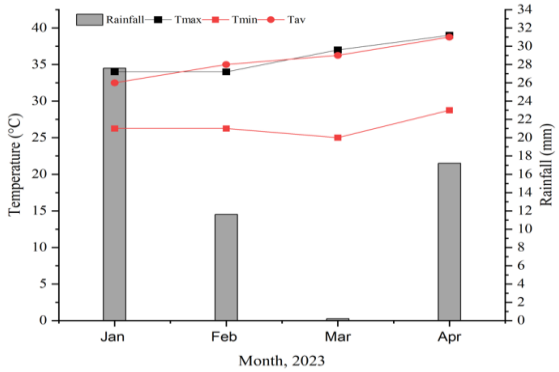


Fig. 1. The daily temperature and rainfall during the growing period.

2.2 Experimental design

The field experiment was a completely randomized design in three replications and consisted of three treatments, namely: (1) Conventional irrigation (CK), (2) Shallow-deep-Shallow irrigation (S), (3) Shallow and Frequent irrigation (Q) with the same soil type and amount of fertilizer. The 15 m × 4.8 m replication plot sizes were separated by soil ridges covered with plastic film to inhibit water and nutrient exchange between plots. Three irrigation regimes with different water controls in different growth stages were included in these experiments, corresponding to seven growth stages of rice: regreening, early tillering, late tillering, jointing-booting, heading-flowering, milk, and yellow ripening. Irrigation was initiated when the water depth reached the lower limit, and excess rainwater was drained to the maximum storage height. For the CK treatment, the field was continuously flooded with 50 mm of surface water at the regreening and early tillering stages, 70 mm at the late tillering stage, and 100 mm at the jointing-booting, heading-flowering, and milk stages. The water was drained out at the yellow ripening stage, and the soil water content was maintained at 70%. For the shallow-deep-shallow treatment (S), the water level was maintained at 30 mm during the regreening and early tillering stages. At the late tillering stage, the field was drained, and the soil water content was lowered to 70%. The field was then flooded with 70 mm of water at the jointing-booting and heading stages and with 30 mm of water at the milk stage. At the yellow ripening stage, the field was drained again, and the soil water content was maintained at 70%. The water level was allowed to fluctuate between 10 mm and 70 mm throughout the S regime. For the shallow and frequent irrigation treatment (Q), the water level was maintained at 30 mm throughout all growth stages, except for the late tillering and yellow ripening stages, when the field was drained, and the soil water content was reduced to 70%. During the late tillering stage, the surface water was removed from the rice crop for 5 to 7 days, allowing the soil to dry, crack, and re-aerate. The water level was allowed to fluctuate between 10 mm and 30 mm in the Q regime, meaning that water was reapplied to the field when the water level reached 10 mm.

2.3 Soil sampling

Soil samples were taken from the experiment site at the depth of 0 to 30 cm to measure soil texture and soil bulk density. The samples were prepared and dried at 105°C for 24h. The bulk density was then calculated as:

$$\rho_b = \frac{m}{v} \qquad \text{(Eq.1)}$$

Soil texture, which is defined by the composition of particle size, namely sand, silt, and clay, will be analyzed in the

laboratory using hydrometer analysis (Barman & Choudhury, 2020; Bouyoucos, 1962). Soil texture was classified using the USDA soil textural triangle, which defines soil types based on the proportions of sand, silt, and clay.

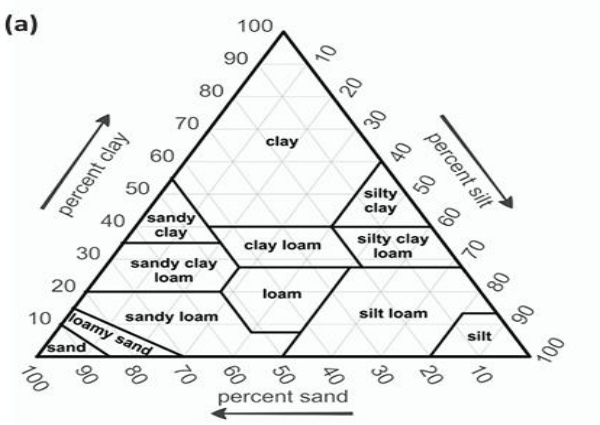


Fig. 2. Soil classifications compared in this study [12].

2.3 Crop measurement

Canopy cover (CC) was measured at 7-day intervals. Canopy cover measurements were taken using a Samsung Galaxy Note 20 smartphone at three points per plot. The images were captured in a resolution of 1080 × 2040 between 9:00 and 11:00 a.m. each week from early tillering to ripening, under clear-sky conditions. Vertical color images of the paddy fields were taken from a constant height of 2 meters, covering an area of 1 m². Canopy cover sampling followed the methodology described by Yu et al. (2020). The fractional green canopy cover was calculated using a mobile application called Canopeo®.

2.4 Field measurement

The water level measuring sticks were used to monitor and measure the water level in each plot. Plastic sheets were installed in the bunds to a depth of 20 cm to prevent lateral seepage between plots with different water regimes. In all sites, the water level was regularly monitored manually; when the water dropped to the critical levels for S, Q, and CK treatments, irrigation was initiated. The plots were irrigated according to the growth stage of the rice, and soil moisture content was managed using a soil moisture test kit. Water needed to be drained from the plot if rainfall caused the water level to exceed 80 mm. Irrigation was supplied through tubes connected to a pump. Due to technical limitations and in the absence of water meters, the irrigation volume was calculated by multiplying the cross-sectional area of a plot by the water height at a reference point, without accounting for the infiltration rate.

2.5 Water productivity

Water productivity is expressed as the sum of total irrigation water productivity (WP) and rainwater productivity (RWP), which together represent the total water used (rain + irrigation). In our study, rice was cultivated during the dry season, with no rainfall occurring throughout the growth period. Water productivities were determined using the following relationships [13]:

$$WP+R= Y/I+R \qquad \text{(Eq.2)}$$

Where: Y is the yield of rice (kg/ha), I is the amount of irrigation water (m3/ha), and R is rainfall (mm).

Water savings were assessed relative to the irrigation water applied, and the difference in irrigation under the three irrigation regimes was calculated compared to the irrigation water used in the CK regime. This is expressed in the following equation:

$$\text{Water saving (\%)} = (\text{Water applied in CK plot} - \text{Water applied in Water saving plot}) \times 100) / (\text{Water applied in CK plot}) \qquad \text{(Eq.3)}$$

2.6 Yield and 1000-grain yield

A 24 m^2 rice plant at physiological maturity was harvested for yield determination. Grain yield was adjusted to 14% moisture content using the formula [8]:

$$Yield = GW \times (100 - GMC)\,\% / (100 - 14)\,\% \qquad \text{(Eq.4)}$$

Where: GW= Grain weight (kg), GMC = Grain moisture content (%).

The dry weight of 1,000 grains from three replicate samples of filled grains per plot was obtained by drying them at 70 °C in the oven for 48 hours until a constant dry weight was achieved.

2.7 Statistical analysis

One-way ANOVA was used to compare the effects of two water-saving irrigation methods (S and Q) on water reduction, grain yield, 100-grain weight, and canopy cover, compared to the conventional method (CK). This method is appropriate for identifying significant differences among treatment means. The LSD test ($p < 0.05$) was applied for pairwise comparisons, and standard deviations were calculated using Origin 2025.

3. RESULTS AND DISCUSSION

3.1 Soil texture and Bulk density

According to Table 1, the topsoil layer (0–30 cm) at the experimental site is classified as sandy loam, with a measured bulk density of 1.66 g/cm³.

Table 1. Soil characteristics of the experimental site at CARDI

Year/Site	Layer	Depth (m)	Texture	Bulk density (g/cm³)
2023 CARDI	1	0–0.30	Sandy loam (62% sand, 25% silt, 13% clay)	1.66 ± 0.02

3.2 Effect of water-saving irrigation on the growth of Canopy cover.

The interaction of the water regime did not significantly influence the time-series canopy cover across all treatments. From the early tillering stage to the ripening stage, canopy cover increased independently of the water regimes. (Fig. 2) illustrates that the mean canopy cover increased significantly from the late early tillering stage (15 days after transplanting) to the middle heading-flowering stage. It then began to decline from the late heading-flowering stage to the ripening stage. In all treatments, the maximum canopy cover (CC) percentages observed in 2023 were 64% for CK, 64% for S, and 66% for Q. Water-saving methods demonstrated similar canopy cover percentages when compared to the conventional method (CK). This finding is consistent with research by [14], which also reported no significant differences in canopy cover between water-saving and conventional irrigation methods.

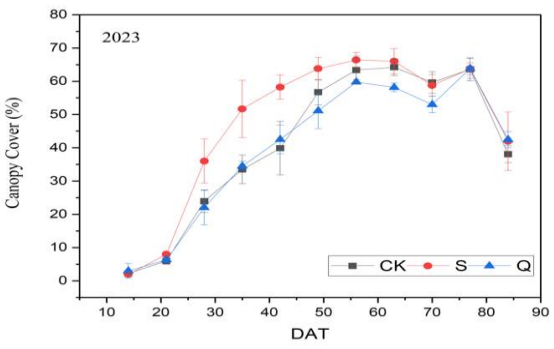


Fig. 3. Time series growth of canopy cover of rice grown under different irrigation methods: Shallow-Deep-Shallow irrigation (S), Shallow and Frequent irrigation (Q), and Conventional irrigation (CK). DAT refers to Days After Transplanting.

3.3 Grain yield

Rice yield exhibited variation across water-saving treatments. However, statistical analysis indicated no significant differences between the water-saving treatments and the Local empirical treatment (CK) (Fig.4). The CK treatment achieved the highest mean yield (3.8 tons/ha), followed by treatment Q (3.4 tons/ha) and treatment S (3.3 tons/ha). Similarly, [3], [14] found yield was statistically the same under the water-saving method compared to the continuously flooded method, across all tested levels of nitrogen for both hybrid and inbred rice varieties.

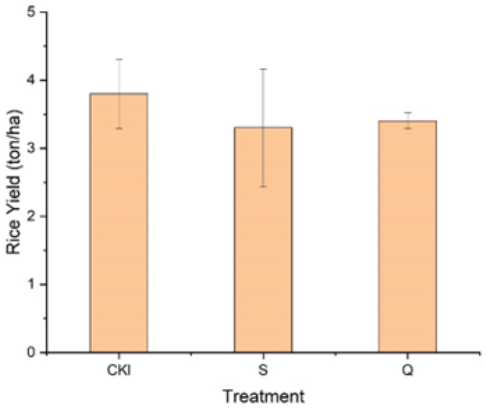


Fig. 4. Rice yield under different irrigation methods: Shallow-Deep-Shallow irrigation (S), Shallow and Frequent irrigation (Q), and Conventional irrigation (CK). DAT refers to Days After Transplanting.

3.4 Total irrigation, water productivity, and water-saving

The results showed that the total amount of water applied using the CK method was 7,898 m³/ha, while the S method used 5,327 m³/ha, and the Q method used 3,660 m³/ha (Fig. 5). Water input was significantly lower for the S and Q methods compared to the CK method during the growth period. Water productivity (WP) was calculated as the total grain yield per unit of total water utilized. The WP values for the irrigation methods (CK),(S), and (Q) were determined to be 0.49, 0.61, and 0.89 kg/m³, respectively (Fig. 5). This indicates that water-saving irrigation management in paddy fields can enhance rice water productivity by up to twofold (as observed in Q) compared to the CK method. The increase in water productivity for the S and Q treatments can be attributed to a significant reduction in total water input. These results are consistent with findings by [15], which indicates that water-saving irrigation methods can enhance water productivity at the farm level due to reduced water input.

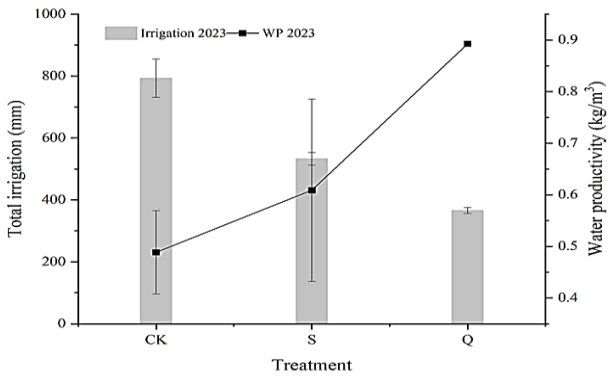


Fig. 5. Total irrigation and water productivity

Water input was significantly lower under the S and Q methods compared to the CK method during the growth period. The S method reduced water use by 32.5%, while the Q method achieved a 53.6% reduction compared to CK.

4. CONCLUSION

In this study, two water-saving irrigation methods (S and Q) were evaluated for their impacts on rice growth, yield, and water productivity. Although rice yield and canopy cover showed no significant differences compared to the conventional irrigation (CK), both water-saving methods substantially reduced water input, with a 32.5% reduction for the S method and a 53.6% reduction for the Q method. These reductions led to higher water productivity, increasing from 0.49 kg/m³ in CK to 0.61 kg/m³ in S and 0.89 kg/m³ in Q. Our findings demonstrate that implementing water-saving irrigation, particularly the Q method is potential for rice cultivation on sandy loam soils by significantly enhancing water use efficiency without negatively affecting rice yield.

ACKNOWLEDGMENTS

We acknowledge the Gansu Academy for Water Conservancy for financial support (21JR7RA766) and technical assistance and the Cambodian Agricultural Research and Development Institute and Faculty of Hydrology and Water Resources and Sub-project HEIP-ITC SGA#20 “Integrated approach of precise irrigation and sustainable soil management to improve crop water productivity in Cambodia through ITC soil laboratory development: the focus on rice farming (PRE-FARM Project)” for providing experimental material to conduct this research.

REFERENCES

[1] N. Monin, The Impacts of Climate Change on Agriculture and Water Resources in Cambodia: From Local Communities’ Perspectives, no. 125. 2021.

[2] D. Sum, “Role of the National and Sub-national Government Organizations for Climate Change Adaptation Plans : Coastal Areas of Cambodia Thesis for the Degree of Master of Arts in Environment and Sustainable Development The Role of the National and Sub-national Gov,” no. August, 2020, doi: 10.13140/RG.2.2.11303.57763.

[3] C. A. Phoeurn, C. Orn, T. Tho, C. Oeurng, A. Degré, and P. Ket, “Assessing the feasibility of alternate wetting and drying (AWD) technique for improving water use efficiency in dry-season rice production,” Paddy Water Environ., no. Maff 2019, 2024, doi: 10.1007/s10333-024-01012-5.

[4] J. Qin, F. Hu, B. Zhang, Z. Wei, and H. Li, “Role of straw mulching in non-continuously flooded rice cultivation,” Agric. Water Manag., vol. 83, no. 3, pp. 252–260, 2006, doi: 10.1016/j.agwat.2006.01.001.

[5] C. Zheng, Z. Zhang, S. Hao, W. Chen, Y. Pan, and Z. Wang, “Agronomic growth performance of super rice under water-saving irrigation methods with different water-controlled thresholds in different growth stages,” Agronomy, vol. 10, no. 2, 2020, doi: 10.3390/agronomy10020239.

[6] Y. Zhuang et al., “Effects and potential of water-saving irrigation for rice production in China,” Agric. Water Manag., vol. 217, no. March, pp. 374–382, 2019, doi: 10.1016/j.agwat.2019.03.010.

[7] A. S. Brar et al., “Water productivity, energy and economic analysis of transplanting methods with different irrigation regimes in Basmati rice (Oryza sativa L.) under north-western India,” Agric. Water Manag., vol. 158, pp. 189–195, 2015, doi: 10.1016/j.agwat.2015.04.018.

[8] Y. Xu et al., “Effects of water-saving irrigation practices and drought resistant rice variety on greenhouse gas emissions from a no-till paddy in the central lowlands of China,” Sci. Total Environ., vol. 505, pp. 1043–1052, 2015, doi: 10.1016/j.scitotenv.2014.10.073.

[9] S. Z. Peng, S. H. Yang, J. Z. Xu, Y. F. Luo, and H. J. Hou, “Nitrogen and phosphorus leaching losses from paddy fields with different water and nitrogen managements,” Paddy Water Environ., vol. 9, no. 3, pp. 333–342, 2011, doi: 10.1007/s10333-010-0246-y.

[10] T. P. Tuong and B. A. M. Bouman, “Field water mangement to save water and increase its productivity in irrigated lowland rice,” Agric. Water Manag., vol. 1615, pp. 1–20, 2000.

[11] N. Sandhu et al., “Root traits enhancing rice grain yield under alternate wetting and drying condition,” Front. Plant Sci., vol. 8, no. October, pp. 1–13, 2017, doi: 10.3389/fpls.2017.01879.

[12] J. M. Moreno-Maroto and J. Alonso-Azcárate, “Evaluation of the USDA soil texture triangle through Atterberg limits and an alternative classification system,” Appl. Clay Sci., vol. 229, no. January, 2022, doi: 10.1016/j.clay.2022.106689.

[13] M. Pourgholam-Amiji, A. Liaghat, M. Khoshravesh, and H. M. Azamathulla, “Improving rice water productivity using alternative irrigation (case study: north of Iran),” Water Sci. Technol. Water Supply, vol. 21, no. 3, pp. 1216–1227, 2021, doi: 10.2166/ws.2020.371.

[14] P. Belder et al., “Effect of water-saving irrigation on rice yield and water use in typical lowland conditions in Asia,” Agric. Water Manag., vol. 65, no. 3, pp. 193–210, 2004, doi: 10.1016/j.agwat.2003.09.002.

[15] R. Joshi, S. C. Mani, A. Shukla, and R. C. Pant, “Aerobic rice: Water use sustainability,” Oryza, vol. 46, no. 1, pp. 1–5, 2009, [Online]. Available: https://www.researchgate.net/publication/215894471.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

The dependency of the planting date on rainfall in paddy fields in the Pursat River Basin

Kimhor Sieng [1], Veasna Ly [3], Sithov Phat[1], Huot Menghour[2], Ty Sok[1,2*]

[1] *Faculty of Hydrology and Water Resources Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[3] *Department of Hydrology and Water Resources, Hohai University, 20th Dormitory of Hohai University JiangNing Campus, Nanjing City, Jiangsu Province, China*

[*] *Corresponding author: sokty@itc.edu.kh*

Abstract: *Rice cultivation is a cornerstone of Cambodia's economy and culture, with rainfed lowland rice dominating agricultural systems. However, its reliance on rainfall makes it highly vulnerable to climate variability. In the Pursat River Basin, a key rice-producing area of Cambodia, planting schedules and yields are deeply affected by temporal changes in rainfall. This study examines the dependency of paddy rice planting based on rainfall dates and proposes the best planting dates in the short future for rains in the intensified agricultural area in the Pursat River Basin. We examined an algorithm to estimate the dependency of the planting date on rainfall in paddy fields using remotely sensed data in the Pursat River Basin, where rain-fed paddies dominate, and identified the paddy rice using the Enhanced Vegetation Index (EVI) through the MODIS satellite data to assess crop growth stages. The study selected a 16-day composite Enhanced Vegetation Index (EVI) product by Moderate Resolution Imaging Spectroradiometer (MODIS), which provides global coverage with 250-m spatial resolution. The satellite data from 01 January 2010 to 31 December 2023 was utilized to determine rice cultivation periods. The planting dates were then estimated from the local minimum of the EVI time series; the lowest EVI value is the day farmers started planting rice between late April and early May. The planting dates in rain-fed paddies started from April to August, with large spatial variations and year-to-year fluctuations. Over the Zone 1(Kandieng), Zone 2 (Pursat), and Zone 3 (Kravanh-Bak Trakoun) of the Pursat River Basin, the zones planted in April and May had positive EVI profiles with rainfall, suggesting that wetter-than-average conditions encouraged farmers to plant earlier in the season. Our result can be a reference for further research to clarify farmers' practical actions regarding their planting date decisions and is important for the agronomic adaptation of planting dates under climate change.*

Keywords: Enhanced Vegetation Index (EVI), Moderate Resolution Imaging Spectroradiometer (MODIS), Pursat River Basin, Planting date, Rainfall

1. INTRODUCTION

There are some 135 million ha of paddy lands in the world, of which 126 million ha, or 93%, are in monsoon Asia, or humid East, Southeast, and South Asia (IRR, 2002). Climate and landforms are the two determinations for this strongly biased distribution of paddy lands. The monsoon climate is characterized by a yearly inversion of wind direction, which accompanies alternating rainy and dry seasons. Therefore, rainfall is concentrated during the rainy season, often bringing more than 1000mm of rain in less than half a year, thus enabling the cultivation of rice crops (Tsujimoto *et al.*, 2021).

Cambodia is part of mainland Southeast Asia and is located in the southern portion of the Indochina Peninsula. It is bordered by Laos, Thailand, and Vietnam and has a coastal area facing the Gulf of Thailand. Its population of 16 million, 77% lived in rural areas as of 2018. The paddy area in 2017 was reported to be 29,510 km^2 (JICA, 2019). The paddy area in 2017 was reported to be 29,510 km^2, occupying 53% of the agricultural land (JICA, 2019). Under the tropical monsoon climate, rice can be cultivated year-round, but approximately

[1*] Corresponding author: Ty Sok
E-mail: sokty@itc.edu.kh; Tel: +855-11 980 698

81% of rice production results from wet-season cultivation(Tsujimoto *et al.*, 2021).

Rain has benefits for agriculture, and Cambodia is one of the nations that is most vulnerable to the effects of climate change. In recent years, the country has experienced an increase in temperature, longer droughts, and more frequent tropical storms(Chhinh and Millington, 2015). Several areas of the country have experienced the consequences of extreme climate conditions, including severe droughts and floods(Talberth and Reytar, 2014). Individuals who depend on agricultural output for their livelihoods and reside in remote locations are especially sensitive to these effects because a decrease in agricultural productivity can increase the risk of child malnutrition while putting millions of people in danger of food insecurity (Talberth and Reytar, 2014). Future climate change might put at least 61% of the country's rural GDP per capita at risk, according to a report (Talberth and Reytar, 2014). Agricultural land occupied approximately 55,660 km2 in Cambodia and accounted for 31% of the entire area in 2018 (FAOSTAT, 2019). The paddy area in 2017 was reported to be 29,510 km^2 (JICA, 2019), occupying 53% of the agricultural land. Under the tropical monsoon climate, rice can be cultivated year-round, but approximately 81% of rice production results from wet-season cultivation. (Tsujimoto *et al.*, 2021)

Rice farmers in Pursat province, face significant challenges in optimizing rice production due to the strong dependence of their rainfed paddy fields on fluctuating rainfall patterns and the timing of rice planting is crucial for maximizing yields, but unpredictable rainfall onset, distribution, and intensity of rainfall, make it difficult for farmers to determine the optimal planting date (Tsujimoto *et al.*, 2021). This uncertainty is exacerbated by potential climate change impacts, which are projected to alter rainfall patterns further. In the current study, we found that local farmers in Cambodia's rainfed paddy fields always modified their agricultural management in response to differences in annual rainfall.(Tsujimoto *et al.*, 2021). In this study, the research examined the existing dependency of the planting date on rainfall amount in paddy fields in the Pursat River Basin, where rain-fed paddies dominate, using remotely sensed data. This necessitates a deeper understanding of the complex relationship between rainfall and optimal planting dates to develop effective strategies that enhance rice productivity and resilience in the face of environmental variability and potential climate change.

The overall goal of this study is to examine the dependency of paddy rice planting on rainfall dates and to propose optimal planting dates for rainfall in the intensified agricultural area of the Pursat River Basin. Furthermore, paddy rice will be identified using the Enhanced Vegetation Index (EVI) derived from MODIS satellite data to assess crop growth stages.

2. METHODOLOGY

2.1 Study Area

The Pursat River Basin is located in the Pursat province in the southern part of the Tonle Sap Great Lake Basin and drains an area of 5,964 km^2 (NCSD, 2019). The Stung Pursat River originates in the drier eastern slopes of Cardamom Mountain and flows for approximately 150 km, ultimately draining into the Tonle Sap Great Lake. Two main tributaries, the Stung Peam and Stung Santre (Prey Khlong) rivers, flow in a northerly direction and meet the Pursat River just above Kravanh-Bak Trakuon. The drainage area at Kravanh-Bak Trakuon, just below the confluence of the Pursat and the two tributaries, is 4,245 km^2, and 4,596 km^2 farther downstream at the Khum Veal gauging station, located near the town of Pursat. The Stung Pursat river basin is shared by six districts: Veal Veng, Kravanh, Pursat, Krakor, Bakan, and Kandieng. People living close to the lake earn their livelihood from fishing, while those further away depend on rice cultivation (NCSD, 2019).

Rainfall around the Pursat River basin area increases with elevation. There are three Zone 1(Kandieng), Zone 2(Pursat), and Zone 3(Kravanh-Bak Trakoun) at the Pursat River Basin are chosen for this study. All three zones are part of the intensified agricultural area but face different water availability challenges and opportunities. The zones represent a clear gradient of rainfall and elevation, from the high-rainfall, high-elevation upstream (Kravanh-Bak Trakoun) to the moderate mid-basin (Pursat) and the lower-rainfall, flood-influenced downstream (Kandieng). This allows for examination of how rainfall dependency changes across the basin. The annual average rainfall spatially distributed over the study area ranges from 1200 mm to 1700 mm, but the annual amount varies considerably from year to year (NCSD, 2019). It has a predominantly tropical monsoon climate, characterized by distinct wet and dry seasons, with most rainfall occurring between May and October. Temperatures remain relatively warm throughout the year, with average highs ranging from 30°C to 35°C. Humidity is generally high throughout the year, but it significantly increases during the rainy season, and during the rainy season (May to October), humidity can reach up to 90% (NCSD, 2019). The hot season also has high humidity levels, which can make the high temperatures feel even more intense. Pursat's climate is characterized by hot and humid conditions, especially during the rainy season. The dry season provides relatively hotter and less humid weather (NCSD, 2019).

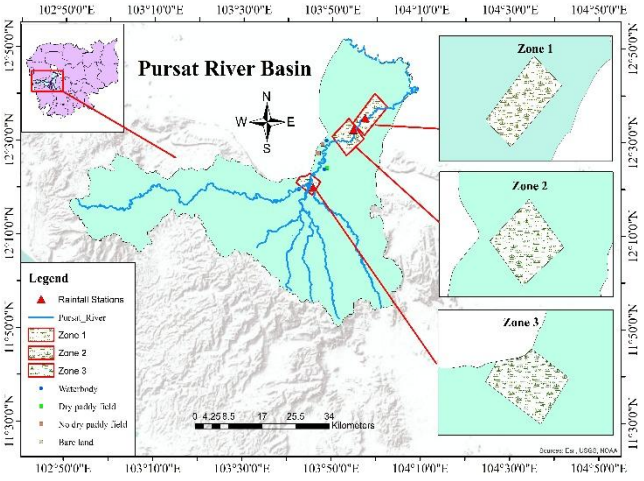


Figure 2.1 Map of the paddy field at the Pursat River Basin

2.2 The Framework of the Study

The general framework of this study is shown in Figure 2.2. The required input data are observed rainfall from 2010-2023 in three zones at Pursat River Basin, The Vegetation Indices (VI) 16-Day L3 Global 250 m SIN Grid V006 from a MODIS instrument onboard the Terra Satellite (MOD13Q1 product) and that from a MODIS onboard the Aqua Satellite (MYD13Q1 product) are obtained from the Web site of the United States Geological Survey (Center) for this study. These products provide EVI values at 250 m pixels over the entire globe as 16-day composites. After the calculation, EVI data is provided by ArcGIS. The annual rainfall and monthly rainfall an analyzed to determine high and low precipitation. The output result and the relationship between monthly rainfall and the estimated planting date.

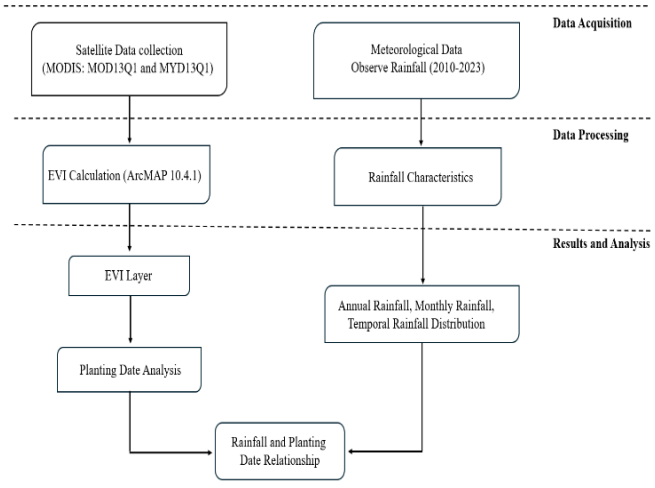


Figure2.2 Framework of this study

2.3 Description of MODIS's EVI Datasets

As mentioned in section 2.2 above these products provide EVI values at 250-m pixels over the entire globe as 16-day composites. Since the 16-day period for the MYD13Q1 has an 8-day lag to the MOD13Q1, data from MOD13Q1 and MYD13Q1 are treated all together and regarded as 8-day interval data.

$$EVI = \left\{ G \times \left(\rho'_{NIR} - \rho'_{red} \right) \Big/ \rho'_{NIR} + C_1 \times \rho'_{red} - C_2 \times \rho'_{blue} + L \right\} \quad \textbf{(Eq. 6)}$$

Where ρ'_{NIR} and ρ'_{red} are the surface bidirectional reflectance factor for near-infrared (NIR) and red MODIS bands, respectively, and ρ'_{NIR}, ρ'_{red}, and ρ'_{blue} are atmospherically corrected surface reflectance factors for NIR, red, and blue MODIS bands, respectively. L (=1) is the canopy background adjustment that addresses nonlinear, differential NIR, and red radiant transfer through a canopy. C_1 (=6) and C_2 (=7.5) are the correction coefficients for the aerosol influences. G (=2.5) is the gain factor.

2.4 Algorithm to Estimate Planting Date in Paddy Fields

Planting is the time when the time-series EVI is at its lowest level (Sakamoto *et al.*, 2005). This shows that the amount of vegetation doesn't drastically decrease during planting and gradually rises after planting. Paddy fields are often fertilized and plowed before being planted. Following this heading, the rice plant likewise changes from the vegetative development stage to the reproductive growth stage, which results in a decrease in vegetation due to the withering and loss of leaves.

3. RESULTS AND DISCUSSION

3.1 Rainfall Characteristics

3.1.1 Monthly Rainfall

The dry season extends from October to March, and the wet season is from May to September Figure 3.1. The graph illustrates the mean rainfall in a month for a period of a year in the Pursat River Basin. There is a result of the average rainfall show fluctuation throughout the 12-month period in 2010-2023 in each of the three zones. The most noticeable feature of the graph is that the rainfall reaches the first peak in October and the second in September. Fig 3.1 shows the rainfall in Pursat River Basin in 2010-2023 in three zones. October has the heaviest rainfall on Fig3.1 Zone 1 available in 487.53mm, Zone 2 in 492.90mm and zone 3 in 517.90mm respectively in 2010, 2016 and 2020.

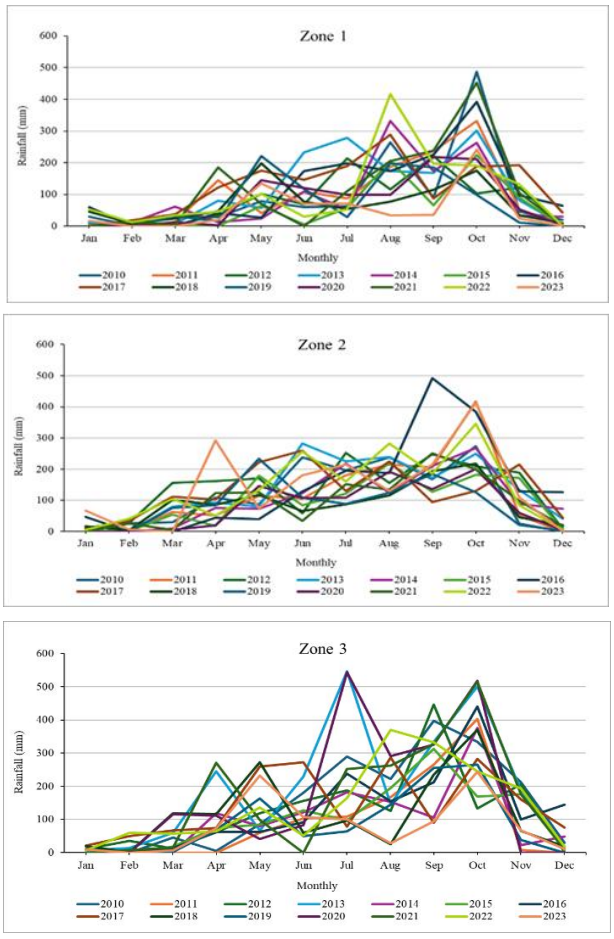


Figure 3.1 Monthly Rainfall in each of the three Zones

The bar chart in Fig 3.2 illustrates the compare mean rainfall in a month for a period of a year in the three zones. Overall, there is a result of the average rainfall show fluctuation throughout the 12-month period. Overall, there is a result of the average rainfall show fluctuation throughout the 12-month period.

First of all, the most noticeable feature of the bar graph is that the rainfall reaches the first peak in October and the second in September (just above 343 mm and well over 267 mm, respectively) in Kravanh-Bak Trakoun (Zone 3). The next striking feature of this graph is that rainfall has its lowest point in April and May which be approximately 96 mm for the wet season. To compare the rainfall over the time frame, it is apparent that rainfall in April is lower than that in October (around 96 mm and 343mm, respectively). Turning to the bar graph in Kandieng (Zone 1), we can see that the peak of mean rainfall of the year is in October, which is the overall highest peak, and in September, with the figure 3.2 around 261 mm and 198 mm correspondingly. The lowest rainfall of the year (2010-2023) falls in April with just under 56 mm in the rainy season. Not forgetting to mention that there is less rainfall in the first six months of the year than in the months after July.

The last study area is Pursat is similar to Kandieng (Zone 1) and Kravanh-Bak Trakoun(Zone 3) in that it has higher rainfall in October, about 258mm, than it does in September, which is similar to the two locations, with 258mm for the month. If we consider Cambodia's rainy season, the lowest rainfall occurs in April at just 94mm. All provinces have the same peaks and lows.

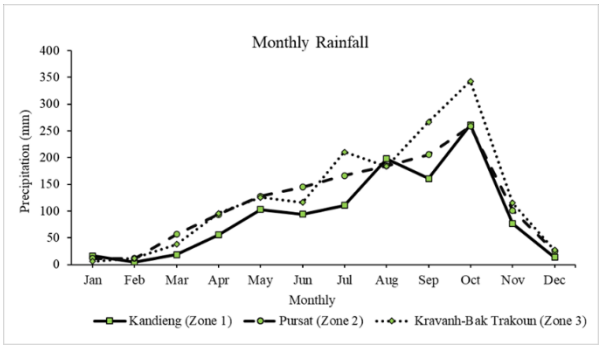


Figure 3.2 Average monthly rainfall of three zones

3.1.2 Annual Rainfall

Rainfall variability in each time scale is an applicable feature of the meteorological cycle and climate in Zone 1, Zone 2, and Zone 3. Figure 3.3 shows the average annual rainfall from 3 stations from 2010 to 2023. The chart illustrates the highest rainfall in Zone 3, and then Zone 2 also has a high rainfall after Zone 3. The highest average rainfall of the three zones, Zone 1, Zone 2, and Zone 3, was 1486.69mm (2017), 1773mm (2016), and 2348.39mm (2013), respectively. Additionally, the lowest average rainfall was (Zone 1) in 2023, 652.63 mm, (Zone 2) in 2020, 978mm, and (Zone 3) in 2023, 971.8mm.

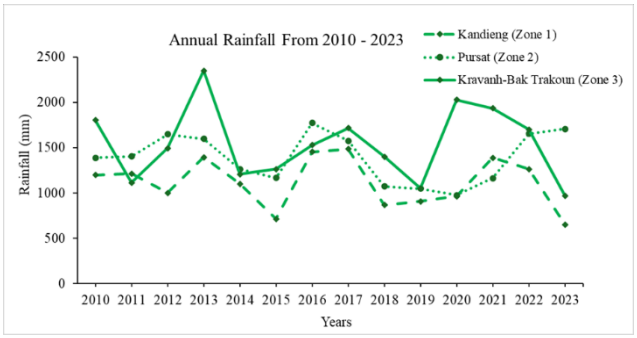


Figure 3.3 Annual rainfall at the three Zones in the Pursat River Basin between 2010-2023

The results indicate irregular variations in annual rainfall across these three study areas, posing a significant challenge for farmers in accurately determining optimal rice planting

dates. Consequently, these findings are instrumental in identifying future rice planting dates and forecasting water availability.

3.1.3 Monthly planted grid

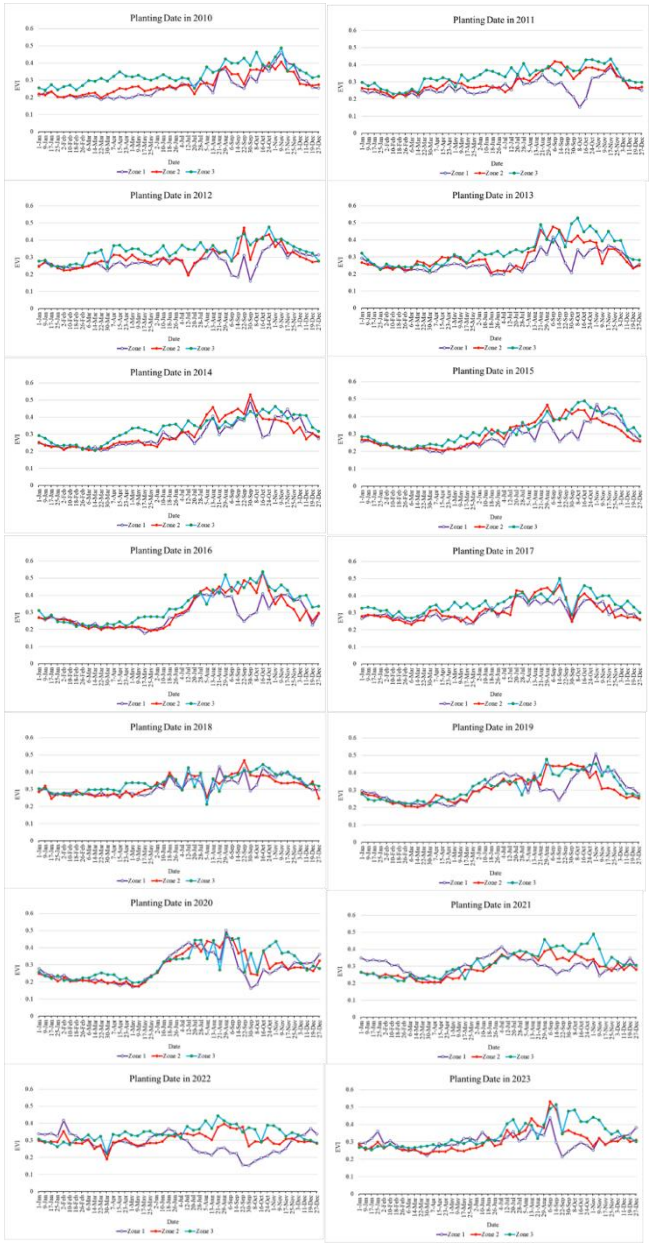
According to the graph in Fig. 3.4 that calculated from EVI values, shows that in Zone 1, the highest amount of rice was planted in June, while in Zone 2 the highest number of rices was planted in April, and in Zone 3 it was planted in April. Most in April compared to 2010-2023.



Figure 3**Error! Use the Home tab to apply 0 to the text that you want to appear here.**4 Monthly planted grids in three zones 2010-2023

3.1.4 Planting date and EVI over the study Area Through Satellite Image

In each graph used to illustrate smooth EVI profiles, it shows the temporal pattern characteristics of rice crops throughout the 2010 to 2023 growing season Figure 3.5. In general, the cropping calendar of rice crops in the study region varied from the rainy season to the end of the event. This seasonal information helped determine crop growth conditions and confirmed the value of utilizing MODIS vegetation indices to follow rice crop phenological occurrences. In these zones, most farmers start planting rice in early April.



Based on the EVI values for the 14 years shown in the graph, we can estimate the relative planting periods for the three zones.

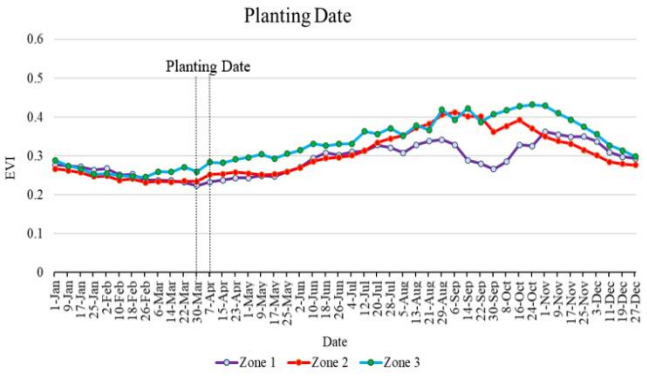
In Zone 1, the lowest EVI price of 0.23 on 30 March is a price that indicates the day when farmers in Zone 1 prepare to start farming, and the people in the area start plowing. About a month later, on July 20, the price of EVI rose as the rice crop in the area grew strongly due to the heavy rains, which allowed people to grow water-dependent rice. All EVIs at all events peaked at 0.34 on July 20, the day between the start of farmers' harvest days, while from this on EVI prices trend decreased. In Zone 2 is Similar to Zone 1; Zone 2's EVI prices begin to increase noticeably around late March by 0.23 and

Figure 3.6 Planting Date in three zones

Early April by 0.25. This indicates that the planting activities in Zone 2 probably took place around the same time as in Zone 1. Rice was planted on April 7, as we observed that from today onwards, EVI changes because crops in that area began to grow after sowing. From the day of planting, the increase in EVI indicates that the rice in the lowland area is growing in the field. 0.41 is the highest EVI compared to other EVI values on September 06. In terms of harvest, it is a bit different from Zone 1, as it is reduced in EVI between October and December.

The last Zone shows the earliest increase among the three Zones, starting in early April. This suggests that planting in Zone 3 will likely begin earlier in the season compared to Zone 1 and Zone 2.

The results in Fig.3.6 indicate the Planting Date between EVI values and the time series of three zones in the Pursat River Basin. The three provinces are slightly different, but it can be summarized that people can farm from the end of the rainy season and the beginning of April because there may be rain, and although April is the dry season, in some parts of the



Pursat River Basin, it also rains.

The Analyzing EVI time series in conjunction with rainfall data provides valuable insights into planting date, the success of crop establishment, and the overall impact of water availability on vegetation growth in agricultural regions.

4. CONCLUSIONS

In this study, we examined the dependency oof planting date on rainfall in paddy fields to quantify farmers' empirical knowledge regarding their decision regarding the planting date, especially in rain-fed paddies. The planting dates in rain-fed paddies ranged widely from April to September in 2010-2023, and the planting date was most frequently recorded from June to August during the full-monsoon season. The planting areas in April were found to have positive EVI values, with rainfall suggesting that wetter-than-average conditions prompted farmers to plant larger areas in that season.

ACKNOWLEDGMENTS

The author would like to acknowledge the HydroMet and Disaster Management Lab at ITC for guidance and funding support.

REFERENC

Center, N. L. D. a. t. U. E. United States Geological Survey (USGS)

Chhinh, N., Millington, A. J. C. 2015. Drought monitoring for rice production in Cambodia. *3*(4), 792-811.

FAOSTAT, F. (2019). Food and Agriculture Organization of the United Nations-Statistic Division https://www. fao. org/faost at/en/# data. In: QC.

IRR. 2002. RRI (International Rice Research Institute). 2002. World rice statis- tics, 2002. Los Bañes (Philippines): IRRI.

JICA. 2019. Japan International Cooperation Agency (JICA)

Figure 3**Error! Use the Home tab to apply 0 to the text that you want to appear here.**5 Planting date in three zones 2010-2023

(2019) Irrigation development in Cambodia, status as of March 2019.

NCSD. 2019. Pursat River Basin Profile.

Sakamoto, T., Yokozawa, M., Toritani, H., Shibayama, M., Ishitsuka, N., Ohno, H. J. R. s. o. e. 2005. A crop phenology detection method using time-series MODIS data. *96*(3-4), 366-374.

Talberth, J., Reytar, K. J. N. Y. T. U. S. A. f. I. D. 2014. Climate change in the Lower Mekong Basin: an analysis of economic values at risk.

Tsujimoto, K., Ono, K., Ohta, T., Chea, K., Muth, E.-N., Hor, S., . . . Environment, W. 2021. Multiyear analysis of the dependency of the planting date on rainfall and soil moisture in paddy fields in Cambodia, 2003–2019. *19*(4), 635-648.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Hydraulic Propagation of Pesticides at the Watershed Scale

Ech Chin[1], Khy Eam Eang[2*], Sombath KEO[3], Paul Baudron[4]

Graduate School, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia
[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[3] *Cambodia Development Resource Institute, P.O. Box 622, Phnom Penh, Cambodia*
[4] *UMR G-EAU, French National Research Institute for Sustainable Development, Montpellier, France*

[*]*Corresponding author: khyeam@itc.edu.kh*

Abstract: *Pesticides are chemical compounds used in agriculture to control pests and boost crop productivity particularly for rice. However, their toxicity poses risks to both the environment and human health, as they can leach, evaporate, and contaminate water sources, spreading to other areas. The transport mechanisms of soluble pesticides is crucial because their ability to readily dissolve in water significantly influences their movement through aquatic systems and their potential to impact a wider range of environmental compartments and human populations. This study investigates how soluble pesticides are hydraulically transported from agricultural areas (mostly rice) through the Bassac River (BSR), the Tonle Sap river (TSR) and in the Mekong River (MKR). Existing pesticide data from previous publications (from MKR and BSR) were reinterpreted within the scope of hydraulic propagation (specifically highly-soluble compounds and focusing on the evolution along water flow). New water samples for pesticide analysis were also collected at 16 points (MKR, BSR and TSR), 3 meters from the riverbanks and 50 cm depth, in March 2025 (dry season) and they are under analysis in the laboratory, while physico-chemical parameters are already available. The results provide a series of highlights. Firstly, regarding previous data, while highly-soluble pesticide levels were specifically expected to be higher during the rainy season, with runoff from farms being the main way pesticides enter the watershed, some locations and compounds showed higher values in the dry season, potentially due to their accumulation in low-flow transects or transport through groundwater. Secondly, preliminary data from the new campaign evidence the hydraulic connection patterns between TSR, MKR, BSR: BSR is only hydraulically connected to MKR and disconnected from TSR, thereby establishing a critical foundation for pesticide propagation evaluation. Thirdly, this data highlights the potentially high contribution of sewage water at some location of the BSR, thus having a potentially high effect on the origin and fate of measured pesticide compounds. The findings highlight the need for action to protect the river's ecosystem, and the need to assess hydraulic context before interpreting pesticide data in watersheds. This research will provide ideas for better water management in the river system and other similar areas.*

Keywords: Pesticides, Transportation, Tonle Sap River, Water Management

1. INTRODUCTION

Pesticides are defined as any chemical substances, mixtures of substances or natural agents that are used in agriculture to control or kill pests, insects, rodents, weeds, and diseases. While they offer advantages like boosting productivity, safeguarding harvests, preserving food, and preventing the spread of vector-borne illnesses[1], their widespread use can negatively impact crop yields (mainly by harming beneficial soil organisms), public health, and the environment Insecticides, fungicides, rodenticides, and herbicides are among the most commonly used pesticides globally [2]. Pesticides are categorized based on the pest they target, their toxicity level, and their chemical family or composition. The hazard of pesticide was classified into extremely hazardous (Ia), highly hazardous (Ib), moderately hazardous (II), slightly hazardous (III), and unlikely to present acute hazard (U)m [3].

