The 5th International Conference

Vol.3

Informatics, Agriculture, Engineering, Sciences and Technology

May 29 - 31, 2024

Deevana Plaza Krabi Aonang, Krabi Province, Thailand
Oraganized by King Mongkut's Institute of Technology Ladkrabang,
Prince of Chumphon Campus, Chumphon Province, Thailand

Conference Proceedings The 5th

International Conference on Informatics, Agriculture, Engineering, Sciences and Technology: IAMBEST 2024

Organized by:
King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus and University Network

1st Published: May 29th, 2024.
Cover Page Design: Mr. Sarawut Sawasdee

Published by:
King Mongkut's Instituteof Technology Ladkrabang Princeof Chumphon Campus,
Chumphon 17/1 Moo 6 Chumko Pathio Chumphon 86160 Thailand 86160
Tel: (+66) 77 506 - 410

MESSAGE FROM THE EXECUTIVE VICE PRESIDENT KMITL, PRINCE OF CHUMPHON CAMPUS

The 5th International Conference on Informatics, Agriculture, Management, Business Administration, Engineering, Science and Technology (IAMBEST 2024) along with the 9th National Conference on Informatics, Agriculture, Management, Business Administration, Engineering, Science and Technology are the conference for academics, experts, and researchers in the eight fields. The conferences are hosted by the King Mongkut's Institute of Technology Ladkrabang (KMITL), Prince of Chumphon Campus during the 29th – 31th of May 2024. The aim of this conference is the exchange stage of idea, knowledge, and research between the researchers from various fields. The continue development of the research provides the exchange the knowledge, connection, collaboration, and integration each other's. All above make the development the community, society, and country.

As the chairman of the IAMBEST 2024**,** I sincerely thank all committee members for your time and determination of organizing this wonderful conference. I thank all attendees and guests for sharing your research and innovation. The success of this conference is from all of you. I wish exchanging of idea and sharing of knowledge from this conference will give you another step to advance your knowledge and technology and to benefit your communities and countries.

Best wish to all of you.

(Assoc. Prof. Dr. Kamronwit Thipmanee)
Executive Vice President,
KMTIL, Prince of Chumphon Campus
Chairman of the Organizing Committee

IAMBEST
KMITL PRINCE OF CHUMPHON
29-31 MAY 2024

Keynote Speaker

Topic: Driving Technology and Innovation with Bio-Circular-Green Economy

Prof. Dr. Apanee Luengnaruemitchai

CHAIRPERSON, Energy, Chemicals and Biomaterials
Program Board from Program Management Unit
for Competitiveness (PMUC)

Contents

Program
The 9th National Conference and The 5th International Conference 2024 on Informatics, Agriculture, Management, Business Administration, Engineering, Sciences and Technology (IAMBEST 2024)
May 29–31, 2024
Deevana Plaza Krabi Aonang Hotel, Krabi Thailand
Organized by King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus

May 29, 2024		
13.00 – 17.30	Installation of poster and soft file presentation	Hornbill Ballroom 1&2
May 30, 2024		
08.00 – 08.50	Registration and Installation of poster and soft file presentation	Hornbill Ballroom 1&2
09.00 – 09.20	**Open Ceremony** The 9th National Conference and The 5th International Conference 2024	Hornbill Ballroom 1&2
09.20 – 09.35	**Coffee break**	Foyer
09.35 – 10.10	Keynote Speaker **Topic: Driving Technology and Innovation with Bio-Circular-Green Economy** **Prof. Dr. Apanee Luengnaruemitchai** CHAIRPERSON, Energy, Chemicals and Biomaterials Program Board from Program Management Unit for Competitiveness (PMUC)	Hornbill Ballroom 2
10.15 – 12.00	**Oral Presentation International** Session I: Social Science & Humanities, Management and Business Administration	Hornbill Ballroom 2
	Oral Presentation national Session 1: Management and Business Administration Session 2: Engineering	Broadbill Meeting Room Deevana Krabi Meeting room
	Poster Presentation national Sciences, Agriculture, Technology, Business Administration, Engineering, Social Science & Humanities and Management	Foyer
10.30 – 12.00	Session: PMUC meet & consult by PMUC	Starling Lobby Lounge
12.00 -13.00	**Lunch break**	Kingfisher Restaurant

May 30, 2024		
12.30 – 12.45	Installation soft file presentation	
13.00 – 14.30	**Oral Presentation International**	
	Session II: Technology and Engineering	Hornbill Ballroom 2
	Poster Presentation International	
	Session V: Sciences and Business Administration	Foyer
	Oral Presentation national	
	Session 3: Engineering	Hornbill Ballroom 1
	Session 4: Sciences	Broadbill Meeting Room
	Sciences 5: Social Science & Humanities	Deevana Krabi Meeting room
13.00 - 14.00	Session: PMUC meet & consult by PMUC	Starling Lobby Lounge
14.30 – 14.45	**Coffee break**	Foyer
14.45 – 17.00	**Oral Presentation International**	
	Session III: Informatics and Engineering	Hornbill Ballroom 2
	Session IV: Sciences and Agriculture	Broadbill Meeting Room
	Oral Presentation national	
	Session 6**:** Informatics Technology and Engineering	Hornbill Ballroom 1
	Session 4: Agriculture	Deevana Krabi Meeting room
18.30 - 21.00	**Dinner& Award Announcement**	Hornbill Ballroom 1&2
May 31, 2024	**Morning session**	
09.00 – 10.30	**Oral Presentation International**	
	Session I: Engineering	Hornbill Ballroom 2
	Oral Presentation national	
	Session 1 Management and Business Administration	Broadbill Meeting Room
	Session 2 Engineering	Deevana Krabi Meeting room
	Poster Presentation national	
	Informatics, Sciences, Agriculture, Business Administration Management, Engineering, Sciences, Technology and Social Science & Humanities	Foyer
10.30 – 10.45	**Coffee break**	Foyer
10.45 – 11.30	**Oral Presentation International**	
	Session 1 Technology and Engineering	Hornbill Ballroom 2
	Session 2 Agriculture and Sciences	Broadbill Meeting Room
	Poster Presentation International	
	Sciences and Business administration	Foyer
	Oral Presentation national	
	Session 1 Informatics, Technology and Engineering	Hornbill Ballroom 1
	Session 2 Sciences Agriculture and Social Science & Humanities	Deevana Krabi Meeting room

IAMBEST 2024 PROGRAM
May 30, 2024

Oral International Session II: Technology and Engineering **Hornbill Ballroom 2**

Chairperson	Assoc. Prof. Dr. Apiluck Eiad-Ua Assoc. Prof. Dr. Naruebodee Srisang Assoc. Prof. Dr. Srirat Chuayboon	KMITL KMITL Prince of Chumphon KMITL Prince of Chumphon
Time	**Code**	**Topic**
13.00 - 13.15	OT-12	Techno-economic assessment of domestic technology pico-hydro power turbine *Nitikorn Nanthawirojsiri Netithorn Ditnin Attakarn Jansasithorn and Usa Boonbumroong*
13.15 - 13.30	OT-58	Application of Remote Sensing and Machine Learning for Quantifying Ocean Acidification Parameters in a Marine Protected Area *Joel Colobong Kanchana Nakhapakorn Jirataya Roemmontri Phaisarn Jeefoo and Naruemon Pratanwanich*
13.30 - 13.45	OE-13	A Comparative Study of a Preventive Maintenance Plan Using Historical Data: A Case Study of Skytrain in Thailand *Thirit Lertsumruai Kittchai Sojiphan Nuksit Noomwongs and Sompoap Talabgaew*
13.45 - 14.00	OE-14	Rolling Stock Maintenance Analysis by Using Data-Driven Approach: A Case Study of Green Line Skytrain HVAC System *Suchanet Jaroenjittakam Kittichai Sojiphan, Chayut Ngamkhanong and Sompoap Talabgaew*
14.00 - 14.15	OE-22	Criticality Analysis of Automated People Mover Components by Applying Failure Mode and Effects Analysis *Thanodom Sirisongkram Kittchai Sojiphan Nuksit Noomwongs and Sompoap Talabgaew*
14.15 - 14.30	OE-23	Effects of temperature and heat treatment time on the hardness of stainless steel *Kawin Jewsuwun and Prayoon Surin*
14.30 - 14.45	OE-26	Alteration of Relative Permeability and Investigation of Oil Recovery Mechanism during Low Salinity Waterflooding in Different Rock Formations *Tina Tharn Falan Srisuriyachai and Chandoeun Eng*

IAMBEST 2024 PROGRAM
May 30, 2024

Oral International Session III: Informatics and Engineering **Hornbill Ballroom 2**

Chairperson	Assoc. Prof. Dr. Thatchapol Chungcharoen KMITL Prince of Chumphon Asst. Prof. Dr. Kittisak Phetpan KMITL Prince of Chumphon Asst. Prof. Dr. Warunee Limmun KMITL Prince of Chumphon	
Time	**Code**	**Topic**
15.00 - 15.15	OE-27	Simple and Effective Approach for Fresh or Frozen Fish Classification Using Image Processing and Deep Learning *Taweepol Suesut Chanitnunt Sanbooranapunt and Ari Kuswantori*
15.15 - 15.30	OE-48	Development of Hemagglutination Classification Using Machine Learning Techniques Based on Proposed Shadow Augmentation *Aneesah Satoh Somyot Chirasatitsin and Jermphiphut Jaruenpunyasak*
15.30 - 15.45	OE-77	Hydroxyapatite mixed with polycaprolactone and carboxymethyl cellulose sodium salt crosslinked with acetic acid for wound dressing application *Orranat Eiangmee Jatupat Maikaew Naruebodee Srisang and Siriwan Srisang*
15.45 - 16.00	OE-90	Exploring Underground Railway Structures Using Ground Penetrating Radar: Testing Sites in Bang Klam District, Songkhla Province, Southern Thailand *Nisarat Prangsuwan Kamhaeng Wattanasen Sawasdee Yodkhayan and Thanan Chub-Uppakarn*
16.00 - 16.15	OE-92	Detection of Coconut Black Headed Caterpillar Using Aerial Imagery Combined with Artificial Intelligence *Orapadee Joochim Kridtat Satharanond and Wirachat Kumkun*
16.15 - 16.30	OE-95	A Demonstration of Solar Pyrolysis of Agricultural Waste Biomass for Green Chemical Production *Tawanrung Artphan Tanapat Sartsiri Surakiat Yeamamphan Kanyarat Jitsuwan Rangsima Plapngam Sirapob Boonyawanthang Ratikorn Kulporm and Srirat Chuayboon*
16.30 - 16.45	OI-98	Integrated Landslide Risk Assessment of Kaghan valley in the northern Pakistan *Punyawi Jamjareegulgarn Naseem Ahmad Muhammad Shafique Munawar Shah Bushra Ghaffar Mian Luqman Hussain José Francisco de Oliveira-Júnior and Sarun Duansuwan*
16.45 - 17.00	OI-99	Spatiotemporal variation of water vapor over Northeast Japan retrieved from GNSS stations *Kutubuddin Ansari Janusz Walo Kinga Wezka Dorota Marjanska and Punyawi Jamjareegulgarn*

IAMBEST 2024 PROGRAM
May 30, 2024

Oral International Session IV: Sciences and Agriculture **Broadbill Meeting Room**

Chairperson	Asst. Prof. Dr. Jongjit Jantra Dr. Nannaphat Suwannarat Dr. Monsicha Tipawanna	KMITL Prince of Chumphon KMITL Prince of Chumphon KMITL Prince of Chumphon
Time	**Code**	**Topic**
14.45 - 15.00	OA-60	Effects of edible coating on the Quality and Shelf-life of fresh-cut "Kimju" guava (*Psidium guajava*) *Wisanee Nueamekin Pornteera Rattanarat and Panupat Polkham*
15.00 - 15.15	OS-08	Estimation Methods for the Difference between Variances of Birnbaum-Saunders Distributions Containing Zero Values with Application to Wind Speed Data in Thailand *Natchaya Ratasukharom Sa-Aat Niwitpong and Suparat Niwitpong*
15.15 - 15.30	OS-09	Confidence Intervals for the Coefficient of Variation of Zero-Adjusted Inverse Gaussian Distribution with Application to Rainfall Dispersion Data *Wasurat Khumpasee Sa-Aat Niwitpong and Suparat Niwitpong*
15.30 - 15.45	OS-35	Nanocomposite of carbon from hemp waste and silica source via mechanochemical process *Napat Kaewtrakulchai Kanokwan Wannason Korn Sukphunphoncharoen Sutee Chutipaijit Tawat Suriwong Pramote Puengjinda Gasidit Panomsuwan Nuttapong Ruttanadech Masayoshi Fuji and Apiluck Eiad-Ua*
15.45 - 16.00	OS-36	Nanoporous carbon materials from hemp waste via hydrothermal-carbonization: An effective adsorbent for phenol red *Korn Sukphunphoncharoen Napat Kaewtrakulchai Kanokwan Wannason, Sutee Chutipaijit Tawat Suriwong Pramote Puengjinda Gasidit Panomsuwan, Nuttapong Ruttanadech Masayoshi Fuji and Apiluck Eiad-Ua*
16.00 – 16.15	OS-50	AI-generated animation for improving knowledge and comprehension of triptans for acute migraine treatment in hospital pharmacists *Budsarin Sivathitikul Kemjira Chalermsuk Siranan Klanbut Sumet Klinhom Suteera Boonyapitaks Pattarawit Rukthong and Tulaya Potaros*

IAMBEST 2024 PROGRAM
May 30, 2024

Poster International Session V: Sciences and Business administration

Chairperson	Asst. Prof. Dr. Chodsana Sriket Asst. Prof. Dr. Usa Sukkha Dr.Siwakon Sokjabok	KMITL Prince of Chumphon KMITL Prince of Chumphon KMITL Prince of Chumphon
Time	**Code**	**Topic**
13.00 - 13.10	PB-30	Enhancing Entrepreneurial Capability among Fruit Farmers: Insights from a Study of Entrepreneurial Elements and Styles in Chanthaburi Province, Thailand *Bhawat Chiamjinnawat*
13.10 - 13.20	PS-84	Efficiency of betel Leaf extract against *Escherichia coli* *Naree Phanchindawan Kavalee Lononglin and Kamonwan Chucheep*
13.20 - 13.30	PS-25	Preparation of Amine-Functionalized Rice Husk Ash for Selective Adsorption Polyphenols from Green Tea *Poungmanee Phromchana Nadrada Maneesri Paphada Pathomnatikul Jinda Chuemue and Sakdinun Nuntang*
13.30 - 13.40	PS-34	Development of Probiotic Legume Beverages with their Potential Antioxidant and Antidiabetic Properties *Suree Nanasombat*
13.40 - 13.50	PS-57	Application of Recombinant Human scFv Antibody for Determination the Nodule Occupancy of Rhizobia from Dual Culture Inoculant *Chayanan Singtothong Kuntalee Rangnoi Pakpoom Boonchuen Kamonluck Teamtisong Jenjira Wongdee Teerana Greetatorn Pongpan Songwattana Pongdet Piromyou Monchai Manassila Nantakorn Boonkerd Neung Teaumroong Montarop Yamabhai and Panlada Tittabutr*
13.50 - 14.00	PS-59	Antibacterial Activities of Kombucha Derived from Oolong Tea, Butterfly Pea, Roselle, and Peppermint *Engkarat Kingkaew Jutamas Pracumtong and Duangjai Ochaikul*
14.00 -14.10	PS-62	Empirical Correlations Between Standard Penetration Test N Value and Shear Wave Velocity from Multichannel Analysis of Surface Wave for Earthfill Dam in Northern Thailand *Orawan Lunanta and Siriporn Chaisri*
14.10 -14.20	PS-75	Assessment of Antioxidant Activity, Total Flavonoid Content, and Total Phenolic Compounds from Sequential Exhaustive Extraction of *Wolffia globosa* Cultivated in a Controlled Greenhouse Environment *Pakjira Patchim Chutwadee Krisanapun Sukanya Mingyai Phuwanai Palakachen Arnatchai Maiuthed and Pattarawit Rukthong*
14.20 - 14.30	PS-87	Microwave Pretreatment Of Sugarcane Bagasse For Bioethanol Production By *Saccharomyces cerevisiae* TISTR 5088 and *Pichia stipitis* TISTR 5806 *Engkarat Kingkaew Onwanya Chaimanat Siripak Wongchachom and Duangjai Ochaikul*

Academic Committee

International Academic Committee

Dr. Pietro Luca Borsano	Chulalongkorn University, Thailand
Assoc. Prof. Dr. José Francisco de Oliveira Júnior	Federal University of Alagoas, Brazil
Dr. Sneha Sharma	Independent Scholar, Melbourne, Australia
Mr. K.C. Sumesh	Independent Scholar, Melbourne, Australia
Dr. Prayitno Abadi	Indonesian National Research and Innovation Agency, Indonesia
Asst. Prof. Dr. Munawar Shah	Institute of Space Technology Islamabad, Pakistan
Dr. Kutubuddin Ansari	Integrated Geoinformation (IntGeo) Solution Private Limited, India
Dr. Albert Abad Braga	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr.Raymund Tapacal Palayon	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Sampad Kumar Panda	Koneru Lakshmaiah Education Foundation, India
Dr. Sylvain Rodat	Materials and Solar Energy Laboratory, PROMES-CNRS, France
Asst. Prof. Dr. Muralidharan Nagarajan	Tamil Nadu Fisheries University, India
Dr. Md Abdul Momin Sumon	TN Technical University, USA
Prof. Dr. Gerhard Fink	Vienna University of Economics and Business, Austria