Pesticide use in Cambodia began in the 1960s and has generally increased over time. Between 1980 and 1994, the percentage of farmers using pesticides ranged from 7% to

[1]* Corresponding author: Khy Eam Eang
E-mail: khyeam@itc.edu.kh; Tel: +855-123-311-110

49%, and this figure rose to 67% in 2019 [4]. By 2022, this had further increased to 73%, with the Ministry of Agriculture, Forestry and Fisheries (MAFF) estimating a total pesticide use of 160,000 tons in 2024, primarily sourced from Thailand and Vietnam, with smaller quantities originating from China (PRC) and the European Union [5]. These pesticides are categorized in four main types: insecticides, herbicides, fungicides, and rodenticides.

Post-application, pesticides can threaten the environment and human health through leaching, evaporation, and water source contamination. With thousands of varying pesticide compounds classified by water solubility [6], understanding their fate and transport is crucial. Hydraulic propagation of pesticides through a watershed can be described by focusing on water-soluble compounds [7],[8] as they have a high mobility. In this study, we present i) a reinterpretation of high-soluble pesticides data from a previous study in the MKR and BSR areas [9] and ii) preliminary data from a targeted sampling campaign conducted in March 2025, which successfully extended coverage beyond MKR and BSR areas to also include the emblematic TSR.

2. METHODOLOGY

2.1 Data reinterpretation

A screening of a series of pesticide compounds along the MKR and BSR was recently published by [9]. Discussion was centered on describing the overall occurrence of pesticides in soils and water. Although the original study did not focus neither on hydraulic transport nor on highly soluble compounds, this paper offers a new interpretation of the data within that specific context.

2.2 Sample collection

Water samples were collected along Mekong River and Bassac River during the dry season in March 2025 from sixteen different locations, all at 3 meters from the riverbanks and 50 cm depth. In-situ measurements such as electrical conductivity (EC), pH, and temperature were measured while water flow and water level can be calculated from the existing data of Mekong River Commission (MRC). One-liter polypropylene bottles were used for collection, and samples were taken in triplicate. On-site filtration was performed using 2 µm glass microfiber filters (Whatman, GF/CTM) to remove suspended solids and debris. The filtered samples were then transported to the SATREPS laboratory at the Institute of Technology of Cambodia under iced conditions and stored at 4 °C.

2.3 Chemicals and reagents

Sample preparation for GC-MS analysis involved the use of a 1 mol/L sodium phosphate buffer solution (pH 7.0), dichloromethane (ACS BASIC), acetone, n-hexane (HPLC grade), sodium sulfate, nitrogen gas, PLS3, and an AC cartridge.

2.4 Sample preparation

Solid-phase extraction (SPE) was performed on the completely filtered water samples (2 um fiberglass filter). The SPE process began with the activation of the PLS3 sorbent attached to an AC cartridge. This activation involved conditioning by sequentially passing 5 mL of dichloromethane, 5 mL of acetone, and two 5 mL aliquots of purified water through the cartridge. The conditioning step aimed to ensure the cartridge was fully wetted and clean. Subsequently, the pH of each liter of filtered water sample was adjusted to 7.0 by adding 1 mL of a 1 mol/L sodium phosphate buffer solution. The samples were then shaken to ensure homogeneity and allowed to settle for 1 hour. Following this, the buffered water samples were passed through the activated cartridges using a vacuum pump at an approximate flow rate of 15 mL/min. After loading the samples, the cartridges were dried under a nitrogen gas stream for 30 minutes to evaporate any residual water. Next, the cartridges were disassembled. The PLS3 cartridge was washed with 2 mL of acetone and 5 mL of dichloromethane. Additionally, the assembled PLS3 and AC cartridge set was washed by eluting with 5 mL of acetone, with the vacuum pump activated at the end of the elution. The combined eluted solvent was collected in a glass tube and concentrated under a nitrogen gas stream to a final volume of approximately 1 mL. To this concentrated solution, 10 mL of n-hexane was added, and the mixture was dehydrated by passing it through anhydrous sodium sulfate (Na_2SO_4) used as a filter. The resulting solution was then concentrated again under a nitrogen gas stream to a final volume of 1 mL and transferred to a vial for storage at -20 °C prior to GC-MS analysis.

2.5 Calibration for analysis

A 0.1 mL aliquot of internal standard (IS) solution was added to each sample prior to GC-MS analysis. Additionally, external standard solutions were prepared at six different concentrations (0.025 ppm, 0.05 ppm, 0.1 ppm, 0.125 ppm, 0.5 ppm, and 1 ppm) by dilution from a 10-ppm stock standard solution. These external standards will be used to quantify the concentrations of 11 target pesticide compounds, while other detected pesticide compounds will be subject to semi-quantitative analysis.

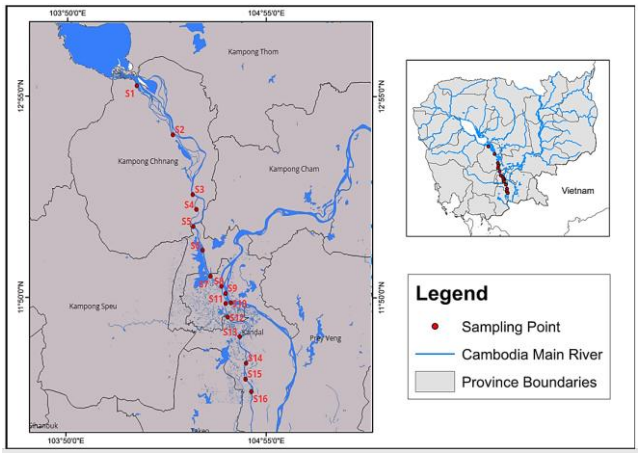


Fig. 1 Mapping of sampling sites at TSR, MKR and BSR.

2.6 GC-MS analysis

Pesticide compounds are being screened using a gas chromatography-mass spectrometry (GC-MS) system, model TQ8040 series (Shimadzu, Japan), utilizing an MS database containing 451 pesticide compounds for identification. A DB-5ms column (30 m length, 0.25 µm film thickness, and 0.25 mm internal diameter) is employed for separation. Samples were injected with a 1 µL volume into the GC-MS system using splitless injection mode. The column oven temperature is initially held at 40 °C for 2 minutes, then increased to 310 °C at a rate of 8 °C/min and held at the final temperature for 5 minutes. Ultra-pure helium is used as the carrier gas at a flow rate of 1.23 mL/min. The ion source temperature was maintained at 200 °C, and the interface temperature was set to 300 °C. Mass spectra are acquired in scan mode over a mass-to-charge ratio (m/z) range of 33-600. Identification of pesticide residues is performed using the system's data processing software, which calculates the monoisotopic mass, predicts the structural formula of compounds, and compares these against the MS database.

3. RESULTS AND DISCUSSION

The next 3 figures illustrate the variation of pesticides or EC along the two or three studied rivers.

3.1 physico-chemical parameters

.

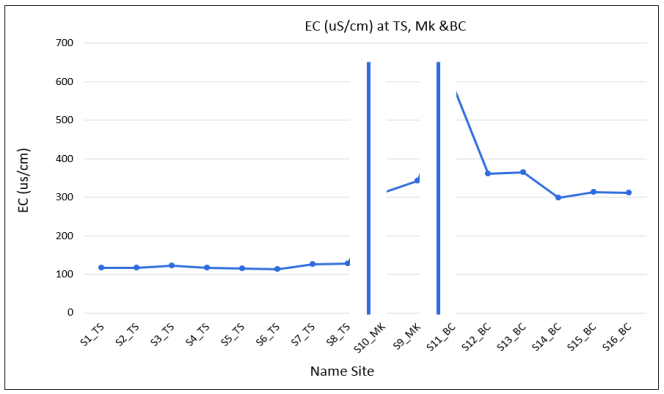


Fig. 2 Electrical conductivity EC (µS/cm).

There are two vertical blue lines on Figure.2. The first is located between sampling points (S8_TS) and (S10_MK), indicating a potential confluence or transition point between the Tonle Sap and Mekong Rivers. The second vertical line is located between sampling points (S10_MK) and (S11_BC), suggesting the confluence or transition point between the Mekong and Bassac Rivers. The electrical conductivity (EC) from sampling point S1_TS to S8_TS remains relatively low and stable, generally fluctuating between around 110 and 130 µS/cm. There is a significant and abrupt increase in electrical conductivity between sampling point S8_TS and S10_MK, where the EC rises from around 130 µS/cm to over 300 µS/cm. It reveals that TSR and MKR have very different EC, thus allowing for an efficient differentiation of waters in the watershed. From S12_BC to S16_BC, the electrical conductivity remains relatively stable, fluctuating between approximately 300 and 360 µS/cm, which is notably higher than the values observed in the initial Tonle Sap River samples.

Based on the Electric conductivity related to water flow from MKR, thus revealing that BSR receives water (and potentially pesticides only from MKR and not from TSR. So, that is a reason why pesticides were not transported from TSR to BSR. Similar to research results from [9] state that concentration from Mekong upstream (M1-M4) is lower than either Mekong downstream (M5-M8) and BSR(B1-B2). According to these arguments, the pesticides weren't transported through water flow; otherwise it moved from nearby agriculture activities. The high EC at sampling point S11_BC, reaching a peak of approximately 600 µS/cm, is potentially due to the contribution of sewage water at some location of the BSR, thus having a potentially high effect on the origin and fate of pesticide compounds.

3.2 Pesticide detected Around MKR and BSR

This section illustrates a re-interpretation from a screening of a series of pesticide compounds along the MKR and BSR was recently published by [9], state that there are 9 pesticide compounds consists of Atrazine-desethyl, Carbendazim Cyproconazole, DMST, Fluroxypyr, Iprobenfos, Isocarboxazid, Tebufenozide and Triadimefon. This research focuses on soluble pesticides and studies the spatial variation of these compounds along river flow and mixing.

Table.1 The soluble value of pesticides (g/L) [8].

No	Pesticides	Soluble in water (g/L)	Level dissolve
1	Tebufenozide	0.001	Very low
2	Fluroxypyr	0.2	Moderate soluble
3	Triadimefon	0.1	Low soluble
4	DMST	0.01	Low soluble
5	Cyproconazole	0.05	Low soluble
6	Carbendazim	0.03	Low soluble
7	Iprobenfos	0.01	Low soluble
8	Isocarbamid	0.05	Low soluble
9	Atrazin	0.033	Low soluble ,20C°

Results are separated between low-concentration (Figure 3A) and high concentration (Figure 3B) pesticides data.

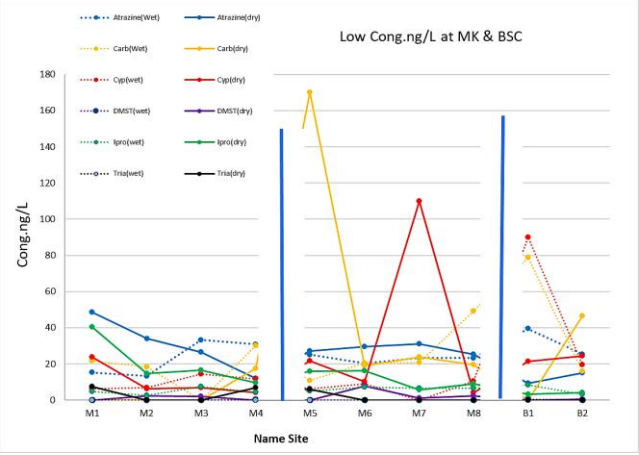


Fig. 3A Low concentration (ng/L) at MKR and BSR.

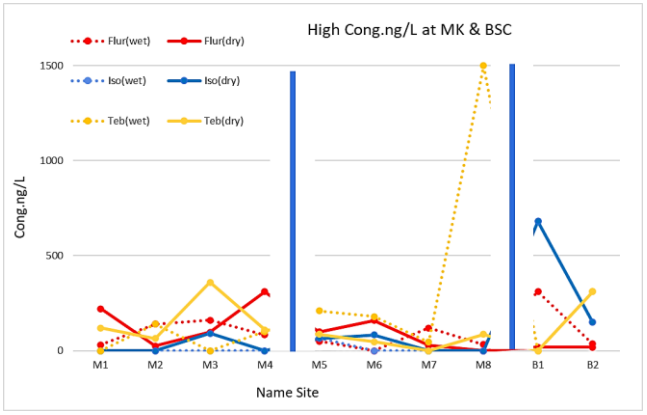


Fig. 3B High concentration (ng/L) at MKR and BSR.

Figure 3A,3B illustrates the concentrations of nine pesticides (Atrazine, Carbendazim (Carb), Cypermethrin (Cyp), Dimethyl Sulfotep (DMST), Iprobenfos (Ipro), Triadimefon (Tria),Tebuconazole(Teb), Isoprothiolane (Iso), Fluroxypyr (Flur) in the Mekong (M1-M8) and Bassac (B1-B2) river systems during both wet and dry seasons, measured in ng/L.

The spatial distribution reveals distinct patterns potentially influenced by hydraulic propagation and seasonal variations. In the dry season along the Mekong River (M1-M8), a general trend of decreasing low-concentration pesticides is observed from upstream (M1) to downstream (M4) from 0.0 ng/L to 110 ng/L respectively. However, after passing Phnom Penh (between M4 and M5), an increase in these concentrations of 360 ng/L is noted. This suggests that the city might be a source of pesticide input during the dry season when dilution from rainfall is minimal. Notably pesticide, Carbendazim (Carb) reaches its highest concentration at M5 (360 ng/L), and Cypermethrin (Cyp) peaks at M7 (210 ng/L) during this season, further supporting the idea of localized inputs or reduced dilution downstream. Fluroxypyr also shows moderate concentrations, particularly at M4 (100 ng/L) in the dry season.

In contrast, Isoprothiolane generally remains at low concentrations across the Mekong sampling points.

The Bassac River (B1-B2) in the dry season shows a substantial peak of Isocarboxazid at B2 (680 ng/L), indicating a potential source or accumulation in this part of the system during the low flow period.

During the wet season, the clear downstream decrease observed in the dry season for low-concentration pesticides in the Mekong is not apparent. This could be due to increased runoff from diffuse agricultural sources throughout the basin, leading to a more homogenous distribution or masking any point-source influences. Interestingly, the Mekong River exhibits higher peak concentrations for several pesticides in the wet season, including Carbendazim and Cypermethrin,

and a particularly dramatic peak of Tebuconazole at M8 (over 1400 ng/L). This suggests that the wet season flow might be mobilizing and transporting significant amounts of these pesticides to the lower reaches of the Mekong.

In the Bassac River during the wet season, Carbendazim and Fluroxypyr show higher levels compared to the dry season, potentially due to increased runoff carrying these compounds into the river. Isoprothiolane remains generally low in the Bassac as well during the wet season.

Overall, Tebuconazole stands out with exceptionally high concentrations at specific locations and seasons, suggesting a significant source or usage pattern impacting both river systems. The contrasting seasonal patterns for Fluroxypyr (higher in the dry season at M4 but higher in the wet season at B1) and Tebuconazole (wet season peak at M8, dry season peak at B2) highlight the complex interplay of application timing, runoff, dilution, and potential point sources influencing pesticide transport and concentration along these river systems. The lack of a consistent seasonal variation across all pesticides and locations underscores the need for further investigation into the specific sources and transport pathways of each compound.

4. CONCLUSIONS

In conclusion, both the Mekong and Bassac rivers are significantly affected by pesticide contamination, with the Mekong River potentially experiencing higher maximum concentrations for certain compounds. The distinct seasonal patterns observed highlight the need for continuous monitoring to understand the dynamics of pesticide pollution in these vital aquatic ecosystems. The findings underscore the importance of implementing effective water management practices to mitigate the risks posed by pesticide runoff and protect both human health and the environment. Further research is essential to develop strategies that address the challenges of pesticide transport and its impact on water quality in the region.

ACKNOWLEDGMENTS

This research was funded by the projects FEF-R Inversap and FEF-R Healthy Rice, two initiatives supported by the French Embassy in Cambodia, and funded by the French Ministry for Europe and Foreign Affairs.

REFERENCES

[1] H. Kaur, R. Kaur, S. Singh, N. Jagota, and A. Sharma, “Pesticide biology in plants: plant uptake, translocation, and accumulation,” in *Pesticides in the Environment*, Elsevier, 2024, pp. 67–86.

[2] K.-H. Kim, E. Kabir, and S. A. Jahan, “Exposure to pesticides and the associated human health effects,” *Sci. Total Environ.*, vol. 575, pp. 525–535, 2017.

[3] I. C. Yadav and N. L. Devi, “Pesticides classification and its impact on human and environment,” *Environ. Sci. Eng.*, vol. 6, no. 7, pp. 140–158, 2017.

[4] V. Preap and K. Sareth, “Current use of pesticides in the agricultural products of Cambodia,” in *Submitted for the FFTC-KU International Workshop on Risk Management on Agrochemicals through Novel Technologies for Food Safety in Asia, November*, 2015, pp. 10–14.

[5] U. Dhar, “Asian Development Bank (ADB), Asian Development Outlook 2020: What Drives Innovation in Asia?,” 2021, *SAGE Publications Sage India: New Delhi, India*.

[6] M. Kah, S. Beulke, and C. D. Brown, “Factors influencing degradation of pesticides in soil,” *J. Agric. Food Chem.*, vol. 55, no. 11, pp. 4487–4492, 2007.

[7] G. Pérez-Lucas, N. Vela, A. El Aatik, and S. Navarro, “Environmental risk of groundwater pollution by pesticide leaching through the soil profile,” *Pestic. misuse their impact Environ.*, vol. 17, pp. 1–28, 2019.

[8] A. Belles *et al.*, “Transfer pathways and fluxes of water-soluble pesticides in various compartments of the agricultural catchment of the canche river (Northern France),” *Water*, vol. 11, no. 7, p. 1428, 2019.

[9] P. Ngin, “Pesticides in Cambodia: usage, fate, and health risk,” 2024, *Umeå University*.

POSTER SESSION

Energy Technology Management (ETM)

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Economic analysis of solar rooftop at Toyota Cambodia Office in Phnom Penh by using PVsyst software: EAC OLD vs New Regulation

Prom Piseth [1*], Vannak Vai [1], Kimsrornn Khon [1,2]

[1] *Department of Electrical and Energy Engineering, Faculty of Electrical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[2] *Energy Technology and Management Unit, Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh 120404, Cambodia*

[*] *Corresponding author: piseth_prom14@yahoo.com*

Abstract: *Rooftop photovoltaic (PV) systems have become increasingly popular due to their technological, economic, societal, and environmental benefits. Business, residential, and industrial owners are eager to invest in solar energy as a way to reduce their electricity bill. The main objective of this paper is to propose 6 scenarios case studies (a grid-connected solar system with a limited peak load demand and available roof space) for Toyota Cambodia's commercial office. The study will compare six different PV system scenarios, with and without storage systems, using PVsyst software to evaluate technical and economical KPI. The objective of study propose 6 scenario as described in methodology below. For grid-connected PV systems, the generated electricity can be sent directly to the load. In contrast, storage PV systems often require energy storage devices BESS, like batteries, to store the electricity generated during the day when there is a surplus energy. To calculate the technical sizing of a grid-connected PV system with/without BESS with different scenarios under the solar PV regulation we have the method apply mathematical formular and PVsyst software, compute the energy production and compare the results with the proposed scenario (S0, S1, S2, S3, S4, and S5). For Economic analysis on the above 6 scenarios, we have method to calculate the NPV and LCOE, compare all scenarios to find which one has a good LCOE, saving, ROI, and payback period. Regarding to the Environmental analysis on the above 6 scenarios we have method to calculate the CO2 emission in all scenarios above by multiplying CO2 emission factor/kWh and compare the CO2 emission in all scenarios above to find which one is the lowest. This stored energy can then be used at night or during cloudy periods. After selecting the optimal solar system, the system will be designed and its maximum output power and efficiency will be evaluated. The first step in the study process is gathering significant data, including site location, energy consumption, meteorological information and space available for installation. The following step, PVsyst will be utilized to choose the best-case study among these two cases study which is mentioned above. Finally, in the economic part, make a comparison based on the COE and surely followed by EAC's regulation.*

Keywords: Photo Voltaic system (PVsyst); Battery Energy Storage System (BESS), Levelized Cost of Electricity (LCOE); Total Net Present Cost (NPC); Return on Investment (ROI); Total Revenues (TR); Lifecycle cost of PV project (LCC PV project); Project Lifetime (PL); Initial costs (Init costs).

1. INTRODUCTION

The energy market will become ever more important for Cambodia's economy as it continues to develop, with rapidly increasing demand for electricity. The Cambodia Energy Outlook estimates an increase by 8.63 times the last 15years [1], [2]. According to Cambodian authorities, electricity demand in the country is expected to grow from its current 1.5 GW to 3.9 GW by 2020, and 5.04 GW megawatts by 2024. So far, power imports from Laos, Vietnam and Thailand are helping the country to meet its growing demand. In Cambodia, electricity generation is dominated by hydro power and coal power plants. Power outages are especially common in the dry season, making Cambodian businesses dependent on diesel back-up systems to ensure their electricity supply. Both, high electricity costs and power

[1*] Corresponding author: Piseth Prom
E-mail: piseth_prom14@yahoo.com; Tel: +855-16 710 959

outages lead to an increasing attractiveness of investments in photovoltaic systems. Looking for ways to make energy more sustainable, solar PV systems for on-site electricity generation can be a solution to provide an alternative to conventional fossil fuels, lower electricity costs, and to increase electricity supply security. Solar PV systems for on-site electricity generation can be a solution to lower electricity costs and to increase electricity supply security.[1], [2]

2. METHODOLOGY

2.1 Flow chart of Methodology for Solar Roof Top Exiting project

The flowchart in Figure 1**.** shows how each stage of the process needs to be completed. Firstly, we start with the site assessment then moving to the analysis of the site information. Last but not least Design PV system for all scenarios. Finally, the best scenario is considered by the cost of efficient (LCOE) which is less than.

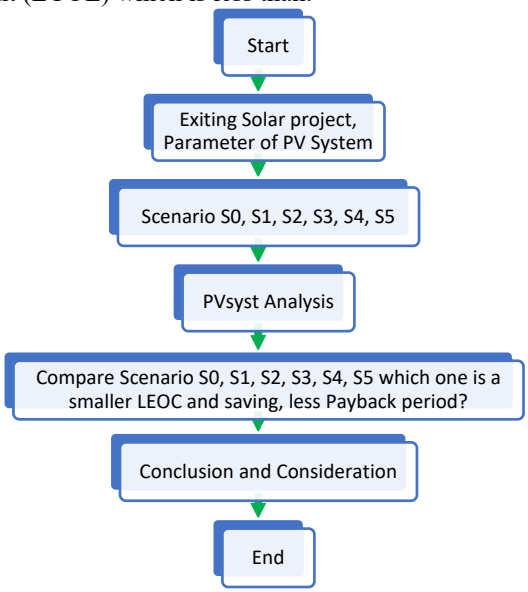


Fig. 1. Flow chart of Methodology for Solar Roof Top Exiting project

The flowchart in figure 1 shows how each stage of the process needs to be completed. First, we start the site assessment then moving to the analysis of the site information. Last but not least Design PV system and simulation with PVsyst for Scenario 0, 1, 2, 3, 4, 5. Finally the best scenario is considered by the cost of efficient (LCOE) which is less than. The step-by-step process for optimal sizing for PV systems + BESS is explained through the flowchart, as shown in Figure 1.

- Scenario 0: 100% purchase from EDC
- Scenario 1: EDC + PV (exiting) + Old Regulation
- Scenario 2: EDC+PV (exiting) + New Regulation
- Scenario 3: EDC+PV (exiting)+ BESS + New Regulation
- Scenario 4: EDC + PV (exiting)+ BESS=50% +New Regulation
- Scenario 5: EDC + PV (exiting)+ BESS=70% +New Regulation

2.2 Exiting Regulation Solar roof Top in Cambodia

Table 1. Old Regulation on Energy Tariff for consumers with PV system:

Type of Consumer and Conditions	Unit	Tariff
5. Commercial, Administration and Other Consumers Connecting to MV Distribution Line		
• For Consumers with PV System - Maximum Power Capacity Rate	$/kW/ month	5.80
- Electricity Rate (24H)	$/kWh	0.1500
➢ The Maximum Capacity of the Solar Power Project, as mentioned on AC side at the output of inverter based on rated inverter capacity, shall not be more than 50% of the Contract Demand in kW (if Contracted Demand is stated in kVA, it will be converted to kW at normative Power Factor of 0.90), as stated in the Power supply Agreement with the consumer.		

2.3 New Regulation on Energy Tariff for consumers with PV system

Condition 1:
Electricity users who have received permission from the Electricity Authority of Cambodia to install rooftop solar power systems and then use electricity from the national grid Combined with their roof top solar power to meet their full fill consumption needs must pay a compensation tariff to the national grid based on the number of kilowatt-hours of energy from the rooftop solar system, according to the capacity requirements as specified table below.

Table 2. New Regulation on Energy Tariff for consumers with PV system

Solar Rooftop size Capacity		Compensate Tariff
Small size	Not Over 10kWAC	No need to pay compensation tariff
Medium size	Over 10kWAC to 50kWAC	0.037 $/kWh
	Over 50kWAC to 100kWAC	0.047 $/kWh

In case the electric user has installed the BESS system combined with Solar Energy Rooftop on their own roof to fulfill the need of electric used and the Energy quantity (kWh) that discharge from the BESS system in which month have amount from 50% of Energy quantity which generate from that Solar rooftop. This is assumed that month the electric user has used their own Energy from BESS system to fulfill the power need for their electric used for at that month the electric user is not required to pay the compensate tariff as above. In contrast in which month the Energy quantity (kWh) which generate from the BESS system is less than 50% of the Energy quantity that generated from the Solar Rooftop system so, the electric user must pay the compensate tariff to the Energy quantity which is not fulfill, follow the formular below:

Energy Payment= {Energy quantity generate from Solar Rooftop system (kWh) – [2 x Energy generated from BESS system (kWh)]} x Compensate Tariff as table above.

2.4 Load Profile and PV system sizing

The Load Profile estimates the total power usage for a project over a particular time period. The greater the accuracy of the load profile, the more effective the PV system. In Cambodia, the on-grid PV system is required three important factors: Available space, as we have seen in the line graph of Toyota office the load operation almost flat, the peaking hour is around **190kW**. So, this solar system is exiting solar project designed PV system and installed is (100KW) based on which is chosen in the load calculation 50% of peak load demand capacity 200kW.

Battery capacity required is known as Ampere-hour **(**Ah).

The total ampere-hour can be obtained using the following formula:

$$C_{BAH} = \frac{Day\ excess\ Energy(kWh)}{DOD * \eta_{BAH} * V_B} \quad (1)$$

Where is the daily energy required from the battery, DOD is the permissible depth of discharge? The values of the factors considered while calculating for battery size are battery loss 0.85, depths of discharge, DOD – 0.8, and battery's nominal voltage. From our battery specification, the nominal voltage is 26V.

2.5 Economic of PV system

2.5.1 Project costing

- Total Net Present Cost (NPC)

The NPC is the present value of all the system's costs over its lifetime, minus the present value of all the revenue it earns. Costs include capital costs, replacement costs, O&M costs. Revenues include salvage value and grid sales revenue [3]

$$C_{NPC,total} = \frac{C_{ann,total}}{CRF(i,n)} \quad (2)$$

Where:

$CNPC,total$ is the total net present cost ($)

$Cann,total$ is the total annualized cost $/year

$CRF(i,n)$ is the capital recovery factor that represent a series of equal annual income and can be defined by:

$$CRF(i,n) = \frac{i(1+i)^n}{1+i^n-1} \quad (3)$$

n is the project life time(year)

i is the real discount rate % and can be defined by:

$$i = \frac{i'-f}{1+f} \quad (4)$$

i' is the nominal interest rate %

f is the annual inflation

- Discount Factor

$$Discount\ Factor = \frac{1}{1+i^N} \quad (5)$$

N is duration of a project lifetime

2.6 Project Profit

- Levelized Cost of Electricity (LCOE)

The levelized cost of electricity (LOCE), also referred as levelized cost of energy (LCOE), is used for electrical systems to calculate the cost of producing one unit of electricity (USD/kWh). The operational lifespan cost includes the initial cost of capital, operations, and maintenance and the cost of replacement (i.e., batteries and electronic equipment). With regard to the second assumption, time value of money is considered and LCOE is calculated using a discount rate. The LCOE at discount rate is calculated as follows [4]

where, N is the project lifetime and DR is the discount rate.

Considering economic aspects, the per unit cost of electricity for PV systems significantly depends on several factors including the geographical locations and discount rates [5]. For PV projects in Asia countries, this value is seen to range between 4 and 8%[6]. As an average, discount rate of 12% is used for Cambodia.

$$LCOE = \frac{Ini\cos ts + \sum_{n=1}^{N} \frac{O\&M(\$)}{(1+DR)^n} + \sum_{n=1}^{N} \frac{\mathrm{Re}\,p\cos t(\$)}{(1+DR)^n}}{\sum_{n=1}^{N} \frac{Annual\ PV\ System(kWh)}{(1+DR)^n}} \quad (6)$$

Where:

Init costs: Initial costs ($)

Rep: Replacement cost ($)

Annual PV system: Annual electricity supply to the load by PV system (kWh)

2.7 Payback Period

The Payback Period is the period required to recoup the funds expended in an investment, starting from the investment year. The payback period for Solar roof top PV systems can be calculated by following [7]:

$$Payback\,Period = \frac{LCC(\$)}{TR(\$)/\mathrm{P}\,L(year)} \quad (7)$$

Where:
PP: Payback Period
LCC: Lifecycle cost of PV project ($)
TR: Total Revenues ($)
PL: Project Lifetime (year)

2.8 Return on Investment (ROI)

Return on Investment (ROI) is a popular profitability metric used to evaluate how well an investment has performed. ROI for PV system is expressed as a percentage and is calculated by dividing project's net profit by operational lifespan cost of PV project. Mathematically, it can be expressed as follows[4], [5]:

$$ROI = \frac{T\,R(\$) - LCC(\$)}{LCC(\$)} * 100 \quad (8)$$

Where:
ROI: Return of investment (%)
TR: Total Revenues ($)
LCC: Lifecycle cost of PV project ($)
Similar to the LCOE, payback period and return on investment (ROI) for the 4 cases shall be calculated separately with and without the use of the discount rate.

3. RESULTS AND DISCUSSION

3.1. Sizing Methodology PV Systems in PVsyst

From the main result data PVsyst, solar produced energy per year =184448kWh/year, Used Energy per year =636578kWh/year. Performance Ratio=70.81%, Solar fraction =28.97%.

3.2. BESS (kWh) Calculation

By using (1):

- Calculation BESS (kWh) for scenario 3:

From PVsyst total daily Excess Energy (kWh) = 62.4kWh
By using equation (1):

$$C_{BAH} = \frac{62.4(kWh)}{0.85*0.8} = 91.76kWh \approx 100kWh$$

- Calculation BESS (kWh) for scenario 4:

50% Total daily Energy (kWh) selected highest month in June. So 50% daily (kWh) =[(18294/30)*50%]= 304.9 kWh
By using equation (2):

$$C_{BAH} = \frac{304.9kWh}{0.85*0.8} = 448.38kWh \approx 450kWh$$

- Calculation BESS (kWh) for scenario 5:

70% Total daily Energy (kWh) selected highest month in June. So 70% daily (kWh) =[(18294/30)*70%]= 426.86 kWh By using equation (3):

$$C_{BAH} = \frac{426.86kWh}{0.85*0.8} = 627.73kWh \approx 630kWh$$

3.3. Project costing

The project costing of case study 1 is approximately = 138040 x 0.785 = **108361.40$**.

Table 3. Price of Component for PV system

	Equipment	Cost
	PV modules + BOS	0.755 USD/W
	Inverter	0.03 USD/W
	Storage	0.250 USD/W

There for the Total Project costing:

- Scenario S1&S2: The project costing of case study 1 is approximately = 138040 x 0.785 = **108361.40$**.
- Scenario S3: Battery sizing with Total daily Excess Energy (kWh) = 100*1000*0.250 = **2,5000.00 $**
- Scenario S4: Battery sizing with 50% of Total daily PV Energy (kWh) = 450*1000*0.250 = **112,500.00$**
- Scenario S5: Battery sizing with 70% of Total daily PV Energy (kWh) = 630*1000*0.250 = **157,500.00$**

Table 4. Total Project investment cost (USD) for all Scenarios

Scenario	Total Investment Cost (USD)
S0	0.00 USD
S1	108361.40 USD
S2	108361.40 USD
S3	133,361.40 USD
S4	220,861.40 USD
S5	265,861.40 USD

3.4. Comparison of electricity for the 1st year operations expense saving All scenarios

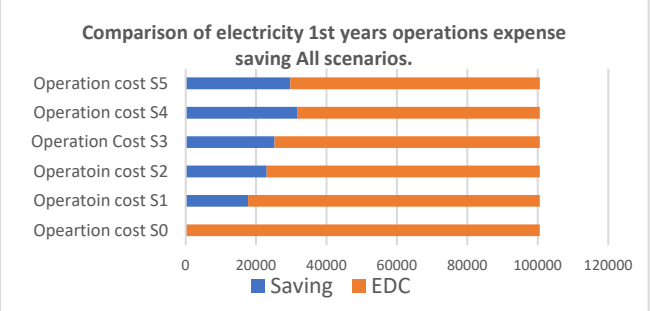


Fig. 2. Electricity 1st year operation expenses and saving all scenario

Base on electrical Tariff above Exiting regulation and new regulation the price for the grid is 0.158$ without solar and grid with solar is 0.15$.
As we have seen in the Figure 2 and Table 5 below, Scenario 4 is saving 31,793.82$ in a 1st year which is accounted for 32% while the other saving in scenario S1=18%, S2=23%, S3=25%, S5=30%.

Table 5. Operation cost for 1st year and saving ($,%)

Operation Cost	Total Expenses ($)	Saving ($)	Saving (%)
S0	100,579.32 $	0.00 $	0.00 %
S1	82,816.16 $	17,763.15 $	18 %
S2	77,533.46 $	23,045.86 $	23 %
S3	75,313.61 $	25,265.71 $	25 %
S4	68,785.5 $	31,793.82 $	32 %
S5	70,846.05 $	29,733.27 $	30 %

3.5. Payback period

The operating and maintenance cost or it was called O &M cost is the total budget that use for operating the whole system.

Every single component in the project, especially, solar PV for 25 years and inverter for 15 years is their limit lifetime during the operating of the system.
Replacement of Inver is doing every 15 years of warranty and it costs $ 4,206.88 also the BESS for Scenario 3 is 25,000.00$ and the BESS scenario 4 is 112,500.00 $ and the BESS scenario 5 is 157,500.00 $. By Using (7):

Table 6. Payback period for all scenarios

Payback Period		
Scenario 1	16.23years	= 16 Years 3 Months
Scenario 2	8.66 years	= **8 Years 8 Months**
Scenario 3	13.49 years	= 13 Years 6 Months
Scenario 4	20+ years	= Over 20 Years
Scenario 5	20+ years	= Over 20 Years

As we have seen the payback period, the case study **Scenario S2** is spending 8 years and 8 months shorter than the case study S0, S1, S3, S4, S5 to paid back the system cost. The case study **Scenario S2** yields is more annual profits than the case study after 8 years of repayment.

3.6. Net Present Cost

This part is about the economic calculation for the lifetime of the project that we counted on 20 Years of the project lifetime. Discounted Rate 12.00% base on investment economic analysis from ADB economic analysis paper [3]. Since the inflation rate is 2% will into account for economic analysis [3].

In summary, we got all parameters that can be done in (3) and (4):

$$i = \frac{0.12 - 0.02}{1 + 0.02} = 9.8\%$$

$$\Rightarrow CRF(0.098, 20yrs) = \frac{0.098(1+0.098)^{20}}{(1+0.098)^{20} - 1} = 0.116$$

We use this discount rate and for calculation the hold project for 20 years of a project lifetime to find NPC, Cann and LCOE.

As we have seen the from Fig. 3 the LCOE for scenario S2=0.086$/kWh is the lowest electricity cost than the other Scenario S0 =0.158$/kWh, S1=0.115$/kWh, S3=0.093$/kWh, S4=0.111$/kWh, S5=0.134$/kWh.

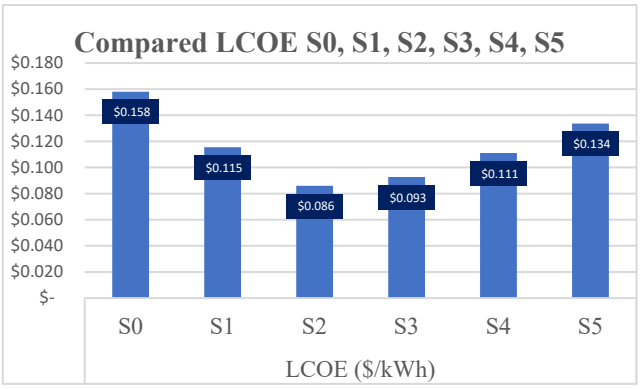


Fig. 3. LCOE comparison for all Scenarios

3.7. Project Evaluation and Discussion

Table 7. Overall system comparison initial cost ($) and O&M cost ($) between scenario S0, S1, S2, S3, S4, and S5:

	PV Capacity & BESS	Initial Cost ($)	O&M Cost ($)
S0	0.00kWp	0.00 $	100,579.32 $
S1	138.04kWp /100kWAC	108361.40 $	82,816.16 $
S2	138.04kWp /100kWAC	108361.40 $	77,533.46 $
S3	138.04kWp /100kWAC +BESS 100kWh	133,361.40 $	75,313.61 $
S4	138.04kWp /100kWAC +BESS 450kWh	220,861.40 $	68,785.5 $
S5	138.04kWp /100kWAC +BESS 450kWh	265,861.40 $	70,846.05 $

Table 8. Overall system comparison NPC cost ($) and LCOE cost ($) and 1st year saving ($) between scenario S0, S1, S2, S3, S4, and S5:

	NPC ($)	LCOE ($/kWh)	1st Year Saving ($)
S0	100,579.32 $	0.158 $/kWh	0.00$
S1	82,816.16 $	0.115 $/kWh	817,763.15$
S2	77,533.46 $	**0.086$/kWh**	23,045.86 $
S3	75,313.61 $	0.093 $/kWh	25,265.71 $
S4	68,785.5 $	0.111 $/kWh	31,793.82 $
S5	70,846.05 $	0.134$/kWh	29,733.27 $

By seeing the summary comparison from Table above, the LCOE of scenario **S2 is 0.086$/kWh** lower than the LCOE of scenario S0, S1, S3, S4, and S5. Besides that, in 1st years of saving cost, the scenario S4 could saving 31, 793.82 $ higher than the rest scenarios. Therefore, the grid-connected system is design with power peak load demand (200kWAC), (Case Study **S2** New regulation Compensate Tariff (0.047$/kWh)) which is 138kWp with an annual production of 184448kWh with 8.66 years payback period including the degradation factor. In contrast, the **S1** Old regulation Capacity Charge (5$/kW) which design the same solar capacity is 138kWp with annual production of 184448kWh with 16.23 years payback more than case S2 above also S3, S4, and S5 are applying with the new regulation are takes so long payback period. Therefore, after the recalculate the exiting solar project KPI for Toyota project the best consideration for client is to consider Case Study **Scenario S2** New regulation compensate tariff (0.047$/kWh) system which could financially profitable than the other scenarios.

4. CONCLUSIONS

This research report study of the Grid-connected photovoltaic system for the Toyota Cambodia head Quatre has been carried out and applied successfully to the clients after the recalculate the exiting solar project KPI for Toyota project the best consideration for client is to consider Case Study **S2** the LCOE of scenario S2 is 0.086$/kWh lower than the LCOE S1 which is 0.115$/kWh and LCOE scenario S3 is 0.0935$/kWh and scenario S4 is 0.111$/kWh and S5 is 0.134$/kWh and compare to EDC S0 LCOE is 0.158$/kWh. Furthermore, we can see the system of scenario S2 is much more saving cost on the capacity charge by replace with the compensate tariff. Besides that, in 20 years of project life time, the scenario S2 could saving **$ 314128.87** compare to the Electricity bill from the grid**.**

References

[1] Sevea Consulting, "giz2021_en_Sector Brief PV_GBN-Cambodia_WEB," 2020, Accessed: Jan. 21, 2025. [Online]. Available: Responsibility for the content of external websites linked in this publication always lies with their respective publishers. GIZ expressly dissociates itself from such content.

[2] "Important developments in the electricity sector of the Kingdom of Cambodia up to December2024"

[3] Homer Pro, "HOMERHelpManual," *HOMER® Pro Version 3.7*, 2016.

[4] A. Ali, M. Volatier, and M. Darnon, "Optimal Sizing and Assessment of Standalone Photovoltaic Systems for Community Health Centers in Mali," *Solar*, vol. 3, no. 3, pp. 522–543, Sep. 2023, doi: 10.3390/solar3030029.

[5] P. M. Khuong, R. McKenna, and W. Fichtner, "A cost-effective and transferable methodology for rooftop PV potential assessment in developing countries," *Energies (Basel)*, vol. 13, no. 10, May 2020, doi: 10.3390/en13102501.

[6] Chun Sing Lai and Malcolm D. McCulloch, "16 ManuscriptAppliedEnergyaccept_2".

[7] F.-E. Riakhi and A. Khaldoun, "Development of a Stand Alone Solar Carport Model Linked to Charging Stations and its Economic Analysis (Case Study)."

Food Technology and Nutrition (FTN)

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Effect of Oven-drying Temperature on Nutrition Value of Dried Nile Tilapia Slices

Sreyroth Eang, Seanghai Hoeun, Sokneang In*

Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia

Corresponding author: in@itc.edu.kh

Abstract: *Nile tilapia (Oreochromis niloticus) is a popular type of fish consumed by the Cambodian population due to its ease of feeding, affordable price, and high nutritional content, especially its higher protein level compared to other fish. Recognizing that fresh fish is prone to quality deterioration, it has been processed into dried fish products, which is a simple and cost-effective method to preserve and extend the product's shelf life. However, the drying temperature can affect the nutritional composition of the final product, such as lipids, proteins, and carbohydrates. Therefore, the objective of this study focused on the effect of drying temperatures on the nutritional composition of dried Nile tilapia slices. The drying kinetics of dried Nile tilapia were studied at various temperatures (60°C, 70°C, and 80°C). After drying, the nutritional value, including ash content, lipid content, protein content, and total carbohydrates, was also conducted by the AOAC method. The study revealed that dried Nile tilapia contained ash (3.90% to 4.48%), lipids (13.33% to 13.39%), protein (70.86% to 73.22%), and carbohydrates (3.09% to 1.33%). In conclusion, temperature had a positive impact on the ash and protein content, while it negatively affected the moisture content and total carbohydrates. The lipid content remained constant. Additionally, based on the results of the drying kinetics of Nile tilapia slices, the optimal drying period was between 7 to 8 hours, as the moisture content of the fish slices became stable during this time.*

Keywords: Nile tilapia; nutritional value; Oven-drying; temperature; drying kinetic

1. INTRODUCTION

Nile tilapia *(Oreochromis niloticus)* is the freshwater fish which is an essential source of food for Cambodians countrywide. It is the most commonly cultivated tilapia species because it is one of the most valuable farmed fish species in the world (Gonzales & Brown, 2006). Nile tilapia may satisfy humans' requirements for protein and omega 3 (Nuryanto et al., 2022). The study of Gonzales & Brown (2006) found that the Nile tilapia contained the essential vitamins such as vitamins A, D, E, and K. It also presented some minerals, including calcium, sodium, phosphorus, magnesium, potassium, zinc, selenium, molybdenum, copper, cobalt, boron, iron, and manganese. Presently, the local product is from Nile tilapia varieties such as whole frozen fish, fish fillets, smoked fish, dried fish, crispy fish and fish powder. Since fresh fish is prone to spoiling, it is processed into dried fish products which is a simple and affordable method to preserve and extend the shelf life of the product. Drying is the process of monitoring the decrease in moisture content in fish slides versus drying time. According to Chukwu (2009) indicated that tilapia fish nutritional qualities are impacted by drying techniques. Thus, the objective of this research is to investigate the impact of drying temperature on the nutritional value of Nile tilapia. The study also evaluated the nutritional components, including ash content, fat content, protein content, and total carbohydrates.

2. METHODOLOGY

2.1 Sample preparation

The fresh Nile tilapia was thoroughly cleaned by descaling, degutting (take only muscle tissue). Then, it was washed again with clean water. After that, the muscle tissue was cut 0.3 to 0.4 cm in length and 10 to 14 mm in thickness before arranging on trays.