National Academic Committee

Assoc. Prof. Dr. Chaiwat Rattanamechaiskul	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Naruebodee Srisang	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Pornprapa Kongtragoul	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Punyawi Jamjareegulgarn	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Raumjit Nokkoul	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Sira Saisorn	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Siriwan Srisang	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Srirat Chuayboon	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Acting Sub Lt. Dr. Kittisak Phetpan	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Chanadda Phawachalotorn	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Chodsana Sriket	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Duangjai Pisuttharachai	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Jongjit Jantra	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Ousanee Sawagvudcharee	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Piyada Tavitchasri	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Sirichatnach Pakdeepromma	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Usa Sukkha	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Asst. Prof. Dr. Warunee Limmun	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Nanaphat Suwannarat	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Naruethai Chanthap	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Punjapha Pitigraisorn	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Rattikorn Sombutkaew	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Ms. Kanchalar Keeratirawee	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Ms. Yenying Chongchit	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Assoc. Prof. Dr. Atiporn Gerdruang	Stamford International University, Thailand

List of Reviewers

Dr. Kritsadi Thetpraphi	Centre National de la Recherche Scientifique (CNRS), France
Asst. Prof. Dr. Lin Min Min Myint	King Mongkut's Institute of Technology Ladkrabang, Thailand
Asst. Prof. Dr. Ploypailin Yongsiri	King Mongkut's Institute of Technology Ladkrabang, Thailand
Asst. Prof. Dr. Sirima Sinthusamran	King Mongkut's Institute of Technology Ladkrabang, Thailand
Asst. Prof. Dr. Sitthipong Nalinanon	King Mongkut's Institute of Technology Ladkrabang, Thailand
Asst. Prof. Dr. Walairat Chandra-Ambhorn	King Mongkut's Institute of Technology Ladkrabang, Thailand
Dr. Nattawut Chaomuang	King Mongkut's Institute of Technology Ladkrabang, Thailand
Asst. Prof. Acting Sub Lt. Dr. Kittisak Phetpan	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Raymund T. Palayon	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Dr. Siwakon Sokjabok	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Ms. Yenying Chongchit	King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Thailand
Prof. Dr. Maurice Yolles	Liverpool John Moores University, the United Kingdom
Asst. Prof. Dr. Yingrak boondam Chuayboon	Mahidol University, Thailand
Assoc. Prof. Dr. Nittaya Junka	Nakhon Pathom Rajabhat University, Thailand
Dr. Buncha Limpabandhu	Panyapiwat Institute of Management, Thailand
Asst. Prof. Dr. Kunlapat Thongkaew	Prince of Songkla University, Thailand
Asst. Prof. Dr. Sutitar Choosawang	Prince of Songkla University, Thailand
Asst. Prof. Dr. Nittaya Ummarat	Prince of Songkla University, Suratthani Campus, Thailand
Dr. Putipong Lakachaiworakun	Rajamangala University of Technology Phra Nakhon, Thailand
Asst. Prof. Dr. Thanwit Naemsai	Rajamangala University of Technology Srivijaya, Thailand
Asst. Prof. Dr. Orawan Aumporn	Silpakorn University, Thailand
Assoc. Prof. Dr. Atiporn Gerdruang	Stamford International University, Thailand
Dr. Pornpong Sakdapat	Thailand Securities and Exchange Commission, Thailand
Asst. Prof. Dr. Munawar Shah	Tongji University, China
Asst. Prof. Dr. Kutubuddin Ansari	Warsaw University of Technology, Poland

Oral Presentation

Integrated Landslide Risk Assessment of Kaghan valley in the northern Pakistan

Punyawi Jamjareegulgarn[1], Naseem Ahmad[2,3], Muhammad Shafique[2,3], Munawar Shah[4], Bushra Ghaffar[4], Mian Luqman Hussain[2,3], José Francisco de Oliveira-Júnior[5], Sarun Duansuwan[1*]

[1] King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon 86160, Thailand.
[2] National Centre of Excellence in Geology, University of Peshawar, Peshawar 25000, Pakistan.
[3] GIS and Space Applications in Geosciences (G-SAG), University of Peshawar, Peshawar 25000, Pakistan.
[4] College of Surveying and GeoInformatics, Tongji University, Shanghai 20092, China.
[5] Universidade Federal do Alagoas (UFAL), Maceio 57072-260, Brazil.

*Corresponding author: Sarun Duansuwan (Email: sarun.du@kmitl.ac.th)

Abstract. This article proposes an integrated approach of both geospatial data and semi-quantitative techniques to evaluate landslide hazard, vulnerability, and risk at Kaghan Valley located in the northern part of Pakistan. The Google Earth Pro, high-resolution DEM and satellite images are used to develop a landslide inventory and assess the landslide hazard and risk assessment. The landslide susceptibility map is then integrated with landslide-triggering factors to derive a landslide hazard index map. A spatial technique is used to investigate vulnerability assessment, and a semi-quantitative technique is also considered the relative risk classes for risk assessment. The landslide risk assessment map is classified in five classes including very low, low, medium, high and very high risks. The findings in this work can show different patterns of high-risk locations that were not clear from hazard or vulnerability analyses alone and the improvement of the region's resilience to landslide-related disasters. However, an obstacle in this study is the lack of temporal data for making comprehensive risk map, so it should be extended in the future.

***Index Terms*—Kaghan valley, Landslide, Northern Pakistan, Risk assessment, Susceptibility, Vulnerability.**

I. INTRODUCTION

LANDSLIDE is a natural hazard leading to several losses to human lives, infrastructure, and the economy around mountainous terrains. In the Himalayan mountainous region, landslides are yearly attributed to around the 200 human casualties and the economic losses of larger than US$ 1 billion [1]. The spatial probability of landslides is impacted by local geological and environmental settings. The temporal distribution and intensity of landslides are largely governed by the earthquakes and deforestation activities. Landslide inventory indicates the location, type, as well as occurrence date of landslides and is a prerequisite for the hazard, vulnerability and risk assessment [2]. So far, the landslide hazard assessment has been mainly conducted by analysing the temporal distribution of landslides and the frequency-magnitude relationship. However, the temporal landslide record is rarely recorded in the data-poor regions, the susceptibility map is finally used as a landslide trigger so as to derive the landslide hazard map. As for landslide hazard assessment, the Weighted overlay Analytical Hierarchy Process (AHP) are effectively applied. The Spatial Multi-Criteria Evaluation (SMCE) model is an efficient technique to assess the landslide hazard based on the combination of the causative and triggering factors.

Landslide vulnerability is the degree to which a community or natural resource is impacted by the landslides. Landslide vulnerability is a complex and dynamic issue with various aspects. Landslide vulnerability assessment involves analyzing the physical (roads, buildings, etc.) and environmental (forests, agriculture, etc.) indicators. The physical vulnerability can be effectively assessed by using the typological information of the buildings including the building material, age, design, foundations and physical conditions [3]. Environmental vulnerability involves the assessment of the impacts of landslide hazards on the environment [4]. Landslide risk can be assessed using quantitative [5] and qualitative techniques [6].

It is well-known that the northern areas of Pakistan, especially the northwestern Himalayas of the Kaghan valley, are highly prone to landslide hazards due to their geological conditions, climate and active seismic zones. These areas were tragically demonstrated in 2005, when the Kashmir earthquake triggered thousands of landslides with 26,000 casualties and significant economic losses to the roads and buildings. Therefore, this study aims to utilize the remote sensing and field data to develop the landslide hazard and risk assessment. To deal with landslide hazard in Kaghan Valley, this research focuses on identifying the landslide-prone locations and assessing the possible effects on vulnerable communities and infrastructure.

II. Data and Methodology

A. Study Area

The Kaghan valley is situated in the Balakot tehsil of Mansehra district, north Pakistan with an area of 12960 km2 (Fig. 1). The Kunhar River originating from the lake Saiful Malook, flows from northeast to southeast and drains the cities of Balakot and Garhi Habibullah, before joining the Jhelum River in Muzaffarabad. This area generally has cold winters and warm summers with temperatures ranging from −1 and 43 °C. The study area receives 1000 and 2000 mm of rainfall per year. The study area is blessed with a beautiful landscape, forested mountains and pleasant weather that pulls millions of foreign and domestic tourists every year. The construction of new hotels and resorts and the improvement of tourism infrastructure are attracting many tourists over the globe. However, the constructions of road by cutting the slope have triggered the widespread landslides in the area. The valley road traverse through the mountain is prone to landslides, and thus in the event of a landslide, the access route to the valley is blocked disrupting the flow of tourists with significant socio-economic consequences [6-7].

B. Data Used

The data used in this work is given in Table I. The landslide inventory was developed from the on-screen digitization on Google Earth using visual interpretation and field verification. The Advanced Land Observing Satellite-1 Phased Array type L-band Synthetic Aperture Radar (ALOS POLSAR) Digital Elevation Model (DEM) at a 12.5 m spatial resolution for the area was acquired from the Alaska Copernicus European Space Agency. The landslide causative factors consisting of the terrain slopes, aspect, curvature and streams to assess landslide susceptibility were computed from the ALOS DEM. The land use/land cover classified data was acquired from the Khyber Pakhtunkhwa Forest department which was developed from the SPOT-6 Imageries of the year 2020. The rainfall data was acquired from the Global Precipitation Measurement (GPM) with a spatial resolution of 10 km and subsequently resampled to 12.5 m for analysis. The Peak Ground Acceleration (PGA) map showing seismic hazards in the area was developed by the Pakistan Meteorological Department. The rainfall and PGA data were further employed as the landslide triggering factors for the landslide hazard assessment. The remote sensing technique was used to develop building footprints for the area, validated to collect the topological information, and utilized for evaluating the physical vulnerability.