2.2 Drying kinetic

* Corresponding author: Sokneang In
E-mail: in@itc.edu.kh; Tel: +855-12 447 898

An amount of 400 g of Nile tilapia was weighed by using electronical balance. Then, it was dried at various temperatures (60°C, 70°C, and 80°C) for 13 hours in the oven, with its weight recorded at one-hour intervals. After the drying process, it was determined the nutritional composition, such as ash content, fat content, protein content, and carbohydrates.

2.3 Ash content

Ash content of dried Nile tilapia was carried out following the AOAC method (950.46). Porcelain crucible was dried in oven for 1 hour before use. Then, it was weighed after cooling in the desiccator. About 3 to 5 g of fish powder was weighed into the porcelain crucible (tare mass) and ignited in the furnace (B-180, Nabertherm, Germany) at 505°C for 3 hours until light gray ash was obtained. Next, it was cooled down in the desiccator and weighed.

$$\% \text{ Ash} = [(M_D - M_C) \times 100] / M_S \qquad \text{(Eq. 1)}$$

Where:
M_S = mass of sample (g)
M_C = mass of empty crucible (g)
M_D = mass of sample after heating (g)

2.4 Fat content

Fat content of dried Nile tilapia was determined by using extraction method (AOAC 963.15). The empty extraction glass beaker was dried in the oven for 1 hour. Then, it was cooled in the desiccator for about 15 minutes and weighed for mass. Next, about 1 g of dried fish was weighed into a cleaned thimble and covered with a small amount of cotton. After that, about 70 ml of n-hexane solvent were poured into glasses. The thimbles and glasses were then inserted into a Soxhlet apparatus (SER-148, VEIP, Italy). The machine was set to: temperature; 60 minutes for immersion; 60 minutes for washing; and 15 minutes for recovery. After the end of the operation, the extraction glass beaker that contained fat were dried in the oven at 105°C for 30 min, then cooled in a desiccator and weighed.

$$\% \text{ Fat} = [(M_A - M_B) \times 100] / M_S \qquad \text{(Eq. 1)}$$

Where:
M_S = mass of sample (g)
M_B = mass of empty extraction glass beaker (g)
M_A = mass of lipid after extraction with glass beaker (g)

2.5 Protein content

Protein content of dried Nile tilapia was analysis by using Kjeldahl method (A.O.A.C 928.08). There are three main steps for determining protein content: digestion, distillation, and titration. For the digestion step, 1 g of dried fish sample was weighed into the digestion tube, followed by 10 g of catalyst and 20 ml of concentrated sulfuric acid (98%). The speed digestion unit (Digester K-439, BUCHI, Switzerland) was then started with a set preheating at 250°C for 60 min, digestion at 420 °C for 95 min and cooling for 30 min. The digestion process was finished when the solution changed to a mint color. Next, the sample was cooled down for distillation. The distillation unit (KjelFlex K-360, BUCHI, Switzerland) was used with some chemical reagents, such as 10 ml of boric acid (4%) with a few drops of methyl red indicator, 90 ml NaOH (32%), and 50 ml of distilled water. After distillation, the solution turned into yellow color. Next, it was titrated with sulfuric acid (0.25 M) until the solution changed to an orange color.

$$\% \text{ Nitrogen} = [(V_S - V_B) \times Z \times C \times F \times M_N \times 100] / (M_S \times 1000) \qquad \text{(Eq. 3)}$$

$$\% \text{ Protein} = \% \text{ Nitrogen} \times \text{Factor of Protein (6.25)} \qquad \text{(Eq. 4)}$$

Where:
V_S = amount of titrant for the sample (ml)
V_B = amount of titrant for blank (ml)
Z = molar reaction factor for H_2SO_4 (2)
C = concentration of titrant (mol/L)
F = factor of titrant (0.9969)
M_N = molecular weight of nitrogen (g/mol)
M_S = mass of sample (g)

2.5 Total carbohydrate

Carbohydrates of dried fish were calculated by subtracting the total moisture, ash, fat, and protein from the total dry mater (A.O.A.C 986.25).

$$\% \text{ Carbohydrate content} = 100\% - \% \text{ (Moisture content + ash content + fat content + protein content)} \qquad \text{(Eq. 5)}$$

3. RESULTS AND DISCUSSION

3.1 Impact of temperature on drying kinetic of Nile tilapia slices

The drying kinetics of Nile tilapia slices at various temperatures (60°C, 70°C, and 80°C) for 13 hours were illustrated in **Fig. 1**. The first phase is the period where the moisture content of the fish decreased steadily, from 0 to 8 hours. During this phase, the moisture content of Nile tilapia at 80°C was observed to decrease most rapidly from 50% to 14%, compared to temperatures of 60°C and 70°C. However, between 11 to 13 hours, it was observed that the moisture in the fish slices reduced to below 15%. For fish samples dried at 70°C and subsequently at 60°C, the decrease in moisture content demonstrated the process of water removal using hot air drying in the oven. Thus, the higher the drying temperature, the faster the moisture removal in the fish samples.

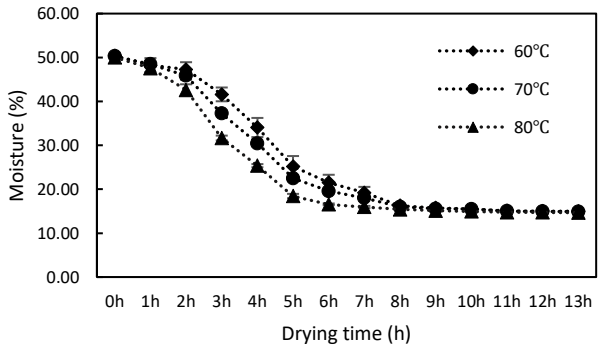


Fig. 1. Drying kinetic of Nile tilapia slices

This phase highlighted the transfer of moisture from the interior surface to the hot air within the oven (Dhahkal, 2022). Additionally, Sidhi et al. (2018) explained that water is released from the samples through heat, leading to a decrease in the moisture rate of fish slices, while the drying time increased. The second phase occurred when the fish's moisture level remained constant for 8 to 13 hours. At the temperature of 80°C, the moisture in Nile tilapia slices dropped to below 15% for 10 hours. At a temperature of 70°C, moisture content fell below 15% for 12 hours, while at 60°C, this occurred for 13 hours. During this phase, the results indicated that when the fish was exposed to heat, the free water molecules in the inner layers of the fish slices gradually migrated to the outer layers. Consequently, the moisture on the surface was transferred to the surrounding hot air. It was noted that higher temperatures result in faster water removal from food products since they enhanced the rate of breaking the bonds of free water molecules within the dried fish slices. Moreover, increased evaporation rates of water molecules from the fish samples helped to reduce the drying time (Matthews, 2010). A study by Jeyakumari et al. (2018), found that the moisture content in dried fish ranged between 15% and 20%. The results of drying kinetics, temperature significantly affected moisture content, as higher temperatures led to greater moisture reduction in dried fish slices. In short, the drying process of fish could be completed within 7 to 8 hours to achieve the moisture content of approximately 15%.

3.1 Nutritional value of dried Nile tilapia slices

As shown in **Table 1**, the ash content in the dried fish samples at temperatures of 60℃, 70℃, and 80℃ was 3.90 ± 0.15%, 4.08 ± 0.06%, and 4.48 ± 0.09%, respectively. Based on the results, the ash content of the dried fish at 80℃ was the highest compared to the samples dried at 60℃ and 70℃, which were similar values. In line with the study by Rasul et al. (2021), ash content of dried freshwater fish ranged from 4.92% to 24.4%. The ash content in dried fish samples can vary depended on different living environments. For farmed fish, this variation may depend on the feed provided by farmers and the living conditions. Generally, the ash content of tilapia increased after drying due to the reduction in moisture content in the samples. On the other hand, drying conditions, including temperature and drying duration, can affect the ash content in dried Nile tilapia samples. Improper temperature selection during the drying process can result in insufficient removal of moisture from the product, and any remaining moisture may impact the results obtained (Maula et al., 2020).

The lipid content of dried Nile tilapia at different temperature conditions, specifically at 60°C, 70°C, and 80°C. According to **Table 1**, the lipid content for samples dried at 60°C (13.33 ± 0.06%), 70°C (13.39 ± 0.26%), and 80°C (13.79 ± 0.14%) was not significantly different. As stated in the previous study by Rasul et al. (2021), dried freshwater fish contained lipid values ranging from 3.10% to 22.7%. On the other hand, Chukwu (2009) conducted a study on the drying of Nile tilapia fish at 110°C for 45 min, which indicated a lipid value of approximately 20.25%, while fresh fish only had a lipid value of 12.85%. This study demonstrated the difference in lipid content before and after drying due to the proportional composition of the components in the sample, including water content and lipid content. When water evaporated, the oil molecules in the sample remained stable and concentrated, resulting in a higher lipid content in dried fish compared to fresh fish (Kibar et al., 2010). Furthermore, Sroy et al. (2023) mentioned that the lipid content increased as the drying temperature raised. However, the obtained result of dried tilapia fish from 60°C to 80°C was not shown any variation in lipid content.

Table 1. Nutritional value of dried Nile tilapia slices at different temperatures (60°C, 70°C, and 80°C)

Parameter	Temperature		
	60°C	70°C	80°C
Ash (%)	3.90 ± 0.15[a]	4.08 ± 0.06[a]	4.48 ± 0.09[b]

Fat (%)	13.33 ± 0.06[a]	13.39 ± 0.26[a]	13.79 ± 0.14[a]
Protein (%)	70.89 ± 0.12[a]	72.89 ± 0.02[b]	73.22 ± 0.30[b]
Carbohydrates (%)	3.09 ± 0.16[a]	2.46 ± 0.23[b]	1.38 ± 0.06[c]

After drying at different temperature conditions, the protein content of dried Nile tilapia was insignificant difference between drying at 70℃ (72.89 ± 0.02%) and 80℃ (73.22 ± 0.30%). Meanwhile, the Nile tilapia dried at a temperature of 60℃ (70.89 ± 0.12%) had the lowest value. The reason for the lower protein content in Nile tilapia dried at 60℃ compared to the other two drying conditions could be due to the moisture content present in the tilapia. According to Aba, (2013), who studied the effects of smoke-drying temperature and time on physical and nutritional quality parameters of tilapia, it was explained that the moisture content decreased as temperature and drying time increased, while the protein content correspondingly increased. Based on the study by Akther S et al. (2018), the protein content in dried fish ranged from 57.76% to 73.32%. Similarly, the study by Chukwu (2009) found that the protein content of dried fish at 110°C for 45 min was about 64.10%. The higher drying temperatures, such as at 80°C, can trigger the Maillard reaction, whereby amino acids react with heat, resulting in a brown color and fragrant aroma. The degree of browning in dried fish varies depending on the drying time. Afolabi, (2014) showed that prolonged exposure to high temperatures intensified the color change in the fish samples. On the other hand, at drying temperatures of 60°C and 70°C, it was observed that fish samples were only slightly brown, which may result in the dried fish having an unpleasant odor (a rancidity). To sum up, the protein content obtained from drying at 80°C is considered better than the two other drying conditions.

The total carbohydrates is a nutrient that humans necessarily need to consume in their daily diet. In general, fish contains a small amount of carbohydrates because protein is the nutrient that is observed to be present in greater amounts. Therefore, if the quantity of carbohydrates is high, the amount of protein may be reduced (Fitri et al., 2022). The total carbohydrates of dried Nile tilapia at temperatures 60℃, 70℃, and 80℃ was shown in **Table 1**. Following the results, the highest carbohydrate value was obtained from Nile tilapia dried at 60℃ which was 3.09 ± 0.16 %. Subsequently, at 70℃, the value of carbohydrates decreased to 2.46%, and at 80℃, the value further dropped to 1.33%. This indicated that higher drying temperatures accelerated the rate of the Maillard reaction, a non-enzymatic browning process that occurred between sugars and amino acids. This reaction can lead to a reduction in the carbohydrate content (Manley, 2011).

4. CONCLUSIONS

According to result of this study, the kinetics of drying are carried out at the temperatures of 60°C, 70°C, and 80°C, starting constant for 7 to 8 h with the moisture content under 10%. Furthermore, the dry temperature of 80°C is suitable for oven drying based on the highest ash and protein content. However, the moisture content and carbohydrates decreased, except for lipids. For further study, it should be evaluated on the physicochemical and microbiological quality of dried Nile tilapia.

ACKNOWLEDGMENTS

This research was supported by the Food Technology, Research, and Innovation Platform (FTRIP) under the framework of the CAPFish-Capture project Funded by the European Union (EU) and implemented by the United Nations Industrial Development Organization (UNIDO).

REFERENCES

AOAC. (2012). AOAC Official Methods of Analysis (J George W. Latimer, Ed,). 19th ed. George W. Latimer, Jr. 1: 43–45.

Aba, P. (2013). Effects of smoke-drying temperatures and time on physical and nutritional quality parameters of Tilapia (Oreochromis niloticus). *International Journal of Fisheries and Aquaculture*, *5*(3), 29–34. https://doi.org/10.5897/IJFA12.078

Afolabi, I. S. (2014). Moisture Migration and Bulk Nutrients Interaction in a Drying Food Systems: A Review. *Food and Nutrition Sciences*, *05*(08), 692–714. https://doi.org/10.4236/fns.2014.58080

Chukwu, O. (2009). Influences of Drying Methods on Nutritional Properties of Tilapia Fish (Oreochromis nilotieus). *World Journal of Agricultural Sciences 5*, *5*(2), 256-258,.

Fitri, N., Chan, S. X. Y., Che Lah, N. H., Jam, F. A., Misnan, N. M., Kamal, N., Sarian, M. N., Mohd Lazaldin, M. A., Low, C. F., Hamezah, H. S., Rohani, E. R., Mediani, A., & Abas, F. (2022). A Comprehensive Review on the Processing of Dried Fish and the Associated Chemical and Nutritional Changes. *Foods*, *11*(19). https://doi.org/10.3390/foods11192938

Gonzales, J. M., & Brown, P. B. (2006). Nile tilapia Oreochromis niloticus as a food source in advanced life support systems: Initial considerations. *Advances in Space Research*, *38*(6), 1132–1137. https://doi.org/10.1016/j.asr.2005.11.002

Kibar, H., Öztürk, T., & Esen, B. (2010). The effect of moisture content on physical and mechanical properties of rice (Oryza sativa L.). *Spanish Journal of Agricultural Research*, *8*(3), 741–749. https://doi.org/10.5424/sjar/2010083-1273

Manley, D. (2011). Sugars and syrups as biscuit ingredients. In *Manley's Technology of Biscuits, Crackers and Cookies: Fourth Edition* (Fourth Edition). Woodhead Publishing Limited. https://doi.org/10.1533/9780857093646.2.143

Matthews, S. (2010). Effect of drying temperature on fuel moisture content measurements. *International Journal of Wildland Fire*, *19*(6), 800–802. https://doi.org/10.1071/WF08188

Nuryanto, N., Afifah, D. N., Sulchan, M., Martosuyono, P., Ihsani, K., & Kurniastuti, P. L. (2022). Potential of Nile

Tilapia (Oreochromis niloticus) as an Alternative Complementary Food Ingredient for Stunting Children. *Open Access Macedonian Journal of Medical Sciences*, *10*(A), 1170–1177. https://doi.org/10.3889/oamjms.2022.9650

Rasul, M. G., Yuan, C., & Shah, A. K. M. A. (2021). Chemical composition and nutritional value of dried fish in bangladesh. *Egyptian Journal of Aquatic Biology and Fisheries*, *25*(4), 379–399. https://doi.org/10.21608/EJABF.2021.189666

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Nutritional analysis of roasted cashew nuts harvested from agroecological and conventional-practice farming in Preah Vihear

Munthae That[1], Nita Suon[1], Peany Houng[1], Sokuntheary Theng [1*]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*]*Corresponding author: theng.sokuntheary@itc.edu.kh*

Abstract: *Cashew nut cultivation involves two main approaches: agroecological systems and conventional chemical farming. The agroecological method incorporates cover crop, which enhances soil fertility by providing nitrogen and organic carbon. In contrast, conventional farming relies on chemical inputs, such as herbicides and insecticides, for pest and weed control.This study compares the nutritional profiles of roasted cashew nuts harvested from trees aged 4 to 6 years, cultivated under both systems in Preah Vihear province, Cambodia. Samples were analyzed for moisture, ash, sugar, fiber, fat, protein, and carbohydrate content using standard AOAC methods, and data were statistically evaluated using one-way ANOVA. Results revealed that agroecologically grown (organic) cashew kernels exhibited lower moisture levels (1.70%, 1.82%, and 1.67%) compared to their conventionally grown counterparts (1.85%, 1.95%, and 1.94%). Ash content was slightly higher in organic kernels (2.76%–3.52%) than in chemical ones (2.76%–2.79%). Reducing sugar levels were similar across both treatments, ranging from 2.50% to 2.73% in organic nuts and 2.47% to 2.56% in chemically treated samples. Organic cashew nuts generally had higher fiber content (3.09%–3.53%) than those grown conventionally (2.13%–3.39%). Similarly, fat content was greater in organic kernels (45.57%–49.05%) compared to chemical treatments (41.11%–47.24%). Notably, protein levels were also significantly higher in organic nuts (18.10%–19.33%) versus chemical ones (18.03%–18.58%). Conversely, carbohydrate content was greater in chemically treated nuts (29.77%–35.56%) than in organic samples (26.51%–31.80%). These findings indicate that agroecological farming practices can enhance the nutritional quality of cashew nuts, particularly in terms of fiber, fat, and protein content. The study highlights the potential health benefits of sustainable agriculture and highlights the value of eco-friendly cultivation methods in improving the nutritional profile of food products.*

Keywords: Cashew Nuts; Agroecology; Conventional Farming; Nutritional Analysis; Sustainable Agriculture

1. INTRODUCTION

Cashew nuts are delicious and nutritious seeds that come from the cashew tree (*Anacardium occidentale*), native to Brazil but now widely grown in tropical regions like India, Vietnam, and Africa (Oliveira et al., 2020). Nutritionally, cashew nuts are rich in lipids, proteins, and essential minerals, making them a valuable dietary component (Griffin & Dean, 2017). However, the nutritional profile of cashew nuts can vary significantly depending on cultivation practices (Azeez & Olabanji, 2024). The cultivation of cashew nuts can be managed through different agricultural approaches, notably agroecological (organic) and conventional (chemical) practices (Awodun et al., 2015). Agroecological practices involve intercropping cashew trees with cover crops such as Stylo, which enrich the soil by fixing nitrogen and adding organic carbon. This natural fertilization supports soil health and potentially improves nut quality(Zephirin Somda et al., 2023). Conversely, conventional farming methods often depend on the application of synthetic chemicals, such as herbicides and insecticides, to manage pests and weeds. This reliance can lead to the degradation of soil health over time and, furthermore, introduce concerns regarding both the quality of crop yields and the long-term sustainability of the environment (Soares et al., 2013).

The research aimed to analyze and compare the nutritional content of roasted cashew nuts produced under agroecological and conventional systems. Key nutritional parameters measured included moisture, ash, reducing sugar, fiber, fat, protein, and carbohydrate content. The study focused on nuts from trees aged 4, 5, and 6 years.

[1*] Corresponding author: Sokuntheary Theng
E-mail:theng.sokuntheary@itc.edu.kh; Tel: +855-87 848 181

2. METHODOLOGY

2.1 Sample collection and preparation

Roasted cashew nuts were collected from cashew trees aged 4, 5, and 6 years in Preah Vihear province. Two cultivation systems were compared: agroecological (organic) practices involving intercropping with cover crops like Stylo, and conventional practices using chemical herbicides and insecticides. Cashew nuts from both agroecological (organic) and conventional practice (inorganic) systems were roasted in an oven at 120°C for 30 minutes. After roasting, the nuts were ground into a fine powder, which was then used for nutritional analysis.

2.2 Nutritional composition analysis

The powdered cashew nuts from both agroecological and conventional farming systems were analyzed to determine their moisture, ash, reducing sugar, fiber, fat, protein, and carbohydrate content.

In the determination of moisture content for roasted cashew kernels, the procedure began by drying an aluminum plate in an oven at 105°C for one hour, followed by placing it in a desiccator for 15 minutes to cool before weighing. Then, 2 grams of milled roasted cashew kernel were weighed on the pre-weighed aluminum plate. The sample was dried in the oven at 105°C for 6 hours to remove moisture. Each sample was analyzed in duplicate to ensure accuracy. This method follows the AOAC official procedure 925.40. The moisture content percentage was calculated using the weight difference before and after drying according to the standard formula below.

$$\text{Moisture content (\%)} = \frac{M_2 - (M_3 - M_1)}{M_2} \times 100$$

Where:

M_1 is the mass aluminum (g)

M_2 is the mass of the sample (g)

M_3 is the mass of the dried sample with aluminum (g)

To determine ash content, porcelain crucibles were dried at 105°C, cooled in a desiccator, and weighed. One gram of milled, roasted cashew kernel was placed in the crucible, then ashed in a furnace at 550°C for 3 hours. After cooling in the furnace and desiccator, the crucibles were reweighed. The analysis was performed in duplicate, following AOAC method 950.49, with ash content calculated from the residual weight after ashing.

$$\text{Ash content (\%)} = \frac{(W_3 - W_1)}{W_2} \times 100$$

Where:

M1 is the mass of the porcelain crucible(g)

M2 is the mass of the sample(g)

M3 is mass of porcelain crucible and ash (g).

To analyze reducing sugar content, a sample solution was prepared by mixing 2g of milled roasted cashew kernel with 20 ml of distilled water and 1 ml of sulfuric acid, then boiling under reflux for 3 hours and neutralizing to a pH of 7-8.5 using 40% NaOH. After dilution to 250 ml and filtration, the solution was titrated against boiling Fehling's solution (5 ml each of Fehling A and B), using 2-3 drops of methylene blue as an indicator to identify the brick-red endpoint. A preliminary titration was first performed to estimate the titrant volume, followed by an accurate titration, adding 1.5 ml less than the preliminary volume before boiling and completing the titration with methylene blue until the endpoint was reached, according to AOAC method 923.09; duplicate analyses ensured accuracy. For standardizing the sugar solution, 4.75g of AR grade sugar was dissolved in 50 ml of distilled water in a 500 ml volumetric flask. After adding 5 ml of concentrated HCl and allowing the solution to stand for 24 hours, it was neutralized with 0.1N NaOH using phenolphthalein as an indicator until a pink color appeared. A 25 ml aliquot of this solution was then transferred to a 100 ml volumetric flask and diluted to volume. Titration was performed using 5 ml each of Fehling A and Fehling B with the standardized sugar solution in the burette. Below is the formula for analyzing the sugar content:

$$\text{Reducing sugar (\%)} = \frac{0.0025 \times V_1 \times V_2 \times 100}{V_3 \times W}$$

Where:

V_1 is the volume of standard sugar

V_2 is the volume of preliminary titration

V_3 is the volume of the final titration

W is the mass of the sample.

To analyze fiber content in roasted cashew kernels, the Manual FibreBag system, following AOAC method 978.10, was employed. First, fiber bags and porcelain crucibles were dried at 105°C. A 1g milled roasted cashew kernel sample was sealed in a fiber bag and digested in 360ml of 1.25% sulfuric acid (H_2SO_4) for 30 minutes using a Fibretherm unit,

maintaining a controlled boiling temperature. After rinsing with hot distilled water, the fiber bag was further digested in 360ml of 1.25% NaOH for another 30 minutes. The bag was rinsed again with hot distilled water, dried at 105°C for 1 hour, and weighed. Finally, the fiber bag and crucible were ashed in a furnace at 550°C for 3 hours, and the final weight was recorded to calculate fiber content. The formulation to analyze the fiber content of roasted cashew kernel is shown below:

$$\%\text{Fiber} = \frac{(W_2 - W_1) - (W_4 - W_3)}{W_0} \times 100$$

Where:

W_0 is the mass of sample (g)

W_1 is the mass of the fiber bag before dry (g)

W_2 is the mass of the dried sample with the fiber bag (g)

W_3 is the mass of the porcelain(g)

W_4 is the mass of the dried porcelain crucible and ash (g).

The fat content of roasted cashew kernels was analyzedusing the AOAC 2003.05 method, suitable for oily seeds and nuts. Before analysis, fat glasses were dried for one hour or overnight at 105°C, cooled in a desiccator for 15 minutes, and weighed. A 1g sample of milled roasted cashew kernel was placed into a thimble sealed with a cotton ball, which was then connected to the solvent extractor. Hexane (70 ml) was added to the fat glass, and the extraction was performed at 130°C with 60 minutes of immersion, 60 minutes of washing, and 15 minutes of solvent recovery to evaporate residual hexane. After extraction, fat glasses were dried at 105°C for one hour, cooled in a desiccator, and weighed to calculate fat content based on the weight difference. The fat content calculated with the following formula below:

$$\text{Fat content (\%)} = \frac{(W_3 - W_1)}{W_2} \times 100$$

Where W_1: is mass of empty glass (g)

W_2: is mass of sample (g)

W_3: is mass of glass with extracted fat

Protein content in roasted cashew nuts was determined using the Kjeldahl method (AOAC 2001.11), involving digestion, distillation, and titration. The process began with digesting a 1g roasted cashew kernel sample in a digestion tube with potassium sulfate (3.5g) and copper sulfate pentahydrate (0.1g) using sulfuric acid (20ml) and hydrogen peroxide (5ml) at 420°C for 2 hours, until the sample turned green. After cooling, the digested sample was distilled using a Kjelflex K-360 into an Erlenmeyer flask containing 20ml of 4% boric acid and a few drops of bromocresol green with methyl red indicator. Finally, the solution was titrated with 0.25M sulfuric acid, and the protein content was calculated based on the volume of sulfuric acid used.

$$W_N = \frac{(V_{sample} - V_{blank}) \times Z \times C \times F \times M_N}{m_s \times 10}$$

$\%N = W_N \times 100$

$\%P = \%N \times P_F$

W_N is the weight fraction of nitrogen

V_{sample} is the volume of the blank (ml)

V_{blank} is the volume of the blank factor for H_2SO_4

Z is the molar reaction factor

C is the concentration of the titrant (mol/l)

F is the factor of titrant (0.9969)

M_N is the mass molecular weight of the nitrogen

%N is the percentage of the weight of the nitrogen

%P is the percentage of protein

P_F is the protein factor (5.3 for cashew nut)

The total carbohydrate content was determined using the AOAC 986.25 method, where the percentages of moisture, ash, protein, fat, and fiber were deducted from 100 to obtain the carbohydrate percentage in the sample.

Carbohydrate (%) = 100 – (Moisture+Ash+Protein+Fiber).

3. RESULT AND DISCUSSION

Cashew nut samples from organic and conventional farming systems were processed into powder form and subjected to nutritional analysis. The resulting data allowed for a comprehensive comparison of their macronutrient content, revealing variations attributable to both farming practices (agroecological versus chemical management) and the age of the sampled trees (4-6 year old plantations).

Management	%M	%A	%Fa	%F	%P	%S	%C
Org4	1.71	3.08	46.27	3.51	19.2	2.68	29.76
Inorg4	1.88	2.87	41.32	3.46	18.5	2.52	35.40
Org5	1.88	2.72	46.10	3.39	18.2	2.54	31.13
Inorg5	1.96	2.62	45.88	3.27	18.0	2.53	31.46
Org6	1.62	3.39	49.00	3.17	19.3	2.55	26.68
Inorg6	1.87	2.76	47.17	2.51	18.2	2.45	29.92

*%M=Moisture, %A=%Ash, %Fa=%Fat, %F=%Fiber, %P=%Protein, %S=%Reducing sugar, %C=%Carbohydrates
**Org4= Organic cashew nut aged 4 years, Inorg4= Inorganic cashew nut aged 4 years, Org5= Organic cashew nut aged 5 years, Inorg5= Inorganic cashew nut aged 5 years, Org6= Organic cashew nut aged 6 years, Inorg6= Inorganic cashew nut aged 6 years.

The comparative analysis of roasted cashew nuts under agroecological (organic) and conventional (chemical) management systems reveals significant differences in nutritional composition. Organic samples consistently exhibited higher fat content (46.10–49.00%) compared to conventional samples (41.32–47.17%), aligning with the hypothesis that agroecological practices enhance lipid synthesis, possibly due to improved soil health from nitrogen-fixing cover crops like Stylo(Zoundji et al., 2023). The protein content of cashew nuts was found to be significantly influenced by the farming system, with agroecological (organic) practices yielding higher protein levels compared to conventional (inorganic) methods. Organic cashew nuts consistently exhibited protein concentrations ranging from 18.20% to 19.30%, whereas conventionally grown nuts contained slightly lower levels, between 18.00% and 18.50%. This difference, though modest, suggests that agroecological farming may enhance the nitrogen availability and protein synthesis in cashew kernels (Department of Horticulture, Federal University of Agriculture, Abeokuta, Nigeria et al., 2018). Conversely, carbohydrate content was notably higher in conventional samples (29.92–35.40%) than in organic ones (26.68–31.13%), likely due to rapid nutrient mobilization from chemical fertilizers favoring carbohydrate production (Mohapatra et al., 2016). Fiber content followed a similar trend, with organic samples showing higher values (3.17–3.51% vs. 2.51–3.46%), indicating that agroecological systems may promote structural plant components (Rico et al., 2016). Agroecological cashew nuts exhibited lower moisture levels (1.62%–1.88%) compared to conventionally grown nuts (1.87%–1.97%). This suggests that organic farming may lead to better drying efficiency or reduced water retention, which could enhance shelf stability. Additionally, organic nuts had higher ash content (2.72%–3.39%) than their conventional counterparts (2.62%–2.87%), indicating a richer mineral profile (Stéphane et al., 2020). This difference likely stems from the natural soil enrichment provided by cover crops like Stylo in agroecological systems, which enhance nutrient availability for cashew trees. Additionally, the results underscore the importance of considering farming practices in food production, as they have a tangible impact on the nutritional profile and quality of the final product. This evidence supports the broader adoption of agroecological methods, both for their positive environmental effects and for their ability to produce more nutrient-dense foods.(Olatidoye, n.d.).

4. CONCLUSION

This study reveals that agroecological farming yields roasted cashew nuts with better nutritional quality, containing higher protein, healthy fats, fiber and minerals compared to conventional methods. In contrast, conventional farming produced nuts with higher carbohydrate content, likely due to different nutrient management approaches. These findings demonstrate how sustainable farming practices can enhance both the nutritional value of crops and environmental health. By adopting agroecological methods, farmers could offer more nutrient-dense foods while supporting sustainable agriculture. The results provide valuable insights for consumers, farmers and policymakers seeking to improve food quality through environmentally-friendly practices.

ACKNOWLEDGMENTS

This research was made possible through the support received from the Faculty of Chemical and Food Technology's laboratory at the Institute of Technology of Cambodia as part of the ASSET (Agroecology and Safe Food System Transitions) initiative.

REFERENCES

Awodun, M. A., Osundare, O. T., Oyelekan, S. A., & Okonji, C. J. (2015). Comparative effects of organic and inorganic soil amendments on the growth of cashew nut (Anacardium occidentale l.) seedlings. Journal of Agricultural Biotechnology and Sustainable Development, 7(4), 37–42. https://doi.org/10.5897/JABSD2015.0239

Azeez, K. K., & Olabanji, O. P. (2024). Cashew trees cultivation and management practices in Nigeria. 6(1).

Department of Horticulture, Federal University of Agriculture, Abeokuta, Nigeria, Olubode, O. O., Joseph-Adekunle, T. T., Department of Horticulture, Federal University of Agriculture, Abeokuta, Nigeria, Hammed, L. A., Department of Horticulture, Federal University of Agriculture, Abeokuta, Nigeria, Olaiya, A. O., & Department of Plant Physiology and Crop Production, Federal University of Agriculture, Abeokuta, Nigeria. (2018). Evaluation of production practices and yield enhancing techniques on productivity of cashew (Anacardium occidentale L.). Fruits, 73(2), 75–100. https://doi.org/10.17660/th2018/73.2.1

Griffin, L. E., & Dean, L. L. (2017). Nutrient Composition of Raw, Dry-Roasted, and Skin-On Cashew Nuts. Journal of Food Research, 6(6), 13. https://doi.org/10.5539/jfr.v6n6p13

Mohapatra, A., Dash, D. K., & Tripathy, P. (2016). Comparative analysis of various organic amendments on tree growth and nut yield of cashew (Anacardium occidentale L.). International Journal of Agriculture, Environment and Biotechnology, 9(2), 225. https://doi.org/10.5958/2230-732X.2016.00031.0

Olatidoye, O. P. (n.d.). Effect of temperature and time combinations on colour characteristics, mineral and vitamin content raw and roasted cashew kernel. 554.

Oliveira, N. N., Mothé, C. G., Mothé, M. G., & De Oliveira, L. G. (2020). Cashew nut and cashew apple: A scientific and technological monitoring worldwide review. Journal of Food Science and Technology, 57(1), 12–21. https://doi.org/10.1007/s13197-019-04051-7

Rico, R., Bulló, M., & Salas-Salvadó, J. (2016). Nutritional composition of raw fresh cashew (Anacardium occidentale L.) kernels from different origin. Food Science & Nutrition, 4(2), 329–338. https://doi.org/10.1002/fsn3.294

Soares, D. J., Vasconcelos, P. H. M. D., Camelo, A. L. M., Longhinotti, E., Sousa, P. H. M. D., & Figueiredo, R. W. D. (2013). Prevalent fatty acids in cashew nuts obtained from conventional and organic cultivation in different stages of processing. Food Science and Technology, 33(2), 265–270. https://doi.org/10.1590/S0101-20612013005000050

Stéphane, K. Y., Halbin, K. J., & Charlemagne, N. (2020). Comparative Study of Physical Properties of Cashew Nuts from Three Main Production Areas in Côte d'Ivoire. Agricultural Sciences, 11(12), 1232–1249. https://doi.org/10.4236/as.2020.1112081

Zephirin Somda, N. I., Ugwu, J. A., & Dao, M. C. E. (2023). Comparison of some characteristics of cashew nut production in two agroecological zones of Burkina Faso. Cogent Food & Agriculture, 9(2), 2272489. https://doi.org/10.1080/23311932.2023.2272489

Zoundji, G. C., Vodouhe, F., Houngbèmè, J.-F. D., & Yabi, J. A. (2023). Adoption and Technical Efficiency of Organic and Fair-Trade Cashew Production in Benin (West Africa). In Review. https://doi.org/10.21203/rs.3.rs-2686754/v1

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Determination of Heavy Metals Contamination in Cashew Nuts Collected from Preah Vihear Province

Sreyoun Yan [1], Meymey Heng [2], Marinich Net[1], Peany Houng[1], Kimleang Khoeurn[1*] Florent Tivet[2], Mathide Sester[2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[*]*Corresponding author: khoeurn@itc.edu.kh*

Abstract: *Cashew nut is one of the famous nuts eaten as a snack or used as ingredient in different cuisines around the world. A potential health risk may result from eating of nuts contaminated with heavy metals such as arsenic (As), iron (Fe), and manganese (Mn). This study aims to analyze the levels of heavy metals, including As, Fe, and Mn in cashew nuts collected from organic and conventional treated plantations in Rovieng District, Preah Vihear Province, Cambodia. The heavy metals such as Fe and Mn content were detected by using photometer (AL 400, Germany) and As content was tested by Arsenic test kit. The results showed that As content in cashew nuts are significantly different between cashew nuts grown through organic and conventional treatments. The As levels ranged from 0.0033 to 0.00435 mg/kg in organic-treatment cashew nut, while the conventional-treatment cashew nuts have higher As content from 0.00395 to 0.00425 mg/kg. For Fe content, the levels ranged from 0.31 to 0.66 mg/kg in organic-treatment cashew nuts and ranged from 0.335 to 0.665 mg/kg in conventional-treatment cashew nuts. While, Mn level in cashew nuts varied from 0.025 to 0.435 mg/kg in organic treatment, and from 0.065 to 0.85 mg/kg in conventional treatment. In overall, the levels of As, Mn, and Fe in conventional treatment are higher than in the organic treatment cashew nuts, however, the amount of heavy metals in both organic-treatment and conventional-treatment cashew nuts are still in the range of permissible limits. The findings suggest that the type of fertilizer used in agriculture can influence the accumulation of certain heavy metals in cashew nuts.*

Keywords: Heavy metals; Cashew nut; Conventional treatment; Organic treatment

1. INTRODUCTION

Cashew tree was introduced to Asia and Africa during the sixteenth century from brazil. Cashew nut is one of the significant agricultural crops in Cambodia, as well as, a famous nut eaten as a snack or used as ingredient in different cuisines. Cashew tree were plants in some provinces of Cambodia such as Ratanakiri, Mondoulkiri, Kompong Thom, Kampong Cham, Tbong Khmum, Kompong Chhnang, Kratie, Stung Treng, Siem Reap and Preah Vihear with an estimated 60, 000 tons of raw cashew nuts produced year from about 120,000 hectares of land [1].

There are seven villages (Oh, Bac Dong, Boh , Bhal hal, Kompongcham) located in Rovieng district, Preah Vihear province. Cashew nut is grown using two main method conventional-treatment and organic treatment. For organic treatment, compose, manure, or bone meal are used to nourish the soil and tree. In addition, cashew tree is grown with different cover crop which can protect, reduce soil erosion during fallow period in annual cropping system, and increase soil productivity [2]. Cover crop can take up nutrient, especially nitrogen (N), phosphorus (P) to cashew tree [3]. In contrast, conventional cashew farming utilizes synthetic chemical fertilizers for a rapid nutrient boost. While these fertilizers can increase yields, their overuse can harm the environment through soil degradation, water pollution, and potential health risks [4].

Cashew nuts are type of seed which are source of the protein, healthy fat, dietary fiber and also contain some several minerals such as iron, zinc, magnesium, and copper that good for health [5]. However, a potential health risk such as carcinogenic effects, neurological impairment, and organ poisoning may pose a serious concern by eating of nuts that contaminated with heavy metals [6]. Cashew nuts are susceptible to heavy metal accumulation due to a variety of

1* Corresponding author: Khoeurn Kimleang
E-mail: khoeurn@itc.edu.kh; Tel: +855-86 455 931

factors, including pesticide use, soil contamination, and industrial activity. To guarantee food safety and safeguard the public's health, it is essential to analyze the heavy metal content of Cambodian cashew nuts. Numerous analytical methods, including inductive photometry, atomic absorption spectroscopy, are used to quantify the concentration of heavy metals in cashew nuts [7]. Despite the importance of cashew nuts in Cambodia's agriculture, there is limit research on the analysis of heavy metals, particularly manganese, iron, and arsenic, in these nuts. The lack of comprehensive data on the levels of these metals in Cambodian cashew nuts poses a significant challenge in assessing the potential health risks associated with their consumption. To address this gap, it is essential to conduct a thorough analysis of manganese, iron, and arsenic levels in cashew nuts produced in Rovieng district, Preah Vihear province, Cambodia. This analysis will provide valuable insights into the extent of heavy metal contamination in the country's cashew nut supply and help identify potential sources of contamination. Furthermore, the findings from this research can contribute to the development of regulatory frameworks and guidelines for acceptable levels of heavy metals in cashew nuts, ensuring the safety of consumers and promoting sustainable agricultural practices in Cambodia. The purpose of this study is to determine the concentration of As, Mn, and Fe of 18 samples of cashew nuts collected from the conventional-treatment and organic-treatment located in Rovieng District, Preah Vihear Province.

2. METHODOLOGY

2.1 Study area description

Preah Vihear Province especially Rovieng District is a significant cashew-producing area in Canbodia with around 9,452 hectares of cultivation as of 2019. Its strategic location near Thailand and Laos make it susceptible to potential environmental concentration from cross-border industrial activities and mining, which could lead to heavy metal accumulation in local crop, including cashew nuts. The ecological setting, characterized by the Dangrek Mountains and a tropical climate with distinct rainy seasons, affects soil and water quality, influencing heavy metal uptake by cashew plants. The focus on manganese, iron, and arenic is crucial due to the risks posed by industrial runoff and agricultural practices, including the use of fertilizers. This study on heavy metals in Rovieng District will contribute valuable data for environmental management, public health protection, and policy development, ensuring the sustainability of cashew production in the region.

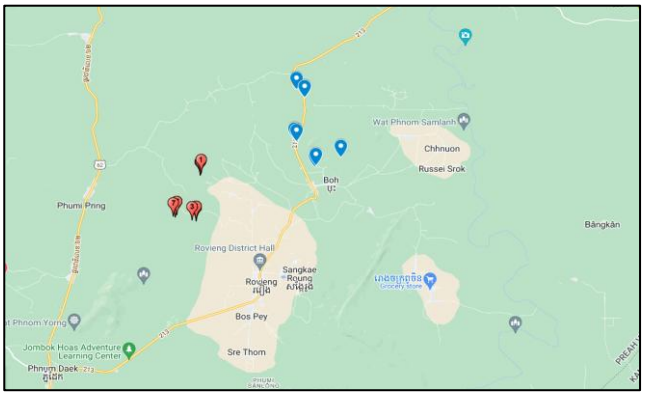


Fig. 1. Accessibility and location and map of the study area.

2.2 Sample collection

The 18 samples of cashew nuts, planting using conventional and organic treatments were collected from six villages such as Oh, Bac kdong, Dong, Boh, Bhal hal, and Kompongcham, located in Rovieng district, Preah Vihear province. The detail information of the cashew nut sample is shown in **Table 1.**

Table 1. Show the name given from farm rename to the code sample use in his report.

No.	Plot code given	Village of owner	Rename code sample	Treatments
1	1-cm-Morm sorm	Oh	Cf1	Chemical
2	1-om-Heng Pea	Oh	Om1	Organic
3	1-ol-Da Ley	Oh	Ol1	Organic
4	3-cm-prom sun	Bac kdong	Cf 3	Chemical
5	3-om-Mey Vanna	Bac kdong	Om3	Organic
6	3-ol-Nart Changly	Bac kdong	Ol3	Organic
7	7-cm-Kem Vannei	Bac kdong	Cf4	Chemical
8	7-om-San Soknorm	Bac kdong	Om4	Organic
9	7-ol-Thiem Rathanak	Bac kdong	Ol4	Organic
10	1-cm-Ny Pheak	Dong	Cf5	Chemical
11	1-om-Kang Senart	Dong	Om5	Organic
12	1-ol-Bun Heng	Dong	Ol5	Organic
13	2-cm-Bor Phern	Kompongcham	Cf2	Chemical
14	2-om-Rouen Saem	Boh	Om2	Organic
15	2-ol-Bun Heng	Dong	Ol2	Organic
16	6-cm-Hea Lanyh	Boh	Cf6	Organic
17	6-om-Bun Bin	Bhal Hal	Om6	Organic
18	6-ol-Sovannaim	Bhal Hal	Ol6	Chemical

2.3 Sample preparation

Samples were firstly dried under the sun for several days to reduce the moisture content. The dried cashew nuts were then stored in an airtight containers and dry place before transporting to the laboratory for quality analysis. Next, the cashew nut was roast in the oven at 140 ℃ for 30 min. The nuts were finally crushed with a mortar and pestle to powder form.

2.4 Digestion and analysis

To prepare the cashew sample for heavy metal analysis, 1 gram of the ground sample was placed in a beaker and mixed with 10 ml of 65% nitric acid (HNO_3). The mixture was heated to 170°C for 15 minutes under a watch glass and in a fume hood with airflow on. After cooling for 10 minutes, 5 ml of nitric acid was added and the mixture was re-heated at 200 to 250°C until the volume reduced to about 5 ml. After cooling, 3 ml of 30% of hydrogen peroxide (H_2O_2) and 2 ml of distilled water were added, followed by re-heating at 250°C for 15 minutes by covering with watch glass. Then, 10 ml of 30% of hydrogen peroxide was added until no more white fumes appeared. In this stage, 1ml of 30% H_2O_2 was added if solution was not digest well and heat until the effervescence noticeable. After that, the solution was then cooled, filtered through a 0.45 µm syringe filter, and diluted to 100 ml with distilled water, making it ready for heavy metal determination [8, 9].

2.5 Sample analysis

Manganese (Mn) and iron (Fe) concentration were determined by photometer (AL 400, Germany). The detection limit of the machine for Manganese is 0.01-0.7 mg/l and the detection limit of iron is 0.02-3 mg/l. For Arsenic (As) analysis was done by arsenic test kit with the detection limit range from 0-0.5 mg/l.

2.6 Statistical analysis

Statistical analyses were performed using One-way analysis of variance (ANOVA). Tukey multiple-range test was used as a post hoc test to investigate significant differences ($p < 0.05$). Data are described as mean ± standard deviation (SD) from duplicate of each experiment.

3. RESULTS AND DISCUSSION

3.1 Determination of Arsenic

Previous studies have found detectable levels of arsenic in cashew nuts are generally low. One study reported that arsenic content in cashew nuts ranging from 0.01 to 0.06 mg/kg [10]. These levels are within the permissible limits set by food safety regulations in many countries.

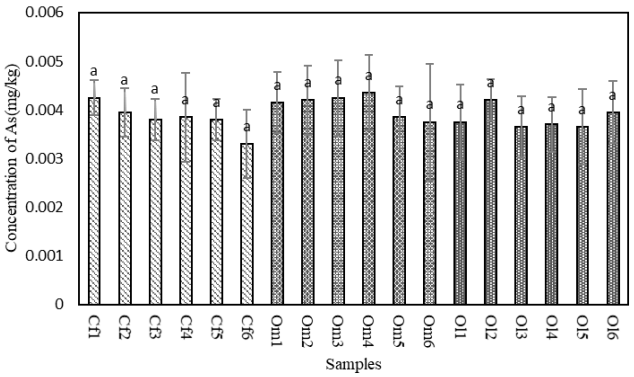


Fig. 2. Result of As concentration of cashew nuts. Graph bars with different superscript letters (a, b) differ significantly from each other ($p < 0.05$).