III. Landslide Susceptibility and Hazard

A. Landslide Inventory

The landslide outlines (Fig. 3A) were manually digitized through the visual interpretations of Google Earth Imageries using their morphological expressions such as a main body or foot in many cases [9]. The process of landslide inventory focuses on identifying areas with soil and rock, and the presence of scarps, depressions, and irregularities (Fig. 2).

TABLE I
DATA USED FROM DIFFERENT SOURCES.

Data	Sources
1. Landslide Inventory	Google Earth Imageries
2. ALOS POLSAR DEM	European Space Agency
3. Geology	Searle et al.(1996) [16]
4. Land Cover	Khyber Pakhtunkhwa Forest Department (SPOT 6 Images).
5. Rainfall Data	Global Precipitation Measurement
6. Peak Ground Acceleration	Pakistan Meteorological Department report
7. Topology	Google Earth Imageries
8. Population Data	Pakistan Bureau of Statistics

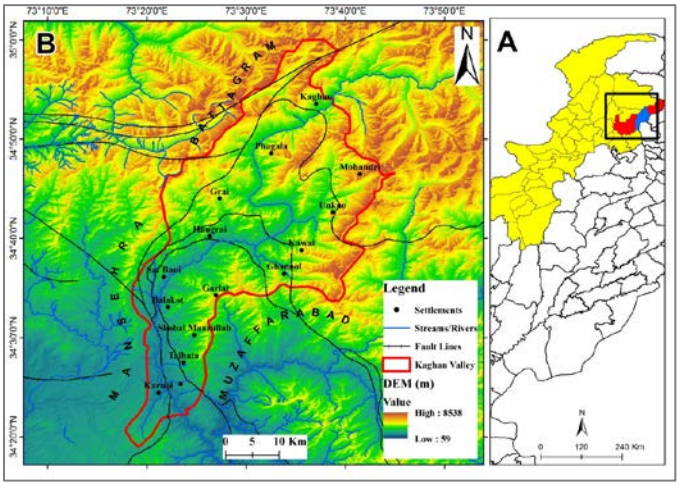


Fig. 1. (A) Geographical boundary of Pakistan, Khyber Pakhtunkhwa and Mansehra district, and (B) Kaghan valley.

Fig. 2. (A) Adobe column and cement block wall proper columns and RCC/BCC linter, (B) Pitched sloping with prefabricated sheets, (C) Wooden wall and roof with sheet, (D) Simple sheets, (E and F) No column or beams, (G and H) Hairline cracks and plaster damages, broken, Seepage, and (I) Destructed shops on October 8th, 2005 Earthquake at Pakistan.

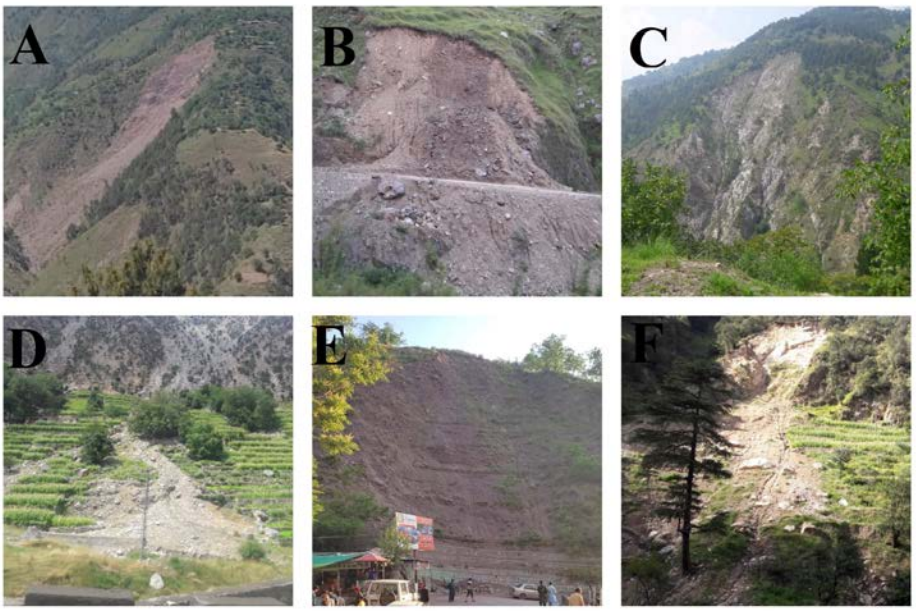


Fig.3. Photos of the different landslides in the study area (A) Debris Slide (B) Landslide (C) Rockslide (D) Debris Flow, (E and F) Mud slid.

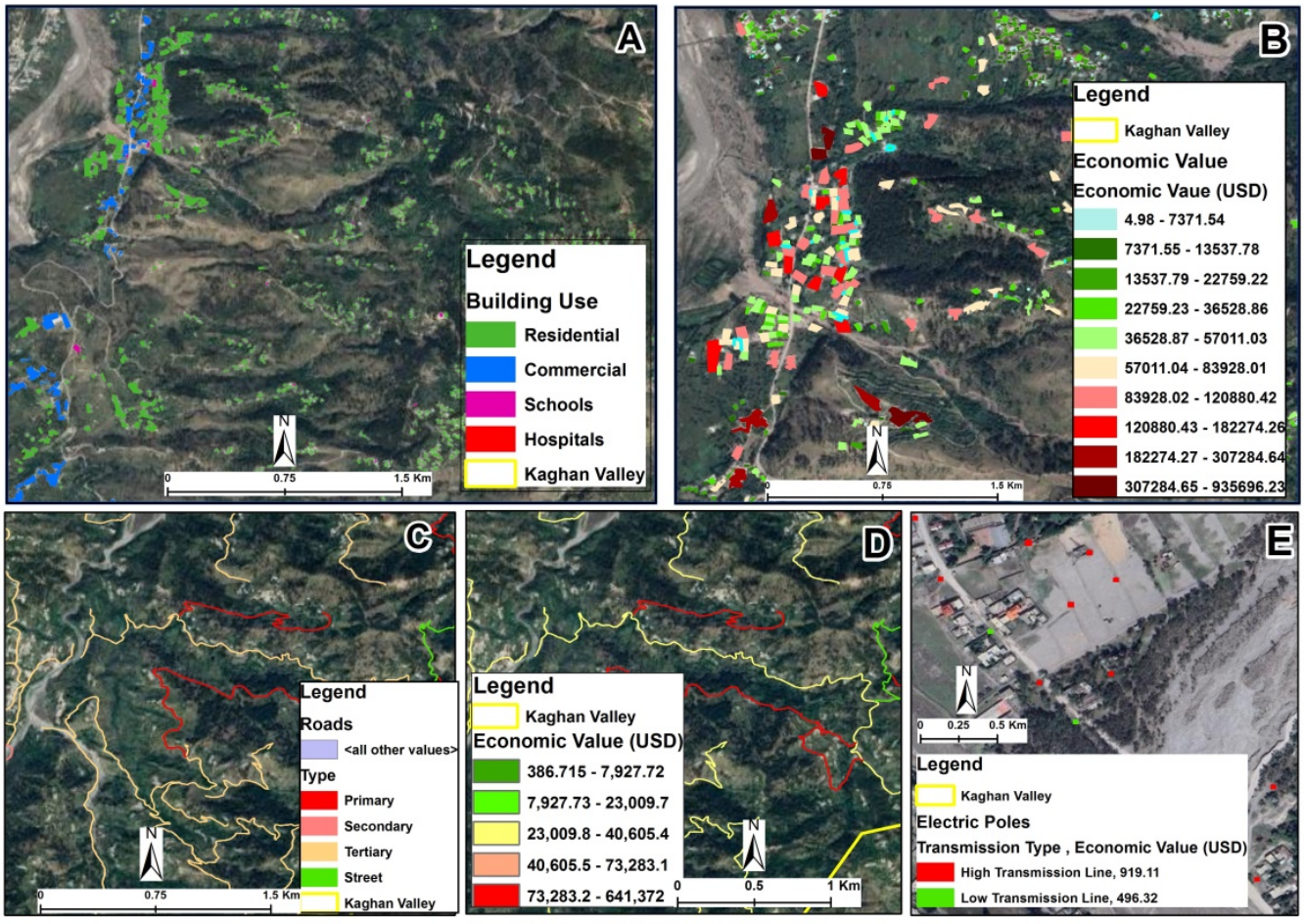


Fig. 4. (A) Type of buildings (B) Economic values of buildings (C) Type of roads (D) Economic values of roads (E) Type and economic values of the electric poles used as economic Indicator for the derivation of landslide vulnerability.