Fig. 2 presents the arsenic (As) testing results from 18 cashew nut samples collected across six villages in Rovieng district, Preah Vihear province. The range of Arsenic in organic-treatment from 0.0033 to 0.00435 mg/kg, while in conventional-treatment range from 0.00395 to 0.00425 mg/kg. Arsenic Levels varies depending on village, fertilizer, and cashew variety. Oh village (Cf1), which uses the M23 type and chemical fertilizer (Cf), has the highest level ever measured (0.0045 mg/kg). There is some heterogeneity in the samples from Bac Kdong village, with the native variety and organic fertilizer (Ol) exhibiting the lowest quantity (0.0033 mg/kg). As was previously mentioned, varied concentrations of arsenic are found naturally in soil. Although future research may need to take soil samples from every town to see whether there is a relationship between soil arsenic and nut arsenic concentration. The information points to a potential connection between increased levels of arsenic and chemical fertilizers (Cf). The cashew plants may acquire trace levels of arsenic from chemical fertilizers that are present as contaminants. The data suggests a possible link between elevated arsenic levels and chemical fertilizers (Cf) [11]. The arsenic levels given are fictitious. The observed patterns may be the result of random variation or other factors not taken into account. The analysis would be strengthened by additional testing with a bigger sample size. Test kits for arsenic offered significant insights into possible fluctuations in cashew nut arsenic levels. The research raises the possibility that chemical fertilizers are linked to elevated levels of arsenic, and it also raises the possibility that cashew nut variety may have an impact. It is advised to conduct additional research, such as soil testing and larger cashew nut sample analysis, to have a more thorough grasp of these variables. If required, cashew nut growers can minimize the amount of arsenic in their products by putting such methods into place.

3.2 Manganese

The analysis determined the concentration of manganese (1.95mg/kg) in cashew nuts. This value will be presented and compared to existing data on manganese content in cashews nuts, in different management.

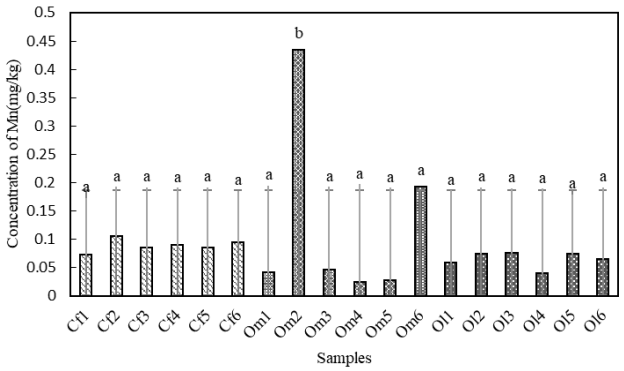


Fig. 3. Result of Mn concentration of cashew nuts. Graph bars with different superscript letters (a, b) differ significantly from each other ($p < 0.05$).

Fig. 3. Demonstrate the manganese (Mn) testing results from 18 cashew nut samples collected across six villages in Rovieng district, Preah Vihear province. The range of Manganese in organic-treatment from 0.025 to 0.435 mg/kg, while in conventional-treatment range from 0.065 to 0.85 mg/kg. There are several reasons for the differences in manganese content between the cashew nut samples. M23 and local were the two distinct types used to grow the cashew nuts. The variations in genetic makeup across the cultivars might affect how well they absorb and accumulate manganese from the soil. In addition, fertilizers either chemical or organic that were used to grow cashew trees can also a source causing manganese contamination to sample. The availability and uptake of manganese by cashew plants can be affected by the fertilizers' chemical makeup and rate of application. The manganese concentration of cashew nuts might vary depending on the soil in each location and water condition. Different levels of manganese uptake by the cashew trees can result from variations in soil composition, water sources, and agricultural practices [12]. The Cf1 sample (0.665 mg/kg) from Oh village, which was grown using chemical fertilizer and the M23 variety, had the greatest manganese concentration. The Om2 sample (0.31 mg/kg) from Boh village, which was grown using organic fertilizer and the M23 variety, had the lowest concentration. These differences finding of manganese in all samples can be ascribed to a confluence of regional, environmental, and fertilizer factors [13] .

3.3 Iron

The presence of heavy metals, including iron (Fe), in cashew nuts is an important food safety consideration. Cashew nuts can potentially accumulate various heavy metals from the environment, such as through soil contamination or the use of certain fertilizers. Understanding the levels of heavy metals in cashew nuts is crucial to ensure the safety and quality of this popular tree nut.

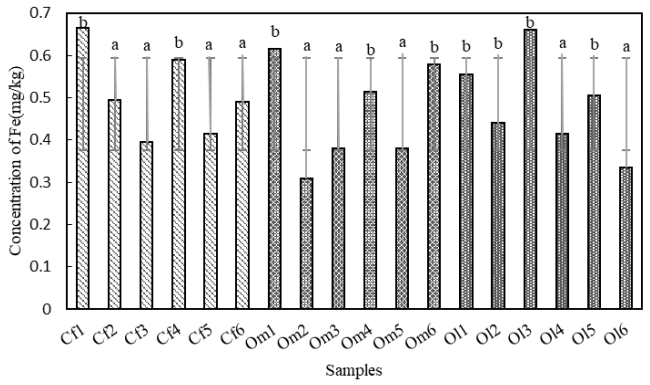


Fig. 4. Result of Fe concentration of cashew nuts. Graph bars with different superscript letters (a, b) differ significantly from each other ($p < 0.05$).

Fig. 4. illustrates the iron (Fe) testing results from 18 cashew nut samples collected across six villages in Rovieng district, Preah Vihear province. The range of Arsenic in organic-treatment from 0.31 to 0.66 mg/kg, while in conventional-treatment range from 0.335 to 0.665 mg/kg. Antioxidant minerals such as iron are abundant in cashew nuts. Iron is a necessary micronutrient that the body uses for a variety of processes, including oxygen transport and cellular function regulation. It was discovered that the average iron concentration of cashew nuts grown both conventionally and organically ranged from 11.0 to 21.4 mg/kg. There are multiple reasons for the differences in iron concentrations found in cashew nut samples taken from various villages in the Rovieng district. M23 and local were the two distinct types used to grow the cashew nuts. The variations in genetic makeup across the cultivars might affect how well they absorb and accumulate manganese from the soil. Fertilizers, either chemical or organic, were used to grow cashew trees. Fertilizer application rates and chemical composition can affect the availability and uptake of Iron by the cashew trees. The Iron concentration of cashew nuts might vary depending on the village's soil and water condition. Different levels of Iron uptake by the cashew trees can result from variations in soil composition, water sources, and agricultural practices. The concentration of iron in samples can cause by several reasons such as type of fertilizer, location and environmental may also influence the iron content in cashew nuts [14].

4. CONCLUSIONS

In conclusion, this study investigated heavy metal content in cashew nuts collected from organic and conventional farming in Rovieng district, Preah Vihear province to measure the levels of manganese (Mn), iron (Fe), and arsenic (As). The results revealed the differences in heavy metals levels between organic and conventional-treatment of cashew nuts sample, with the highest concentration was 0.094 mg/kg for Mn, 0.66 mg/kg for Fe, and 0.615 mg/kg for Arsenic. These values fall within the safety limits for human consumption. While the current study established the levels of Mn, Fe, and As, further research could delve into identifying the specific sources of these heavy metals in both organic and conventional farming systems in the Rovieng district. This could involve analyzing soil, water sources, and fertilizers used in both farming practices to pinpoint potential contamination pathways.

ACKNOWLEDGMENTS

The project was funded by the European Union and the French Development Agency.

REFERENCES

[1] Cambodian Organic Agriculture Association, "Organic Agriculture and Food Processing in Cambodia Status and Potentials," no. April, 2011.

[2] T. Roberts, "Benefits provided by pigeonpea".

[3] T. C. Kaspar and J. W. Singer, "The Use of Cover Crops to Manage Soil," *Soil Manag. Build. a Stable Base Agric.*, pp. 321–337, 2015, doi: 10.2136/2011.soilmanagement.c21.

[4] L. N. Tayoh, "No Title Ecological and Practical Applications for Sustainable Agriculture," *Ecol. Pract. Appl. Sustain. Agric.*, 2020, doi: https://doi.org/10.1007/978-981-15-3372-3_3.

[5] EuroCham Cambodia, "Sourcing from Cambodia - Phnom Penh," 2021, [Online]. Available: https://www.eurocham-cambodia.org/uploads/6dfd3-sourcing-from-cambodia_light-manufacturing-latest.pdf

[6] L. S. Faculty and P. Sciences, *Levels of persistent organic compounds and heavy metals in farmed and wild fish from Zambia and Tanzania , and assessment of possible health effects*. 2023.

[7] H. S. Manzoor *et al.*, "Effect of microwave roasting and storage on the extent of heavy metals present in dry fruits," *Int. J. Chem. Biochem. Sci.*, vol. 3, no.

2013, pp. 74–82, 2013.
[8] N. W. Alcock, “A hydrogen-peroxide digestion system for tissue trace-metal analysis,” *Biol. Trace Elem. Res.*, vol. 13, no. 1, pp. 363–370, 1987, doi: 10.1007/BF02796647.
[9] C. A. Bizzi, E. L. M. Flores, J. A. Nóbrega, J. S. S. Oliveira, L. Schmidt, and S. R. Mortari, “Evaluation of a digestion procedure based on the use of diluted nitric acid solutions and H2O2 for the multielement determination of whole milk powder and bovine liver by ICP-based techniques,” *J. Anal. At. Spectrom.*, vol. 29, no. 2, pp. 332–338, 2014, doi: 10.1039/c3ja50330e.
[10] T. Nguyen-Quang *et al.*, “Classification of Vietnamese Cashew Nut (Anacardium occidentale L.) Products Using Statistical Algorithms Based on ICP/MS Data: A Study of Food Categorization,” *J. Anal. Methods Chem.*, vol. 2023, 2023, doi: 10.1155/2023/1465773.
[11] F. B. Sugar *et al.*, “Common / Popular Tests Allergology”.
[12] Minnesota Department of Health, “Manganese in Drinking Water,” *Manganese Drink. Water*, no. 49, p. 116, 2016.
[13] O. S. O. Ankabi *et al.*, “Growth response of cashew seedlings to applied organic fertilizers of animal and plant origin on a degraded soil in Ibadan , South-West , Nigeria,” *Int. Res. J. Agric. Sci. Soil Sci.*, vol. 3, no. 3, pp. 93–98, 2013.
[14] D. J. Soares *et al.*, “Mineral content, based in the Recommended Daily Intake, in cashew nut obtained from conventional and organic cultivation in different stages of processing,” *Semin. Agrar.*, vol. 33, no. 5, pp. 1869–1876, 2012, doi: 10.5433/1679-0359.2012v33n5p1869.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Nutritional analysis of Cassava varieties KM325, KU50, Rayong 4, Rayong 72 and Rayong 90 in Preah Vihear Province.

Seangleng Thong [1], Somphors Tot [1], Peany Houng [1], Florent Tivet[2], Mathilde Sester[2], Sokuntheary Theng[1*]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2]*CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[*]*Corresponding author: theng.sokuntheary@itc.edu.kh*

Abstract: *Cassava (Manihot esculenta Crantz) is a vital crop in tropical countries, serving as a major source of dietary energy and a promising raw material for industrial starch production. In Cambodia, cassava is a key agricultural commodity, ranking second only to rice in term of production volume. The Cassava sector in Cambodia offers significant opportunities for improving food security, rural livelihoods, and bioenergy development. This study aimed to evaluate the nutritional composition of five cassava varieties commonly cultivated in Cambodia: KM325, KU50, Rayong 4, Rayong 72, and Rayong 90 from four different farms in Phear Vhihea Province. Nutritional parameters including moisture content, ash, fiber, reducing sugars, carbohydrates, and protein were analyzed using standard AOAC methods, while fat content was determined using Soxhlet extraction. Moisture content ranged from 7.35% to 9.70%, with lower moisture levels favoring extended shelf life and improved processing suitability. Ash content, indicative of total mineral content, varied between 0.75% and 2.23%, with the KM325^{S} variety exhibiting the highest ash concentration. Fat content remained low across all samples (0.20%–2.11%), consistent with cassava's known macronutrient profile. Fiber content ranged from 1.57% to 2.49%, with RY90^{A} demonstrating the highest level, suggesting potential for enhanced dietary fiber intake. Protein content was relatively low but varied considerably, ranging from 0.78% to 3.34%; the RY2^{R} variety showed the highest protein concentration, which may contribute to improved nutritional value despite cassava's classification as a carbohydrate-dominant crop. Reducing sugar content remained relatively stable (2.46%–2.79%), supporting the potential application of these varieties in food processing and fermentation industries. Carbohydrate content was high across all varieties, ranging from 83.90% to 88.09%, reinforcing cassava's role as a key energy source. These findings highlight the influence of varietal and environmental factors on cassava's nutritional composition. Varieties such as RY72^{R}, characterized by low moisture content and elevated levels of protein and fiber, emerge as promising candidates for industrial processing and nutritional enhancement. Overall, the results affirm cassava's ongoing importance in food security and energy provisioning, while highlighting specific varieties that may deliver added functional or nutritional benefits.*

Keywords: Cassava variety; nutritional value; Carbohydrate; moisture; protein.

1. INTRODUCTION

Cassava is a key agricultural commodity in Cambodia, ranking second only to rice in terms of production and consumption. The cassava sector in Cambodia offers significant opportunities for improving food security, rural livelihoods, and bioenergy development. However, the sector faces several post-harvest challenges that affect its quality and profitability. One of the main issues is the rapid deterioration of cassava roots after detachment from the plant. Due to limited processing capacity (often at small or medium scale), Cassava produced is observed to be exported at low-value such as chips and starch and in which restricts the potential for value addition and diversification of the industry (Cambodia Development Resource Institute, 2024).

Cassava (Manihot esculenta Crantz) is a drought tolerant, staple food crop grown in tropical and subtropical areas where many people are afflicted with under nutrition and thus, it could be used as potentially valuable food source for developing countries. It is a perennial woody shrub extensively cultivated as an annual crop for its edible starchy tuberous root[1]. Currently, cassava is commonly cultivated without the use of fertilizers, because it is considered well-adapted to low-fertility soils, thereby creating the false impression that cassava does neither require nor respond to fertilizer[2]. Cassava root is a nutritionally rich food,

Corresponding author: Sokuntheary Theng
E-mail: theng.sokuntheary@itc.edu.kh; Tel: +855-87 848 181

primarily composed of carbohydrates. Its carbohydrate content ranges from 32%-35% on a fresh weight basis and 80%-90% on a dry weight basis. Approximately 80% of the carbohydrates in cassava are in the form of starch. One notable characteristic of cassava-based meals is their ability to provide a prolonged feeling of fullness. This character can be attributed to the types and amount of starch present in the flour. Starch (SDS), and resistant starch (RS). RDS is quickly broken down by digestive enzymes, resulting in a gradual and sustained release of glucose into the bloodstream. Resistant starch, as the name suggests, resists digestion in the small intestine and passes into the large intestine. It acts as a prebiotic, stimulating the growth of beneficial bacteria in the gut, which contributes to improved digestive health and a reduced risk of colon cancer. These various proportions of starch in cassava flour are important because of the role it plays on human health. The nutrient composition of the plant differs based on its variety and age of the harvested crop, and soil conditions, climate, and so on. The main aim of this seminar is to review the nutritional value of Cassava and improving its nutrient content through bio-fortification [3].

The nutritional value of cassava roots is important because they are the main part of the plant consumed in developing countries. However, there is much variation in the nutrient quality of the cassava root depends on the several factors, such as geographic location, variety, age of the plant, and environmental conditions[4], [5]. Protein content in dried cassava roots or flour is low, that depending on the form (flour, peels, or leaves), with the root itself closer to 1-3% crude protein[6].

2. METHODOLOGY

2.1 Sample Preparation

Cassava was collected from 4 different farms in Preah Vihear province. All cassava was dried by sunlight for a week before arriving at the Institute of Technology of Cambodia. After cassava arrived at the Institute of Technology of Cambodia, it was separated into zip lock for grinding to powder.

2.2 Moisture Analysis

Firstly, Aluminum plate was placed 105℃ for 1h in oven and then transferred it into desiccator for 15 mins to cool down before weighed(M_1). After that weight the sample 2g and then put it into aluminum plate(M_2) and dry at 105℃ for 3h. After 3h sample were put into desiccator to cool down for 15min at room temperature and note mass after weighted it(M_3).

The moisture content calculated with the following formula below:

$$\text{Moisture content (\%)} = \frac{M_2-(M_3-M_1)}{M_2} \times 100 \qquad \text{(Eq.1)}$$

Where M_1: is mass of empty aluminum plate (g)
M_2: is mass of sample (g)
M_3: is mass of dried sample and aluminum plate after drying (g).

2.3 Ash Analysis

Porcelain cups were dried in oven at 105℃ for 1h. In addition, it was cooled down in a desiccator for 15mins and weighed it(W_1). Then weight 2g of sample put into porcelain cup (w_2) and placed in a furnace machine and heated for 12h at 550℃. After the heating finished, sample in porcelain cup was put into desiccator for 15mins and weighed it (W_3).

The ash content calculated with the following formula below:

$$\text{Ash content (\%)} = \frac{(W_3-W_1)}{W_2} \times 100 \qquad \text{(Eq. 2)}$$

Where W_1: is mass of empty porcelain cup (g)
W_2: is mass of sample (g)
W_3: is mass of dried sample and porcelain cup after heat (g).

2.4 Reducing Sugar Analysis

AOAC Lane and Eynon Method (923.09) is used to determine the reducing sugar content in a sample through titration, based on the principle that reducing sugars reduce Fehling's solution. To prepare the standard sugar solution, 4.75 g of AR grade sucrose is hydrolyzed by adding concentrated hydrochloric acid and standing for 24 hours. The solution is then neutralized using sodium hydroxide with phenolphthalein as an indicator and diluted to a final volume of 500 ml. A 25 ml aliquot is further diluted to 100 ml, giving a solution where 1 ml is equivalent to 2.5 mg of invert sugar. For the sample, 1 g is weighed and boiled with sulfuric acid under reflux for 3 hours, then cooled, neutralized to a pH of 7–8.5, and diluted to 250 ml before filtration. Fehling's solution is prepared by mixing Fehling A (copper sulfate) and Fehling B (alkaline tartrate) in equal volumes. In the preliminary titration, the standard sugar solution is added to the boiling Fehling's mixture with methylene blue indicator until the blue color disappears, indicating the endpoint. This volume is used to estimate the expected titration range. The final titration is performed using the prepared sample solution in the same manner, and the test is done in duplicate.

The reducing sugar content is calculate using the formula:

$$\text{Reducing sugar (\%)} = \frac{0.0025\times V_1\times V_2\times 100}{V_3\times W} \qquad \text{(Eq.3)}$$

2.5 Fat Analysis

VELP extraction cup were cleaned and dried in oven at 105℃ for 1h, the cooled down in the desiccator for 15mins

and weighed it (W_1). Then 1g of sample was weighed in cellulose thimble (W_2) and 70ml of n-hexane was pure into the fat glass for extraction of the hydrolyzed sample at 130℃ for 60mins by crucible immersed in boiling solvent. Then, it was refluxed and washed for 60mins and recovered for 15mins by Soxhlet machine. After that the fat glass was placed into oven at 105℃ for 1h and transferred into desiccator for 15mins and then weighed it (W_3).

The fat content calculated with the following formula below:

$$\text{Fat content (\%)} = \frac{(W_3 - W_1)}{W_2} \times 100 \quad \text{(Eq. 4)}$$

Where W_1: is mass of empty glass (g)
W_2: is mass of sample (g)
W_3: is mass of glass with extracted fat

2.6 Protein Analysis

The protein content was analyzed by using kjeldahl machine. The sample was digested by sulfuric acid to breakdown organically bonds nitrogen into ammonium ions., followed by neutralization distillation, and titration. The nitrogen content was then multiplied by a conversion factor (typically 6.25) to estimate protein content.

The protein is calculate using the formula:

$$\%N = \frac{\left(V_{sample} - V_{blank}\right) \times Z \times C \times F \times M_N}{m_s \times 10}$$

$$\%P = \%N \times P_F \quad \text{(Eq.5)}$$

Where W_N is the weight fraction of nitrogen
V_{sample} is the volume of the blank (ml)
V_{blank} is the volume of the blank factor for H_2SO_4
Z is the molar reaction factor
C is the concentration of the titrant (mol/l)
F is the factor of titrant (0.9969)
M_N is the mass molecular weight of the nitrogen
%N is the percentage of the weight of the nitrogen
%P is the percentage of protein
P_F is the protein factor (6.25 for cassava)

2.7 Fiber Analysis

Crude fiber content was determined using the AOAC method. The sample was digested with acid (H_2SO_4 1.25%) and alkali (NaOH 1.25%). The undigested residue was then filtered, dried in an oven at a specific temperature to obtain a constant weight, and finally ashed in a muffle furnace. The loss in weight after ashing represented the crude fiber content.

The fiber content calculated with the following formula below:

$$\%Fiber = \frac{(W_2 - W_1) - (W_4 - W_3)}{W_0} \times 100 \quad \text{(Eq.6)}$$

Where W_0 is the mass of sample (g)
W_1 is the mass of fiber bag (g)
W_2 is the mass of dried sample with fiber bag (g)
W_3 is the mass of the crucible (g)
W_4 is the mass of the crucible and ash (g)

2.8 Carbohydrates Analysis

Carbohydrate content was calculated by difference. The percentages of moisture, ash, protein, fat and fiber were subtracted from 100 to determine the total carbohydrate percentage in the sample. The formula used was:

$$\text{Carbohydrate (\%)} = 100 - (\text{Moisture} + \text{Ash} + \text{Protein} + \text{Fiber}). \quad \text{(Eq. 7)}$$

3. RESULTS AND DISCUSSION

Dried cassava that collects from Phear Vihea province in four different farmers including Am poeum(A), Kong Kuy(K), Rous Ran(R), and Sem Sokur (S) was analyzed on nutritional value that showed in Table 1.

Table 1. Nutritional value of four differences farm in Phear Vinhea

Parameter	%M	%A	%Fa	%F	%P	%S	%C
$KM325^A$	9.05	2.05	0.46	2.17	1.69	2.59	86.72
$KU50^A$	8.99	1.98	0.61	2.34	2.15	2.56	86.26
$RY4^A$	7.54	1.76	0.71	2.05	2.38	2.51	87.58
$RY72^A$	9.18	2.17	0.60	1.89	1.84	2.69	86.19
$RY90^A$	7.62	1.71	0.72	2.49	2.14	2.57	87.79
$KM325^K$	8.59	1.48	0.44	2.05	1.92	2.51	87.69
$KU50^K$	8.41	0.87	0.76	2.19	1.96	2.56	88.04
$RY4^K$	8.66	0.91	0.36	2.65	1.96	2.46	88.09
$RY72^K$	8.60	1.33	0.71	2.32	2.35	2.59	86.99
$RY90^K$	8.83	1.00	0.58	2.02	1.60	2.55	87.97
$KM325^R$	8.88	1.92	0.81	1.85	2.51	2.62	84.72
$KU50^R$	8.62	1.71	1.12	1.80	2.31	2.70	85.55
$RY4^R$	8.51	1.60	0.58	0.94	3.20	2.79	85.05
$RY72^R$	7.35	1.48	0.45	2.03	3.34	2.52	86.48
$RY90^R$	8.66	0.75	2.11	1.67	2.53	2.66	83.90
$KM325^S$	9.17	2.23	0.71	1.95	0.94	2.61	86.92
$KU50^S$	9.01	1.93	0.57	1.57	1.00	2.57	87.46
$RY4^S$	9.70	1.35	0.82	2.18	0.78	2.51	87.34
$RY72^S$	9.30	1.71	0.20	1.79	1.17	2.64	87.74
$RY90^S$	8.97	1.53	0.50	1.93	1.23	2.56	87.74

*%M=Moisture, %A=%Ash, %Fa=%Fat, %F=%Fiber, %P=%Protein, %S=%reducing sugar, %C=%Carbohydrates
**$KM325^A$= cassava KM325 collect from Am Poeum (A=Am Poeum, K=Kong Kuy, R=Rous Ran, S=Sem Sokur)

The nutritional composition of cassava sample shows significant variation among the different varieties tested, suggesting that genetic and environment factors likely influence nutrient levels. Moisture content ranged from 7.35% in $RY72^R$ to 9.70% in $RY4^S$. Moisture is a critical factor because it determines cassava's shelf life and susceptibility to microbial spoilage. Lower moisture content is preferable for longer storage and for processing into flour and chips. Fresh cassava roots typically contain 60-70% moisture [13]; but after sun drying, moisture content drops considerably to 6.55%±0.07 [7]. The values observed here

reflect processed or partially dried cassava and suggest that $RY72^R$ may be more suitable for industrial application due to its lower moisture content. Cassava roots with lower moisture are preferred for flour and starch processing, as they improve drying efficiency and product stability[8]. Ash refers to the inorganic residue remaining after ignition or complete oxidation of organic matter, it ranges from 0.75% to 2.23% in $RY72^R$ and $KM325^S$. While higher ash content generally indicates a greater presence of minerals, which can be beneficial nutritionally, it is also important to consider the specific minerals present, as not all inorganic residues are beneficial[9]. The elevated ash levels in $KM325^S$ and $KM325^A$ may point to richer soil mineral composition or improved nutrient absorption[10]. Fat content across the samples remained low, in line with cassava's naturally low lipid content. Values ranged from 0.20% in $RY72^S$ to 2.11% in $RY90^R$. The unexpectedly high fat content in $RY90^R$ may suggest external lipid contamination during handing or specific post-harvest practices such as sun-drying on contaminated surfaces[11]. Fiber content, important for digestive health and functional food formulations, ranged from 1.57% in $KU50^S$ to 2.49% in $RY90^A$. High dietary fiber in cassava helps lower cholesterol, stabilize blood sugar, and enhance gut health. Cassava's fiber content is highly influenced by variety and processing, typically ranging between 0.5-3.5%. The elevated fiber content in $RY72^R$ makes it a promising candidate for producing high-fiber snacks or flour blends for health-conscious consumers[12]. Protein content in cassava is generally low, which is a concern in populations where cassava is a primary calorie source. Protein levels in the study ranged from 0.78% in $RY4^S$ to 3.34% in $RY2^R$. Depending on cultivars, the protein content in cassava storage roots ranges from less than 1% to 55% (dry weight)[13]. Although the values are within this range, they remain inadequate to meet daily protein requirements, especially for children and pregnant woman. The higher protein levels in $RY72^R$ may suggest it is better suited for consumption in protein deficient regions or could be blended with legumes to improve protein balance. Reducing sugar content was relatively stable, ranging from 2.46% in $RY4^K$ to 2.79% in $RY4^R$. Reducing sugars such as glucose and fructose play a key role in flavor, fermentation, and browning reaction during cooking. The sugar content in cassava can influence its suitability for ethanol production, animal feed, and snack food processing. High reducing sugar content, as seen in $RY4^R$, can enhance flavor and contribute positively to fermentation processes in biofuel production or traditional fermented foods like *kanji* or *lafun*[14]. Carbohydrate content, the main nutritional asset of cassava, ranged from 83.90% in $RY90^R$ to 88.09% in $RY4^K$. Cassava is known for its high states starch content, making it an essential energy source in tropical regions. Cassava typically contains 80-90% carbohydrates (dry basis), primarily in the form of starch. These results reaffirm cassava's value as an energy-dense food, suitable for both direct consumption and industrial starch extraction. The high carbohydrates levels in $RY4^K$ may favor its use in baked goods, flour production, or processed foods[15].

4. CONCLUSIONS

The nutritional composition analysis of various cassava varieties highlights the significant influence of both genetic and environmental factors on their nutrient profiles. Varieties such as $RY72^R$, characterized by lower moisture and higher fiber and protein contents, are particularly well-suited for industrial processing and can contribute to nutritional enhancement in protein-deficient regions. In contrast, $KM325^S$ and $KM325^A$, with their higher ash content, may offer improved mineral nutrition. The relatively consistent levels of reducing sugars across all samples indicate a stable potential for flavor development. High carbohydrate content reaffirms cassava's role as a key energy source, with varieties like $RY4^K$ being especially favorable for flour production and other processed food applications. Strategic selection of cassava varieties based on their distinct nutritional attributes can enhance their utility in food security, industrial uses, and the development of health-oriented products.

ACKNOWLEDGMENTS

The authors acknowledge the support of the laboratory of Faculty of Chemical and Food technology of Institute of Technology of Cambodia under project ASSET(Agroecology and Safe Food System Transitions).

REFERENCES

[1] A. Bayata, "Review on Nutritional Value of Cassava for Use as a Staple Food," *Sci. J. Anal. Chem.*, vol. 7, no. 4, p. 83, 2019, doi: 10.11648/j.sjac.20190704.12.

[2] G. O. Agbaje and T. A. Akinlosotu, "Influence of NPK fertilizer on tuber yield of early and late-planted cassava in a forest alfisol of south-western Nigeria," *Afr. J. Biotechnol.*, vol. 3, no. 10, pp. 547–551, Oct. 2004, doi: 10.5897/AJB2004.000-2107.

[3] M. Wambua, R. Mulwa, J. M. Nduko, and J. Matofari, "Nutritional Composition, Antinutritive Compounds, and Starch Properties of Flour from Cassava Varieties Grown in Nakuru County, Kenya," Dec. 01, 2023, *In Review*. doi: 10.21203/rs.3.rs-3606733/v1.

[4] A. Bayata, "Review on Nutritional Value of Cassava for Use as a Staple Food," *Sci. J. Anal. Chem.*, vol. 7, no. 4, p. 83, 2019, doi: 10.11648/j.sjac.20190704.12.

[5] H. M., H. L., J.-S. L., A. H. Mansoor, M. H. A. Jahurul, and M. K. Zainol, "Physicochemical and functional properties of cassava flour grown in different locations in Sabah, Malaysia," *Food Res.*, vol. 4, no. 4, pp. 991–999, Mar. 2020, doi: 10.26656/fr.2017.4(4).405.

[6] N. K. Morgan and M. Choct, “Cassava: Nutrient composition and nutritive value in poultry diets,” *Anim. Nutr.*, vol. 2, no. 4, pp. 253–261, Dec. 2016, doi: 10.1016/j.aninu.2016.08.010.
[7] E. A. Nainggolan, J. Banout, and K. Urbanova, “Recent Trends in the Pre-Drying, Drying, and Post-Drying Processes for Cassava Tuber: A Review,” *Foods*, vol. 13, no. 11, p. 1778, Jun. 2024, doi: 10.3390/foods13111778.
[8] M. Priska, N. Peni, L. Carvallo, and V. P. S. M. Wae, “Effect of Drying Techniques on Nutrition Content and Gross Energy of Nuabosi Cassava (Manihot esculenta) Flour,” *Biosaintifika J. Biol. Biol. Educ.*, vol. 12, no. 2, pp. 220–225, Aug. 2020, doi: 10.15294/biosaintifika.v12i2.24184.
[9] A. Monti, N. Di Virgilio, and G. Venturi, “Mineral composition and ash content of six major energy crops,” *Biomass Bioenergy*, vol. 32, no. 3, pp. 216–223, Mar. 2008, doi: 10.1016/j.biombioe.2007.09.012.
[10] B. I. Kayode, R. M. O. Kayode, O. A. Abiodun, C. Nwosu, O. R. Karim, and S. A. Oyeyinka, “Chemical, functional and physicochemical properties of flour from cassava stored under freezing,” *J. Stored Prod. Res.*, vol. 92, p. 101816, May 2021, doi: 10.1016/j.jspr.2021.101816.
[11] E. I. Adeyeye, A. A. Oshodi, and K. O. Ipinmoroti, “Functional properties of some varieties of African yam bean (Sphenostylis stenocarpa) flour II,” *Int. J. Food Sci. Nutr.*, vol. 45, no. 2, pp. 115–126, Jan. 1994, doi: 10.3109/09637489409166150.
[12] A. S. Olanrewaju, “EFFECT OF GARIFICATION (ROASTING) DURATION ON THE QUALITY CHARACTERISTICS OF CASSAVA GARI,” vol. 17, no. 2, 2016.
[13] M. Stupak, H. Vanderschuren, W. Gruissem, and P. Zhang, “Biotechnological approaches to cassava protein improvement,” *Trends Food Sci. Technol.*, vol. 17, no. 12, pp. 634–641, Dec. 2006, doi: 10.1016/j.tifs.2006.06.004.
[14] B. Lukuyu, I. Okike, A. J. Duncan, M. Beveridge, and M. Blummel, “Use of cassava in livestock and aquaculture feeding programs”.
[15] R. Howeler, N. Lutaladio, and G. Thomas, *Save and grow: cassava a guide to sustainable production intensification*. Rome: FAO, 2013.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Determination of Nutritional Composition in Rice Growth through Agroecological and Conventional Practice in Preah Vihear Province.

Sonyta Sek[1], Pisey Lim [1], Peany Houng[1], Florent Tivet[2], Mathilde Sester[2], Sokuntheary Theng [1*]

[1]*Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2]*CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[*]*Corresponding author: theng.sokuntheary@itc.edu.kh*

Abstract: *Rice is a widely cultivated cereal grain and serves as a staple food for over half of the world's population, providing essential nutrients. This study compares the nutritional composition of rice grown under agroecological and conventional farming practices in Preah Vihear province, focusing on three rice varieties: White Phka Rumdoul, Red Phka Rumdoul, and Kroches. Key nutritional parameters analyzed include moisture, ash, fiber, fat, protein, carbohydrates, and reducing sugars to evaluate the impact of cultivation methods. The results showed that conventionally grown rice consistently exhibited higher moisture content, particularly in WPR CT (11.6 ± 1.27%) and RPR CT (10.5 ± 1.53%), which may impact post-harvest storage and shelf stability. In contrast, agroecological practices led to significantly higher ash (KC GM: 1.85 ± 1.78%), fat (RPR GM: 2.57 ± 0.29%), and fiber (RPR GM: 1.04 ± 0.13%) levels, reflecting improved soil health and slower, more natural growth conditions Carbohydrate content was also slightly elevated under agroecological management across most varieties, with WPR GM reaching 81.6 ± 1.46%. Protein content showed variation by variety and treatment: conventional WPR and RPR had slightly higher protein, whereas KC GM had the highest overall protein value (7.16 ± 0.19%). Reducing sugar content remained relatively constant across both systems, suggesting a stronger genetic influence. These findings suggest that agroecological farming can enhance the nutritional density of rice, particularly for fat, fiber, and mineral content, supporting its potential as a sustainable and health-promoting cultivation method.*

Keywords: Rice; Conventional Farming; Agroecological Farming; Carbohydrate; Nutritional composition.

1. INTRODUCTION

Rice is a valuable seed which belongs to grass family where 90% of the world's population relied on rice to fulfill their daily nutritional requirements. Scientifically, rice is called *Oryza sativa* and is widely cultivated in Southeast Asian countries which occupy third place in rice production next to maize and sugarcane [1]. Rice supplies more than 21% of human energy demands and up to 76% of Southeast Asian energy intake, feeding around half of the world's population [2]. It plays important role for supporting in human nutrition and almost 760 million ton of rice were produced over than 100 nations, particularly China around 210 million tons, and India approximately 175 million tons recording for greater than half of the world rice production take place by eight of the top rice producing countries in 2020 involving Indonesia, Bangladesh, Vietnam, Thailand, Myanmar, Philippine, Pakistan and Cambodia [3].

Rice can adapt to moisture-induced anaerobic condition and typically grown in flooded condition. The stem and rood

1* Corresponding author: Sokuntheary Theng
E-mail: theng.sokuntheary@itc.edu.kh; Tel: +855-87 848 181

contain large air spaces that allow for air diffusion throughout the plant [4]. It is a significant source of fiber, calories, minerals, protein, antioxidants, and other macromolecules that may function synergistically to confer health benefits. These nutrients vary depending on rice variety, environmental conditions, and cultivation methods [5], [6].

Farming practices can have a considerable impact on the nutritional quality of crops. Conventional farming systems typically involve synthetic fertilizers and chemical inputs aimed at maximizing yield, while agroecological farming emphasizes environmentally sustainable practices, such as organic fertilization and crop diversification, which may enhance soil health and nutrient retention [7]. This system as agricultural practices which include the use synthesis chemical fertilizer, pesticides, herbicides, and other continual input genetically modified organism, concentrated animal feeding operation, heavy irrigation, intensive tillage, or concentrated monoculture production [8].

This study aims to evaluate and compare the nutritional composition of rice cultivated using agroecological and conventional practices in Preah Vihear province. Three traditional rice varieties including White Phka Rumdoul, Red Phka Rumdoul, and Kroches were selected for analysis. Key nutritional parameters assessed include moisture, ash, fiber, fat, protein, carbohydrates, and reducing sugar content. The findings will contribute to a better understanding of how farming methods influence rice quality and may support more informed agricultural and nutritional decision-making in Cambodia.

2. METHODOLOGY

2.1 Sample Preparation

In this study, three different samples, freshly harvested, were selected for the study and they were collected from Preah Vihear province and package by using paper bag and transported to the laboratory at Institute of Technology of Cambodia. Three varieties of rice White Phka Rumdoul, Red Phka Rumdoul, and Kroches were cleaned by removing the foreign matter such as husk, rough rice and other particles. Next, the cleaned samples were ground in a grinder to get rice powder and storage at room temperature during the experimental work.

2.2 Moisture Analysis

Moisture was determined by standard official methods of analysis of the AOAC with slight modifications. The crucible was thoroughly washed and dried in an oven at 105 for 1 hour and allowed to cool inside desiccators. After cooling, they were weighed using weighing balance, and their various weights recorded as (W1). Then about 2.0 g of the finely ground rice sample was put in crucible (W2). After that, the sample with crucible were placed inside the oven and dried at 105 for 6 hours, until the mass constant, cooled and weighed (W3). Then, the moisture content of the rice sample was calculated from the equation:

$$\text{Moisture content (\%)} = \frac{(M2-M3)}{(M2-M1)} \times 100 \qquad \text{(Eq.1.)}$$

Where M_1: is mass of empty aluminum plate (g)
M_2: is mass of sample (g)
M_3: is mass of dried sample and aluminum plate after drying (g).

2.3 Ash Analysis

Total ash of the rice sample was determined by Furnace Incineration described by AOAC, based on the vaporization of water and volatiles with boric organic substances in the presence of oxygen in the air to CO_2 at a temperature of 550 ℃ with slight modifications. The crucible porcelain was thoroughly washed and dried in a oven at 105 ℃ for 1 hour and allowed to cool inside desicators. After cooling, they were weighed using weighing balance and their various weight recorded as (W1). Then, 1.0 g of the finely ground rice sample was put in the crucible porcelain (W2). After that, the sample plus crucible porcelains were incinerated at 550 ℃ for 4 hours in an incinerating muffler furnace until ash was obtained.

The ash content calculated with the following formula below:

$$\text{Ash content (\%)} = \frac{(W_3-W_1)}{W_2} \times 100 \qquad \text{(Eq. 2)}$$

Where W_1: is mass of empty porcelain cup (g)
W_2: is mass of sample (g)
W_3: is mass of dried sample and porcelain cup after heat (g).

2.4 Reducing Sugar Analysis

AOAC Lane and Eynon Method used to determine the reducing sugar content in a sample through titration, based on the principle that reducing sugars reduce Fehling's solution. To prepare the standard sugar solution, 4.75 g of AR grade sucrose is hydrolyzed by adding concentrated hydrochloric acid and standing for 24 hours. The solution is then neutralized using sodium hydroxide with phenolphthalein as an indicator and diluted to a final volume of 500 ml. A 25 ml aliquot is further diluted to 100 ml, giving a solution where 1 ml is equivalent to 2.5 mg of invert sugar. For the sample, 1 g is weighed and boiled with sulfuric acid under reflux for 3 hours, then cooled, neutralized to a pH of 7–8.5, and diluted to 250 ml before filtration. Fehling's solution is prepared by mixing Fehling A (copper sulfate) and Fehling B (alkaline tartrate) in equal volumes. In the preliminary titration, the standard sugar solution is added to the boiling Fehling's mixture with methylene blue indicator until the blue color disappears, indicating the endpoint. This volume is used to

estimate the expected titration range. The final titration is performed using the prepared sample solution in the same manner, and the test is done in duplicate.

The reducing sugar content is calculate using the formula:

$$\text{Reducing sugar (\%)} = \frac{0.0025 \times V_1 \times V_2 \times 100}{V_3 \times W} \quad \text{(Eq.3)}$$

2.5 Fat Analysis

Total fat of rice sample was determined using Soxhlet Extraction method AOAC with slightly modification. 1.0 g of ground rice sample(W) was weighted and place in cartouche. Then, glass of Soxhlet apparatus was dried in oven about 1 hour. Then, allowed it to cool inside desiccator and weighted (W1). After that, sample plus cartouche was placed in glass of Soxhlet equipment. Add 70 mL of hexane solution, and set the program and start extracting. After the extraction was finished, the glass was dried at 105 ℃ for 1 hour in order to evaporate the remaining hexane solution. After drying, it was cooled in a desicator and weighed (W2).

Then, percentage of fat in the sample was computed using the formula below:

$$\text{Fat content (\%)} = \frac{(W_3 - W_1)}{W_2} \times 100 \quad \text{(Eq. 4)}$$

Where W_1: is mass of empty glass (g)
W_2: is mass of sample (g)
W_3: is mass of glass with extracted fat

2.6 Protein Analysis

The protein content was analyzed by using kjeldahl machine. The sample was digested by sulfuric acid to breakdown organically bonds nitrogen into ammonium ions., followed by neutralization distillation, and titration. The nitrogen content was then multiplied by a conversion factor5.95 for rice to estimate protein content.

The protein is calculate using the formula:

$$\%N = \frac{(V_{sample} - V_{blank}) \times Z \times C \times F \times M_N}{m_s \times 10}$$

$$\%P = \%N \times P_F \quad \text{(Eq.5)}$$

Where W_N is the weight fraction of nitrogen
V_{sample} is the volume of the blank (ml)
V_{blank} is the volume of the blank factor for H_2SO_4
Z is the molar reaction factor
C is the concentration of the titrant (mol/l)
F is the factor of titrant (0.9969)
M_N is the mass molecular weight of the nitrogen
%N is the percentage of the weight of the nitrogen
%P is the percentage of protein
P_F is the protein factor 5.95

2.7 Fiber Analysis

Crude fiber content was determined using the AOAC method. The sample was digested with acid (H_2SO_4 1.25%) and alkali (NaOH 1.25%). The undigested residue was then filtered, dried in an oven at a specific temperature to obtain a constant weight, and finally ash in a muffle furnace. The loss in weight after ash represented the crude fiber content.

The fiber content calculated with the following formula below:

$$\%Fiber = \frac{(W_2 - W_1) - (W_4 - W_3)}{W_0} \times 100 \quad \text{(Eq.6)}$$

Where W_0 is the mass of sample (g)
W_1 is the mass of fiber bag (g)
W_2 is the mass of dried sample with fiber bag (g)
W_3 is the mass of the crucible (g)
W_4 is the mass of the crucible and ash (g)

2.8 Carbohydrates Analysis

Carbohydrate content was calculated by difference. The percentages of moisture, ash, protein, fat and fiber were subtracted from 100 to determine the total carbohydrate percentage in the sample. The formula used was:

Carbohydrate content (100%) = 100% - (% Moisture + Protein %+ Fat % + Ash %) (Eq. 7)

3. RESULTS AND DISCUSSION

Table 1. Nutritional value of rice in Phear Vinhea

Parameter	%M	%A	%F	%Fa	%P	%C	%S
WPR CT	11.6 ±1.2 7	0.84 ±0.2 9	0.76 ±0.2 1	1.59 ±0.4 5	6.68 ±1.2 7	79.6 ±1.7 6	5.01 ±0.1 1
WPR GM	10.0 ±0.7 7	0.77 ±0.4 0	0.24 ±0.0 4	1.60 ±0.7 9	6.01 ±0.3 8	81.6 ±1.4 6	4.99 ±0.1 1
RPR CT	9.52 ±1.9 6	1.00 ±0.1 6	0.40 ±0.2 4	1.9± 0.28	6.95 ±0.1 7	80.5 ±2.1 5	4.92 ±0.0 8
RPR GM	9.17 ±0.3 8	0.93 ±0.2 4	0.74 ±0.2 7	2.49 ±0.3 2	6.44 ±0.8 4	80.9 ±0.8 6	5.03 ±0.2 4
RPR CT	10.5 ±1.5 3	0.80 ±0.5 2	0.79 ±0.2 4	2.27 ±0.2 3	6.77 ±0.7 7	79.6 ±1.0 9	4.96 ±0.1 7
RPR GM	8.72 ±0.2 1	1.12 ±0.6 5	1.04 ±0.1 2	2.57 ±0.2 9	6.10 ±1.0 0	81.4 ±0.9 7	4.84 ±0.1 2
KC GM	9.21 ±0.5 2	1.85 ±1.7 7	0.75 ±0.2 5	2.16 ±0.1 2	7.16 ±0.1 8	79.1 ±2.6 4	4.92 ±0.2 2

KC CT	9.58 ±0.5 5	0.60 ±0.2 7	0.58 ±0.1 9	2.39 ±0.3 6	6.30 ±0.2 9	81.1 ±1.0 6	5.02 ±0.1 7
WPR CT	9.15 ±0.5 4	0.56 ±0.3 5	0.97 ±0.4 6	2.12 ±0.2 5	7.49 ±0.5 0	80.6 ±0.8 5	4.90 ±0.2 4
WPR GM	9.71 ±0.7 7	0.60 ±0.6 0	0.52 ±0.2 4	1.66 ±0.3 5	6.99 ±0.1 2	81.0 ±0.9 0	5.03 ±0.2 0
WPR CT	9.58 ±0.9 6	0.78 ±0.4 0	0.60 ±0.2 7	2.11 ±0.2 2	6.74 ±0.2 8	80.7 ±1.1 3	4.76 ±0.3 9
WPR GM	9.4± 0.45	0.48 ±0.3 8	0.81 ±0.2 7	1.96 ±0.6 4	6.79 ±0.2 2	81.2 ±1.0 6	4.90 ±0.1 4
KC CT	9.80 ±0.2 3	0.30 ±0.2 0	0.42	1.75	4.71	83.4 ±0.4 3	4.86 ±0.1 5
KC GM	9.68 ±0.1 2	0.90 ±0.3 2	0.97	1.89	6.88	80.6 ±0.2 3	4.85 ±0.1 9

*%M=Moisture, %A=%Ash, %Fa=%Fat, %F=%Fiber, %P=%Protein, %S=%reducing sugar, %C=%Carbohydrates
**CT=conventional treatment, GM=green manure
***WPR= White phka romdoul, RPR= Red phka romdoul, KC= Kroches

The nutritional composition of three rice varieties, White Phka Rumdoul (WPR), Red Phka Rumdoul (RPR), and Kroches (KC) cultivated under agroecological and conventional practices in Preah Vihear province revealed significant differences in rice nutritional composition based on both farming practices and varietal characteristics. Moisture content was consistently higher in conventionally grown rice across all varieties, with WPR CT ($11.6 \pm 1.27\%$) and RPR CT ($10.5 \pm 1.53\%$) showing the greatest values. These elevated moisture levels may result from synthetic nitrogen inputs and irrigation, which promote higher water uptake [9]. In contrast, agroecological samples tended to have lower moisture, enhancing storage stability and reducing post-harvest spoilage risks [10]. Ash content, indicative of mineral accumulation, was enhanced in agroecologically grown rice, especially in KC GM ($1.85 \pm 1.78\%$) and RPR GM ($1.12 \pm 0.65\%$). This may reflect improved soil biodiversity and nutrient cycling associated with organic fertilization [11]. Similar trends were observed for fiber content, which was notably higher in RPR GM ($1.04 \pm 0.13\%$) and KC GM ($0.75 \pm 0.26\%$), underscoring the potential of green manure systems to support dietary fiber accumulation [12]. Fat content was consistently elevated under agroecological conditions. RPR GM exhibited the highest fat concentration ($2.57 \pm 0.29\%$), suggesting that longer maturation periods or reduced chemical stress may favor lipid biosynthesis [13]. Carbohydrates, a key energy source, also trended higher in GM samples WPR GM and RPR GM reached $81.6 \pm 1.46\%$ and $80.9 \pm 0.87\%$, respectively. Protein levels displayed variety-specific responses. While conventional treatments yielded slightly more protein in WPR and RPR, agroecological KC recorded the highest protein value ($7.16 \pm 0.19\%$), possibly due to slower nitrogen release from organic inputs allowing for extended uptake. Reducing sugar content remained stable across all treatments (range: 4.84–5.03%), suggesting a stronger genetic influence than environmental.