Vegetation can provide valuable clues to the presence of landslides as it facilitates the identification of the landslide in optical data through visual interpretation and field observations (Fig. 3B). The eroded deposits are another indicator of landslides, as material that has moved downslope was often accumulated at the base of the slope or along the sides of valleys and streams (Fig. 3). Once landslides have been demarcated, they are further classified after Hungr and Leroueil. In the developed inventory, 70% of the landslides are used for the derivation of the LSI map and 30% for the validation.

B. Landslide Causative Factors and Susceptibility

The DEM calculated by topographic parameters influences the spatial distribution of landslides including the terrain slope, aspect, curvature, and proximity to streams. Landslides due to steep slopes are caused by shear pressures operating on hill slopes, which are determined by the terrain slope (Fig. 4A) [10]. Terrain aspects including the vegetation cover, moisture retention, and soil strength, eventually influence the likelihood of landslides. (Fig. 4B. Distance to streams (Fig. 4F) is another critical parameter that contributes to slope stability through infiltration, saturation, and slope toe-cutting. High levels of water runoff can cause saturation of the soil, lessening the soil strength and leading to landslides [11]. Other than topographic factors, the road distance (Fig. 4E), land cover (Fig. 4G), distance to faults (Fig. 4D) and lithology of the area also play pivotal roles in the assessment of the landslide susceptibility.

A bivariate statistical technique called Weights-of-Evidence (WofE) is employed in the GIS environment to evaluate the susceptibility of landslides. Assessing the impact of particular causative variables on the landslides includes integrating various conditioning factor maps with the landslide inventory map. By superimposing each conditioning factor map with the landslide inventory map, the landslide density in each class of all the landslide factor maps is obtained. Additionally, it offers a quantitative evaluation of landslide susceptibility that can be used to rank sites for additional research. It is crucial to remember that the model's assumptions, as well as the quality and resolution of the input data, determine how accurate the findings will be.

C. Landslide vulnerability and risk assessment [12]

The vulnerability of an area to landslides was assessed by using the physical and environmental indicators as shown in Fig. 4. Physical vulnerability is usually assessed by the potential for damage to buildings as well as utilities, and can also be visualized in the form of a vulnerability index map through the use of the SMCE model. Social vulnerability considers the potential for fatalities or injuries, and environmental vulnerability assesses the impact of landslides on land cover such as water bodies, forests, and agriculture. Understanding the vulnerability of an area to landslides is important for assessing risk, as well as planning for preventive and mitigation measures.

IV. Results and Discussion

A. Landslide Inventory Classification

A comprehensive database of 454 distinctive polygons illustrating mapped landslides offers extensive information about the size, distribution and type of slides (Fig. 3). The total area of the landslides is 73.08 km^2, with the smallest slide of 0.0005 km^2 and the largest of 13.8 km^2. Among the mapped slides, there are 48 slides (Fig. 2B), 140 debris flow (Fig. 3D), 19 debris slides (Fig. 3A), 9 mud flow events, 189 mud slides (Figs. 2E and 3F), 37 rockslides (Fig. 3C), and 8 scree occurrences (Fig. 3). For the accuracy assessment, the field validation survey confirmed 136 landslides out of the total 454 landslides, i.e. around 30% of the inventory. A large proportion of the mapped landslides are shallow translational landslides. Spatially, the landslides are mainly concentrated in the less vegetated areas. Based on the Google Earth imagery from 2021, the inventory offers a snapshot of landslide activity in the Kaghan Valley, although it does not include data on past or present activity of landslides.

B. Landslide Susceptibility and Hazard

The landslide susceptibility index for the area was determined by this study using the WofE (Weight of Evidence) model. The natural break approach was used to further categorize the susceptibility index values to create a zonal map for further investigation. The Kaghan Valley's northern, southern, northeastern, and southeastern regions were mostly where the research area's high-susceptibility landslide zones were concentrated. These landslides mostly happened amid hill slopes and along the Kunhar River. Slope bodies are unstable due to the Kaghan Valley's steep topography and significant development along the leading edges of exposed surfaces. Loose materials on slopes tend to slide down the surface where they meet the bedrock below, making the slope less stable and encouraging the formation of landslides. Moreover, the Kaghan Valley region between Balakot and Garhi Habibullah is distinguished by its steep slopes and dense population.

In some areas with cliffs, the collapses frequently occur due to the stress unloading on the rock mass. In addition to causing surface erosion on slopes and the downward movement of loose materials along slope surfaces, the heavy rainfall frequently causes surface erosion, which could under some events lead to the debris flows. As a result, there are significant debris build-ups close to gully bases, offering enough solid substrate for landslide formation. Additionally, a thick layer of loosely accumulated residual quaternary sediments has been deposited on several of the slope surfaces. The chance of landslide disasters gradually increases with greater human engineering operations, justifying the inclusion of these disaster-prone sites in risk preventive and management programs.

C. Discussion

This study is intended to evaluate the landslide risk in the complex topography of Kaghan valley located in the northern part of Pakistan. The existing literature on landslide hazard and risk assessment is rare because of incomplete recording of the events and the unavailability of high-

resolution remote sensing data between the academia and administrative departments. The preparation of landslide inventory is the preliminary step towards the assessment of the landslide-based hazard and risk assessment. By using the high resolution Google Earth Pro data instead of the commercial satellite images in this work, we are able to fill in this gap and produce an extensive and detailed landslide inventory over the whole Kaghan valley. As for the assessment of the landslide susceptibility, the different numerical models are used using causative factors derived from DEM i.e. slope, aspect, curvature, distance to stream, distance to fault, distance to road, and lithology data.

Our studied results give good agreements to the previous research. For example, the typology by digitizing each building in a certain part of district Ghizer, Gilgit Baltistan were studied and the landslide vulnerability by assigning weights to the typology parameters and creating detailed landslide vulnerability map were also assessed by Shah et al. [14]. The most important findings in this work show different patterns of high-risk locations that were not clear from hazard or vulnerability analyses alone. The region's resilience to landslide-related disasters can be enhanced by this thorough risk assessment, allowing for better prioritization of mitigation measures and emergency response planning. However, an obstacle in this study is the lack of temporal data to provide a comprehensive risk map.

V. Conclusion

This study presents a detailed insight into the landslide susceptibility, hazard, vulnerability and risk assessment of the Kaghan valley using open-source remote sensing and field data. A 66282-house footprint data was used for the quantitative assessment of the landslide vulnerability and risk. We find that the Kaghan Valley poses a major landslide-induced threat to the community living due to its complex geology and un-engineered structures as well as no proper use of material. Most of the villages in Kaghan valley are exposed to landslide risk. The landslide vulnerability-based studies both on regional and local scales can be made to provide some detailed insight into the considered problems and a suitable mitigation strategy can be employed into practice to avoid the damages due to the landslides. The awareness campaigns may be conducted to educate the local community about the landslide impact to daily lives. To improve the risk mapping accuracy, future studies should be conducted to increase the availability of data and quantify the risk index.

Acknowledgment

The authors acknowledge the financial support for this research from Network Grant of KMTIL Research Fund (Grant Number: KREF026701).

References

[1] T.H. Tran, *et al.*, “GIS-Based Soft Computing Models for Landslide Susceptibility Mapping: A Case Study of Pithoragarh District, Uttarakhand State, India,” Mathematical Problems in Engineering, 2021(9914650), 2021.
[2] F. Ardizzone, *et al.*, “Geomorphological landslide inventory map of the Daunia Apennines, southern Italy,” Earth System Science Data, 15(2), pp. 753-767, 2023.
[3] I. Permanajati, A.H. Suranda, J.A. Zaenurrohman, “Assessment of Landslide Susceptibility in the Pagentan Area, Banjarnegara Regency: A Spatial Multi-Criteria Evaluation Approach,” Riset Geologi dan Pertambangan-Geology and Mining Research, 33(1), 2023.
[4] F. Guzzetti, “Landslide hazard and risk assessment,” Ph.D. Thesis, 2006.
[5] Y. Andersson-Sköld, S. Falemo, M. Tremblay, “Development of methodology for quantitative landslide risk assessment-Example Göta river valley,” Natural Science, 6(3), pp. 130-143, 2014.
[6] F. Dai, C.F. Lee, Y.Y. Ngai, “Landslide risk assessment and management: an overview,” Engineering geology, 64(1), pp. 65-87, 2002.
[7] Muhammad, S., et al., Spatial distribution of heavy metal and risk indices of water and sediments in the Kunhar River and its tributaries. Geocarto International, 37(20), pp. 5985-6003, 2022.
[8] F. Zamir, F. Hanif, S. Naz, “Extreme rainfall frequency analysis for Balakot, Pakistan, using Gumbel’s distribution,” Arabian Journal of Geosciences, 14, pp. 1-10, 2021.
[9] F. Guzzetti, *et al.*, “Landslide inventory maps: New tools for an old problem,” Earth-Science Reviews, 112(1-2), pp. 42-66, 2012.
[10] A. Maqsoom, *et al.*, “Landslide susceptibility mapping along the China Pakistan Economic Corridor (CPEC) route using multi-criteria decision-making method,” Modeling Earth Systems and Environment, 2021.
[11] B. Barasa, *et al.*, “Natural hazards magnitude, vulnerability, and recovery strategies in the Rwenzori Mountains, Southwestern Uganda,” Remote Sensing of African Mountains: Springer, pp. 83-116, 2022.
[12] E.A. Castellanos Abella, “Multi-scale landslide risk assessment in Cuba,” Ph.D. Thesis, 2008.
[13] L., Natsvlishvili, N. Jorjiashvili, V. Kochoradze, “Development of a PostGIS-based method for creating risk maps of natural disasters using the example of Georgia,” Geodesy and Cartography, 48(2), pp. 70–77, 2022.
[14] N.A. Shah, *et al.*, “Integrated Approach for Landslide Risk Assessment Using Geoinformation Tools and Field Data in Hindukush Mountain Ranges, Northern Pakistan,” Sustainability, 15(4), p. 3102, 2023.
[15] X. Liu and C. Miao, “Large-scale assessment of landslide hazard, vulnerability and risk in China,” Geomatics, Natural Hazards and Risk, 9(1), pp. 1037-1052, 2018.
[16] M.P. Searle and M.A. Khan, "Geological Map of North Pakistan and Adjacent Areas of Northern Ladakh and Western Tibet," Western Himalaya, Hindu Kush, 1997.

Spatiotemporal variation of water vapor over Northeast Japan retrieved from GNSS stations

Kutubuddin Ansari[1], Janusz Walo[1], Kinga Wezka[1], Dorota Marjanska[1], Punyawi Jamjareegulgarn[2*]

[1] Faculty of Geodesy and Cartography, Warsaw University of Technology, Warsaw, Poland
[2] King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon, 86160, Thailand

*Corresponding author: Punyawi Jamjareegulgarn (Email: kjpunyaw@gmail.com)

Abstract. The lowest part of the atmosphere known as the troposphere extends from the surface of the Earth to an approximate altitude of 10 km. This is the atmospheric densest layer which covers almost all water vapor, precipitation and cloud. The study investigated precipitable water vapor (PWV) variation over Northeast Japan by using more than two decades (2000 to 2022) of measurements. The PWV variation shows that the northern part of Northeast Japan which is located at higher latitudes showed a smaller magnitude compared to PWV magnitude in the southern part of low latitudes. The southern part where the correlation coefficients are very high (more than 0.80), indicates good agreements between GNSS PWV and ERA5 values and can be considered well-sampled observations. Moreover, the correlation coefficients drop the values approximately 0.40–0.60 for poorly sampled observations on the northern coast, pointing out interannual inconsistency of PWV measurements which are loosely related to ERA observations.

***Index Terms*— ERA5, Japan, NE, PWV, Water vapor.**

I. INTRODUCTION

Water vapor which is the fundamental component of the troposphere plays an important role in keeping the air temperature above the freezing point by absorbing thermal radiation and solar energy on the surface of the Earth [1, 2]. Water vapor condensation to ice or liquid phase is responsible for most important weather elements such as snow, cloud, rain, and precipitation [3, 4]. The measurement of water vapor usually is used to express the term known as precipitable water vapor (PWV). The PWV is the total atmospheric water vapor (mm) of a unit cross-sectional area in a vertical column that extends from the surface of the Earth to the highest point of the troposphere.

The measurement obtained from ground-based GNSS stations can be taken for all weather situations, with a high consistency and time-based resolutions (only a few minutes). GNSS observations post-processing provides the results of comparable accuracy with the measurements of traditional PWV sensors [5] due to neutral atmospheric effects on the propagation of signals (zenith angle mapped), mean ground air pressure values and atmospheric temperature [6, 7]. As a result, tropospheric delay estimated by GNSS measurement can be applied to quantify the PWV (unit mm) or equivalent to integrated water vapor (unit kg/m2) by utilizing the other metrological data such as temperature and pressure [8]. In the current study, we investigated PWV variation in NE Japan by using 36 GNSS station geodetic measurements from 2000 to 2022 (Fig. 1). We tried to perform an intercomparison of the variability and climatology of the PWV derived from Japanese GNSS data from 2000 to 2012 with ERA5.

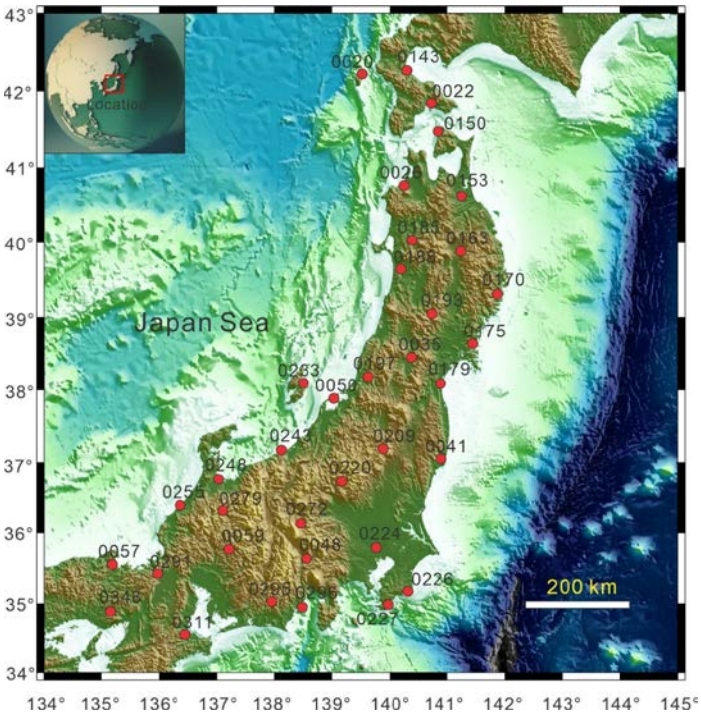


Fig. 1. Location of 36 GNSS stations in NE Japan during 2000-2022 for the PWV variation.

II. Annual Mean PWV variation

We studied the annual PWV spatial annual variation by taking the mean of daily PWV variation from 2000 to 2022 and interpolating with Thin-pale spline as shown in Fig. 2. Note that the annual mean is an average of the whole PWV values in each year of each GNSS station. The area outside of the Japan boundary is based on extrapolation, this should be ignored. The contour plot of the figure depicts the latitudinal and longitudinal distribution of PWV concentration in the Japanese atmosphere. As we can see from the figure the PWV range variation lies between 6 to 13 mm. The northern part of NE Japan which is located at higher latitudes showed a smaller magnitude and variation of PWV (7-9 mm) compared to the PWV (10-13 mm) magnitude and variation in the southern part of low latitudes. This is possible because the cold airs are drier and the warm airs hold more moisture in the lower southern part. We also tried to understand the longitudinal variation of PWV, which is not very significant. This is notable that the sites located in coastal areas have higher PWV compared to the continental region. This is because the sites located in coastal areas generally show higher PWV annual variations compared to the sites located in continental inland areas [9].

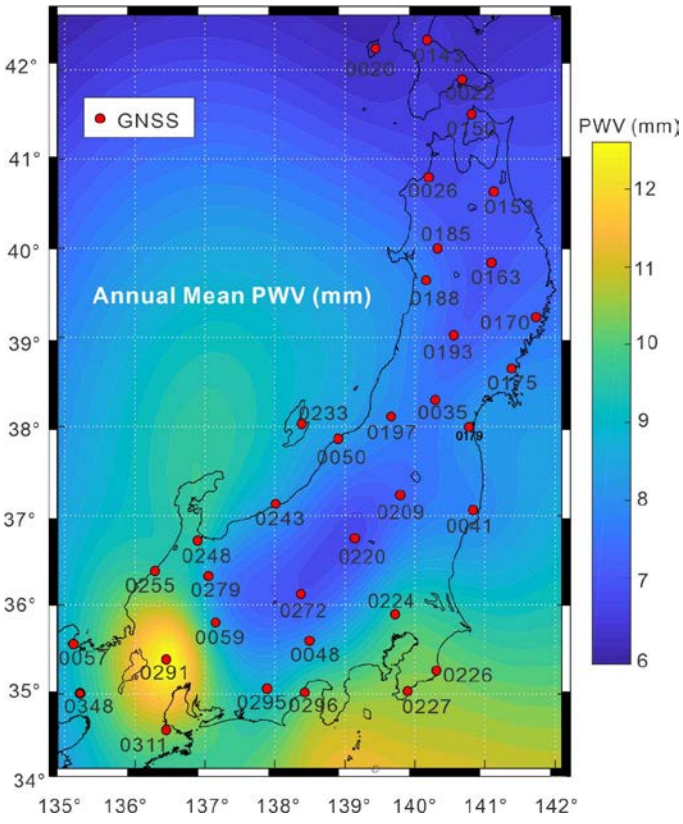


Fig. 2. PWV spatial annual variation by taking the mean of daily PWV variation from 2000 to 2022 and interpolating with Thin-pale spline interpolation