In summary, agroecological farming practices enhanced several nutritional components particularly ash, fat, fiber, and carbohydrates without compromising protein quality. These results highlight the nutritional advantages of sustainable rice farming, offering important implications for food security, dietary quality, and eco-friendly agricultural policy in Cambodia.

4. CONCLUSIONS

The determination of the nutritional composition in rice clearly demonstrates that cultivation practices have a significant impact on the nutritional profile of rice in Preah Vihear province. Across the three rice varieties examined White Phka Rumdoul, Red Phka Rumdoul, and Kroches conventional farming generally resulted in higher moisture and fiber contents, while agroecological practices enhanced fat and carbohydrate levels. Variations in protein and ash content were variety-specific, suggesting an interaction between genotype and farming method. Although reducing sugar content remained consistent across all treatments, the observed differences in other nutrients provide valuable insights into how cultivation methods can be tailored to meet specific nutritional or functional goals. These findings not only support the promotion of agroecological farming for its potential health and sustainability benefits but also offer guidance for food product developers, nutritionists, and agricultural planners in optimizing rice quality through strategic farming choices.

ACKNOWLEDGMENTS

The authors acknowledge the support of the laboratory of Faculty of Chemical and Food technology of Institute of Technology of Cambodia under project ASSET (Agroecology and Safe Food System Transitions).

REFERENCE

[1] L. Devraj, A. Panoth, K. Kashampur, A. Kumar, and V. Natarajan, “Study on physicochemical, phytochemical, and antioxidant properties of selected traditional and white rice varieties,” *J. Food Process Eng.*, vol. 43, no. 3, pp. 1–13, 2020, doi: 10.1111/jfpe.13330.

[2] B. O. Juliano and A. P. P. Tuaño, “Gross structure and composition of the rice grain,” *Rice Chem. Technol.*, no. 1, pp. 31–53, 2018, doi: 10.1016/B978-0-12-811508-4.00002-2.

[3] S. Sen, R. Chakraborty, and P. Kalita, “Trends in Food Science & Technology Rice - not just a staple food : A comprehensive review on its phytochemicals and therapeutic potential,” *Trends Food Sci. Technol.*, vol. 97, no. January, pp. 265–285, 2020, doi: 10.1016/j.tifs.2020.01.022.

[4] M. Zhao, Y. Lin, and H. Chen, “Improving nutritional quality of rice for human health,” *Theor. Appl. Genet.*, vol. 133, no. 5, pp. 1397–1413, 2020, doi: 10.1007/s00122-019-03530-x.

[5] S. Zafar and X. Jianlong, “Recent Advances to Enhance Nutritional Quality of Rice,” *Rice Sci.*, vol. 30, no. 6, pp. 523–536, 2023, doi: 10.1016/j.rsci.2023.05.004.

[6] K. T. Akhter *et al.*, “Variations in the Major Nutrient Composition of Dominant High-Yield Varieties (HYVs) in Parboiled and Polished Rice of Bangladesh,” *Foods*, vol. 12, no. 21, pp. 1–12, 2023, doi: 10.3390/foods12213997.

[7] A. Gamage *et al.*, “Role of organic farming for achieving sustainability in agriculture,” *Farming Syst.*, vol. 1, no. 1, p. 100005, Apr. 2023, doi: 10.1016/J.FARSYS.2023.100005.

[8] V. Quintarelli *et al.*, “Cover Crops for Sustainable Cropping Systems: A Review,” *Agric.*, vol. 12, no. 12, 2022, doi: 10.3390/agriculture12122076.

[9] Y. Xing, X. Zhang, and X. Wang, “Enhancing soil health and crop yields through water-fertilizer coupling technology,” *Front. Sustain. Food Syst.*, vol. 8, no. November, pp. 1–23, 2024, doi: 10.3389/fsufs.2024.1494819.

[10] J. Brinkhoff, B. W. Dunn, and T. Dunn, “The influence of nitrogen and variety on rice grain moisture content dry-down,” *F. Crop. Res.*, vol. 302, p. 109044, Oct. 2023, doi: 10.1016/J.FCR.2023.109044.

[11] W. I. Susanti, S. N. Cholidah, and F. Agus, “Agroecological Nutrient Management Strategy for Attaining Sustainable Rice Self-Sufficiency in Indonesia,” *Sustain.*, vol. 16, no. 2, pp. 1–29, 2024, doi: 10.3390/su16020845.

[12] Y. Arévalo-aranda *et al.*, “Impact of Green Manuring and Nitrogen Fertilization on Rice Cultivation : A Peruvian Amazon Forest Study in San Martín Province,” vol. 12, no. version 1, 2024.

[13] J. Han and R. J. Kaufman, “The role of ER stress in lipid metabolism and lipotoxicity,” *J. Lipid Res.*, vol. 57, no. 8, pp. 1329–1338, 2016, doi: 10.1194/jlr.R067595.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Determination of Heavy Metals Contamination in Rice Collected from Preah Vihear Province

Tivorn SREANG [1], Sreyna HAK [1], Marinich NET [1], Peany Houng[1], Kimleang Khoeurn[1*], Florent TIVET[2], Mathide SESTER[2]

[1] *Faculty of Chemical and Food Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *CIRAD, UPR Agroecologie et Intensification Durable des cultures Annuelles (AIDA)*

[*]*Corresponding author: khoeurn@itc.edu.kh*

Abstract: *Rice is an important source of energy, vitamins, and minerals that make up over 70% of the daily caloric intake of most people worldwide, especially those in developing Asian countries. Heavy metals from natural and anthropogenic sources accumulate in soil and plants, and the presence of heavy metals above thresholds raises concerns about food safety and adverse health effects. The purpose of this study was to ensure the safety of the rice grains and to determine the amount of selected heavy metals (Fe, Mn, and As) in a total of 48 samples harvested from green manual (GM) and conventional tillage (CT) systems. Then, all samples were digested using the nitric acid-hydrogen peroxide digestion method, and both manganese (Mn) and iron (Fe) were determined by photometer, whereas arsenic (As) was determined by an arsenic test kit. The results showed that As, Mn, and Fe content in rice ranged between 1.4 and 150.5 mg/kg, 0.0175 and 0.53 mg/kg, and 0.0035 and 0.0049 mg/kg in CT practice, respectively. In addition, the levels of As, Mn, and Fe in rice grown through GM practice ranged between 4.75 and 137.7 mg/kg, 0.014 and 0.2775 mg/kg, and 0.0011 and 0.00365 mg/kg, respectively. The concentration of Mn and As in rice samples collected from the CT system was higher than in the GM system, while the Fe concentration of the CT system was lower than the GM system. According to IBM SPSS Statistics 27, by utilizing univariate variance analysis, the results were significantly different ($p > 0.05$) within the two systems for Fe, Mn, and As. Thus, rice samples that were treated with GM had a lower concentration than those with the CT system, except for Fe concentration, which was lower in the CT system. According to the results in this study, the concentrations of Mn, Fe, and As were found to be within the standard permitted levels set by the WHO. For further study, more research is required to thoroughly examine pertinent mechanisms contributing to heavy metal content in rice plants, such as the bioaccumulation of heavy metals, influencing factors, and minimizing measures by heavy metal remediation techniques to improve the quality of rice in Cambodia. Additionally, soil management and agricultural techniques should be promoted more to ensure the uniform quality and safety of soil and agricultural products.*

Keywords: Heavy metal; Rice; Arsenic; Manganese; Iron.

1. INTRODUCTION

Rice, a cereal of choice in Asia, is a staple food crop, and 70% is consumed globally across the world [1]. However, rice is possibly contaminating of heavy metal and other toxic substance which is a grave concern for food safety and human health. International Organization for Standardization of Contamination (MAC) was set the limitation of exposure to heavy metals in rice and plants to protect safety of human health, but even below MAC levels, heavy metals can still lead to some disease such as high blood pressure disease, nervous disease, and cancer [2]. The heavy metal that accumulate in rice and plants by 2 ways are nature and anthropology. Anthropogenic and natural sources are the main sources of heavy metals pollutants, and heavy metals content affects microorganisms, soil chemical and physical properties, and soil pH [3]. Heavy metals are taken up by plants and deposited in human cells, affecting human health and the environment. The most significant heavy metal pollutants are Cd, As, and Pb, and these are among the top 20 poisonous elements [4]. These pollutants affect human beings, the environment, and rice plants, but in lesser quantities. Heavy metals such as Zinc (Zn), iron (Fe), manganese (Mn) are essential plant micro-nutrients that improve their chemical and biological characteristics. Improving the safety of rice is necessary in order to protect consumers from harms and risks related to heavy metal contamination [5]. Therefore, the purpose of this study to

[1]* Corresponding author: Khoeurn Kimleang
E-mail: khoeurn@itc.edu.kh; Tel: +855-86 455 931

determine the concentration of heavy metal, Fe, Mn and arsenic (As) in rice sample from green manual (GM) and conventional tillage (CT) agriculture systems and to compare quantity of concentration those heavy metal between both systems.

2. METHODOLODY

2.1. Rice Collection

A total of 48 rice samples were collected from Rovieng district, Preah Vihear province.There are three varieties of rice that was planted using two different systems (system 1, conventional tillage (CT), and system 2, green manual (GM)) and four treatments in each system, such as control (1-T0), biochar of HUSK Venture: 2 tons/ha (2-R2), biochar of HUSK Venture: 5 tons/ha (3-R5), and ash of rick husk from Amru rice: 5 tons/ha (4-AHS5). Thus one rice code would consist of 8 conditions (CT-1-T0 meaning the rice is grown in the field of conventional system with control treatment, CT-2-R2, CT-3-R5, CT-4-AHS5, GM-1-T0 meaning the rice is grown in green manual system with control treatment, GM-2-R2, GM-3-R5, GM-4-AHS5). For each condition, only 1 sample was collected, so 1 rice code would have 8 samples. Thus, for the 3 types of rice: white Phka Romdoul (3 codes) consists of 24 samples, Red Phka Romdoul (2 codes): 16 samples, and Kroches (1 code): 8 samples; in total there were 48 samples.

2.2. Rice preparation

Rice samples with husk and bran were dried in an oven at 65°C for 24 h [6]. After that, samples were husked by using a husker machine to remove the bran. Samples were then grinded by mortar and pestle until become powder form and stored in zip lock for digestion procedure before heavy metal analysis.

2.3. Digestion method

1 g of rice powder was weighed using an electric balance and was transferred into a 100 ml beaker. Following, 10 ml of 65% nitric acid was carefully added to the beaker. The mixture was heated at 170ºC for 15 minutes. When completing the heating, the beaker was allowed to cool down at room temperature for about 10 minutes, and 10 ml of nitric acid was added again. The beaker was then continued to heat at 200ºC until the mixture evaporated to 5 ml, allowing it to cool down for 10 minutes. Next, 3 ml of H_2O_2 with 2 ml of distilled water was added to the mixture, and was continue heated at 200ºC for 15 min. In addition, 10 ml of H_2O_2 was poured into the solution and continue heated until the white smoke of the solution was exhausted. Then, 1 ml of the 30% H_2O_2 concentration was gradually dropped and was heated until the solution in the beaker effervescence was noticeable. Next, the solution was cooled down, and was filtrated using a syringe filter with a diameter of 0.45 micrometers. Finally, the filtrate was diluted with 100 ml of distilled water and stored in the centrifuge tube Samples were finally stored in a refrigerator at 4ºC for further analysis [7] .

2.4. Analysis procedure

The concentrations of Mn and Fe in digested rice samples were measured using a photometer (M242 method for Mn) [8]. This method has the detection at a range from 0.01 to 0.7 mg/l. For Mn analysis, 10 mL of the digested sample was transferred into a 24 mm vial, while a separate vial containing 10 mL of deionized water served as a blank. Vario ascorbic acid powder was added to both vials, followed by swirling to mix. Next, 15 drops of alkaline reagent were added and mixed, followed by the addition of 21 drops of Vario Pan indicator. After a 2-minute reaction time, the blank vial was inserted into the photometer to zero the instrument. The sample vial was then measured, and Mn concentration was recorded.

For Fe measurement, 10 mL of the digested sample was placed in a 10 mm vial and inserted into the photometer. After pressing the zero key, the Vario Ferro F10 reagent pack was added, the vial was sealed and swirled to mix, and then reinserted into the chamber. After a 3-minute reaction period, Fe concentration was measured [9].

As was analyzed using a commercial arsenic test kit [10]. A 60 mL sample was transferred into a reaction bottle. Arsenic Reagent 1 was added and mixed, followed by one red spoon of Reagent 2 and one green spoon of Reagent 3, with mixing after each addition. A test paper was then inserted into the bottle cap holder and allowed to react for 15 minutes before colorimetric result interpretation.

3. RESULTS AND DISCUSSION

3.1. Concentration of Iron in rice sample

The result of Fe concentrations in Rice White Phkar Romdoul growth with the CT and GM systems were demonstrated in Fig.1**.** There was significantly difference ($p<0.05$) of Fe content in rice samples growth using CT and GM system. According to the result in Fig. 1 (a), the Fe concentration of D11 variety planted using GM system was higher than CT system among four treatments, except 1-T0 treatment. The Fe concentrations for CT systems were between 7.95 mg/kg and 42.1 mg/kg, while those for GM systems were between 9.5 mg/kg and 55.65 mg/kg. Fig. 1 (b) illustrates that the result of Fe concentration of rice D22 variety was higher in GM system, except 3-R5, and 4-AHS5 treatments, ranging from 17.575 to 33.375 mg/kg in CT and GM was 4.75 to 137.7.mg/kg. Based on Fig. 1 (c), the Fe

concentration in CT of rice D23 variety ranged between 35.75 to 150.5 mg/kg and GM system was 9.35 to 81.25 mg/kg.

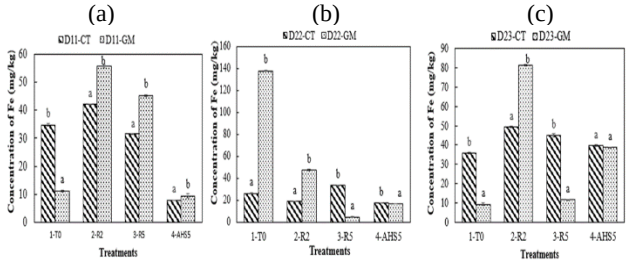


Fig. 1: Concentration Fe in White Phkar Romdoul rice: (a) in field 1 at demo 1 (code - D11), (b) in field 2 at demo 2 (code - D22), (c) in field 3 at demo 2 (code - D23). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

Fig. 2 demonstrates the result of Fe concentrations in Red Phkar Romdoul growth with the CT and GM systems. According to Fig. 2 (a & b), the Fe concentration of D12 in GM system was significantly difference (p<0.05) in CT system among all treatments. The Fe concentration of D12-CT was 12.35 to 38 mg/kg, while D12-GM was 6.4 to 43.75 mg/kg. For D13 variety, the Fe content using CT system was 4.5 to 32.255 mg/kg and 22.25 to 67.25 kg for GM system. The result in this study vary depend on the history of sample collection site and the soil quality before applying cover crop or usage biochar and ash [11].

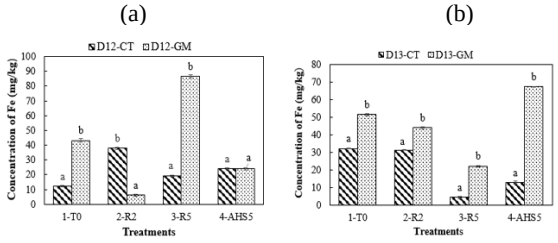


Fig. 2: Concentration Fe in Red Phkar Romdoul rice: (a) in field 2 at demo 1 (code - D12), (b) in field 3 at demo 1 (code – D13). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

For kroches rice variety, the concentration of Fe in CT was significantly higher (p<0.05) than GM, except the treatment 4-AHS5. The Fe concentrations for GM systems were between 4.75 mg/kg and 137.7 mg/kg while those for CT systems were between 1.4 mg/kg and 49.4 mg/kg. Based on this study, cover crops in farming helps in increasing the crop diversity and environmental impact by reducing erosion, promoting infiltration, retaining soil moisture, increasing nutrients, improving soil health, and suppressing weed growth. A number of heavy metal contaminants such as iron present in soil can be remediated using biochar derived from rice husks. Iron is vital in maintaining the human and plant body but in excess amounts, iron accumulates and becomes detrimental to both structure and function [12].

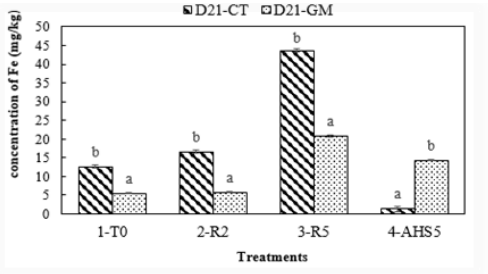


Fig. 3: Concentration Fe in Kroches rice in field 2 at demo 1 (code - D12). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

3.2. Concentration of Mn in rice sample

According to the Phkar Romdoul rice, the concentrate of Mn in the fig.4 is significantly difference (p<0.05) in both systems. For Fig.4 (a) of D11 is, the concentration of Mn in CT was 0.0475 to 0.53 mg/kg, while GM was 0.0275 to 0.2775 mg/kg. For Fig.4 (b) in D22 system, the concentrate of Mn of CT is higher than GM except treatment 1-T0 is almost the similar was 0.0425 to 0.19 in CT and 0.029 to 0.0675 mg/kg in GM. Fig.4 **(c)** in D23, the concentration of Mn is lower than CT except 3-R5 which the concentration is almost the same was 0.0415 to 0.0715 in CT and GM was 0.0425 to 0.055 mg/kg . Recording to Fig.4 (a), (b), and (c) in treatment 3-R5 showed that Fig.4 (a) and Fig.4 (c) have a concentration of Mn in CT similar to the GM system, except for 3-R5 in Fig.4 (b) having a concentration of Mn in CT higher than the GM system, and treatment 1-T0 showed that (a) and (c) have a concentration of Mn in CT higher than the GM system, except for 1-T0 in Fig.4 (b) being similar to the GM system.

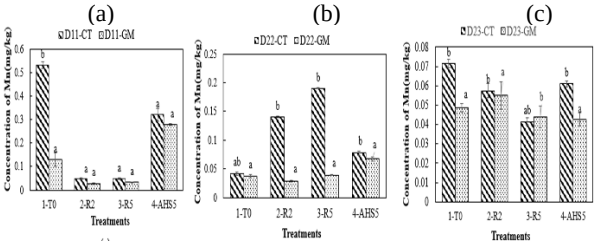


Fig. 4: Concentration of Mn in Phkar Romdoul rice: (a) in field 1 at demo 1 (code - D11), (b) in field 2 at demo 2 (code - D22), (c) in field 3 at demo 2 (code - D23). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

The result of Mn in D12 and D13 in CT and GM of Fig**.** 5**.** of Red Romdoul. Fig.5 (a), in CT the concentrate of Mn was 0.0175 to 0.0775 mg/kg, while Mn in GM was 0.014 to

0.0585 mg/kg. Fig.5 (b), Mn was 4.5 to 32.255 mg/kg in CT and GM was 0.019 to 0.092 mg/kg.

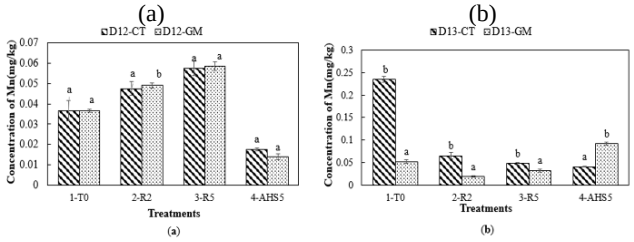


Fig.5: Concentrate of Mn in Red Phkar Romdoul: (a) in field *2* at demo 1 (code - D1*2*), (b) in field *3* at demo *1* (code – D*13*). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

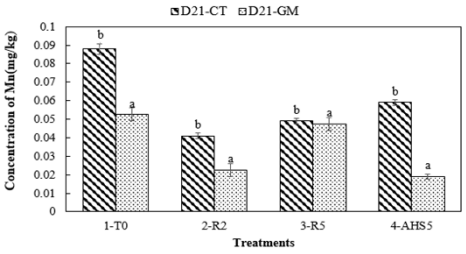


Fig.6. Concentration of Mn in Kroches in field 1 at demo 2 (code – D21). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

In Fig.6 the concentration of Mn in D21 show that the Mn in GM is smaller than CT for all treatments. Comparison between the CT system and the GM system in Fig.4 and Fig.5 observed that rice samples with GM-system concentrations were less than those with the CT system. Mn concentrations in the CT systems have ranged from 0.0035 mg/kg to 0,32mg/kg, and the concentration of Mn in the GM system has ranged from 0.0014 mg/kg to 0.0585 mg/kg. In this study, rice grain samples collected from the Convectional system heavy metal content in the rice grain samples had Mn concentrations ranging from 0.73 mg/kg to 18mg/kg [13]. Rice samples with GM system concentrations were less than those with CT system application because of the capacity of GM to enhance rice plants' growth and tolerance to heavy metal absorption. Moreover, the mechanism of GM-system accumulation in rice plants and environmental stressors are also factors that can decrease heavy metal uptake. A recent study showed that supplementation. Implementing cover crops helps prevent soil loss and retain organic matter, conserving water, nutrients, and aeration. Continuous use of cover crops can help mitigate the crisis. successfully decreases the Mn concentration in the shoots but increases Mn in the roots of rice under high Mn stress, alleviating Mn toxicity [14]. Based on research, Rice Husk Biochar application in agriculture can reduce the availability of heavy metals in contaminated soils, especially Fe, and Mn [15].

3.3. Concentration of As in rice sample

It is a very hazardous element that is found in many different species. The toxicity of arsenic varies depending on the species. The pH, redox circumstances, surrounding mineral composition, and microbial activity determine arsenic's form (inorganic or organic) and oxidation state, and the maximum permissible concentration of arsenic for rice is 0.1 mg/kg by the WHO [16]. The result of As in Phkar Romdoul that according to the Fig.7**.** in CT-GM system. In CT system of D11, the concentration of As was 0.00335 to 0.00485 mg/kg, while in GM was 0.0019 to 0.0031 mg/kg. For D22 concentration of As is 0.0041 to 0.048 mg/kg in CT and GM was 0.0175 to 0.0325 mg/kg. For D23, the concentration between, 0.0036 to 0.0049 mg/kg in CT and GM was 0.0011 to 0.00325 mg/kg. For Fig.7 (a), (b), and **(c)** the concentration of As in CT system is larger than GM for all treatments.

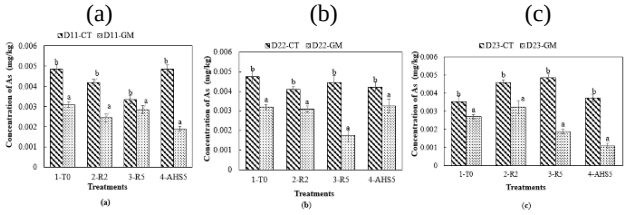


Fig.7. Concentration of As in Phkar Romdoul: (a) in field 1 at demo 1 (code - D11), (b) in field 2 at demo 2 (code - D22), (c) in field 3 at demo 2 (code - D23). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

The result of Fig.8. show about the As content in Red Romdoul rice following standard deviation and average concentration. In Fig.8 (a), concentration of As was between 0.0046 to 0.0049 mg/kg in CT, while GM was 0.0022 to 0.00355 and Fig.8 (b), the concentration of As was 0.00175 to 0.0028 mg/kg in GM, and in CT was 0.0035 to 0.00455 mg/kg. The concentration of As is significant different ($p < 0.05$) in GM system is lower than CT for all treatments.

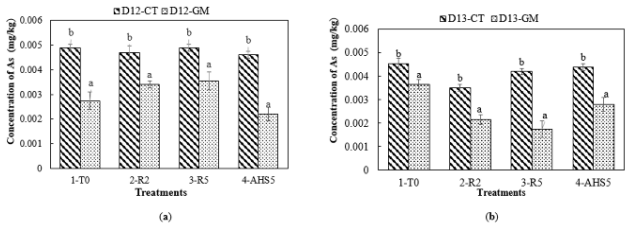


Fig.8. As concentration in Red Romdoul rice: (a) in field 2 at demo 1 (code - D12), (b) in field *3* at demo 1 (code – D13). Letter (a and

b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

Fig.9**.** show that the average concentration and standard deviation of Kroches rice is different that CT is higher if compare with GM. Contrasting the CT and GM systems revealed that the rice samples stemming from the GM system had lesser concentrations of As when compared to those from the CT system. The As concentrations for CT systems were between 0.0035 mg/kg to 0.00485 mg/kg, while those in GM systems were within the bracket of 0.0011 mg/kg and 0.00365 mg/kg. A study noted that the samples of rice grains originating from CT systems had As levels in the range of 0.03 mg/kg and 0.34 mg/kg and some of these exceed the WHO limit of 0.1 mg/kg set for grains [17]. Studies suggest that the use of cover crops can improve agricultural biodiversity and the environment by mitigating erosion and runoff, enhancing soil moisture, improving soil health and nutrients, preventing weed growth, and reducing heavy metals in contaminated soils [18]. Tamed plants as defined by Van der Ploeg et al. are capable of enhancing agriculture. Cover crops physically block weeds and affect the processes that take place in the soil.

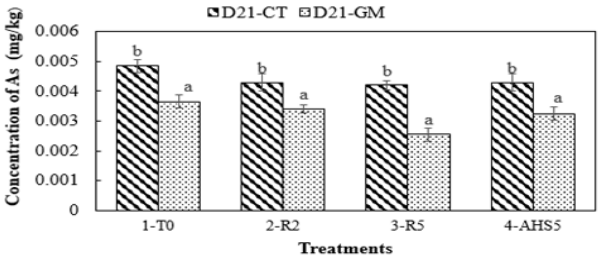


Fig.9. Concentration of As in Kroches rice in field 1 at demo 2 (code – D21). Letter (a and b) defines the significantly difference of Fe concentration of all the treatments between GM and CT.

3. CONCLUSIONS

The concentrations of Fe, Mn, and As in rice grains cultivated in Rovieng District, Preah Vihear Province were investigated in this study. The results showed that Mn and As concentrations were higher in rice samples grown under the CT system, whereas Fe concentration was higher in samples from the GM system compared to the CT system. Investigated heavy metals were found in all rice samples in both agricultural systems, the concentration remained below the maximum permissible limits set by WHO and FAO. These findings underscore the importance of comprehensive assessments agricultural practices to ensure food safety and environmental health. The study also suggests that using GM techniques for rice plating can help reduce heavy metal accumulation in rice grains.

ACKNOWLEDGMENTS

The project was funded by the European Union and the French Development Agency.

REFERENCE

[1] G. Trébuil, “Rice production systems in Asia: The constant presence of an essential cereal on a continent in mutation,” *The Asian side of the world: Editorials on Asia and the Pacific 2002-2011*, no. July, pp. 1–6, 2012, [Online]. Available: https://agritrop.cirad.fr/559930/1/document_559930.pdf

[2] M. Kou, Z. Liu, Z. Wang, J. Cai, and W. Tan, “Heavy metal concentrations in rice that meet safety standards can still pose a risk to human health,” pp. 1–9, 2023, doi: 10.1038/s43247-023-00723-7.

[3] A. Alengebawy, S. T. Abdelkhalek, S. R. Qureshi, and M. Q. Wang, “Heavy metals and pesticides toxicity in agricultural soil and plants: Ecological risks and human health implications,” Mar. 01, 2021, *MDPI AG*. doi: 10.3390/toxics9030042.

[4] Atsdr, “TOXICOLOGICAL PROFILE FOR LEAD,” 2007.

[5] R. Lohry, “Micronutrients: Functions, Sources and Application Methods.”

[6] P. Vargas-Escobar, O. Flórez-Acosta, and L. L. Corrales-García, “Renewing the potential of rice crop residues as value-added products in the cosmetics industry,” *Heliyon*, vol. 10, no. 7, p. e28402, Apr. 2024, doi: 10.1016/J.HELIYON.2024.E28402.

[7] U. Environmental Protection Agency, “METHOD 3050B ACID DIGESTION OF SEDIMENTS, SLUDGES, AND SOILS 1.0 SCOPE AND APPLICATION.”

[8] K. , et al. , Goto and Talanta, *Determination of Manganese LR with Vario Powder Packs*. 2024.

[9] D. F. Boltz and J. A. Howell, “A New Field Method for Determining the Levels of Iron Contamination in Oilfield Completion Brine,” *SPE International Symposium*, vol. 8, 1978.

[10] T. Htike Oo, T. Hata, K. Zaw Htun, and K. Marumo, “IMAGE ANALYSIS BY USING SMARTPHONE FOR THE APPLICABILITY OF ARSENIC TEST KIT.”

[11] H. Ali, E. Khan, and I. Ilahi, “Environmental chemistry and ecotoxicology of hazardous heavy metals: Environmental persistence, toxicity, and bioaccumulation,” 2019, *Hindawi Limited*. doi: 10.1155/2019/6730305.

[12] R. Daulta, T. Sridevi, and V. K. Garg, "Spatial distribution of heavy metals in rice grains, rice husk, and arable soil, their bioaccumulation and associated health risks in Haryana, India," *Toxin Rev*, pp. 1–13, 2020, doi: 10.1080/15569543.2020.1793781.

[13] J. Ouyang, Z. Liu, L. Zhang, Y. Wang, and L. Zhou, "Analysis of influencing factors of heavy metals pollution in farmland-rice system around a uranium tailings dam," *Process Safety and Environmental Protection*, vol. 139, pp. 124–132, Jul. 2020, doi: 10.1016/J.PSEP.2020.04.003.

[14] L. Clay, K. Perkins, M. Motallebi, A. Plastina, and B. S. Farmaha, "The perceived benefits, challenges, and environmental effects of cover crop implementation in south carolina," *Agriculture (Switzerland)*, vol. 10, no. 9, pp. 1–14, Sep. 2020, doi: 10.3390/agriculture10090372.

[15] Y. Yulnafatmawita, S. Yasin, and L. Maira, "Effectiveness of rice husk biochar in controlling heavy metals at polluted paddy soil," in *IOP Conference Series: Earth and Environmental Science*, IOP Publishing Ltd, Oct. 2020. doi: 10.1088/1755-1315/583/1/012008.

[16] T. Kormoker *et al.*, "Presence of toxic metals in rice with human health hazards in Tangail district of Bangladesh," *Int J Environ Health Res*, vol. 32, no. 1, pp. 40–60, 2022, doi: 10.1080/09603123.2020.1724271.

[17] A. Yao, "| 983."

[18] Rafenomanjato Antsa, "Weed management using a no-till system with Stylosanthes guianensis cover crop in upland rice-based cropping systems in the Mid-West of Madagascar. Pisa : Scuola Superiore Sant Anna," *Agrobiodiversity*, 2018.

Mechatronic and Information Technology (MIT)

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Design and Implementation of 12-Lead ECG Data Acquisition using AD1298 Sensor

Sukun THANNI [1*], Rithea SUM [1], Rithea NGETH [2],

[1] *Department of Electrical and Energy Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Department of Telecommunication and Network Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: e20201164@itc.edu.kh*

Abstract: *This paper presents an electrocardiogram (ECG) waveform data acquisition system. A medical grade ECG sensor, AD1298 sensor, is employed in this work. The system uses an STM32 microcontroller to read ECG data from the ADS1298 sensor through SPI communication. STM32 then sends this data to an ESP32 microcontroller using USART. ESP32 microcontroller is used to store the data in a buffer and to send them to a database on a local server. An ECG simulator, Contac MS400, is used to generate ECG waveform. The result shows that the error between the ECG data received by the ESP32 and the data generated by the ECG simulator is acceptable. The results prove that the system works reliably and maintains data accuracy. This work aims to provide an affordable system that can help make health monitoring more accessible, especially in areas with limited resources.*

Keywords: ECG waveform; ECG simulator; ESP32; ADS1298; STM32

1. INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of death worldwide, accounting for approximately 17.9 million lives annually a staggering figure representing 32% of all global deaths [1]. Alarmingly, over 80% of CVD-related fatalities result from heart attacks and strokes, with one-third occurring prematurely in individuals under 70 years of age. In this critical context, the electrocardiogram (ECG) serves as an indispensable diagnostic tool, providing essential insights into cardiac electrical activity and enabling early detection of life-threatening arrhythmias and other abnormalities [2]. Given this paradigm, the timely identification and analysis of arrhythmias through ECG monitoring are vital for implementing preventive measures and improving treatment outcomes in CVD management.

Several prior studies have explored ECG data acquisition and transmission systems. In the work of [3] developed an IoT-based solution where ECG data was uploaded to a cloud platform using HTTP and MQTT protocols for storage and real-time monitoring. In contrast, the work of [4] implemented a Bluetooth Low Energy (BLE) gateway to stream ECG data from a smart sensor to a smartphone, leveraging GATT profiles and WebSocket protocols for near real-time cloud transmission. Meanwhile, the work of [5], utilized an STM32L431RCT6 microcontroller to interface with the ADS1298 analog front-end, enabling local ECG visualization via an OLED display and remote monitoring through Bluetooth connectivity.

This paper presents an ECG data acquisition system designed to monitor cardiac rhythms and store recordings in a local database. The local database implementation demonstrates proof-of-concept functionality for real-time ECG data transmission via HTTP protocols. This foundational work establishes the technical framework for future expansion to cloud-based servers, with the ultimate goal of enabling long-range telemedicine applications for remote cardiac monitoring.

The remainder of this paper is structured as follows: Section 2 details the system methodology, Section 3 presents

[1*] Corresponding author: Sukun THANNI
E-mail: e20201164@itc.edu.kh; Tel: +855-60 907 332

and discusses the experimental results, and Section 4 provides concluding remarks.

2. METHODOLOGY

2.1 System architecture

The system architecture, shown in Fig.1, is designed for collecting data on heart electrical activity, transmitting ECG data to a local database via the Internet, and retrieving data from the Local database to plot the graph in Python. Our system utilizes the ADS1298, an advanced 8-channel, 24-bit analog front-end, to capture ECG signals from the MS 400 simulator. The STM32 microcontroller is responsible for sending initialization commands to the ADS1298 and receiving ECG data at a sampling rate of 250 samples per second via SPI communication. Once the data is processed, the STM32 transmits the formatted ECG data to an ESP32-S3 microcontroller via UART communication.

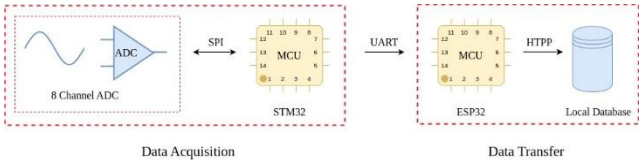


Fig. 1. System Architecture

2.2 Hardware

2.2.1 ADS1298

ADS 1298 is a low-power, 8 channel, 24-bit analog frontend for biopotential measurement. With high levels of integration and exceptional performance, the ADS129x and ADS 1298xR enables the development of scalable medical instrument systems at significantly reduced size, power, and overall cost [6]. The application of ADS1298 include measurement of ECG, EMG, EEG, Patient Monitoring, Holter, Event, Stress, and Vital Signs Including ECG, AED, Telemedicine Bispectral Index (BIS), Evoked Audio Potential (EAP), Sleep Study Monitor [6]. Following are the features of ADS 1298:

- Eight Low-Noise PGAs and Eight High-Resolution ADCs
- Low Power: 0.75 mW/channel
- Data Rate: 250 SPS to 32 kSPS
- CMRR: –115 dB
- Programmable Gain: 1, 2, 3, 4, 6, 8, or 12
- Built-In Right Leg Drive (RLD) Amplifier, Lead-Off Detection, Wilson Center Terminal, Pace Detection, Test Signals
- Built-In Oscillator and Reference
- SPI Compatible Serial Interface

The ADS129x have a highly-programmable multiplexer (mux) that allows for temperature, supply, input short, and RLD measurements. Additionally, the mux allows any of the input electrodes to be programmed as the patient reference drive. The PGA gain is chosen from one of seven settings: 1, 2, 3, 4, 6, 8, or 12. The ADCs in the device offer data rates from 250 SPS to 32 kSPS. Communicate with the device by using an SPI-compatible interface. The device provides four GPIO pins for general use [3].

2.2.2 STM32

The STM32F103C8T6 Blue Pill board is a popular development board based on the ARM Cortex-M3 microcontroller. It is widely used in embedded systems and IoT projects due to its affordability and versatility. Below are the key specifications of the STM32F103C8 Blue Pill board:

- Clock Speed: 72 MHz maximum operating frequency
- Flash Memory: 64 KB (program storage)
- SRAM: 20 KB (data storage)
- GPIO Pins: 37 general-purpose input/output pins
- Communication Interfaces:
    - 2x Serial Peripheral Interface (SPI)
    - 2x Inter-Integrated Circuit (I2C)
    - 3x Universal Synchronous/Asynchronous Receiver/Transmitter (USART)
    - 1x Universal Serial Bus (USB)
- Analog-to-Digital Converter (ADC): 2x 12-bit ADCs with up to 10 channels
- Timers: 4x 16-bit timers and 1x advanced control timer
- Power Supply: 3.3V to 5V input voltage range[7].

The STM32F103C8 Blue Pill board is ideal for applications requiring high performance, low power consumption, and a wide range of peripherals, making it suitable for health monitoring systems, robotics, and automation projects.

2.2.3 ESP32

ESP32-S3-WROOM-2 are powerful, generic Wi-Fi + Bluetooth Low Energy MCU modules that have a rich set of peripherals. They provide acceleration for neural network computing and signal processing workloads. It is used for receiving ECG data from STM32 via UART and transmit the receiving data to local database.

2.2.4 Multiparameter Simulator MS 400

The Multiparameter Simulator MS 400 is a kind of small, lightweight portable device. It can accurately simulate the 12-lead ECG, respiration, temperature and IBP channels-1, 2, 3 and 4. ECG includes about 30 kinds of arrhythmia, normal sinus rhythm, performance testing wave, pacemaker, and ST segment. The simulator generates a variety of respiratory rates, 4 kinds of impedance and apnea of 12, 22 and 32 seconds. The simulator can simulate four kinds of temperature and four-channel IBP. In addition to built-in functions, the simulator can also use the external ECG data.

2.3 Software

2.3.1 STM32CubeIDE

STM32CubeIDE is an advanced development platform for STM32 microcontrollers, combining peripheral configuration, code generation, compilation, and debugging in a single tool. Built on Eclipse®/CDT™ with GCC and GDB, it integrates STM32CubeMX for seamless project setup and code initialization. Users can easily configure peripherals, regenerate code without affecting user logic, and benefit from features like memory usage analysis and stack monitoring. The IDE also offers powerful debugging tools, including register views, live variable watch, and fault analysis, making it a comprehensive solution for embedded system development.

2.3.2 FLASK

Flask is a lightweight and flexible web framework for Python, widely used for developing web applications and APIs. It is designed to be simple and easy to extend, making it ideal for small to medium-scale projects. In this work, Flask is used to create a local server that stores ECG waveform data received from the ESP32 microcontroller. Its built-in development server and routing capabilities enable efficient data handling and communication between the embedded system and the local database.

2.2.3 ESP IDF

ESP-IDF is the official development framework for programming ESP32 microcontrollers. It provides a complete set of tools for firmware development, including drivers, libraries, and a real-time operating system (RTOS) for building reliable and efficient embedded applications. In this project, ESP-IDF is used to program the ESP32 to receive ECG data from the STM32 microcontroller via USART, temporarily store it in a buffer, and transmit it to a local Flask-based server. ESP-IDF supports robust communication protocols, efficient memory management, and integration with Wi-Fi, making it ideal for IoT and medical data acquisition systems.

2.2.4 ADS1298 Software

The ADS1298 software is a dedicated interface designed to configure, control, and visualize data from the ADS1298 analog front-end (AFE), a high-performance, low-noise, 8-channel ECG analog-to-digital converter (ADC). Its primary functions include setting sampling rates, gain adjustments, lead configuration, and real-time signal monitoring. The software enables communication with the ADS1298 via SPI protocol, allowing for precise data acquisition from multiple ECG leads. It also provides tools for noise filtering and signal calibration to ensure medical-grade accuracy. Additionally, the software supports data logging, enabling raw ECG waveform storage for further analysis. In this work, the ADS1298 software serves as a reference to validate the acquired signals against the system's output, ensuring fidelity in the recorded cardiac data for diagnostic reliability.

2.4 System design

2.4.1 Hardware setup

The figure shows the hardware implementation of our ECG signal acquisition and classification system. Since we are not testing with real human subjects, the MS400 simulator is used to generate ECG signals, which are transmitted to the ADS1298 via a 15-pin shielded connector. The ADS1298 does not have a built-in power supply, so it is powered by the MMBO, which is supplied with 6V from a battery pack.

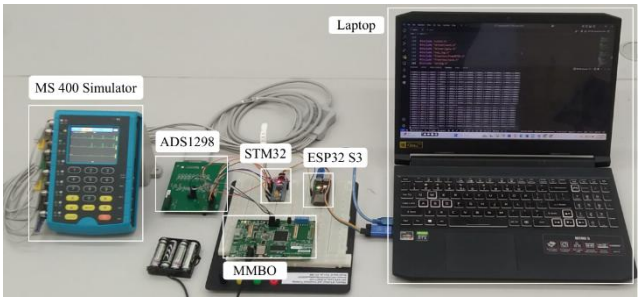


Fig2.Hardware implementation

2.4.2 Data acquisition

The process of reading ECG data can be done as shown in Fig.3. The STM32 microcontroller is responsible for interfacing with the ADS1298 by managing control signals such as START, RESET, DRDY, and SPI communication to read ECG data. After processing the raw ECG signals, the STM32 transmits the data to the ESP32-S3 using UART communication with a baud rate of 460800, ensuring fast and reliable data transfer.

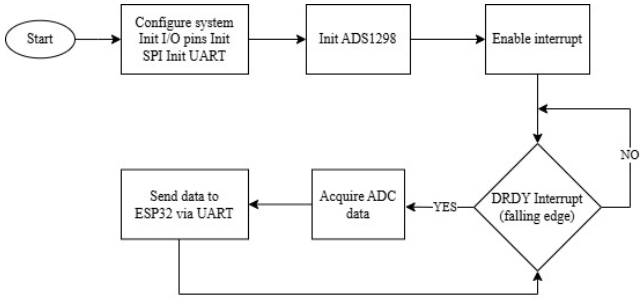


Fig3. Process of reading ECG data

2.4.3 Data transfer

ESP32 collects data from the STM32 and send to the queue. The queue acts as a buffer, ensuring data is not lost if the HTTP is busy. Then the data is format to JSON, and sends it to the database via internet using HTTP.

3. RESULTS AND DISCUSSION

3.1 Results

By conducting the experiment with the circuit connections shown in Fig.2, the system successfully transmits ECG data and store int the database as in Fig.4. The recorded signal is then plotted in the time domain using the stored data, as illustrated in Fig. 5.

Database Structure | Browse Data | Edit Pragmas | Execute SQL

Table: telemetry_data

	time	status	channel1	channel2	channel3	channel4	channel5	channel6	channel7	channel8
1	2025-05-01 02:41:19	9	1F5	5A	2B89	FFE01C	FFED84	FFF99B	FFEB86	FFFE8F
2	2025-05-01 02:41:19	9	128	FFFF84	2AD2	FFE412	FFF4A5	FFFCC1	FFEAEE	260
3	2025-05-01 02:41:19	9	831	5BF	31A3	FFEE5C	6D9	AA4	FFF2BB	88D
4	2025-05-01 02:41:19	9	15BE	1099	3F15	FFFA70	1F31	22D2	43E	E0F
5	2025-05-01 02:41:19	9	288B	1F8E	51DA	64D	3628	3DB1	1B68	1286
6	2025-05-01 02:41:19	9	3D66	2EEE	66AB	F13	4C50	5AD3	3501	1555
7	2025-05-01 02:41:19	9	516A	3DF2	7AD7	E8C	5720	7022	4D4B	1304
8	2025-05-01 02:41:19	9	5C81	4615	85EE	FFFE41	4868	6E60	56A6	71F
9	2025-05-01 02:41:19	9	5819	42B9	8175	FFE137	2358	5871	4C32	FFF3EA

Fig4. ECG data store in local database.

The extracted data from the local database was compared with the original ECG readings obtained from the ADS1298 software, as illustrated in Fig.6. The Mean Squared Error (MSE) method was used to quantify the discrepancy between the two datasets, with the results summarized in Table 1.