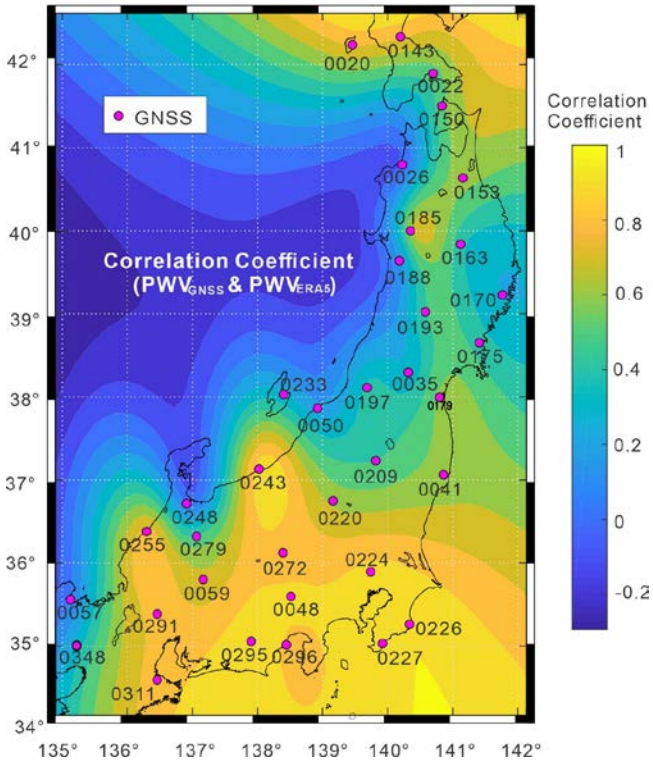


Fig. 3. The correlation coefficient between GNSS PWV and ERA5 measurements over NE Japan.

III. GNSS PWV and ERA5 Correlation Coefficient

The correlation coefficient between GNSS PWV and ERA5 measurements has been carried out to understand the PWV annual variability sampling impact over NE Japan. Fig. 3 shows computed correlation coefficients spatial distribution of the mean of GNSS PWV and ERA5 observations from 2000 to 2022. If we ignore the extrapolation values outside of the Japan boundary, the output depicts positive relationships of correlation coefficients on a spatial scale. The correlation coefficient variation based on geographical area is visible which quantifies the consistency between GNSS PWV and ERA5 and visualizes the effects in sampling. The southern part where the correlation coefficients are very high (more than 0.80), indicates good agreements between GNSS PWV and ERA5 values and can be considered well-sampled observations. These kinds of high correlation coefficients improve the result suitability and confirm the reliability of both the measurements. The appearance of the correlation coefficient decreases from the southern to the northern part. Moreover, the correlation coefficients drop the values approximately 0.40–0.60 for poorly sampled observations on the northern coast, pointing out interannual inconsistency of PWV measurements which are loosely related to ERA observations. This seems that ERA5 estimations have less agreement with GNSS PWV for several stations. These are probably some limitations of ERA5 reanalysis which is unable to present the high temporal and spatial variability of PWV in the humid regions

and complex terrains within the study area [10]. Therefore, we can easily conclude that interannual variation of PWV is affected seriously by sampling in weak observational regions. The regions that have cloudy and rainy weather generally ERA5 observations show their worse performance [11].

IV. Conclusion

The study investigates the annual and seasonal variation of PWV over Northeast Japan by using the GNSS measurements from 2000 to 2022. The results show that PWV annual range variation lies between 6 to 13 mm. The northern part of NE Japan which is located at higher latitudes showed a smaller magnitude and variation of PWV (7-9 mm) compared to the PWV (10-13 mm) magnitude and variation in the southern part of low latitudes because of the cold air those are drier and warm airs hold more moisture in the lower southern part. The correlation coefficient between GNSS PWV and ERA5 measurements in the southern part where the correlation coefficients are very high (more than 0.80), indicates good agreements between GNSS PWV and ERA5 values and can be considered well-sampled observations. The appearance of the correlation coefficient drops the values approximately 0.40–0.60 for poorly sampled observations on the northern coast, pointing out interannual inconsistency of PWV measurements which are loosely related to ERA observations.

Acknowledgment

The research was funded by the Warsaw University of Technology within the Excellence Initiative: Research University (IDUB) programme. The authors acknowledge the financial support for this research from Network Grant of KMTIL Research Fund (Grant Number: KREF026701).

References

[1] B. Adeyemi and S. Joerg, “Analysis of water vapor over Nigeria using radiosonde and satellite data,” Journal of applied meteorology and climatology, 51(10), pp. 1855-1866, 2012.
[2] K. Mohanakumar, “Structure and composition of the lower and middle atmosphere. In Stratosphere Troposphere Interactions: An Introduction,” Springer: New York, NY, pp. 8–10, 2008.
[3] M.T. Chahine, “The hydrological cycle and its influence on climate,” Nature, 359(6394), pp. 373-380, 1992.
[4] K. Ansari, T.S. Bae, J. Lee, “Spatiotemporal variability of total cloud cover measured by visual observation stations and their comparison with ERA5 reanalysis over South Korea,” International Journal of Climatology, 41, pp. E1757-E1774, 2021.
[5] C. Satirapod, S. Anonglekha, Y.S. Choi, H.K. Lee, “ Performance assessment of GPS-sensed precipitable water vapor using IGS ultra-rapid orbits: a preliminary study in Thailand,” Engineering journal, 15(1), pp. 1-8, 2011.
[6] E. Priego, A. Seco, J. Jones, M.J. Porres, “Heavy rain analysis based on GNSS water vapour content in the Spanish Mediterranean area,” Meteorological Applications, 23(4), pp. 640-649, 2016.
[7] M. Bevis, S. Businger, T.A. Herring, C. Rocken, R.A. Anthes, R.H. Ware, “GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system,” Journal of Geophysical Research: Atmospheres, 97(D14), pp. 15787-15801, 1992.
[8] J. Shi, C. Xu, J. Guo, Y. Gao, “Real-time GPS precise point positioning-based precipitable water vapor estimation for rainfall monitoring and forecasting,” IEEE Transactions on Geoscience and Remote Sensing, 53(6), pp. 3452-3459, 2014.
[9] S. Choy, C.S. Wang, T.K. Yeh, J. Dawson, M. Jia, Y. Kuleshov, “Precipitable water vapor estimates in the Australian region from ground-based GPS observations,” Advances in Meteorology, 2015(956481), 2015.
[10] R.C. Ssenyunzi, B. Oruru, F.M. D’ujanga, E. Realini, S. Barindelli, G. Tagliaferro, A. von Engeln, N. van de Giesen, “Performance of ERA5 data in retrieving Precipitable Water Vapour over East African tropical region,” Advances in Space Research, 65(8), pp. 1877-1893, 2020.
[11] H. Jiang, Y. Yang, Y. Bai, H. Wang, “Evaluation of the total, direct, and diffuse solar radiations from the ERA5 reanalysis data in China,” IEEE Geoscience and Remote Sensing Letters, 17(1), pp. 47-51, 2019.

Effects of edible coating on the Quality and Shelf-life of fresh-cut 'Kimju' guava (*Psidium guajava*)

Wisanee Nueamekin[1, a)], Pornteera Rattanarat[1, b)] and Panupat Polkham[1, c)]

[1]Thai Packaging Centre, Thailand Institute of Scientific and Technological Research (TISTR)
196 Phahonyothin Road, Chatuchak, Bangkok 10900, Thailand
Tel: +662 5791121 ext.3205

[a)] Corresponding author: Wisanee@tistr.or.th
[b)] Pornteera@tistr.or.th
[c)] Panupat@tistr.or.th

Abstract. Changes in appearance and physiochemical of fresh-cut fruits significantly impact consumer acceptance. This study explored the influence of edible coatings derived from 1% silk sericin with 2% ascorbic acid (SE), Aloe vera gel with 2% ascorbic acid (AA), and a commercial coating (ED) on the quality attributes of fresh-cut 'Kimju' guava (*Psidium guajava*) during storage at 8 ± 2 °C and 75% relative humidity for 16 days, with uncoated fresh-cut guava serving as the control. The application of SE, AA and ED coatings effectively maintained weight loss and firmness of fresh-cut guavas. SE and AA coated samples exhibited elevated total phenolic content and antioxidant activity compared to the control. Moreover, SE and ED coatings resulted in reduced bacterial, yeast, and mold counts compared to both the control and AA coated samples. In summary, SE emerges as a promising edible coating for fresh-cut guava, demonstrating its capacity to preserve quality and extend the shelf life of fresh-cut guava for 16 days.

***Index Terms*—Guava, Edible coating, Postharvest, Fresh cut**

I. INTRODUCTION

Fresh-cut or minimally processed fruits are very popular in Thailand. They are highly attractive to customers because of their high nutritious value, low price, and good flavor. However, minimally processed fruit or fresh-cut product is very susceptible to moisture loss, discoloration, and microbial spoilage, which are causes of shelf-life limitation [1]. Nowadays, the fresh-cut industry is growing rapidly and continues to expand because of the huge increase in demand for fresh-cut products.