Fig5. ECG signal in time domain

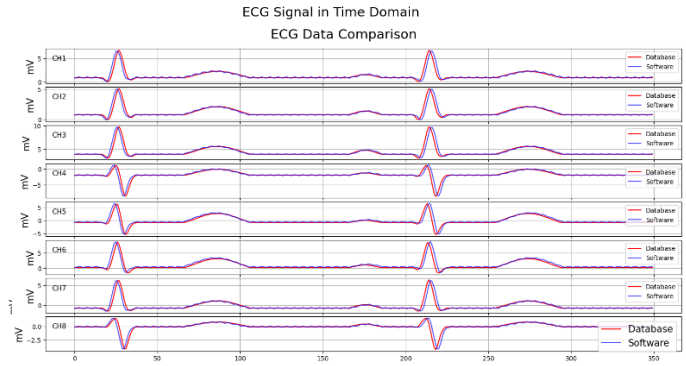


Fig6. Data comparation between the data store in Local database and ADS1298 software.

Table 1. MSE result of the data comparation between the data store in Local database and ADS1298 software.

Method	Ch1	Ch2	Ch3	Ch4	Ch5	Ch6	Ch7	Ch8
MSE	0.15	0.085	0.14	0.30	0.47	0.38	0.22	0.11
Average of MSE	0.23							

3.2 Discussion

The resulting ECG waveform in Fig.5 shows clear, consistent signals across multiple channels, indicating successful data storage and retrieval. The comparison between the local database and ADS1298 software data in Fig.6 reveals a close alignment of waveforms, with minor deviations, suggesting high fidelity in data transmissions. The MSE results in Table 1 quantify this accuracy, with values ranging from 0.085 to 0.47 across channels, all below 1, confirming low error rates. The average MSE across all channels was 0.232, indicating high fidelity in data transmission and storage compared to the original ADS1298 signals.

4. CONCLUSIONS

This paper presents a 12-lead ECG data acquisition system that employ ADS1298 sensor, STM32 and ESP32 microcontrollers. The data are read by STM32 and transferred to ESP32 in order to insert into local database. The measured ECG data are compared and there is small error, that can be acceptable**,** making it suitable for affordable health monitoring in resource-limited settings.

ACKNOWLEDGMENTS

This work is supported by Takahashi Industrial and Economic Research Foundation from April 2024 to March 2025.

REFERENCES

[1] World Health Organization, “cardiovascular diseases (cvds),” https://www.who.int/news-room/fact-sheets/detail/ cardiovascular-diseases-(cvds), 2020, accessed: 15-Feb-2025.

[2] A. H. Association, “Ecg as a diagnostic tool,” Circulation, vol. 140, no. 9, pp. e563–e588, 2019. [Online].Available:https://www.ahajournals.org/doi/full/10.1161/CIRCULATIONAHA.118.036146

[3] Z. Yang, L. Lei, Q. Zhou, K. Zheng, and X. Shen, "An IoT-cloud based wearable ECG monitoring system for smart healthcare," Oct. 2016

[4] J. Granados, T. Westerlund, L. Zheng, and Z. Zou, "IoT platform for real-time multichannel ECG monitoring and classification with neural networks," Shanghai, China, Oct. 2017

[5] R. Kher, B. Thakker, N. Gandhi, and J. Patel, "Ambulatory ECG recording system based on ADS1298 and STM32L431xx microcontroller,"

[6] Texas Instrument “ADS129x Low-Power, 8-Channel, 24-Bit Analog Front-End for Biopotential Measurements datasheet (Rev.K), Texas, America, August 2015.

[7] STMicroelectronics, “STM32F103x8/STM32F103xB: Medium-density performance line Arm®-based 32-bit MCU with 64 or 128 KB Flash, USB, CAN, 7 timers, 2 ADCs, 9 communication interfaces – Datasheet, production data”, accessed May 2, 2025. [Online]. Available: https://www.st.com/

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Enhancing Facial Emotion Classification and Recognition in VR Environments Using Machine Learning Models

Vesal KHEAN [1*], Sokkhey Phauk [1], Ahsan Aziz [2], Yunyoung Nam [3]

[1] *Research and Data Analytics Laboratory, Department of Applied Mathematics and Statistics , Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Department of ICT Convergence, Soonchunhyang University, Asan, 31538, Republic of Korea*
[3] *Emotional and Intelligent Child Care Convergence Center, Soonchunhyang University, Asan, 31538, Republic of Korea*

[*] *Corresponding author: khean.vesal@itc.edu.kh*

Abstract: *Facial emotion recognition plays a pivotal role in human-computer interaction, affective computing, and various other domains. In this paper, we propose an innovative approach to facial emotion classification and recognition by leveraging facial key points captured from a VR device. We utilize machine learning methods, such as Support Vector Machine (SVM) and K-Nearest Neighbors (KNN), to categorize facial emotion using these facial points. The dataset used for training and evaluation is derived from a CSV dataset collected through a virtual reality device. Our findings showcase the value of the suggested approach, having achieved a classification accuracy of 99.1% using the SVM model. This promising accuracy underscores the potential of using facial key points for emotion recognition, offering an efficient means of understanding human emotional states. This research contributes to the development of more natural and intuitive human-computer interfaces, as well as applications in fields such as healthcare, education, and user experience enhancement. Further exploration and refinement of this approach could lead to more accurate and versatile facial emotion recognition.*

Keywords: Virtual Reality, Machine Learning, Emotion Classification, Emotion Recognition

1. INTRODUCTION

Facial expressions play a pivotal role in communication, encompassing both verbal and non-verbal forms of interaction. Non-verbal communication heavily relies on facial expressions, making Facial Expression Recognition (FER) a critical component of robotic vision. Moreover, there is a research paper, using deep learning for facial emotion recognition using the emotion dataset, which includes ten target emotions. Data preprocessing is done to convert video data into images (Agung et al., 2024). Identification of facial emotional expressions is crucial for everyday social functioning. Impairments in facial affect recognition have been found among patients with neurological and psychiatric disorders (Geraets et al., 2021). In recent years, FER has garnered substantial attention from researchers in the fields of computer vision and artificial intelligence, owing to its significant impact on both commercial and academic arenas. Its applications span a wide spectrum, including human-robot interaction, healthcare, augmented reality, road safety, virtual reality, deception detection, and surveillance. Emotion recognition is crucial in cognitive psychology research, but measuring emotions is challenging. Deep learning (DL) techniques have been used for facial expression recognition (FER), but they face challenges like overfitting and complications like occlusion, posture, illumination, and identity bias. This study surveys DL-based methods contributing to FER, analyzing preprocessing, feature extraction, and classification. It explores databases relevant to FER and discusses the current scenario of FER approaches and future directions for facial emotion recognition by machines (Karnati et al., 2023). In this paper, we aim to explore the potential of facial key points for emotion recognition and evaluate the effectiveness of Machine Learning (ML) models such as, Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT). This could benefit areas such as human-computer interaction, healthcare, education, and user experience.

[1*] Corresponding author: Vesal Khean
E-mail: khean.vesal@itc.edu.kh; Tel: +855-96 40 75 885

2. METHODOLOGY

We proposed ML models to classify emotions based on human facial points. There were three significant blocks: preprocessing, feature extraction and ML classification. First, we prepared and clean data such as, labeling data, missing data and data scaling. Second, demonstrated feature extraction technique to extract meaningful features from the facial key points and the last one applied ML models to classify emotions. Each block of the proposed method is briefly discussed in the following in **Figure 1**.

We implemented our emotion recognition system using Python and the scikit-learn library for machine learning tasks. Data was collected from real human subjects using the Meta Quest Pro headset, which captures detailed facial key points in a virtual reality environment. Training and evaluation were conducted on a desktop computer (ID: 1377rq1), which was sufficient for our current scale.

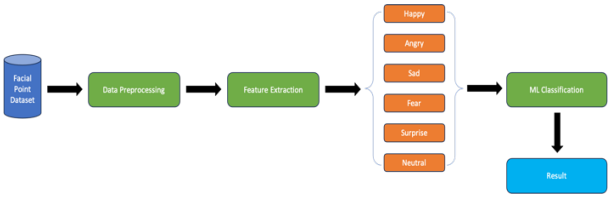


Fig. 1. Our proposed whole flowchart

2.1 Data Preprocessing

Labeling: Assign emotion labels to each data point in the dataset. These labels indicate the emotional expression associated with the corresponding set of facial key points (subject self-reporting).

Data Scaling: This step is essential to ensure that all facial key points have a consistent scale, preventing certain key points from disproportionately influencing the ML model due to their numeric values. Z-score scaling facilitates effective learning based on the relative positions of these key points, regardless of their initial numeric ranges, and is particularly important when applied various ML models.

2.2 Feature Extraction

In the feature extraction step, one effective method for capturing facial expressions is computing the distances between specific facial key points, such as the eyes, nose, and mouth as demonstrated in **Figure 2**. These distances provide meaningful geometric features that reflect changes in facial muscle positions associated with different emotions. For example, the distance between the corners of the mouth can indicate whether a person is smiling. These distances help capture subtle changes in facial geometry, making them powerful indicators for emotion classification using machine learning algorithms.

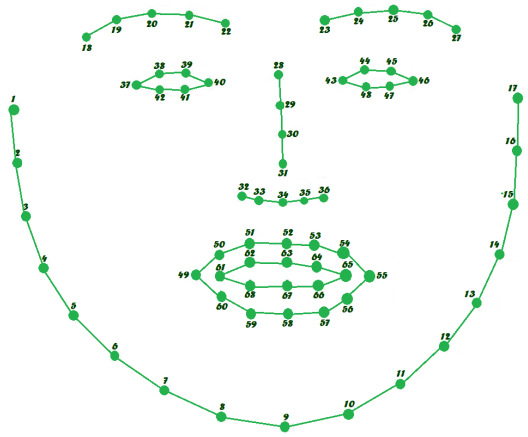


Fig. 2. The 68 facial landmarks

2.3 ML Classification

In this section, begin by preparing a dataset with labelled facial key points. These points represent various expressions for each of the six emotions, including, happy, angry, sad, surprise, fear and neutral. Various ML models were chosen for its effectiveness in multiple classes classification tasks. It's trained on the preprocessed data to learn patterns in the facial key points associated with different emotions including, happy, angry, sad, surprise, neutral, and fear.

3. RESULTS AND DISCUSSION

The dataset collected from a Virtual Reality (Meta QuestPro) device, from student in Soonchunhyang University. In our dataset, we have 8 subjects. We have recorded the emotions for each of the subjects with the following emotions: angry, surprise, sad, happy, neutral, and fear. Then save the facial points in a CSV file for each emotion and each subject separately. Each row within this dataset encapsulates numeric values representing facial points, while the dataset encompasses six distinct emotion categories. This dataset provides a rich source of information for developing and training ML models to recognize facial expressions, offering valuable insights into human emotional states within immersive virtual environments.

In our experiments, we achieved an impressive accuracy rate of 99.1% when applying our machine learning model to the dataset collected from the VR device. This high level of accuracy demonstrates the model's effectiveness in recognizing and classifying emotions based on facial key points. Such results indicate promising prospects for employing this approach in various applications, including

human-computer interaction. The result evaluates metrics of a proposed classification model and confusion metrics of the evaluate metrics can be found in **Table 1**.

Table 1. Summary of ML models classification results

Classifier	Acc	Recall	Precision	F1
KNN	98.2%	99.1%	99.08 %	99.09%
SVM	**99.1%**	99.1%	99.08%	99.09%
DT	68.4%	68.42%	69.8%	69.10%

4. CONCLUSIONS

In this paper, we propose a machine learning-based approach to emotion recognition using a facial key point dataset collected from real human subjects. The dataset was recorded in a controlled virtual reality (VR) environment, where participants were asked to exhibit various emotional expressions including happiness, anger, sadness, fear, surprise, and neutrality. The data, stored in CSV format, contains precise facial landmark coordinates that serve as the foundation for feature extraction. Specifically, we computed geometric features such as Euclidean distances between key points (e.g., between eyes and mouth, or between brows and eyes) to quantify changes in facial structure related to emotional states. We evaluated several classification models and found that the Support Vector Machine (SVM) achieved the best performance, with an average accuracy of 99.1%. The strong performance of SVM highlights its robustness and effectiveness in handling high-dimensional spatial data from real-world subjects. These results demonstrate the strength of our preprocessing and feature extraction strategy, and the model's ability to generalize well across subjects with natural emotional expressions.

This work contributes to the field of affective computing by providing a reliable and interpretable emotion classification pipeline based on real facial behavior. The use of real-subject data increases ecological validity and provides valuable insight for developing emotion-aware applications in VR, human-computer interaction, and digital therapy. In future work, we plan to expand the dataset by collecting data from a larger and more diverse population to improve the model’s generalizability. Additionally, we aim to enhance the system by integrating real-time multimodal signals, such as voice and gesture, and compare our approach with deep learning-based techniques for further performance improvements.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to the UC Lab at Soonchunhyang University for providing the necessary hardware devices and lab materials that were essential for data collection during this research. The support from the lab significantly contributed to the successful completion of the project. I am especially thankful for the well-equipped environment and technical assistance that facilitated smooth experimentation and data acquisition. The availability of these resources allowed me to conduct thorough and accurate work.

I would also like to extend my heartfelt thanks to Professor Yunyoung Nam for their continuous support and guidance throughout the research period. Professor Nam's expert advice, encouragement, and insightful feedback greatly enriched this work.

REFERENCES

[1] Oloyede, M. O., & Hancke, G. P. (2016). Unimodal and multimodal biometric sensing systems: a review. IEEE access, 4, 7532-7555.

[2] Geraets, C. N., Tuente, S. K., Lestestuiver, B. P., van Beilen, M., Nijman, S. A., Marsman, J. B. C., & Veling, W. (2021). Virtual reality facial emotion recognition in social environments: An eye-tracking study. Internet interventions, 25, 100432.

[3] Shan, C., Gong, S., & McOwan, P. W. (2009). Facial expression recognition based on local binary patterns: A comprehensive study. Image and vision Computing, 27(6), 803-816.

[4] Karnati, Mohan, Ayan Seal, Debotosh Bhattacharjee, Anis Yazidi, and Ondrej Krejcar. "Understanding deep learning techniques for recognition of human emotions using facial expressions: A comprehensive survey." IEEE Transactions on Instrumentation and Measurement 72 (2023): 1-31.

[5] Agung, Erlangga Satrio, Achmad Pratama Rifai, and Titis Wijayanto. "Image-based facial emotion recognition using convolutional neural network on emognition dataset." Scientific reports 14, no. 1 (2024): 14429.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Turning hyperparameters in Machine Learning using Simulated Annealing Algorithm

Vinha CHHEANG [1], Sothea HAS [1,2]

[1] *Departments of Applied Mathematics and Statistic, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*
[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: chheang.vinha@itc.edu.kh*

Abstract: *Tuning hyperparameters is essential for optimizing machine learning tasks, whether using Decision Trees, Random Forests, Ensemble models, or building Deep Learning models. However, identifying the optimal hyperparameters remains a significant challenge. Various methods, such as Random Search and Grid Search, have been widely used to explore the hyperparameter space and help identify suitable configurations that enhance model performance. However, these approaches have several limitations, including the prior knowledge of hyperparameter grids and the naïve nature of the searching process. To overcome these challenges, we propose in this study a probabilistic optimization known as* ***Simulated Annealing Algorithm*** *to efficiently search for the best hyperparameters of various Machine Leaning models. The efficiency of the proposed method is illustrated in several numerical experience carried out on difference simulated dataset.*

Keywords: Hyper-parameter, Random Search, Grid Search, Simulation Annealing Algorithm

1. INTRODUCTION

Hyperparameter optimization is a fundamental aspect of improving the performance and generalization capability of machine learning models. Unlike model parameters, which are learned during training, hyperparameters must be set before the learning process begins, and their values can significantly influence the behavior and accuracy of a model. In tasks involving Decision Trees, Random Forests, Support Vector Machines, Ensemble methods, and Deep Learning architectures, selecting appropriate hyperparameters is often the key to achieving optimal performance.

Conventional methods such as Grid Search and Random Search are widely used due to their simplicity and ease of implementation. Grid Search exhaustively evaluates a manually specified subset of the hyperparameter space, while Random Search samples randomly from predefined distributions. Despite their popularity, both methods suffer from critical limitations. Grid Search becomes computationally expensive as the number of hyperparameters increases, and Random Search often fails to efficiently explore the search space, especially when the optimal region is small or sparsely distributed.

To overcome these limitations, probabilistic optimization techniques have emerged as powerful alternatives. In this study, we focus on the Simulated Annealing Algorithm (SAA)—a metaheuristic inspired by the annealing process in metallurgy. Simulated Annealing offers a balance between exploration and exploitation by probabilistically accepting both better and worse solutions, allowing the algorithm to escape local minima and potentially converge to a global optimum.

The main objective of this study is to investigate the effectiveness of Simulated Annealing in hyperparameter tuning across various machine learning models. Through a series of numerical experiments on different simulated datasets, we aim to demonstrate that this method provides a more efficient and flexible approach to finding optimal hyperparameters compared to traditional search techniques.

[1]* Corresponding author: Chheang Vinha
E-mail: chheang.vinha@itc.edu.kh; Tel: +855-96 69 48 937

2. METHODOLOGY

2.1 Objective function

To evaluate the performance of turning hyper-parameter, we define an objective function refer to cost function base on K folds cross-validation. The training dataset is partitioned into K subsets. For each fold k, training set $\mathcal{D}_{train}^{k}$ and validation set $\mathcal{D}_{test}^{k}$ where $\mathcal{D}_{train} = \mathcal{D}_{train}^{k} \bigoplus \mathcal{D}_{test}^{k}$. We define $l_i(\theta)$ be loss function and the model $f(X_i, \theta)$

$$l_k(\theta) = \frac{1}{\| X_i \in \mathcal{D}_{test}^{k} \|} \sum_{X_i \in \mathcal{D}_{test}^{k}} l(y_i, f(X_i, \theta))$$

$$\mathcal{L}(\theta) = \frac{1}{K} \sum_{k=1}^{K} l_k(\theta)$$

2.2. Grid Search

Grid Search is an exhaustive hyperparameter tuning strategy that methodically investigates a predefined set of hyperparameter configurations by assessing the performance of a machine learning model across every possible combination. This approach is especially effective when the hyperparameter space is discrete and of manageable size, as it guarantees that all specified parameter values are systematically considered during the optimization process.

Let Θ be a hyper-parameter space of a discrete grid where each hyper-parameter θ_i.

$$\Theta = [\theta_1 \times \theta_2 \times \ldots \times \theta_p]$$

The optimization can be expressed as:

$$\theta^* = argmin_{\theta \in \Theta}(\mathcal{L}(\theta))$$

Algorithm: Grid Search CV

Function: $GridSearchCV(\theta, \mathcal{L})$:
1. $Grid\ \Theta = [\theta_1 \times \theta_2 \times \ldots \times \theta_p]$
2. **For** $\theta \in \Theta$:
3. $L = \mathcal{L}(\theta)$
4. **End**
5. $\theta^* = argmin_\theta\ (L)$
9. **Return** θ^*

2.3. Random Search

Random Search is a hyperparameter tuning method that involves selecting hyperparameter values randomly from a predefined search space. In contrast to Grid Search, which exhaustively evaluates all possible combinations, Random Search samples points at random, enabling a more efficient exploration of the hyperparameter space.

Let Θ represent the hyperparameter space where each parameter θ_i from the probability distribution p_i:

$$\theta_i \sim p_i$$

Given n random samples from the hyper-parameters space the goal is to fine the optimal parameters θ^* that minimize the loss function can be expressed as:

$$\theta^* = argmin_{\theta \in \Theta}(\mathcal{L}(\theta))$$

Algorithm: Random Search CV

Function: $RandomSearchCV(\theta, \mathcal{L}, \alpha, \beta, T)$:
1. Hyper-parameter space $\Theta = [\theta_1 \times \theta_2 \times \ldots \times \theta_p]$
2. $\theta \longleftarrow A\ initial\ state\ from\ \Theta$
3. **For** $i \longleftarrow range(1, N)$:
4. $\theta_{new} \longleftarrow Random\ (\Theta)$
5. **If** $\mathcal{L}(\theta_{new}) < \mathcal{L}(\theta)$:
6. Update $\theta \longleftarrow \theta_{new}$
7.
8. **Return** θ

2.4. Simulated Annealing Algorithm

To solve optimization problems with large or complex search spaces, the Simulated Annealing (SA) algorithm is often employed due to its ability to escape local minima by accepting worse solutions with a certain probability. The following algorithm outlines the implementation of a simulated annealing approach, where the parameter θ is iteratively updated to minimize the loss function $\mathcal{L}(\theta)$ under a controlled cooling rate of temperature β. To be explore the new neighbors of θ in this study was applied Uniformly Random $\mathcal{U}[a, b]$ where interval $[a, b]$.

Algorithm: Simulated Annealing Algorithm

$\boldsymbol{Function}: SimulatedAnnealingCV(\theta, \mathcal{L}, \alpha, \beta, T)$:
1. $\theta \leftarrow Initail\ State$
2. $\boldsymbol{For}\ T \leftarrow \beta T$:
3. $\quad \boldsymbol{If}\ T = 0\ then\ \boldsymbol{return}\ \theta$
4. $\quad \theta_{new} \leftarrow \mathcal{U}[(1-\alpha) * \theta,\ (1+\alpha) * \theta]$
5. $\quad \Delta E = \mathcal{L}(\theta_{new}) - \mathcal{L}(\theta),\ p \in \mathcal{U}[0,1]$
6. $\quad \boldsymbol{If}\ \Delta E < 0 \quad or \quad p < e^{-\frac{\Delta E}{T}}$:
7. $\quad\quad$ Update $\theta \leftarrow \theta_{new}$
8.
9. $\boldsymbol{Return}\ \theta$

2.5. Hyper-parameter used in this study

In the case study are selected many hyper-parameters in Decision Tree for both Classification and Regression Model.

- **Criterion (criterion):** Defines the metric ('gini' or 'entropy') uses to assess the quality of split for the classification tasks.
- **Minimum Sample Leaf (min_sample_leaf):** Defines the minimum number of samples required to split an internal node which can be defined in integer or fraction of sample.
- **Random State (random_state**): Sets the value of seeds for random number generation for maintaining reproducibility.
- **Maximum Features (max_features):** Specifies the maximum number of features considered when looking for the best split at each node.
- **Minimum Number of Samples Split (min_samples_split):** is the smallest number of observations in the parent node that could be split further.

2.6. Common use hyper-parameters for Decision Tree

The experiments **[2]** were performed considering the hyperparameter tuning of two DT induction algorithms: the 'J48' algorithm a WEKA10 **[3]** implementation of the C4.5 algorithm; and the part implementation of the CART **[4]** algorithm. Experimental methodology of 'J48 ' used to adjust DT hyper-parameters. The tuning is conducted via nested cross-validation: 3-fold CV for computing fitness values and 10-fold CV for assessing performances.

Symbol	Hyper-parameter	Range	Type	Default
cp	complexity parameter	(0.0001,0.1)	real	0.01
minisplit	minimum number of instances in a node for a split to be attempted	[1,50]	integer	20
minbucket	minimum number of instances in a leaf	[1,50]	integer	7
maxdepth	maximum depth of any node of the final tree	[1,30]	integer	30
usesurrogate	how to use surrogates in the splitting process	{0,1,2}	factor	2
surrogatestyle	controls the selection of the best surrogate	{0,1}	factor	0

3. RESULTS AND DISCUSSION

3.1. Hyper-parameter space

In this experiment, we observe that the hyper-parameter results from the Simulated Annealing Algorithm are acceptable, as all values fall within a suitable range. We assume that the grid of the hyper-parameter space has the same size as the space selected by the Simulated Annealing Algorithm. **Table 1** below shows the empirical average values of the hyper-parameters in the Decision Tree model:

Symbol	Range	Type
min_sample_leaf	[2,120]	integer
max_features	[2,20]	integer
min_samples_split	[2,20]	integer

Table 1: Hyper-parameter space

For other two hyper-parameters such that 'criterion' and 'random_state', are assigned for a fix value. In this study we assign criterion= 'squared_error' or 'entropy' for Regression and Classification model respectively. One another hand, for random_state = 42 for both Decision Tree Model.

3.2. Update Loss function

According to the results in **Table 1**, the hyper-parameters span a wide range of values, making it difficult to determine the optimal combination manually. Fortunately, the Simulated Annealing Algorithm is highly flexible in updating to new neighboring hyper-parameter values. As a result, it is both fast and adaptable in converging toward the best set of hyper-parameters. As example for 'Abalone dataset', by split into cross validation 10-CV and 150 iteration of randomly selected new neighbor hyper-parameter with cooling down linear of temperature $\beta = 0.8$.

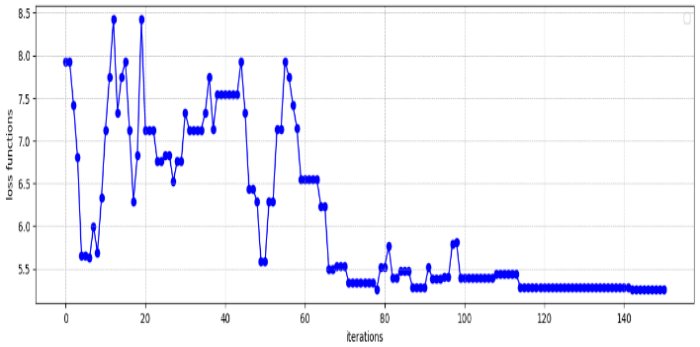


Fig. 1. Loss function MSE with iteration for Regression

We observe that the loss function fluctuates at the beginning due to the random initialization of hyper-parameters. As the number of iterations increases beyond 100, the loss function begins to converge to a smaller value and remains nearly constant after 140 iterations. This indicates that the loss has reached its minimum. Therefore, we chose to apply 150 iterations to find the optimal hyper-parameters using the Simulated Annealing Algorithm.

3.2. Regression Dataset

This table demonstrate the performance of each algorithm such that Grid Search, Random Search and Simulated Annealing Search to turning the hyper-parameter for Regression Tree model. Following the result look like similar the loss function MSE approximate the same for both datasets 'Housing data' and 'Abalone data'. However, the propose method Random Search is gain time compare to two another methods Grid Search and Simulated Annealing algorithm.

Data	Time		
	Grid Search	Random Search	SAA Search
Housing	38s	6.2s	19s
Abalone	6m48.8s	14.2s	45.9s

Data	MSE		
	Grid Search	Random Search	SAA Search
Housing	1.6524	3.5021	2.9908
Abalone	2.3417	2.3159	2.3652

3.3. Classification Dataset

This table illustrate the performance of each algorithm such that Grid Search, Random Search and Simulated Annealing Search to turning the hyper-parameter for Classification Tree model. Base on the result the Accuracy approximate same for both datasets 'Heart Disease data' and 'Wine Quality data'. This result similar with Regression.

Data	Time		
	Grid Search	Random Search	SAA Search
Heart Disease	3m40.6s	6s	21s
Wine Quality	3m18.6s	15.5s	32s

Data	Accuracy		
	Grid Search	Random Search	SAA Search
Heart Disease	0.6721	0.6557	0.7377
Wine Quality	0.5428	0.5265	0.5182

3.4. Discussion

Overall, the Simulated Annealing approach perform well regardless of the initial state of the hyper-parameters, whether it leads to a high or low loss. This is a crucial advantage, as it eliminates the need for prior knowledge of optimal hyper-parameter values. The exploration process depends on two main factors: the cooling rate (β) and the number of iterations. If the algorithm starts from a poor initial state, a slower decrease in temperature allows for broader exploration of neighboring solutions. Conversely, if the initial state is already near optimal, a faster cooling schedule can save time while still refining the solution.

On the other hand, Simulated Annealing contrasts with traditional methods like Grid Search and Random Search, which require predefined ranges and specific values for

tuning hyper-parameters. If the selected hyper-parameter intervals are poorly chosen—falling into less optimal regions—the loss function may not improve significantly. Therefore, having accurate prior knowledge about the appropriate intervals is crucial for the effectiveness of Grid and Random Search. Moreover, for Grid search if we grid is too fine, the search space becomes very large, leading to high computational cost and longer exploration time.

Nevertheless, there are some drawbacks to the Simulated Annealing Algorithm, particularly when dealing with categorical hyper-parameters. For example, in a classification task, if the criterion is set to 'entropy' and not randomly changed to other options (e.g., 'gini'), the algorithm will continue optimizing numerical hyper-parameters based on that fixed categorical choice. If the categorical value is later changed, it can significantly affect the previously optimized settings, causing the loss function to fluctuate and, in some cases, preventing it from converging to a minimal value. This sensitivity highlights a limitation of Simulated Annealing when handling mixed types of hyper-parameters

4. CONCLUSIONS

In this study, we proposed and implemented the Simulated Annealing Algorithm, inspired by discrete search methods, to effectively explore neighboring solutions in the hyper-parameter space. When applied to the Decision Tree model, the algorithm demonstrated strong performance with lower computational time compared to Grid Search. However it take more than Random Search. One of the key advantages of this approach is its flexibility in starting from any initial hyper-parameter values, while still being capable of converging to an optimal solution. This study highlights the potential of Simulated Annealing for enhancing hyper-parameter optimization in machine learning models. Furthermore, future work will focus on adapting and extending this method for deep learning models to handle more complex search spaces efficiently.

REFERENCES

[1]. A. K. Pathak, M. Chaubey, and M. Gupta, "Randomized-Grid Search for Hyperparameter Tuning in Decision Tree Model to Improve Performance of Cardiovascular Disease Classification," arXiv, 2024. [Online]. Available: https://arxiv.org/abs/2411.18234.

[2]. R. G. Mantovani, T. Horváth, A. L. D. Rossi, R. Cerri, S. Barbon Junior, J. Vanschoren, and A. C. P. L. F. de Carvalho, "Better trees: an empirical study on hyperparameter tuning of classification decision tree induction algorithms," Data Mining and Knowledge Discovery, vol. 38, no. 3, pp. 1364–1416, May 2024.

[3]. Witten IH, Frank E (2005) Data Mining: Practical Machine Learning Tools and Techniques, 2nd edn. Morgan Kaufmann, San Francisco.

[4]. Breiman L, Friedman J, Olshen R, et al (1984) Classification and Regression Trees. Chapman & Hall (Wadsworth, Inc.)

[5]. P. Liashchynskyi and P. Liashchynskyi, "Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS," arXiv, Dec. 2019. [Online]. Available: https://arxiv.org/abs/1912.06059.

[6]. A. Askarzadeh, M. R. Haghifam, and M. R. Kamalabadi, "A population-based simulated annealing algorithm for global optimization," IEEE International Conference on Systems, Man, and Cybernetics, 2016, pp. 2461–2466. doi: 10.1109/SMC.2016.784496

.

Materials Science and Structure (MSS)

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Petrography and Petrogenesis of Sandstone at Kulen National Park, Siem Reap Province, Cambodia

LY RATHDANE [1*], NI SOKMENG [1], PECH SOPHEAP [1], ENG CHANDOEUN [1]

[1] *Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[*] *Corresponding author: rathdanely@gmail.com*

Abstract: *Kulen National Park is formed of sandstone, located in Siem Reap province, Cambodia. A Cambodian history, cultural values and ongoing research to reveal new aspects of its past such as archaeological, geological environment, and cultural aspects. For field studies, texture is the most significant property. The proportions of quartz, feldspar, lithic fragment, iron oxide content and matrix support are used to define sedimentary rock at the research area. This research study applies several methods to verify the hosted rock by using Petrography Analysis, and X-Ray Fluorescence (XRF). XRF analysis on oxide minerals to define the sandstone and also provide the mineral composed as well. The result of petrography analysis shows, sandstones are characterized by variable grain sizes, sorting, and rounding, suggesting diverse sedimentary processes. Quartz is the dominant mineral phase, accompanied by feldspars and lithic or rock fragments. There are 3 types of quartz grains: Quartz can occur as monocrystalline grains, polycrystalline grains, microcrystalline grains or called chert. The petrographic data provides important insights into the provenance, transport, and depositional settings of these sedimentary rocks. Laterite exhibits a complex mineralogy, dominated by iron oxides, lithic fragments, cementation, and quartz. Indicate that the laterites formed through intense chemical weathering of the underlying parent rock, influenced by fluctuating groundwater conditions periods of tropical weathering. Based on Petrography and XRF analysis, Sandstone was characterized as "Fe-Sand, Depletion Na, Sublith-Arenite, and Quartz Arenite Sandstone" showing high content of SiO_2 (94.47~49.29%, average 83.03%) content has wide range. The contents for other major elements are as following: AL_2O_3 (6.32~0.79, average 5.22), Fe_2O_3 a sample (29.94~0.10, average 6.48), CaO (0.82~0.01%, average 0.13), TiO_2 (0.29~0.05, average 0.11), K_2O (0.47~0.04, average 0.23), P_2O_5 (9.12~0.02, average 1.45) in various form, 4-7% Feldspar, and 10-15% of Lithic Fragment and Matrix support below 15% and "Weak Latarisation" shows, 53.14% of Quartz, 32.28% of Iron Oxide, and 14.57% of low Aluminum Oxide.*

Keywords: Kulen National Park, Sandstone, Conglomerate, Laterite, Petrography Analysis, X-Ray Fluorescence (XRF)

1. INTRODUCTION

The sandstones within Kulen National Park, situated in Siem Reap Province, Cambodia, represent a significant geological formation with a potentially complex history of deposition and diagenesis. The sandstone formation at Kulen National Park is abundantly of Jurassic Cretaceous Sandstone and its role in the region's historical and cultural development [1] **(Fig. 1)**. It provides a fascinating landscape of geological history, Cambodian history, ancient places such as quarry sites and ancient temples, and cultural value for ongoing research to reveal new aspects of its past such as archaeological, geological environment and cultural aspects. Field observation, most of the study area has similar sedimentary structures such as bedding, cross-bedding, vivid cross-bedding (due to iron oxide expose), ripple and planar cross bedding.

At Kulen National Park, sandstone formations are characterized by diverse lithofacies, include a normal grading bedding, transitioning from coarser grain at the base which are graded upward into progressively finer one. Petrographic analysis reveals variations in mineral composition such as quartz arenite, and sublith-arenite, across different zone within the Kulen National Park. These studies also highlight diagenetic process such as quartz overgrowth, compaction, cementation, and replacement.

The petrogenesis of these sandstones and sedimentary structure observed in the region is linked to understand the depositional environments, including planar cross-bedding, lamination and bedding. Moreover, identifying the diagenetic features, such as cementation by silica, carbonate, or iron oxides, and the presence of replacement textures, will tell the depositional history and the factors that have influenced the sandstone's current physical and chemical properties.

[1*] Corresponding author: Ly Rathdane
E-mail: rathdanelyony@gmail.com; *Tel: +855-86 896 907*

2. GEOLOGICAL SETTING

2.1 General Geology of Cambodia

Cambodia is the part of the large crustal block known as Indosinia. Triassic sedimentary units are widespread in northern and southern of Cambodia with both marine and continental sequences. The Upper Jurassic-Cretaceous Conglomerate and Sandstone are widespread in northern and western of Cambodia, and constitute the south margin of the Khorat Plateau and Kulen National Park in Siem Reap and Preah vihear province. In northern Cambodia, this Jurassic unit is covered by thin, unconsolidated Quaternary deposits this Jurassic sandstone is typically exposed in numerous riverbeds in Siem Reap and Preah Vihear provinces [2]. The study area at Kulen National Park, which is located in Siem Reap province, Cambodia **(Fig. 1)**.

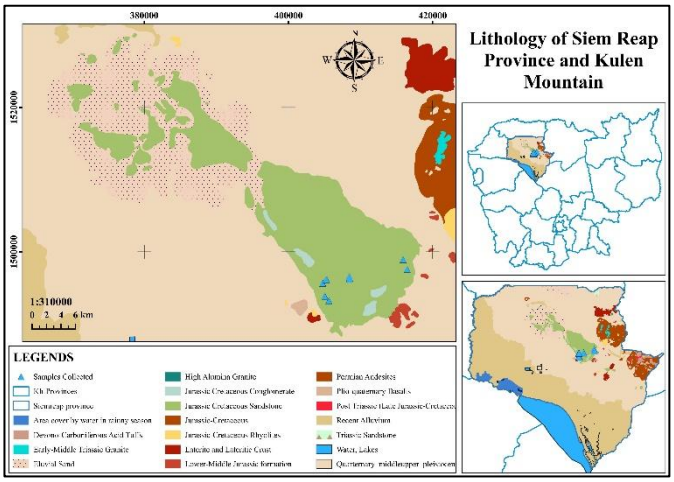


Fig. 1. Study Area is located on Kulen National Park, Siem Reap, Cambodia.

It is considered as an ancient place with many different Cambodian history and cultural value for ongoing research to reveal new aspects of its past such as archaeological, geological environment and cultural aspects. Moreover, the study area has high formation of Jurassic Cretaceous Sandstone **(Fig. 2)**. The physical, chemical and biological properties of sedimentary rocks are strongly influenced by the nature of sediment source area and the conditions of the depositional environment. The characteristics of source areas and depositional environments, in turn, are the result of the tectonic and geologic history of the origin in which the sediments accumulate. For example, volcanic source rocks originate mainly within magmatic arc settings, plutonic igneous rocks are more characteristic of continental block provenances, and metamorphic and sedimentary source rock typically occur in orogenic belts characterized by collision tectonics.

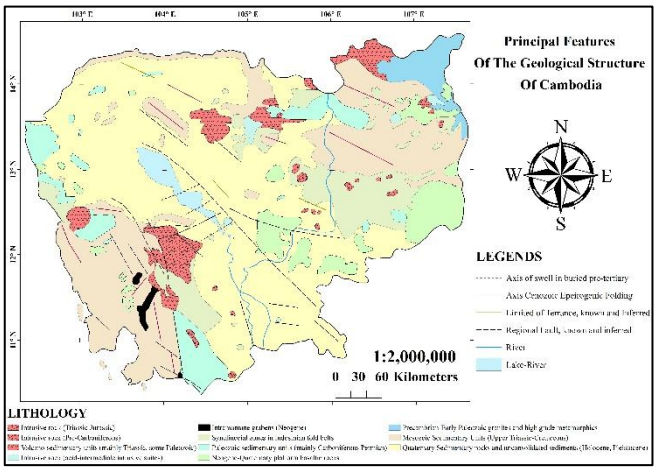


Fig. 2. Principle Features of the Geological Structure of Cambodia [3].

3. SAMPLING AND METHOD

3.1 Petrography

Thin sections were prepared 6 sandstone samples were selected for modal analysis, that is, 2 samples from Peung Vat Chas (VC-3, VC-2), 2 samples from Wat Preah Kral (PK-1, PK-2), 2 samples from Peung Tuk Preah (PTP-1 Bottom, PTP-2 Top). Petrography is a method to analyze samples based on the color observation of minerals, size, shape, orientation under the Optical Microscope. This method provides qualitative information about mineral compositions, texture, sedimentary textures, relative porosity, and diagenetic processes. The thin sections of 6 rock samples taken from the field were prepared using struers specifics resin, blue-dyed powder mixed with curing agent in the ratio 10:1. Petrographic thin sections of sandstones are conducted under Optical Microscope (Model: ECLIPSE CI-POL.B) and determine mineral percentage based on comparison chart estimation [4].

3.2 X-Ray Fluorescence (XRF)

XRF spectrometry is one of the most widely used and versatile of instrumental analytical techniques. An XRF spectrometer uses primary radiation from an X-ray tube to excite secondary X-ray emission from a sample. The radiation emerging from the sample includes the characteristic X-ray peaks of major element present in the sample. The height of each characteristic X-ray peak relates to the concentration of the corresponding element in the sample, allowing quantitative analysis of the sample for the most elements with concentration as low as 1ppm. Approximately 50g of material was cut from each rock sample. The cut pieces were crushed in a tungsten carbide barrel for 2min. Samples were analyzed for major elements using Phhilips PW 1480 automatic XRF spectrometer with Rh-anode X-ray tube [5].

4. RESULTS

4.1 Lithology

3 representative out-crops **(Fig. 7),** shows the presence of conglomerates with polymict clasts within a mud to grain support suggests high-energy fluvial or alluvial environments. Additionally, the study highlights the complexity of the geological landscape, with the presence of laterite formed through intense chemical weathering and the formation as follow (Cohesion: TY Dane, 5th years student);

(A) Peung Vat Chas (VC) - The rock outcrop displays a distinct stratified appearance, characterized by a series of horizontal to sub-horizontal layers that indicate a history of sequential deposition. These layers exhibit noticeable variations in color, ranging from darker reddish-brown hues to lighter yellowish or tan tones, suggesting differences in mineral composition or oxidation states between the depositional events. The thickness of the layers is also variable, with some appearing as thin bands while others form more substantial units. The overall texture of the outcrop is rough and weathered, making it challenging to discern fine details, but there are indications of textural differences between layers, potentially reflecting variations in grain size or compaction. Differential weathering is evident, with some layers protruding slightly due to higher resistance to erosion, while others are recessed, highlighting differences in rock hardness and composition.

(B) Sras Domrei (SDR) - Predominantly sandstone with interbedded gravel conglomerate (angular to sub-angular). There are visible cross-bedded within the sandstone, suggesting deposition in a dynamic environment, likely influences by water or wind. Slightly pink color, highly oxidation caused by weathering process, coarse to medium grains, and poorly sorted.

(C) Kbal Spean (KS) - Sandstone[1]: grain size ranges from very coarse to fine grain, color is from light pink, slightly weathering color is slightly red, strongly weathering. It consists of planar cross bedding, dip of bedding ranges from 20 to 25-degree, ripple mark and massive structure. Sequence of sedimentary is normal graded bedding, transitioning from coarser grain at the base to finer grains toward the top. For Conglomerate[2] (thickness 1-0.3) consist of mix quartz and other fragments of shale (laminated), along with other grain size of pebble in shape from angular to sub-angular, with very poor shorted. Sandstone[3]: thickness approximately 15m, fine grain pinkish, ripple structure, and well sorted.

4.2 Hand specimens Description

Sandstone from various will be describe detail as 6 samples including, VC-3 (Laterite) **(Fig. 4A)**, VC-2 (Iron) **(Fig. 4B)**, PK-2 (Medium) **(Fig. 4C)**, PK-1 (Very Coarse) **(Fig. 4D)**, PTP-2 (Top) **(Fig. 4E)**, PTP-1 (Bottom) **(Fig. 4F)** are the representative sample. The rocks are commonly Sandstone containing mainly of quartz, iron oxide contents, and some of magnetite grain, along with the bedded of conglomerate formation. Grain size are considered as very fine to fine, mostly fine to medium **(Fig. 4A, B, E, & F)**, and some were medium to coarse **(Fig. 4C & D)**, depending on its origin whether its depositional environment is near to the source. There are sparse granular of quartz containing in PK-1 (Very Coarse). Moreover, 5 of these 6 samples were considered as well poorly sorted with grain support and on the other hand, VC-3 was known as iron supported laterite, appear with reddish-orange color

4.3 Petrography Analysis

The rock possibly contains most of quartz mineral (65-95% in some sample) with the accessory of feldspar and lithic or rock fragment (5-10%), mica (<3%) and other type of minerals appear as cemented. The sandstones classified into "Lithic Arenite, Sublith-Arenite, and Quartz Arenite" **(Fig. 5)** and as well as "Week Laterization" **(Fig. 6)** by depending on its mineral content and thin section photomicrograph images **(Fig. 8)**. There are commons mineral of sedimentary rocks such as Various of Quartz grains, Feldspar, Mica, Chert, and Lithic fragment. Key petrographic observations include the identification of quartz grains, feldspar, chert and lithic or rock fragments as well as the presence of cementing materials such as silica and calcite.

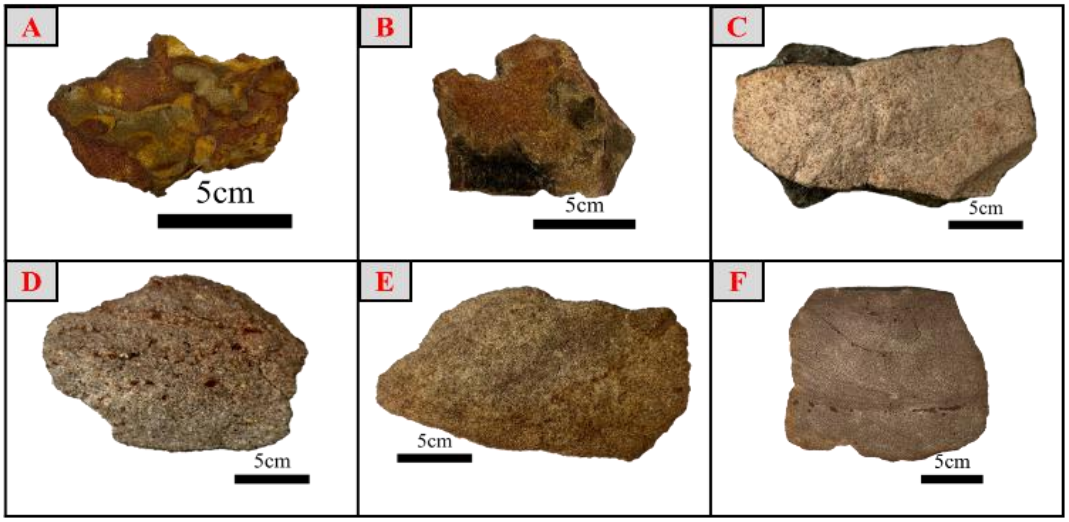


Fig. 4. Photographs of representative sample selection's sandstone classified as similar grain size fine to medium A) VC-3, B) VC-2, E) PTP-2 (Top), F) PTP-1 (Bottom). Selection sample of C) PK-2 is a medium grain and D) PK-1 are very coarse grain.

4.4 X-Ray Fluorescence (XRF) Analysis

Major element content data for the sandstone samples are listed in (**Table 1**). In all the studied samples, SiO_2 (94.47~49.29%, average 83.03%) content has wide range. The contents for other major elements are as following: AL_2O_3 (6.32~0.79, average 5.22), Fe_2O_3 a sample (29.94~0.10, average 6.48).

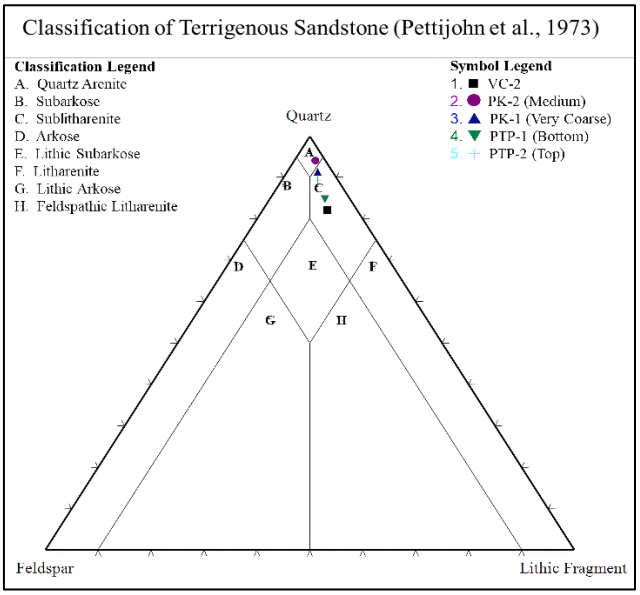


Fig. 5. Classification of Terrigenous sandstone [6], showing different type of sandstone where the matrix support is below 15% as VC-2, PK-1 (Very Coarse), PK-2 (Medium), and PTP-2 (Top) were classify as " Sublith-Arenite" and PK-2 (Medium) was classified as "Quartz Arenite".

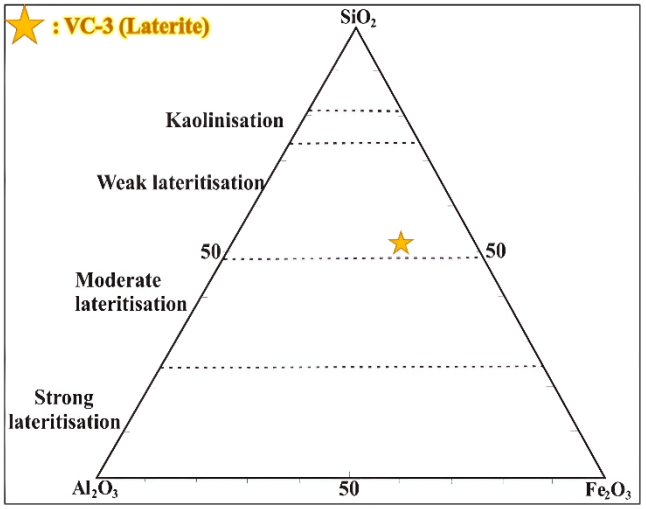


Fig. 6. Triangular diagram of [7] showing the classification of 'laterite' on granite and granitic gneiss into zones of kaolinization and weak, moderate and strong 'lateralization'.

The geochemical classification of discrimination of **(Fig. 9a)** following (**Table 2**) showing the "Sublith-Arenite" for both samples Bat Cave Iron and PT-2 Bottom. PTP-2 Top and VC-3 both also classify into the same "Depletion in Na". For **(Fig.9b)**, "Quartz Arenite" represented for Bat Cave Sandstone. Both of the sample, PK-1 Very Coarse and PK-2 Medium are plotted as "Fe-sand". After calculating the result from XRF, for (**Fig. 6**) VC-3 show as a result "Weak Laterization". For **(Fig. 9a)** Bat Cave Sandstone, PK-2 Medium, PK-1 Very Coarse and **(Fig.9b)** Bat Cave Iron, PTP-2 Top, and VC-3 are not plotted because some of its major oxide concentrations are NA, which is caused the problem, not fit in the Geochemical classification terrigenous sands.

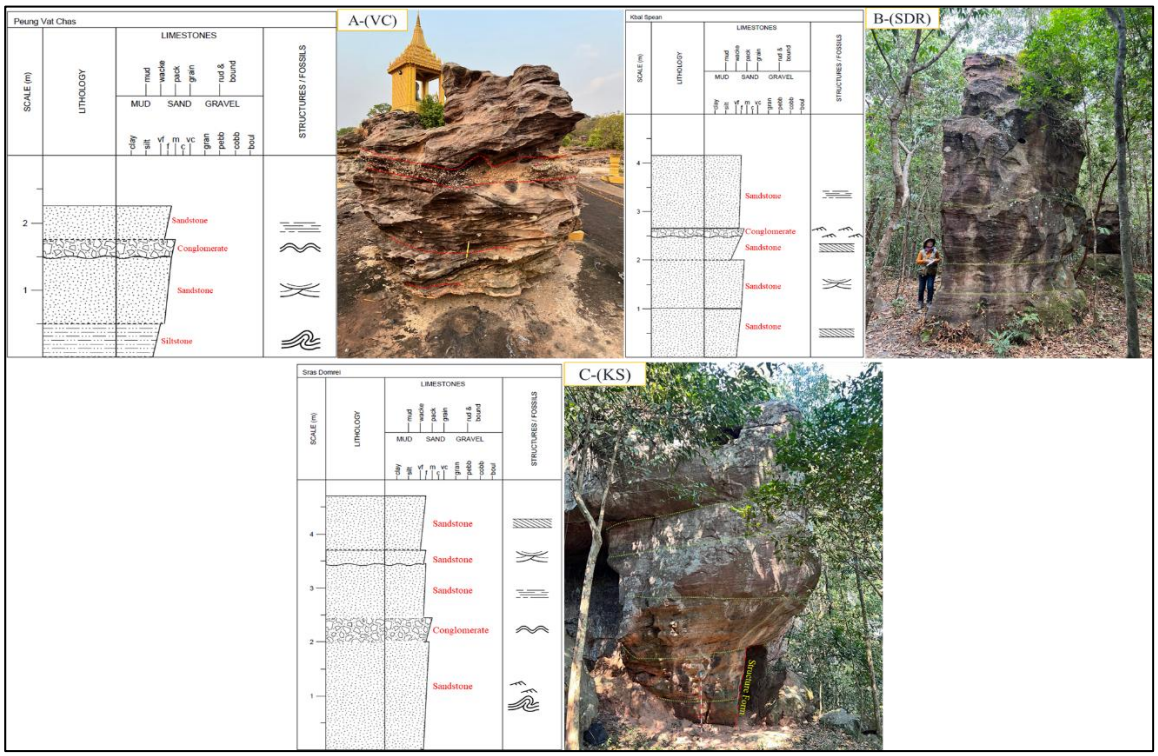


Fig. 7. Photographs of Out-Crops representative at Kulen National Park, showing medium-coarse grained sandstone along with conglomerate and siltstone formation beneath the massive bedded sandstone.

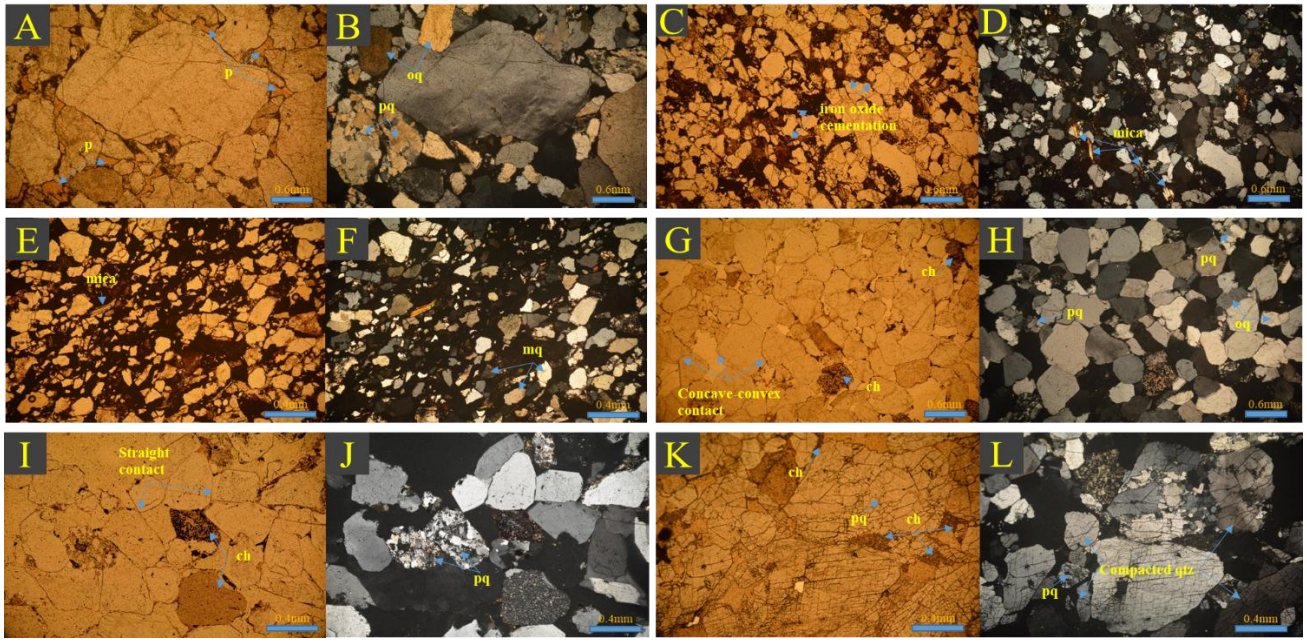


Fig. 8. Photomicrograph representative of sample (PK-1 Very Coarse_A&B), showing three poly-crystalline quartz grain and a single overgrowth quartz. (VC-2 Iron_C&D). Iron Cementation (VC-3 Laterite_E&F). (PTP-02 Top_ G,&H), this image accompanies the sight of well sorted sandstone and lithic sandstone. This photomicrograph of lithic arenite (sandstone) of (PTP-1 Bottom_I&J), and as well as the formation of polycrystalline quartz (PK-Medium_K&L). Abbreviation: qtz=quartz, ch=chert, mq=monocrystalline quartz, pq=polycrystalline quartz, oq=overgrowth quartz, p=pore.

Table 1. Major oxide concentration in mass% of sandstones from Kulen National Park.

Samples	SiO_2	AL_2O_3	Fe_2O_3	MgO	CaO	Na_2O	TiO_2	K_2O	MnO	P_2O_5	Cr_2O_3	LOI	SUM
UNIT	%	%	%	%	%	%	%	%	%	%	%	%	%
ANALYTES													
Bat Cave Iron	84.4319	3.454	8.5917	0.0314	0.8201	0.0162	0.0573	0.0845	N/A	0.722	N/A	2.4938	98.28
Bat Cave Sandstone	81.5844	0.7989	5.3515	0.0175	0.0118	N/A	0.145	0.4735	0.0213	9.1258	0.0105	2.3869	97.57
PK-2 Medium	94.471	4.0961	0.2112	0.0165	0.012	N/A	0.082	0.0468	N/A	0.0574	N/A	0.9639	99.01
PK-1 Very Coarse	94.2798	3.7767	0.3079	N/A	0.0155	N/A	0.0424	0.0482	N/A	0.0483	0.0095	1.4243	98.54
PT-2 Bottom	94.149	4.6255	0.1025	0.0451	0.0194	0.0513	0.1184	0.1239	N/A	0.0294	0.013	0.6539	99.31
PTP-2 Top	90..7581	6.3217	0.8959	0.1449	0.0307	0.0208	0.0913	0.3916	N/A	0.0755	0.0131	1.1875	98.7833
VC-3	49.2908	13.5163	29.9447	0.1302	0.0243	0.0223	0.2996	0.4445	0.0157	0.1364	0.0305	5.9437	93.90

Table 2. Log ratios for SiO2/Al2O3 and Fe2O3/K2O, CaO concentrations (%) for Kulen National Park's Sandstone.

Samples	Log(Na2O/K2O)	Log(Fe2O3/K2O)	Log(SiO2/Al2O3)
ANALYTES			
Bat Cave Iron	-0.7173	2.0072	1.3882
Bat Cave Sandstone	N/A	1.0532	2.0091
PK-2 Medium	N/A	0.6544	1.3629
PK-1 Very Coarse	N/A	0.8054	1.3973
PT-2 Bottom	-0.3830	-0.0823	1.3087
PTP-2 Top	-1.2748	0.3594	1.1571
VC-3	-1.2996	1.8284	0.5619

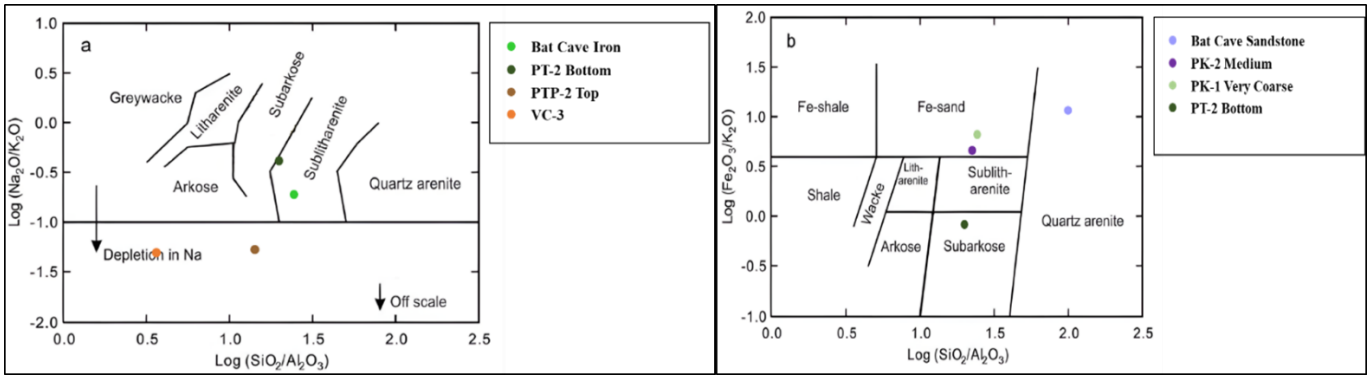


Fig. 9. The formations from which the sandstones were collected are listed below a) Na_2O/K_2O versus SiO_2/Al_2O_3 discrimination diagram [6]; b) Fe_2O_3/K_2O versus SiO_2/Al_2O_3 diagram [8].

5. DISCUSSION

This study of the sandstone formations at Kulen National Park, Siem Reap, Cambodia, has provided valuable insights into the petrography. The photomicrograph image analysis display well sorted grain and composed mainly quartz, monocrystalline quartz, polycrystalline quartz, less quartz cement, iron oxide, chert, mica, and lithic or rock fragment. Quartz is the dominant mineral phase, accompanied by mica, chert and lithic fragments, with the presence of monocrystalline, polycrystalline, and microcrystalline quartz grains suggesting multiple sources [9, 10]The XRF analysis showing a wide range of SiO_2 content and varies concentrations of other major oxide such as AL_2O_3, Fe_2O_3, Na_2O, K_2O, and CaO. The result of discrimination diagram based on the log ratios of Fe_2O_3/K_2O versus SiO_2/Al_2O_3 and CaO value on the sand class, classified sandstone as "Lithic arenite, Sublith-Arenite, and Quartz Arenite" [6, 8]. The sediment admixed with sandstone during diagenesis are geochemically classified using a sand class system [11].

6. CONCLUSIONS

Kulen National Park in Siem Reap, Cambodia, is abundantly of Jurassic Cretaceous Sandstone. The Upper Jurassic-Cretaceous Conglomerate and Sandstone are widespread in northern and western Cambodia, forming the south margin of the Khorat Plateau and Phnom Kulen in Siem Reap and Preah vihear province. The thin section photomicrographs of sandstone sample are composed mainly quartz, monocrystalline quartz, polycrystalline quartz, iron oxide, chert, mica, and lithic or rock fragment. Furthermore, sandstone sample were geochemically analyzed in order to classified lithic arenite, sublith-arenite, and quartz arenite. These may be confirmed to high SiO_2 value of samples. It indicates they are rich of quartz.

ACKNOWLEDGMENTS

I would like to express my gratitude to my Professor Dr. Pech Sopheap and Dr. ENG Chandoeun, for their guide, encouragement and feedback throughout this study.

REFERENCES

[1] H. Hamaguchi, R. Kuroda, N. Okuma, K. Kawabuchi, T. Mitsubayashi, and K. Hosohara, "The geochemistry of tin," *Geochimica et Cosmochimica Acta,* vol. 28, no. 7, pp. 1039-1053, 1964.

[2] E. Uchida and I. Shimoda, "Quarries and transportation routes of Angkor monument sandstone blocks," *Journal of Archaeological Science,* vol. 40, no. 2, pp. 1158-1164, 2013.

[3] D. R. Workman, "Geology of Laos, Cambodia," *South Vietnam and the Eastern Par t of Thailand. Submitted to HM, Stationery Office, United Kingdom,* 1977.

[4] R. L. Folk, *Petrology of sedimentary rocks*. Hemphill publishing company, 1980.

[5] S. Baraka-Lokmane, I. Main, B. Ngwenya, and S. Elphick, "Application of complementary methods for more robust characterization of sandstone cores," *Marine and Petroleum Geology,* vol. 26, no. 1, pp. 39-56, 2009.

[6] F. Pettijohn, P. Potter, and R. Siever, "Sand and Sandstone. Springer-Verlag, Berlin Heidelberg New York," ed, 1972.

[7] W. SCHELLMANN, "A new definition of laterite," *Memoirs of the Geological Survey of India,* vol. 120, pp. 1-7, 1986.

[8] M. M. Herron, "Geochemical classification of terrigenous sands and shales from core or log data," *Journal of Sedimentary Research,* vol. 58, no. 5, pp. 820-829, 1988.

[9] S. W. Young, "Petrographic textures of detrital polycrystalline quartz as an aid to interpreting crystalline source rocks," *Journal of Sedimentary Research,* vol. 46, no. 3, pp. 595-603, 1976.

[10] E. F. McBride, "Quartz cement in sandstones: a review," *Earth-Science Reviews,* vol. 26, no. 1-3, pp. 69-112, 1989.

[11] K. Babu, R. Prabhakaran, P. Subramanian, and B. Selvaraj, "Determination of Sandstone classification by using Geochemical aspect: A Case Study of Garudamangalam Formation, Cretaceous of Ariyalur, Tamilnadu, India."

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Geological Formation at Western Belt Range of S-type Granitoid, Mineralization Series and Feasibility of extracting valuable mineral by using Separation Methods, Kampus Kejuruterran, Universiti Sains Malaysia (USM), Peninsular Northern Malaysia

NI SOKMENG[1,2*], MUHAMMAD NASRUL AMIN BIN MOHD ZAMZURI[2], SITI FATIMAH BINTI RAHIM[2], HASHIM HUSSIN[2], KU ESRYA HANI KU ISHAK, AZRUL[2], ZULKUNIAN[2]

[1] *Faculty of Geo-Resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *School of Materials and Mineral Resources Engineering, Kampus Kejuruteraan, Universiti Sains Malaysia, Nibong Tebal, Pulau Pinang, Malaysia*

[*] *Corresponding author: Sokmeng584@gmail.com*

Abstract: *In the ilmenite-series of magma, tin maybe in the divalent state and has no suitable substitution site in the rock-forming minerals. Tin can accumulate in the residual. It considered that the presence of the ilmenite-series granitoids is a prerequisite in major tin fields. The ilmenite grains in both valleys are not homogeneous and are associated with other heavy minerals that include monazite, zircon, rutile, and cassiterite. Common textures of the ilmenite grains include replacement by leaching, intergrowths and exsolutions. Alluvial sample contaminated of Tin metal associated to Ilmenite mineralization of the study area located in between of main range belt of granite, Pulau Pinang state of Northern Malaysia. Concentration involves the separation of the valuable minerals from the other raw materials. Minerals that are economic value can classify as, metallic and non-metallic, denser or less dense which can extract for commercial use and it involved in the separation has its different properties with different advantages. These properties can be color or size (microscope), density (gravity separation), and magnetism (magnetic separation). The process starts with the coning and quartering method, to reduce the sample size without creating a systematic bias. Sieving method, to determine the particle size and liberate of each size. Spirals, involved in separate the particle according to its density. Magnetic separation, to separate the magnetic and non-magnetic minerals of magnetite and silica. XRF analysis reliable data on the elemental composition, the concentrates from the raw materials, spirals, and magnetic separation run by the grinding methods to achieve a smooth and uniform surface finish, which is essential for many applications. Through non-destructive analysis, these devices enable geologists and mining professionals to swiftly ascertain the percentages of various elements present in a sample. The recovery of spiral concentrator gives the result 86.12 % higher than magnetic separator, which was 64.30 %. The spiral concentrator effectively separates the denser ilmenite and Magnetite particles from the lighter gangue minerals based on gravity, ilmenite and magnetite have higher specific gravity which were 5.5 and 5.3 respectively compared to the silica or quartz has lower specific gravity which was 2.4. The separation between valuable and gangue minerals done easily causing the higher recovery percentage compared to the magnetic separator.*

Keywords: X-ray Florescence (XRF), Mineral Processing, Spiral Separator, Magnetic Separator, Particle Size Analysis (PSA), Microscopic and Petrographic Analysis, Magnetite and Ilmenite-Series Granitoid

1. INTRODUCTION[1]

In the south-eastern Australia the importance of source rock composition and separation of melt and refractory residue in characterizing various suites led to “S” type granite classification. In this scheme, the S-type result from melting and residue of sedimentary or meta-sedimentary source material in which the source was taken to lie within continental crust. The western belt granite of peninsular Malaysia has been regarded to be constituted by exclusively “S” type granites in contrast to the eastern belt of granites which is dominated by “I” type with subordinate “S” type granites. The western belt granites of peninsular Malaysia is characterized by a huge mountain range extending from Malacca in the south to Thailand in the north. The classification by opaque phases mineral was based on the occurrence of Ilmenite and Magnetite in granitic rock. In this classification, magnetite and ilmenite series can be distinguished by their different magnetic susceptibilities; rock of the magnetite series showing high values (more than 100 X 10 emu/g) while those ilmenite series showing lower than 100 X 10 emu/g.

In mineral processing, a major division in the science of Extractive Metallurgy. Extractive metallurgy has been defined as the science and art of extracting metals from their ores through several methods is majority’s process to get high value of concentrate after the process. Concentration means

[1*] Corresponding author: SOKMENG NI
E-mail: Sokmeng584@gmail.com; Tel: +855-88 4697 707

the increase of the percentage of the valuable mineral in the concentrate. Within this research of the experimental study is concentration involves the separation of the valuable minerals from the other raw materials. Alluvial sample contaminated of Tin metal associated to Ilmenite mineralization of the study area located in between of main range belt of granite, Paulu Pinang state of Northern Malaysia.

To identify the recovery of valuable mineral from alluvial sample throughout several methods including mineral characterizations and based on physical processing of hand specimens, coning and quartering, humprey spiral separator, particle size distribution analysis, drying, ore microscope and petrography, magnetic separation, grinding and x-ray florescence (XRF) analysis. The ilmenite grains in both valleys are not homogeneous and are associated with other heavy minerals that include monazite, zircon, rutile, and cassiterite. Common textures of the ilmenite grains include replacement by leaching, intergrowths and exsolutions.

2. GEOLOGICAL SETTING

2.1 Granitoid in Peninsular Malaysia

About ninety granitoids were examined by the magnetic susceptibility meter. Among three major intrusive belts of Permo-Triassic age, the main range granite belt of [1],almost all the Main Range Granites and Eastern belt are predominant in Peninsular Malaysia. Sporadically distributed magnetite-series ones in these belts are generally weakly magnetic, except for one pluton northwest of Malaca.

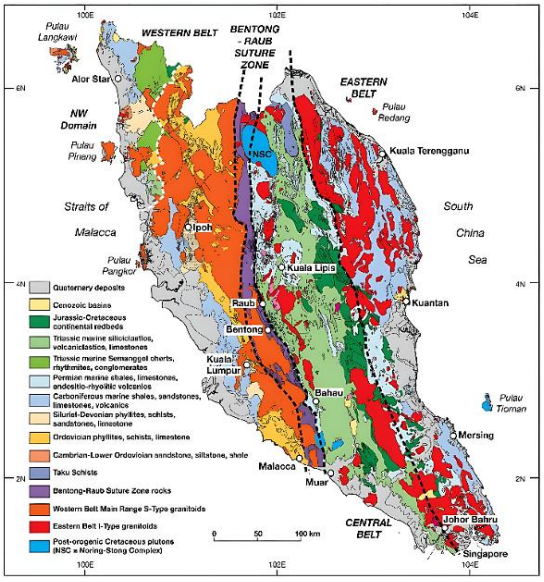


Fig.1. Distribution of the magnetite-series and Ilmenite-series granitoids of the Peninsular Malaysia. Map of the Peninsular Malaysia showing the Western Belt granite in relation to other granite batholiths.

The magnetite-series granitoids seem to occur dorminantly in the Central intrusive belt, although the number of analyses is small (**Fig. 1**). A reported opaque oxide minerals higher than 0.2 vol.% on the Bukit Besar igneous complex, which should belong to the magnetite-series [2]. The retaceous pluton, such as Mount Ophir and Pulau Tioman, are also composed of the magnetite-series granitoids. Thus, aabundance of the magnetite-series or ilmenite-series granitoids in the Peninsular Malaysia is different among three geotectonic units of [1] and depending upon age of the granitoids.

Magnetite-bearing character of the Central Intrusion Belt is in harmony with its island arc nature [1], because typical island arc volcano-plutonic association has magnetite-bearing character [3]. The tendency that magnetite-free rocks predominate in the early Mesozoic Terrance but are scarce in the late Mesozoic one is the same as that found in the southern part of Korean Peninsula [4], which is different from the patterns observed in the Japanese island arc environment.

2.2 S-Type Western Belt

The S-type granite is characterized by restricted SiO2 content between 65% to 75%, which is characteristic of rocks derived from SiO2 rich sources. The granite are low in Na, Ca, and Sr, which are lost during the conversion of feldspar to clay mineral by weathering. High K2O/Na2O is the "S" type rocks is explained by the fact that potassium is incorporated into clay during the chemical weathering to produce sedimentary rocks, whereas sodium is removed in solution along with Ca, Sr, and Pb. The "S" type rocks are significantly lower than "I" type. As a result, of the lower Na and Ca, "S" types are always corundum normative or peraluminous and become more strongly so as the rocks become more primitive.

2.3 Magnetite and Ilmenite Series Granitoid

The significance of opaque oxides of granitoids for tin mineralization was first pointed out by [5], who noted that granitoids accompanying tin deposits hae little magnetite, less than 0.02 wt%, but "tin-barren granitoids" contain an average of 1.26 wt.% of magnetite. A similar conclusion was obtained indepently from studies of granitoids of W-Sn and Mo metallogenic provinces in southwest japan [6], and the genetic beqaring of magnetite in granitoids was discussed in terms of oxygen fugacity of the granitic [3] & [4].

Tin can have either tetravalent or divalent state in magmas [7]. In the magnetite-series magma, tin can be tetravalent and is consumed in common rock-forming minerals, such as sphene, magnetite, ilmenite, epidote-group minerals and hornblende, in the early crystallization stage of magmatic differentiation. Thus, this type of magma cannot have a large concentration of tin in the residual.

In the ilmenite-series of magma, tin maybe in the divalent state and has no suitable substitution site in the rock-forming minerals. Tin can accumulated in the residual. If the original magma in an apical part of the granitic magma. It considered that the presence of the ilmenite-series granitoids is a prerequisite in major tin fields.

The magnetite-series and ilmenite-series granitoids can be identified easily in the field by magnetic susceptibility meter, assemblage of ferro-magnesian silicate, color of biotite and other method [3].

2.4 Relation of Tin Mineralization

Three north-south belts characterize the Malay Peninsula: the Western, Central, and Eastern belts (**Fig. 2**). More than 90% of the plutonic rocks in Peninsular Malaysia are granitic. The granitoids can be divided into two belts, a West Malaya Main Range S-Type group of granitoids that yield Late Triassic to earliest Jurassic, and an eastern Malaya group of dominantly I-Type granitoids with a range of ages from early Middle Permian to early Late Triassic. A significant Late Cretaceous tectono-thermal event affected the Peninsula with major faulting, granitoid intrusion, and re-setting of palaeomagnetic signatures [8].

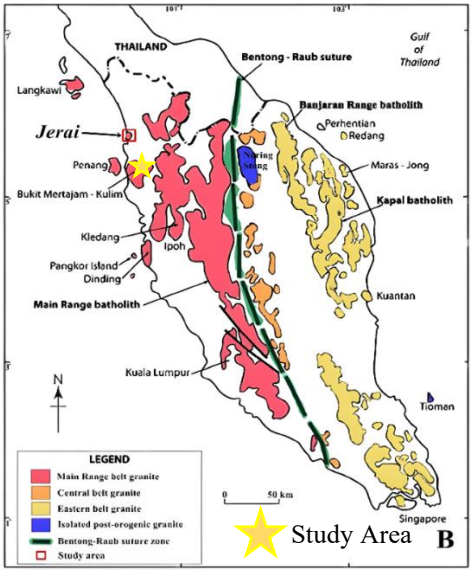


Fig.2. Location area of Ilmenite mineralization with alluvial deposit, northern part of Malaysia.

The study area also falls within the Western Belt of Peninsular Malaysia with a main Range S-Type group of granitoids from the Late Triassic to earliest Jurassic periods. There are conjugate sets of brittle fractures affecting the surrounding plutons, one of which is parallel to the cleavage and may reflect the accommodation of deformation at two different structural depths [9]. Moreover, the geomorphology of the area consists of hilly terrains and undulating plateaus. The geology of the area mainly consists of Quaternary and Devonian granite. Distribution of tin placer deposits was shown in (**Fig. 2**). The tin deposits seem to distributed around the ilmenite-series granitoids. The fact that ilmenite is the most predominant heavy mineral in these deposits supports the above conclusion.

3. METHODOLOGY

For this mineral processing experiment involve in several methods to liberate the valuable and non-valuable minerals from the alluvial feed and to estimate the recovery of the concentration inside the feed and the product. After proceed with each method including Cone and Quartering, Humprey Spiral Separator to separate between Concetration; Middling; and Tailing, Dry process to extract the moisture, Magnetic Separator using magnetic to separate between Silica and Metallic, Particle Size Analysis (PSA), Ore microscopic and Petrography, Grinding in finer molecule, and X-Ray Fluorescence (XRF) analysis. To estimate the recovery of metallic valuable mineral with the formula calculation based on Sieving (PSA).

3.1 Cone and Quartering

The reduction in size of a granular or powdered sample by forming a conical heap which is spread out into a circular, flat cake. The cake is divided radially into quarters and two opposite quarters are combined. The other two quarters are discarded. The process is repeated as many times as necessary to obtain the quantity desired for some final use (e.g. as the laboratory sample or as the test sample). 50 kg of the alluvial sample was given to students to do the experiment, Coning and Quartering is being involved in mixed the sample and after the process is done, 20kg of the sample was taken for the next experiment and the other was discharge.

3.2 Particle Size Distribution (PSD)

Particle size analysis, particle size measurement, is the collective name of the technical procedures, or laboratory techniques which determines the size range, and/or the average, or mean size of the particles in a powder or liquid sample. The particle size measurement is typically achieved by means of devices, called Particle Size Analyzers (PSA), which are based on different technologies, such as high definition image processing, analysis of Brownian motion, gravitational settling of the particle and light scattering (Rayleigh and Mie scattering) of the particles.

Sieve analysis is a process used to separate the particle size in order and after spiral process, the left over piece of the sample of alluvial took to do PSA. Sieving analysis, in order to separate the particle size of our sample such as 3.33mm; 2.8mm; 2.0mm; 1.0mm; 600 µm; 425 µm; 355 µm; 250 µm; 150 µm; 75 µm; and below 75 µm. After the particle size by using sieve-shaking monitor already, it needs to be measuring the sample weight from each tray and after measuring, keep it in a plastic bag and label per each.

3.3 Humprey Spiral Separator

The spiral concentrator is a modern high capacity and low-cost device. It developed for concentration of low-grade ores and industrial minerals in slurry form. It works on a combination of solid particle density and its hydrodynamic dragging properties. The spirals consist of a single or double

helical conduit or sluice wrapped around a central collection column. The device has a wash water channel and a series of concentrate removal ports placed at regular intervals.

Separation achieved by stratification of material caused by a complex combined effect of centrifugal force, differential settling, and heavy particle migration through the bed to the inner part of the conduit. Extensive application is the treatment of heavy mineral beach sand consisting of monazite, ilmenite, rutile, zircon, garnet, and upgrade chromite concentrate. Before putting any sample into the spiral machine, students need to fill-up the reservoir with water below the spiral concentrator about 80-100liter and 20kg of the alluvial sample.

A few minute later, the sample is mixed and it's time to separate the concentration, milling and tailing and after collected the concentration sample, student have to clean up the spiral's machine that contained the left-over of milling and tailing.

3.4 Ore Microscopic and Petrography

The variety of commercially available reflected-light microscopes tends to mask the basic similarities between them in terms of the arrangement of light source, lenses, diaphragms, reflector, objectives, and oculars. Some of this variety is evident in (**Fig. 4**), which shows research and student model microscopes. Each manufacturer incorporates unique design features into the ore microscopes they produce, and it is necessary for the reader to refer to the instruction manual accompanying a particular microscope for the exact placement and employment of the components described below and for information regarding other accessories.

To analyst the type of the ore mineral, there are 6types of sieve sample that need to identify the ore and gangue mineral such as 425 µm, 355 µm, 250 µm, 150 µm, 75 µm and below 75 µm.

3.4 Magnetic Separator

Magnetic separation based on the differing degrees of attraction exerted on various minerals by magnetic fields. Success requires that the feed particles fall within a special size spectrum (0.1 to 1 millimetre). With good results, strongly magnetic minerals such as magnetite, and ilmenite can be removed from gangue minerals by low-intensity magnetic separators. High-intensity devices can separate oxide iron ores as well as iron-bearing manganese, titanium, and tungsten ores and iron-bearing silicates.

After drying the sample in the oven from the spiral process, we took the sample to do the separation of concentration where it could possibly contained Milling and Tailing known as Gangue mineral.

3.5 Grinding Process and X-Ray Fluorescence (XRF) Analysis

XRF analysis need powder sample to analyst, there are 3 samples of plastic bags were taken to do the grinding such as 1 sample from the raw feed (>15kg) and the other 2 samples are after magnetic separation (Concentration and Tailing within >15kg). The process won't take long around 10mn, all the grinding samples are done and need to confirm with the XRF.

The purpose of an XRF analysis is to identify and determine the chemical compositions of recovery of the concentration in feed and concentrate of the product (**Table 3**).

4. RESULT OF THE INTERPRETATION

4.1 Particle Size Distribution (PSD)

The data of the sieving process will accumulate in the table (**Table 1**) as shown. The values of percentage's cumulative passing, which are the diameters that correspond to the percentile of 25%, 50% and 75% respectively, can be determined from the grain-size distribution curve.

Table1. Sieve Data of Particle Size Analysis (PSA).

Particle size, um	Mass (g)	comulative undersize, g	comulative oversize, g	percentage undersize, %	percentage oversize, %
>3350	14.2	1931.05	14.2	99.27	0.73
>2800	17.05	1914	31.25	98.39	1.61
>2000	55.31	1858.69	86.56	95.55	4.45
>1000	402.73	1455.96	489.29	74.85	25.15
>425	777.81	678.15	1267.1	34.86	65.14
>355	120.38	557.77	1387.48	28.67	71.33
>250	224.72	333.05	1612.2	17.12	82.88
>150	240.3	92.75	1852.5	4.77	95.23
>75	85.35	7.4	1937.85	0.38	99.62
<75	7.4	0	1945.25	0.00	100.00
Total	1945.25	8828.82	10623.68		

The grain-size distribution of the ore sample can be obtained by plotting the cumulative percentage passing against the particle size in the semi-log graph paper (**Fig. 3**), Three intersection lines were plotted to determine the particle size percentage and 25%, 50% and 75% respectively. From the (**Fig. 3**), as shown below:

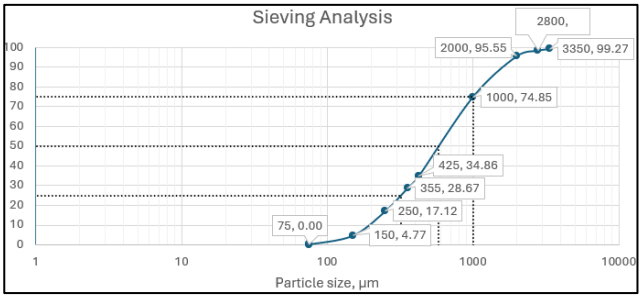


Fig.3. Grain Size Distribution Graph of Mesh size and Cumulative Passing percentage.

4.2 Ore Microscopic and Petrography

The samples were going to observed under microscope based on its size range. There are total 6 samples show the particles with size 425 µm, 355 µm, 250 µm, 150 µm, 75 µm and below than 75 µm (**Fig. 4**). There are three minerals had

been found in the samples which were quartz, ilmenite and hematite. Based on the figure above there are three type of minerals presence in the samples which were hematite, ilmenite, and silica. The characterization of each minerals are shown in (**Table 2**) below:

Table 2. Mineralogy and Mineralization Properties.

Mineral	Chemical Formula	Type of Mineral	Colour
Quartz	SiO2	Gangue Mineral	White
Ilmenite	FeTiO3	Valuable Mineral	Black
Magnetite	Fe3O4	Valuable Mineral	Metallic Gray

Then the particles size above 425 µm had being determine only a few valuable mineral's particle presence. Hence it can be concluded that, the sample which was from the alluvial deposit have the valuable minerals start to liberate from the size particles 355 µm and below.

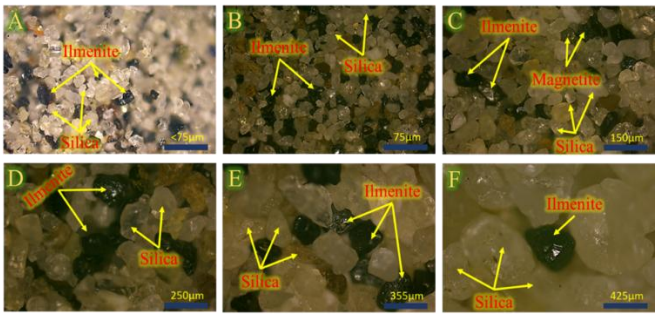


Fig. 4. Petrography Microscopic of representative in mineralogy of Silica and mineralization of Magnetite and Ilmenite inside the alluvial sample. Each of the different title represent for different range of sieve sizes A) below 75 µm, B) 75 µm, C) 150 µm, D) 250 µm, E) 355 µm, and F) 425 µm.

4.3 Spiral Separation

The feed of the ore sample which has weight of 20 kg mixed with 80 litre of water in the reservoir of spiral concentrator give the percentage solid. By the equation of Percentage solid equal to mass of solid (20kg) multiple by 100% and divided by mass of solid (20kg) plus mass of water (80litre), where it equal 20% from its result.

The spiral concentrator operates 20kg of feed in 1.5mn then, the equation formula of feed rate is equaled to mass of solid (20kg) divided by the time taken (1.5mn) equal to 13.33kg/mn.

Based on the feed and concentrate take for 20second, the equation for the mass of feed is equaled to time (20sec) multiple by mass of solid (20kg) and divide by time taken of spiral (1.5mn=90sec) and we receive the number of mass of feed is about 4440g or 4.44kg.

From mass of concentration, we receive about 1052.53g after sieve analysis. Within the XRF analysis on (**Table 2**) below; the Grade of valuable mineral in feed's spiral concentrator = 4.95 + 10.66 = 15.61%. Grade of valuable mineral in concentrate's spiral concentrator = 23.00 + 33.71 = 56.71%. To determine the Recovery of the concentration, it is equal to mass of concentrate (C) 1052.53g multiply by grade of concentrate (c) 56.71% and 100%, divided by mass of feed (F) 4440g and grade of feed (f) 15.61%. The result for the recovery would be 86.12%.

4.4 Magnetic Separation

After the process of the separation by magnetic, the product from the operation, we receive the mass of feed's magnetic separator about 984g, Concentrate's magnetic separator 390g, Tailing's magnetic separator 558 g. Based on the XRF analysis (**Table 2**) below, Grade of valuable minerals from feed = 23.00 + 33.71 = 56.71%. Grade of valuable minerals from concentrate = 37.41 + 51.23 = 88.64%.

The equation of the recovery is equal Concentrate's magnetic separator 390g multiply by Grade of valuable minerals from concentrate about 88.64% and 100% divided by mass of feed's magnetic separator about 984g multiply by Grade of valuable minerals from feed 56.71%. The result of recovery from the magnetic separator equalled to 64.30%.

4.5 X-Ray Fluorescence (XRF) Analysis

The grade of the minerals content in the sample had been measured using XRF machine and type of minerals is set up as oxide minerals. In this analysis only major mineral presence in the samples that will be taken because of the XRF analysis which were silica/quartz, ilmenite and magnetite.

Table 3. Data Interpretation of Ilmenite and Magnetite Concentrate based on X-ray Florescence (XRF) Analysis

CHEMICAL COMPOSITION	SAMPLE RAW (WT %)	SPIRAL'S CONCENTRATE (WT %)	MAGNETIC SEPARATOR'S CONCENTRATE (WT %)
MgO	0.2145	0.0122	0.0045
Al2O3	8.0122	5.0113	1.4653
SiO2	74.9722	37.3166	9.8121
SO3	0.0091	0.0285	0.0137
K2O	0.029	0.019	0.019
CaO	0.7874	0.5274	0.0001
TiO2	**4.9534**	**23.0002**	**37.4143**
Cr2O3	0.0973	0.0983	0.0012
MnO	0.1059	0.0406	0.0006
Fe3O4	**10.6561**	**33.7121**	**51.2311**
NiO	0.0221	0	0
ZnO	0.0012	0.1491	0
SrO	0.0211	0.0215	0.0002
ZrO2	0.0671	0.0611	0

5. DISCUSSION

The geological framework of the western belt range of S-type with magnetite and ilmenite-series granitoid has been reviewed using XRF analysis and Petrography Microscopic Optical observation. The chemical composition of SiO2 content is 74.9722%, restricting between 65%-75%, making it an S-type Granitoid with the appearance of Magnetite and Ilmenite mineralization observed in the microscopic.

Before concentrating valuable minerals from the ore sample, coning and quartering sampling methods are used to ensure even mixing and produce a reference sample for

storage or analysis. Sieving analysis is useful to determine the particle size distribution, with 25% and 75% of samples being medium-sized particles, while the median is below 580 μm. This helps determine the type of deposit based on the physical properties of the ore sample. In mineral observation under microscope or petrography analysis, three types of minerals were identified: quartz, ilmenite, and magnetite. By observing their physical properties, valuable minerals and gangue minerals can be identified more precisely. Ilmenite has black color and metallic or submetallic luster, while hematite has steel-grey color and metallic to dull luster. The unit operation to concentrate valuable minerals can be decided between spiral concentrator and magnetic separator. Both ilmenite and magnetite have similar characteristics, such as heavy minerals with specific gravity of 5.5 and 5.3 respectively, and magnetic particles with strong magnetic characterizations. Separating these minerals from non-magnetic minerals like silica or quartz can affect recovery efficiency.

Spiral concentrators have a higher recovery percentage than magnetic separators, with 86.12% higher recovery than magnetic separators at 64.30%. The spiral concentrator effectively separates denser ilmenite and hematite particles from lighter gangue minerals, such as silica or quartz, based on gravity. The difference in specific gravity between valuable minerals and gangue minerals allows for easier separation, resulting in higher recovery percentages.

However, separating magnetic particles from non-magnetic particles can be difficult due to ilmenite being not too strong magnetic. To ensure accurate picture and color grading of minerals, camera settings such as exposure time, contrast, and brightness need to be set correctly. During gravity separation using spiral concentrators, gangue mineral can enter the concentrate fraction due to mineral particles preventing the gangue mineral from moving to the outer spiral. Lastly, the feed rate of magnetic separators should not be too high, as it may cause magnetic particles to escape from the magnet's attraction to the concentrate fraction.

6. CONCLUSION

The "S" type granite classification in south-eastern Australia is based on source rock composition and separation of melt and refractory residue. This classification is dominated by the western belt granite of peninsular Malaysia, which is characterized by a large mountain range from Malacca to Thailand. The classification is based on the occurrence of Ilmenite and Magnetite in granitic rock, with magnetite series showing high values and ilmenite series showing lower values. This classification helps distinguish between different suites of granites.

Concentration is known as the process to separate the gangue from the valuable mineral in mining industrial. There are three minerals had been found in the samples which were quartz, ilmenite and hematite. Ilmenite is a valuable mineral that associated with the silica mineral know as gangue mineral. Ilmenite also seem to be a very small particle deposited in the alluvial and it contained with Titanium Dioxide (TiO2), where Titanium was one of the naturally metal product used in every industrial. The experiment for the concentration of valuable mineral involved with mineral characterizations based on physical processing of hand specimens, coning and quartering, humprey spiral, particle size analysis, drying, ore microscope and petrography, magnetic separation, grinding and x-ray florescence (XRF) analysis.

AKNOWLEDGEMENT

I would like to express my sincere gratitude and respect to all of the members of the SCHOOL OF MATERIALS AND MINERAL RESOURCES ENGINEERING, which is situated in the Engineering Campus of Universiti Sains Malaysia, including the doctors, technicians, and international friends, for the success of this experiment and research journal on mineral processing. for giving me this significant chance to participate in this project during my spring semester break in 2024 in order to enhance and expand my knowledge and abilities in mineral processing. In addition, I would want to express my gratitude to Dr. Eng Chandoeun, my dean faculty, for providing me with this amazing exchange program experience.

REFERENCES

[1] C. S. Hutchison, "Granite emplacement and tectonic subdivision of Peninsular Malaysia," 1977.

[2] B. Jaafar, "The geology and mineral resources of the Karak and Temerloh areas, Pahang," *Geological Survey of Malaysia,* vol. 15, p. 138, 1976.

[3] S. Ishihara and S. Terashima, "The tin content of the Japanese granitoids and its geological significance on the Cretaceous magmatism," *Jour. Geol. Soc. Japan,* vol. 83, pp. 657-664, 1977.

[4] S. Ishihara, "Metallogenesis in the Japanese island arc system," *Journal of the Geological Society,* vol. 135, no. 4, pp. 389-406, 1978.

[5] P. Aranyakanon, *The cassiterite deposit of Haad Som Pan, Ranong Province, Thailand*. Royal Department of Mines, 1962.

[6] S. Ishihara, "Modal and chemical composition of the granitic rocks related to the major molybdenum and tungsten deposits in the Inner Zone of Southwest Japan," *Jour. Geol. Soc. Japan,* vol. 77, pp. 441-452, 1971.

[7] H. Hamaguchi, R. Kuroda, N. Okuma, K. Kawabuchi, T. Mitsubayashi, and K. Hosohara, "The geochemistry of tin," *Geochimica et Cosmochimica Acta,* vol. 28, no. 7, pp. 1039-1053, 1964.

[8] I. Metcalfe, "Tectonic evolution of the Malay Peninsula," *Journal of Asian Earth Sciences,* vol. 76, pp. 195-213, 2013.

[9] B. Sautter *et al.*, "Late Paleogene rifting along the Malay Peninsula thickened crust," *Tectonophysics,* vol. 710, pp. 205-224, 2017.

THE 14TH SCIENTIFIC DAY OF ITC

Leveraging R&D for Innovation and Growth

Implementation of Reusing Soil Waste from Low Lying Plain as A Backfill of MSE Geogrid Reinforced Vertical Slopes

Chetra HAM [1*], Chanvisal CHAK [1], Puthpisakh CHEA [1], Visal NUTH [1,2,3], Thaileng OENG [1]

[1] *Faculty/Department of Civil Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia*

[2] *Research and Innovation Center, Institute of Technology of Cambodia, Russian Federation Blvd., Phnom Penh P.O. Box 86, Cambodia*

[3] *LEMTA (UMR 7563) CNRS, Université de Lorraine, Nancy, France*

[*] *Corresponding author: hamchetra18@gmail.com*

Abstract: *Soil waste management is a critical concern in road construction, particularly with the rapid infrastructure development in Cambodia. One major issue is slope stability during pavement construction. This research aims to evaluate the effectiveness of lime and lime-cement stabilization techniques in enhancing the strength and stability of soil waste for slope reinforcement. The study involves preparing soil samples with varying contents of lime (optimal lime plus 1%, 2%, and 3% additional increments) and lime-cement (optimal lime combined with 1%, 2%, and 3% cement by dry mass) to assess the performance of different mixture compositions. The physical and mechanical properties of the treated soil are examined through laboratory tests, including the standard Proctor test, unconfined compressive strength (UCS) test, direct shear test, and permeability test. Additionally, simulations using PLAXIS 3D are conducted to analyze slope stability under different conditions and evaluate the performance of stabilized soil waste. The model geometry follows conventional slope construction design, with structural modifications based on stabilization requirements. Soil characteristics are derived from experimental data to reflect real-world conditions. Results indicate that both stabilization methods enhance soil cohesion and friction angles. Cohesion is influenced by the mixture composition, while friction angles are more dependent on soil type. Although lime-cement treatment increases UCS, the dry density of lime-cement-treated clay soil is lower than that of lime-treated soil. Moreover, steeper slope angles contribute to a lower safety factor. A case study will be presented to illustrate soil reinforcement challenges and the effectiveness of various lime and lime-cement stabilization techniques at a real-world infrastructure site.*

Keywords: slope stability; lime and lime-cement stabilization; optimal; PLAXIS 3D; soil waste

1. INTRODUCTION

While soil waste management in Cambodia is not yet perceived as a major issue, like any other developing country, as Cambodia continues to experience urbanization and economic development, it is essential to take proactive measures to address soil waste before it becomes a critical challenge [1, 2].

This paper proposed the reusing of excavated soil from the same site with the same moisture content as Mechanically Stabilized Earth (MSE) backfill, particularly addressing the weak soil conditions and the potential for stabilization using lime and lime-cement. The research's objective is to study the effectiveness of lime and lime-cement stabilization techniques in enhancing the strength and stability of soil waste for slope reinforcement.

Soil stabilization with lime has been proven effective in increasing cohesion, friction angle, compressive strength, plasticity, and workability of cohesive soils [3, 4]. Similarly, lime-cement stabilization involves adding cement in conjunction with lime, which can further improve the stabilization process by introducing a hydraulic bonding mechanism [5]. The combination of lime and cement led to a significant improvement in the CBR values compared to using lime or cement individually [6]. Likewise, the compacted dry densities of the lime-cement were lower than stabilized with cement, but durability did not affect [7].

[1*] Corresponding author: Chetra HAM
E-mail hamchetra18@gmail.com; Tel: +855-78 24 90 44

This case study tackles a slope along a segment of National Road No. 4 (PK 101+000 - PK 143+000) located in Ti Prammuoy village, Yeay Mao Pich Nil commune, Phnom Sruoch district, Kampong Speu Province, as shown in fig. 1.

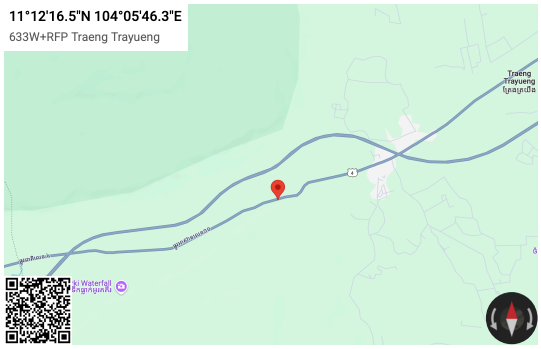


Fig. 1. Location of the study site (Google Map)

2. METHODOLOGY

The methodology consists of two analyses. The lab experimental to obtain the soil characteristic and the simulation to analyze the soil behavior in the large-scale scenario.

2.1 Experimental analysis

The study involves preparing soil samples with different lime contents (optimal lime plus additional increments of 1%, 2%, and 3%) and lime-cement mixtures (optimal lime combined with 1%, 2%, and 3% cement by dry mass) to evaluate the performance of various mixture compositions. Laboratory tests, such as the standard Proctor test, unconfined compressive strength (UCS) test, direct shear test, and permeability test, are performed to assess the physical and mechanical properties of the treated soil.

2.2 Numerical analysis

Once the soil parameters had been determined, the slope stability analysis was conducted using the finite element method with PLAXIS 3D. PLAXIS [8] approaches the shear strength parameters, which are gradually reduced until the structure fails. It can be written as

$$SF = \frac{\tan \varphi_{input}}{\tan \varphi_{reduced}} = \frac{c_{input}}{c_{reduced}} \qquad \text{(Eq. 1)}$$

The strength of interfaces and structural components (plates and anchors), if considered, is reduced in the same manner. Brinkgreve and Bakker [9] found this strength reduction method (phi c-reduction) produced safety factors similar to those obtained from traditional stability analysis using the Limit Equilibrium Method (LEM) and slip-circle analysis.

The modeling process is shown in fig. 2. The Mohr-Coulomb and Hoek-Brown models were applied respectively to soil and rock material. The analysis was performed under dry conditions. The parameters for this model are provided in table 1.

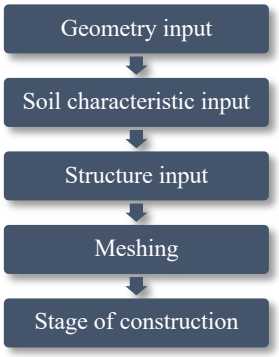


Fig. 2. PLAXIS 3D input workflow

Table 1. Parameters required for modeling in PLAXIS 3D

Properties	Name	Unit
General		
Unsaturated unit weight	γ_{unsat}	kN/m^3
Saturated unit weight	γ_{sat}	kN/m^3
Initial void ratio	e_{init}	-
Criterion: Mohr-Coulomb		
Young's modulus	E'_{ref}	kN/m^2
Poisson's ratio	ν(nu)	-
Cohesion	c'_{ref}	kN/m^2
Friction angle	φ	deg
Dilatancy angle	ψ	deg
Criterion: Hoek-Brown		
Rock Mass Modulus	E_{rm}	kN/m^2
Poisson's ratio	ν(nu)	-
Uniaxial compressive strength	$\lvert\sigma_{ci}\rvert$	kN/m^2
Material constant	m_i	-
Geological strength index	GSI	-
Disturbance factor	D	-

2.2.1 Analysis of the Existing Structure

Since the soil investigation report provided unjustified

Table 2. Available soil/rock parameters

Properties	Name	Unit
Dry density	ρ_D	kN/m^3
Bulk density	ρ	kN/m^3
Specific gravity	G_s	-
Voids Ratio	e	-
Plasticity Index	PI	-
Clay content	$M\&C$	%
Rock-quality designation	RQD	%

and inconsistent data, the empirical formulation was adopted from the properties of clay soil and sandstone that were available in table 2.

Parameter estimation

Unit weight

Assuming that $\gamma_{unsat} = \rho$, and $\gamma_w = 10\ \mathrm{kN/m^3}$, then

$$\gamma_{sat} = \frac{G_s + e}{1 + e} \cdot \gamma_w \qquad \text{(Eq. 2)}$$

Parameters of clay soil

According to Tchakalova and Ivanov [10], cohesion given in eq. 3 appeared to be appropriate to a range of clays with a PI of 11% to 78%,

$$c = 0.776PI + 0.8476 \qquad \text{(Eq. 3)}$$

For friction angle, Winn *et al.* [11] found it decreased with increasing clay content with a value of φ obtained mostly at 20° and 37°, it was expressed as

$$\varphi = 0.14(234.5 - \%\mathrm{Clay}) \qquad \text{(Eq. 4)}$$

The value of Young's modulus for clay is from table 3 where its description base on cohesion (Table 4) and PI (Table 5).

Table 3. Typical values of Young's modulus (MPa) for cohesive material

USCS	Description	Very soft to soft	Medium	Stiff to very stiff	Hard
ML	Silts with slight plasticity	2.5 - 8	10 - 15	15 - 40	40 - 80
ML, CL	Silts with low plasticity	1.5 - 6	6 - 10	10 - 30	30 - 60
CL	Clays with low-medium plasticity	0.5 - 5	5 - 8	8 - 30	30 - 70
CH	Clays with high plasticity	0.35 - 4	4 - 7	7 - 20	20 - 32
OL	Organic silts	-	0.5 - 5	-	-
OH	Organic clays	-	0.5 - 4	-	-

Obrzud and Truty [12] after from Kezdi [13], and Prat *et al.* [14]

Table 4. Classification of cohesion values for soil

Soil condition	Very soft	Soft	Medium	Stiff	Very stiff
Cohesion (kPa)	< 24	24-48	48-96	96-192	192-383

After Rendana *et al.* [15] extract from Terzaghi *et al.* [16]

Table 5. Plasticity index in a qualitative classified by Burmister [17]

PI	Description
0	Non-plastic
1-5	Slightly plastic
5-10	Low plasticity
10-20	Medium plasticity
20-40	High plasticity
>40	Very high plasticity

Dilatancy angle for clay, $\psi \approx 0$, as Houlsby [18] explained. Turk and Dearman [19] estimated Poisson's ratio by

$$\nu = (1 - \sin\varphi)/(1 + \sin\varphi) \qquad \text{(Eq. 5)}$$

Parameter of sandstone

According to Hoek [20], the material constant for sandstone is 15. Considering the slope construction with excavation machinery, $D = 0.7$ was chosen [21].

Hoek *et al.* [22] considered

$$GSI = \begin{cases} RMR - 5, & RMR > 23 \\ RMR, & RMR \le 23 \end{cases} \qquad \text{(Eq. 6)}$$

where RMR is the Rock Mass Rating developed by Bieniawski [23]. By not having access to the other properties, the only approach to determine RMR was direct comparison with RQD (Table 6).

Raj and Pedram's [24] correlation between direct and indirect strength test methods of sandstone is

$$E = 2.17(UCS)^{0.57} \qquad \text{(Eq. 7)}$$

where UCS was estimated from table 6.

Table 6. Relation between RQD, UCS, and RMR

Notation	RQD (%)	UCS (kPa)[1]	RMR (%) [2]
Excellent	100	28,000	100
Good	90	19,000	80
Passable	75	11,000	60
Bad	50	6,000	40
Very bad	25	2,800	20
-	0	900	0

[1] after Prakash and Sharma [25]
[2] after Bieniawski [23]

PlAXIS 3D modeling

The study simulated the structure in fig. 3. The structure's model and generation of mesh are shown in fig. 4.

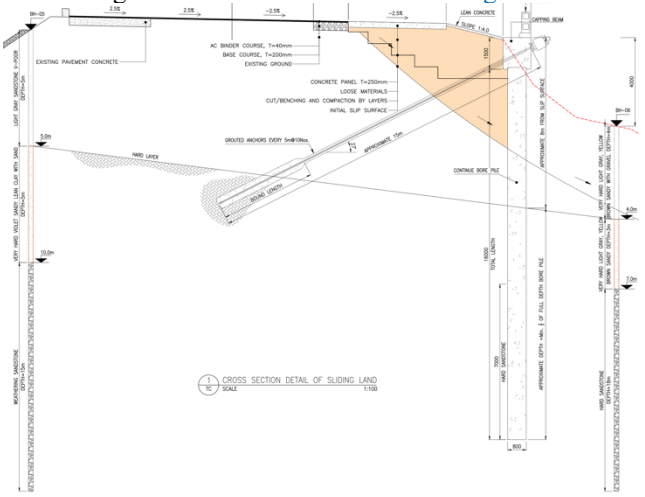


Fig. 3. Cross-section drawing of the studied structure

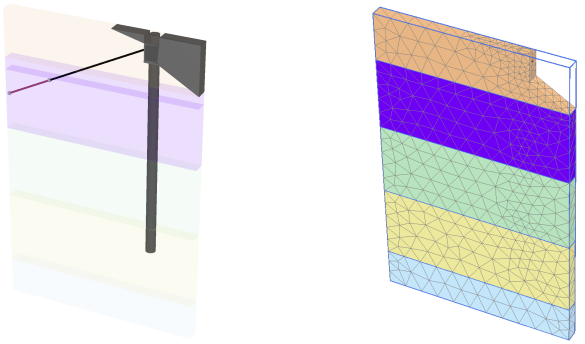


Fig. 4. Structural model and mesh generation

2.2.2 Analysis of the Redesigned Slope

The continuous bore piles and ground anchors were not considered. The redesign approached various slope steepness (See fig. 5) under the untreated and treated soil in which 4m of the existing soil was replaced.

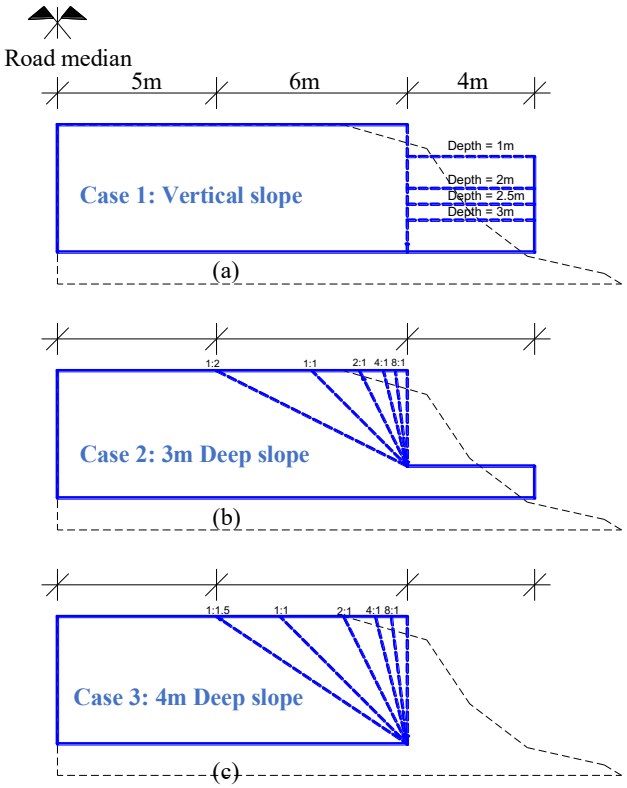


Fig. 5. Each case of the redesigned slope: (a) Case 1: Vertical slope, (b) Case 2: 3m Deep slope, (c) Case 3: 4m deep slope

3. RESULTS AND DISCUSSION

After analysis, the lab testing and the safety factor analysis of each case is shown and compared as following.

3.1 Experimental analysis

Experimental analysis has shown a significant increase in Young's modulus, cohesion, and friction angle for the treated soil.

3.2 Analysis of Existing Structure

Analysis shown $SF = 2.6$ from the existing structure with its failure surface in fig. 6. Without the ground anchor and continuous bore pile, it gave $SF = 2.5$ and 2.2, respectively.

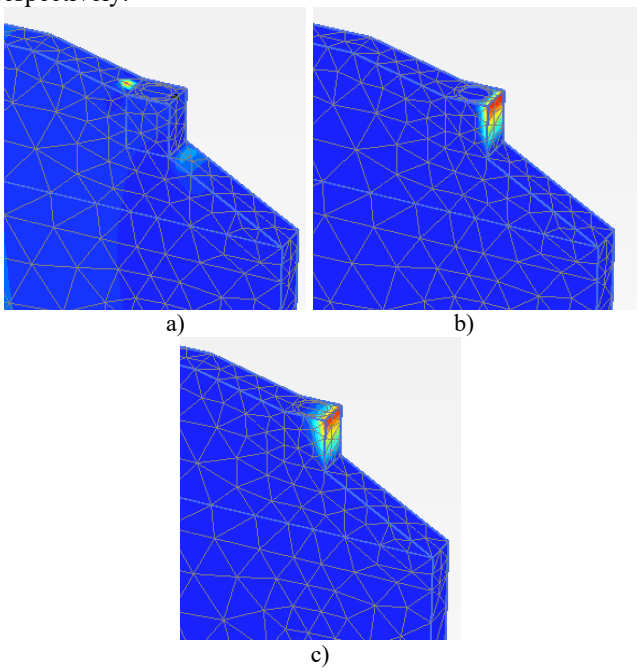


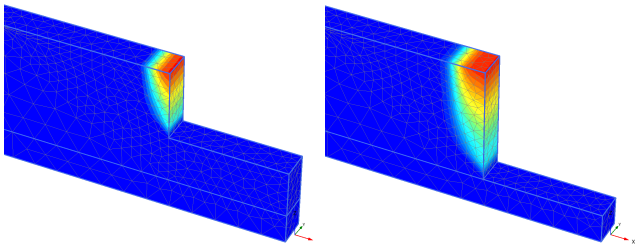
Fig. 6. Failure surface of the structure: a) ordinary, b) without ground anchor, c) without CBP

3.3 Analysis of untreated soil

The safety factors of all cases are shown in table 7 whereas fig. 7-9 show the failure surface of selected cases.

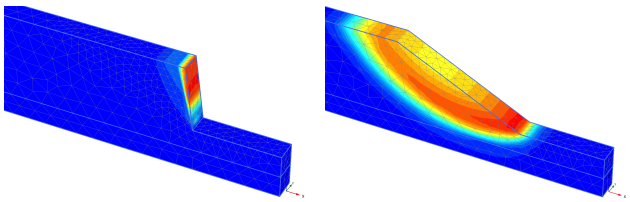
Table 7. PLAXIS 3D's redesigned slope safety factor

Case 1					
Vertical Depth	1m	2m	2.5m	3m	
Safety Factor	4.5	2.2	1.4	fail	
Case 2					
Slope Ratio	8:1	4:1	2:1	1:1	1:2
Safety Factor	1.5	1.7	2.2	2.9	3.7
Case 3					
Slope Ratio	8:1	4:1	2:1	1:1	1:1.5
Safety Factor	1.2	1.4	1.8	2.3	2.9



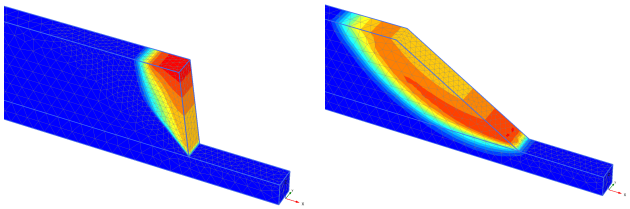
a) b)

Fig. 7. Case 1 Failure surface: a) Depth = 2.5m, b) Depth = 3m



a) b)

Fig. 8. Case 2 Failure surface: a) slope ratio 8:1, b) slope ratio 1:2



a) b)

Fig. 9. Case 3 Failure surface: a) slope ratio 8:1, b) slope ratio 1:1.5

4.3 Analytical discussion

The experimental and numerical analysis conducted in this study demonstrates that both lime and lime-cement stabilization methods significantly improved the mechanical properties of soil, such as Young's modulus, cohesion, and friction angle. The soil treated with lime-cement showed the highest increase in unconfined compressive strength (UCS), though the dry density of lime-cement-treated soil was slightly lower than that of lime-treated soil. This suggested that while lime-cement enhances soil strength, it may lead to a reduction in compactness, which could have implications for certain construction applications, particularly those requiring high-density materials.

Moreover, steeper slopes with untreated soil tended to exhibit a lower safety factor, highlighting the importance of soil treatment in ensuring the stability of high-angle slopes. The case study highlighted in the results demonstrated that soil stabilization methods are necessary for enhancing the overall safety and longevity of slopes in road construction.

4. CONCLUSIONS

This study successfully demonstrated the effectiveness of lime and lime-cement stabilization in improving the mechanical properties and stability of soil waste used for slope reinforcement. The results indicate that both methods enhance cohesion and friction angles, contributing to improved slope safety factors. While lime-cement treatment showed promising results in terms of UCS, further investigation is needed to refine the stabilization methods for specific soil types and slope conditions.

These improvements are particularly relevant in the context of slope stabilization for infrastructure projects in Cambodia. As Cambodia continues its infrastructure development, the use of locally sourced soil waste combined with stabilization techniques could significantly reduce construction costs and improve the sustainability of road projects.

Future research should focus on the long-term durability of stabilized soils under varying climatic conditions and explore alternative stabilizing agents that could further enhance soil properties while maintaining environmental sustainability.

ACKNOWLEDGMENTS

The authors appreciate the financial support of JAPENEL HOME (CAMBODIA) Co., LTD., the technical advising of construction team from Yonemori Construction Company, Prof. Kazunari SAKO (Kagoshima University), and Prof. Hossein NOWAMOOZ (INSA Rennes).

REFERENCES

[1] F. Lord, "Transformation to sustainable and resilient urban futures in Southeast Asia," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 6, pp. 43-50, 2020.

[2] PPCA, IGES, Nexus, UN-Environment and CCCA, "Phnom Penh Waste Management Strategy and Action Plan 2018-2035," Phnom Penh Capital Administration, Phnom Penh, Cambodia, 2018.

[3] D. N. Little and S. Nair, Recommended Practice for Stabilization of Subgrade Soils and Base Materials, Washington, DC: The National Academies Press, 2009.

[4] J. Szendefy, "Impact of the soil-stabilization with lime," in *Proceedings of the 18th International Conference on Soil Mechanics and Geotechnical Engineering*, Paris, 2013.

[5] S. O. Manzoor and A. Yousuf, "Stabilisationof Soils with Lime: A Review," *Journal of Materials and Environmental Science*, vol. 11, no. 9, pp. 1538-1551, 2020.

[6] D. VENKATESWARLU, M. KUMAR, G. P. RAJU and D. PRASAD, "A study on the lime-Cement stabilized marine clay," *Asian Jr. of Microbiol. Biotech. Env. Sc.*, vol. 16, no. 2, pp. 439-444, 2014.

[7] A. Ullah, D. Boumezerane and F. Ahmad, "Subgrade improvement with mixed lime and cement as additives," in *Materials Today: Proceedings*, 2023.
[8] PLAXIS, Reference Manual 3D, Bentley, 2024.
[9] R.B.J Brinkgreve and H.L. Bakker, "Non-linear finite element analysis of safety factors," in *Proc. 7th Int. Conf. on Comp. Methods and Advances in Geomechanics*, Cairns, Australia, pp. 1117–1122, 1991.
[10] B. Tchakalova and P. Ivanov, "Correlation between effective cohesion and plasticity index of clay," *GEOLOGICA BALCANICA*, vol. 51, no. 3, pp. 45-49, 2022.
[11] K. Winn, H. Rahardjo and S. Peng, "Characterization of Residual Soils in Singapore," *Journal of the Southeast Asian Geotechnical Society*, pp. 1-13, 2001.
[12] O. R. and A. Truty, "HE HARDENING SOIL MODEL - A PRACTICAL GUIDEBOOK Z Soil.PC 100701 report," Zace Services Ltd, Software engineering, Préverenges, 2018.
[13] A. Kezdi, Handbook of Soil Mechanics, Amsterdam: Elsevier, 1974.
[14] M. Prat, E. Bisch, A. Millard, P. Mestat and G. Cabot, La modelisation des ouvrages, Paris: Hermes, 1995.
[15] M. Rendana, W. M. R. Idris, S. A. Rahim, Z. A. Rahman and T. Lihan, "Rubber bark dust-zeolite composite improves mechanical strength and microstructural characteristics of soft paddy soil," *Sains Tanah Journal of Soil Science and Agroclimatology*, vol. 18, no. 2, pp. 194-202, 2021.
[16] K. Terzaghi, R. B. Peck and G. Mesri, Soil mechanics in engineering practice, 3rd ed., John Wiley & Sons., 1996.
[17] D. M. BURMISTER, "Principles and Techniques of Soil Identification," in *Annual Highway Research Board Meeting*, Washington, D.C., 1949.
[18] G. T. Houlsby, "HOW THE DILATANCY OF SOILS AFFECTS THEIR BEHAVIOUR," European Conference on Soil Mechanics and Foundation Engineering, Oxford, 1991.
[19] N. Turk and W. Dearman, "Estimation of Friction Properties of Rocks From Deformation Measurements," in *Proceedings of the 27th US Symposium on Rock Mechanics.*, 1986.
[20] E. Hoek, "Estimating Mohr-Coulomb Friction and Cohesion Values from the Hoek-Brown Failure Criterion," *Int. J. Rock Mech. Min. Sci & Geomech. Abstr.*, vol. 27, no. 3, pp. 227-229, 1990.
[21] PLAXIS, Material Models Manual 3D, Bentley, 2024.
[22] E. Hoek, P.K. Kaiser and W.F. Bawden, Support of underground excavations in hard rock, CRC Press, 200.
[23] Z. T. Bieniawski, "Rock mass classification in rock engineering," in *Exploration for rock engineering, proc. of the symp.*, Cape Town, 1976.
[24] K. Raj and R. Pedram, "Correlations between direct and indirect strength test methods," *International Journal of Mining Science and Technology*, vol. 25, pp. 355-360, 2015.
[25] S. Prakash and H. D. Sharma, PILE FOUNDATIONS IN ENGINEERING PRACTICE, Toronto: John Wiley & Sons, Inc., 1990.

ACKNOWLEDGEMENTS

The Organizing Committee of the 14th Scientific Day of the Institute of Technology of Cambodia (ITC) would like to acknowledge the financial support provided by Japan International Cooperation Agency (JICA) and the DX.sea project; our *Platinum Sponsor*— Panasonic; *Premium Sponsors*— Platform for Research and Training on Power System of the EDC-AFD-EU Project; Institut de Recherche pour le Développement (IRD); TEM Trading M&E Product Co., Ltd.; and Food Technology, Research and Innovation Platform (FTRIP); *Gold Sponsors*—Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD); DKK-TOA Corporation & Grande Solid Co., Ltd.; Grande Solid Co., Ltd. & Kyoto Electronics Manufacturing Co., Ltd.; and LHSBI Insurance; *Silver Sponsors*—The Capacity for Cambodia Energy Efficiency (CapCEE); DENSO (Cambodia) Co., Ltd.; Dynamic Scientific; SNP-PT International Co., Ltd., Thai-YAZAKI Electric Wire Co., Ltd., and Yonemori Construction Co., Ltd.; *Internet Sponsors*—Today Internet; and NTC NeocomISP Limited; and *In-kind sponsor*—V.N.C CORPERATION CO., LTD.

The Organizing Committee of the 14th Scientific Day of ITC would like to convey their utmost sincere thanks to the panelists: Prof. HENG Kreng – Head of Management Committee of Department of Research and Innovation, and Deputy Secretary General of Secretariat General of MoEYS; Dr. BUN Long – Deputy Director General of ITC; Mr. Hisakazu MAEDA – Chief Representative of Panasonic Cambodia; Dr. KHUN Kimang – Chief Technology Officer of AI Farm Robotics Factory.

Our great appreciation goes to the advisory boards, and key organizer team for spending their valuable time advising on the organization and organizing this event: H.E. Prof. Dr. PO Kimtho, Mr. SOY Ty, Assoc. Prof. Dr. BUN Kim Ngun, Asst. Prof. Dr. NGUON Kollika, Dr. BUN Long, Assoc. Prof. Dr. OR Chanmoly for their great advice on the organization of the event. The utmost appreciation goes to the core organizers for their valuable time organizing the event including Assoc. Prof. Dr. OR Chanmoly, Asst. Prof. Dr. TAN Reasmey, Asst. Prof. Dr. YOS Phanny, Dr. KRET Kakda, Asst. Prof. Dr. PHAT Chanvorleak, Asst. Prof. Dr. VALY Dona, Asst. Prof. Dr. PENG Chanthol, Dr. DOUNG Piseth, Dr. HENG Sounean, Ms. KAING Vinhteang, Ms. LY Soda, Dr. BUN Polyka, Mr. PLACK Sokhit and to all supporting organizers.

We respectfully appreciate for exhibition, sound, and order committee including Asst. Prof. Dr. PENG Chanthol, Mr. PENH San, Mr. KEO Chhomsethy, and the supporting organizers. The Organizing Committee of the 14th Scientific Day of ITC would like to

acknowledge the communication and protocol committee including Asst. Prof. Dr. YOS Phanny, Dr. CHUNHIENG Thavarith, Asst. Prof. Dr. SIM Tepmony, Asst. Prof. Dr. IN Sokneang, Asst. Prof. Dr. CHHUON Kong, Asst. Prof. Dr. ENG Chandoeun, Asst. Prof. Dr. HAN Virak, Mr. LAY Heng, Asst. Prof. Dr. PHUN Vengkheang, Dr. CHRIN Phok, Asst. Prof. Dr. SRENG Sokchenda, Asst. Prof. Dr. CHAN Sarin, Mr. LY Hav, Mr. HASH Chanly, Asst. Prof. Dr. LIN Mongkolsery, Mrs. SREY Malis, Mrs. KHEMTRAN Krasel, Mr. SO Phea, Asst. Prof. Dr. PHAUK Sokkhey, Asst. Prof. Dr. SRANG Sarot, Asst. Prof. Dr. YIN Molika, Mrs. HANG Vichothy, Mr. SIEANG Phen, Mrs. KOY Sophary, Mr. SOK Kimheng, Asst. Prof. Dr. BUN Saret, Asst. Prof. Dr. HOUNG Peany, and Ms. LY Soda.

We sincerely thank to registration and publication team including Asst. Prof. Dr. VALY Dona, Dr. HENG Sounean, Ms. KAING Vinhteang, Asst. Prof. Dr. SUONG Malyna, Dr. SANG Davin; to chairs and co-chairs for their great contribution to the event, to all reviewers for their great contribution to improving the scientific quality; to technical supports: Mr. KHIEV Samnang, Mr. SIENG Chamroeun, Mr. LUN Chanmaranden, and the IT team for their IT technical supports; to all volunteer for their supports without them the event cannot reach to the success. The family of the 14th Scientific Day of ITC organizers sincerely congratulates all presenters, delegates, and participants for their strong support and active participation in this event.

PROMOTIONAL PAGE

24-hour nanoeX Protection

nanoeX Clean 'n Care function

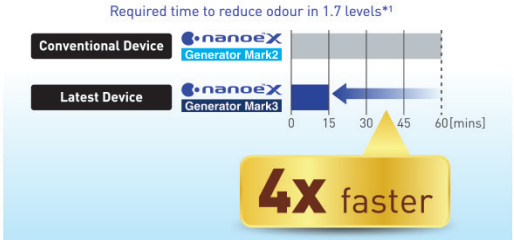
Scan to know more



High Efficiency of nanoe X Generator Mark 3

nanoe™ X technology with mark 3 generator contains 100 times more hydroxyl radicals compared to nanoe™. This helps to effectively inhibit pollutants and deodorise odours at a faster pace.

4x Faster in Deodorisation vs nanoe X Generator Mark 2



7 Effects of nanoe™ X

Inhibits 5 Types of Pollutants



Deodorises Odours

Moisturises Skin & Hair

nanoe™ X effectively deodorise and inhibit pollutants **in air and on surfaces*2**

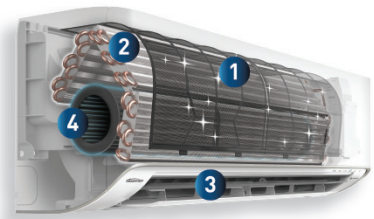


*1 nanoe X Generator Mark 2- •Testing organisation: Panasonic Product Analysis Center •Test subject: Adhered nonenal (Body odour) •Test volume: Approx. 23m³ laboratory •Test result: Odour intensity reduced by 1.7 levels in 60 minutes •Report No.: Y18HM059 (Difference in odour intensity from the blank is greater than 1.7 and the time until odour intensity is less than 2.6).
nanoe X Generator Mark 3- •Testing organisation: Panasonic Product Analysis Center •Test subject: Adhered nonenal (Body odour) •Test volume: Approx. 23m³ laboratory •Test result: Odour intensity reduced by 1.9 level in 15 minutes •Report No.: H21HM005-1

*2 •For more details on nanoe™ X test report, you can visit https://www.panasonic.com/kh/nanoe/all/verification/evidence.html

Inside Cleaning On Demand

Newly improved inside cleaning on demand provides the convenience to activate this function on a need basis. Inside cleaning works inside of an indoor unit by removing moisture, and releases nanoe™ X to inhibit various pollutants. Improved cross flow fan with anti-dust coating prevents dust adherence.



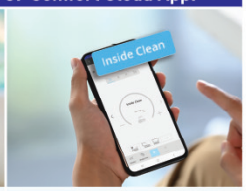
1. Filter
2. Evaporator
3. Air Outlet
4. Cross Flow Fan

INSIDE CLEANING ON DEMAND

By Remote Controller or Comfort Cloud App:



Turn off the AC and press the nanoe™ X button for more than 3 seconds to activate inside cleaning.



Turn off the AC and tap the 'Inside Clean' icon in Panasonic Comfort Cloud App to activate inside cleaning*3.

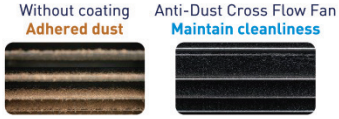
Effectiveness of Inside Cleaning

Filter, Evaporator, Air Outlet

Proven nanoe™ X inhibition on



Cross Flow Fan



The amount of dust may change depending on the usage frequency and environment. Anti-dust cross flow fan is available for X-PREMIUM INVERTER only.

Notes: (1) Air filter routine maintenance is required to ensure optimal performance.
(2) Illustrations of app screens may differ from the actual screen appearance.
(3) Inside Cleaning On Demand function is applicable to X-PREMIUM INVERTER, PREMIUM INVERTER and DELUXE INVERTER.

*3 Turn on inside cleaning on demand features via Comfort Cloud App available in X-PREMIUM INVERTER only.

*4 •Test Organisation: Laboratory for mold prediction •Test Subject: Mould (Aspergillus penicillioides K-712) •Test Result: Mould germination and growth from spores are inhibited inside the indoor unit under the inside cleaning operation •Test Report No: 220207 (nanoe™ effectiveness on mould inhibition is applicable to X-PREMIUM INVERTER only.)

*5 •Testing Organisation: Protectea,Ltd •Test Subject: Bacteria (Escherichia coli NBRC3301) and Virus (Escherichia coli NBRC320012) •Test Result: 99% inhibited (compared to non-operation) •Test Report No: PR190803 (Filter & Evaporator); PR191102 (Air Outlet)

Better Living with Energy-Saving and Eco-Friendly Technologies



Enjoy intelligent cooling with eco-friendly features for better savings and reduce energy consumption.

Environmental Friendly Refrigerant

Panasonic air conditioners are equipped with the environmentally friendly R32 refrigerant which efficiently carries heat, leading to greater energy and cost-effectiveness with lower environmental impact.

R32 REFRIGERANT

HIGHER COOLING CAPACITY

REDUCE ENERGY COST

ECO FRIENDLY



Powerful Performance with Efficiency

Panasonic's Inverter air conditioners integrate DC motor to offer better performance in terms of energy efficiency, quiet operation and wide output range. Inverter technology minimises temperature fluctuation to save energy without compromising comfort.

INVERTER

Scan to know more



Save on Your Electricity Bill

Varying the speed of the compressor according to the temperature changes with the aim of minimising the temperature fluctuations so you can enjoy save energy while cooling comfort.



Quiet Operation

Smooth operation and lower noise as low as 19dB(A)*.



Wide Output Range

Balance comfort level depending on the number of occupants in a room.

* 19dB(A) applicable to X-PREMIUM INVERTER (1.0HP- 1.5HP) and PREMIUM INVERTER (1.0HP- 1.5HP).

Intelligently Balances Energy Savings and Comfort

Every room in the house has a different temperature and it varies throughout the day. ECO mode with A.I. learns and judges the optimal ECO level to provide a good balance between cooling comfort and energy savings.

AI ECO

Scan to know more



*1 Comparison of ECO Mode & normal mode by using 1.5HP INVERTER model.

Food Technology, Research and Innovation Platform (FTRIP)
A Catalyst for Innovation and Applied Research in Cambodia's Food Sector

The Food Technology, Research and Innovation Platform (FTRIP) was officially established in February 2022 as a multi-stakeholder initiative designed to promote applied research, innovation, and technology transfer in Cambodia's food production sector. The platform was initiated under the CAPFISH Capture – Post-Harvest Fisheries Development Project, which is part of the "Cambodia Programme for Sustainable and Inclusive Growth in the Fisheries Sector" (ACA/2019/408-424), co-funded by the European Union. The platform aligns with the Royal Government of Cambodia's long-term vision for social and economic development, food security, and poverty reduction. The United Nations Industrial Development Organization (UNIDO) is delegated by the EU and the Fisheries Administration of MAFF to implement the project.

FTRIP was designed following the European Food Technology Platform model, bringing together key stakeholders from academia, government, and the private sector—namely Cambodia's Higher Education Institutions (HEIs), the Department of Innovation of the Ministry of Industry, Science, Technology and Innovation (MISTI), and the Fisheries Administration (FiA). Its core mission is to strengthen HEI capacities in applied research and promote innovation in food processing and post-harvest fisheries, enabling the private sector—particularly SMEs—to access demand-driven technology solutions.

Hosted by the Institute of Technology of Cambodia (ITC), FTRIP has achieved remarkable results in its first two years of implementation. Highlights include:

- **12 national seminars and webinars** organized to engage researchers and stakeholders in food technology and innovation;
- **Two Grant Management Subcommittee (GMSC) meetings** held to evaluate 47 submitted research proposals;
- **23 applied research projects awarded and implemented**, involving 54 researchers from five universities and 95 students, including 50 final-year interns;
- **Support provided to 10 enterprises**, including **7 women-led businesses**, resulting in improvements to **34 food products** and **36 processing methods**;
- **11 dissemination workshops** conducted to transfer research findings and technologies to SMEs;
- **One international symposium and one national conference** organized, with more than **150 participants**, to showcase the outcomes of FTRIP research.

FTRIP's contributions extend to academic curriculum development, having successfully developed two teaching modules—one on **Quality and Trade** and another on **Hazard Analysis Critical Control Point (HACCP)** for fish products. These modules were shared with four HEIs for formal adoption in August 2024.

In support of national infrastructure, FTRIP played a key role in facilitating the establishment of the **Cambodia Packaging Center (CPC)** hosted by the Royal University of Agriculture (RUA). This included supporting feasibility assessments, drafting the CPC charter, conducting **six stakeholder workshops**, and organizing **two international study visits**.

FTRIP has also actively supported University-Industry Linkages (UIL) and Intellectual Property (IP) awareness in close collaboration with MISTI. These efforts aim to promote knowledge exchange, enhance collaborative research practices, and strengthen innovation management among Cambodian stakeholders.

For regional collaboration, FTRIP signed an MoU with the **Asian Institute of Technology (AIT)** on 6 June 2024. The partnership commenced with a **Technical Training Workshop on Shelf-Life Prediction, Accelerated Shelf-Life Testing, and Food Packaging**, benefiting **25 lecturers and researchers** from five Cambodian universities.

Additionally, FTRIP, with support from the UNIDO CAPFISH project and in partnership with Ubon Ratchathani University (Thailand), organized a **training program on food processing and product development**. Scheduled for **7–10 May 2025**, the program includes hands-on factory site visits and aims to strengthen the capacities of HEIs in supporting Cambodian entrepreneurs.

Financial Support by

Collaborative Partnerships

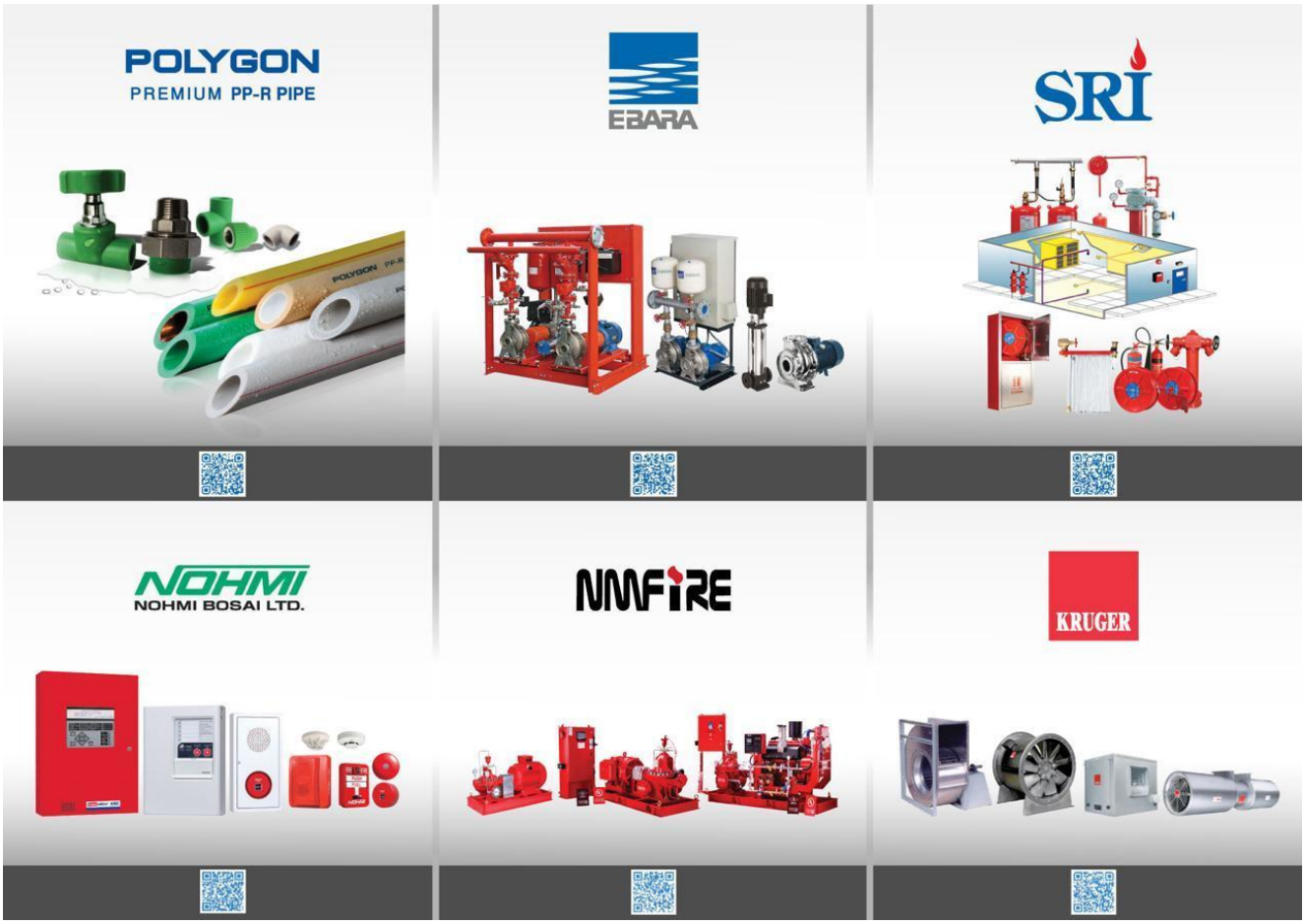
Platform for Research and Training on Power System of the EDC-AFD-EU Project

Our primary goal is to advance grid modernization through capacity building supported by high-quality research and training in power system at the Master's and PhD levels.



Platform for Research and Training on Power System with the financial support from AFD and EU and with the partnership of Electricité du Cambodge (EDC), Institute of Technology of Cambodia (ITC) and Institute Polytechnique de Grenoble – Université Grenoble Alpes (Grenoble INP – UGA) in France.

afd.eu.1211gedc@gmail.com


POLYGON
PREMIUM PP-R PIPE
EBARA
SRI
NOHMI
NOHMI BOSAI LTD.
NMFIRE
KRUGER

IRD is the **French National Research Institute for Sustainable Development.**
With activities aligned with the 17 SDG goals, our activities in Cambodia aim at :

- enhancing research capabilities through research practice;
- supporting the structuration of research laboratories
- facilitating the establishment of research infrastructures

ITC is **IRD**'s main partner in Cambodia.

IRD is permanently represented at **ITC** since 2019.

- 5 of our 9 senior researchers in Cambodia are hosted at **ITC**;
- we finance the work of some 15 research team members ;
- a Representative Office was established in 2020, with 4 headcount of 4 ;

Since 2019 our collaboration proactively structured into 4 joint-laboratories, co-directed by **ITC** and **IRD** researchers :

- HealthyRICE : Rice health and quality
- Coastal and Wetland Environmental Research lab : water resources
- ECO-SOL: Sustainable soil ecology
- KHEOBS - Khmer Earth Observation Laboratory : Satellite Remote sensing of Environment and Climate



These labs also support the training of young researchers through research initiatives.

In support of these ITC laboratories, **IRD** implemented various initiatives :

- Young research teams (JEAI): 3 years, allowing for the formation of ITC research teams and the deployment of scientific activities ;
- LMI : mixed international laboratory: Long term objective (5-10 year) to connect multiple laboratories in a common goal ;
- workshops, trainings and cofounded: 15 in the last 3 years ;
- fellowships : in 2024-2025 we funded 9 Master and 4 PhD ;
- senior researchers also contribute to undergraduate **ITC** education.

We greatly value our collaboration with ITC. We look forward to further developing this partnership and remain committed to working together with **ITC** in the years ahead.

luc.le-calvez@ird.fr
cambodge@ird.fr

THE 14TH SCIENTIFIC DAY

Institute of Technology of Cambodia



ORGANIZED BY

SUPPORTED BY

COOPERATED WITH

PLATINUM SPONSOR

PREMIUM SPONSOR