'Kimju' guava (*Psidium guajava*) is considered as a popular fruit in Thailand due to its thin peel crispy flesh and good flavor. It is a highly palatable fruit with a rich source of vitamin C, dietary fiber and bioactive compound including antioxidants. In term of physical changes, guava is a climacteric fruit with high physiological metabolism such as, high respiratory rate and rapid maturation, resulting in a short period of commercialization. Generally, guava undergo rapid ripening in 3-4 days at room temperature. Rapid ripening is a main problem affecting quality of 'Kimju' guava fruit due to the fruit softening, chlorophyll content decreases and carotenoid content increases, skin brightness decrease, the development of skin browning and rot [2].

Edible coatings are low-cost, successful alternatives that create a modified atmosphere by forming a semi-permeable barrier against O_2, CO_2, and moisture which can reduce respiration, water loss, and oxidation rates [3], [4]. Edible materials such as proteins, polysaccharides and lipids that are applied on the surface of fresh-cut fruits to provide protection and prolong the shelf life of products. Sericin is a water-soluble globular protein derived from silk industry, was investigated for edible coating material. It has the potential in reducing the weight loss of fruits, color and texture improvements while inhibiting the enzyme polyphenol oxidase activity [5]. It also provides antimicrobial properties.

Aloe vera is a well-known plant for its marvelous medicinal properties. Recently, researchers have developed the edible coating for fruits from this gel due to its film-forming properties, prevent loss of moisture and firmness and antimicrobial properties [6].

However, there is a lack of data on the protective function of edible coating from sericin and aloe vera gel on fresh-cut guavas stored at lower temperatures. Thus, the objective of this study was to assess the influence of sericin and aloe vera gel coating, combined with ascorbic acid on the physicochemical, bioactive and microbiological properties of fresh-cut guavas during storage.

II. METHODOLOGY

Plant material

'Kimju' guavas utilized in this study were procured from the 'Or Tor Kor' market in Bangkok, Thailand. Guavas were meticulously chosen based on their size, ensuring the exclusion of fruits exhibiting physical injuries and physiological disturbances. Subsequently, the selected fruits underwent a thorough washing process with tap water, followed by immersion in a 100-ppm sodium hypochlorite

(CARLO ERBA Reagents, Germany) solution for 10 min and were air-dried at room temperature.

Preparation of Edible Coating Solution

The sericin edible coatings (SE) were prepared by dissolving 1% (w/v) sericin (NTGS Corporation Co., Ltd., Thailand) in deionized water. Subsequently, a 2% ascorbic acid (CARLO ERBA Reagents, Germany) solution (w/v) was added to the mixture in a 1:1 ratio.

For the preparation of aloe vera gel edible coatings (AA), the following procedure was followed: Aloe vera leaves were obtained from Prachuap Khiri Khan, Thailand. Initially, the leaves underwent washing with tap water followed by immersion in a 200-ppm sodium hypochlorite solution for 10 min. Subsequently, the aloe vera gel matrix was separated from the outer cortex of the leaves, resulting in a colorless hydro parenchyma which was then cut into cubes. These cubes were washed to remove mucilage and then blanched in hot water before being cooled using cold water. The aloe vera was then finely pulped in a blender. The resulting gel was filtered to remove any fibrous fraction. Fresh aloe vera gel was pasteurized at 75 °C for 30 min, rapidly cooled to room temperature, and subsequently frozen at -20 °C until required. Prior to use, the pasteurized gel was thawed to room temperature and mixed with a 2% ascorbic acid solution in a 1:1 ratio.

The commercial edible coating (ED) was purchased from Eden Agritech, Thailand.

Sample preparation and coating application

The guava was cut into uniform pieces after removing their seed portion and dipped for 3 min in the edible coating solution. The treated guava pieces were then air-dried, packed in 9.5×9.5×4.0 cm plastic trays, and top-sealed with Nylon/PE plastic film (100 g/tray with dimensions of 9). The samples were stored at 8±2 °C and analyzed on days 1, 2, 4, 6, 8, 14, and 16 after storage. An uncoated sample (C) was also prepared for comparison purposes.

Weight loss, Firmness and Color

To evaluate weight loss, separate samples in 3 replicates of each treatment were used. Fresh weight of fresh-cut guava in each tray was measured every sampling day. Weight loss of each sample was calculated as percentage with respect to the initial weight.

To evaluate firmness, samples were cut into rectangular shape with 1.5×1.5 cm (width×thickness). Firmness was measured as the maximum force (N) reached during cutting sample in the central zone with a Warner-Bratzler blade set. The cross-head speed was 2 mm/s using a texture analyzer (Lloyds, LS1, England).

The color of the guava pulp was assessed using a colorimeter (Konica Minolta, CR-400, Japan). The CIE colour system was used to evaluate the *L** (luminosity), *a** (redness) and *b** (yellowness).

Ascorbic acid, titratable acidity and soluble solids content

Ascorbic acid, titratable acidity and soluble solids content was determined in the juice extracted from ground samples, with three replicates of each treatment. Ascorbic acid content was measured using 2,6-dicholorophenol indophenols method described by AOAC (1990) [7]. A sample (30 g) was homogenized in 30 mL of 5% meta-phosphoric acid. The homogenate was filtered and ascorbic acid in the filtrate was determined. The results are expressed as mg/100 g fresh weight.

Titratable acid (TA) of 5 mL guava juice was evaluated through titration with 0.1 N NaOH (UNIVAR®, USA), which phenolphthalein used as an end-point indicator, and expressed as grams of citric acid/100 g fresh weight.

Soluble solids content was determined using Atago PAL-1 digital refractometer (Atago Co. Ltd., Tokyo, Japan) at 25 °C and expressed as °Brix.

Total phenolic content and antioxidant activity

After firmness analysis, guava was cut into small pieces and homogenized in a grinder, and 20 g of ground guava were grinded again in 20 mL of 80% (v/v) Ethanol (CARLO ERBA Reagents, Germany) and then filtered by using Whatman No.4 filter paper. The filtrate was used to determine total phenolic content and antioxidant activity. All tests were conducted in triplicate.

Total phenolic content (TPC) assessed by Folin-Ciocalteu assay with some modification. 20 µL of the extract was mixed with 2.5 mL of 10% (v/v) Folin–Ciocalteu reagent (CARLO ERBA Reagents, Germany) and then 2 mL of saturated 7.5% (w/v) Na_2CO_3 (CARLO ERBA Reagents, Germany) solution was added after 10 min. The mixture was left at room temperature for at least 60 min before measuring absorbance at 760 nm wavelength. The results of TPC were expressed as mg gallic acid equivalent per g of fresh weight DPPH scavenging ability was performed to evaluate antioxidant activity. The reduction of DPPH was measured by a UV–VIS spectrophotometer by recording the absorbance at 517 nm of the mixture of 20 µL extracts and 3 mL of 0.1 mM DPPH (CARLO ERBA Reagents, Germany) in Ethanol immediately after 60 min of incubation in dark room at ambient temperature. The results were expressed as %DPPH radical scavenging calculated by following equation.

$$\%DPPH\ Radical\ scavenging = \frac{A_{contol} - A_{sample}}{A_{control}} \times 100 \quad (1)$$

Whereas; $A_{control}$ = Absorbance of DPPH solution + Ethanol

A_{sample} = Absorbance of DPPH solution + sample extract

Microbiological analysis

Total aerobic plate count was measured following method of FDA BAM (chapter 3) [8]. Yeasts and molds content was measured following method described by A.O.A.C 997.02 (2023) [9]. All tests were conducted in three replicates.

Statistical analysis

The data were reported as the mean ± standard deviation and subjected to the one-way analysis of variance (ANOVA) using SPSS software (version 12.0, SPSS Inc., US). The differences between each treatment and days of storage were determined by using Duncan's multiple-range tests at a 95 % confidence level.

III. Results and Discussion

Weight loss, Firmness and Color

The experimental findings indicated a consistent rise in weight loss among fresh-cut guavas across all experimental groups throughout the storage period. Notably, starting from the 6th day onward, a considerable contrast was observed, with uncoated fresh-cut guavas displaying a substantially higher percentage of weight loss compared to the other coated samples. This observed difference was found to be statistically significant (p<0.05) (Fig. 1a). Significant differences in firmness were observed on the 4th day of storage. Coated fresh-cut guavas of all types exhibited higher firmness compared to the uncoated fresh-cut guavas (Fig. 1b). This experimental evidence suggests that the coating solutions tended to control weight loss and maintain the pulp density of fresh-cut guavas. This effect can be attributed to the coating layer covering the fruit surface, which restricts the exchange of oxygen and carbon dioxide between the interior and exterior of the coated fruits. Consequently, this limitation in gas exchange lowers the respiration rate of the fruit, thereby delaying weight loss. This delayed weight loss also contributes to a slower decrease in firmness [10], [11]. The findings of this study are consistent with previous research on tomatoes coated with sericin [12] and table grapes and papaya coated with aloe vera gel [6], [13].

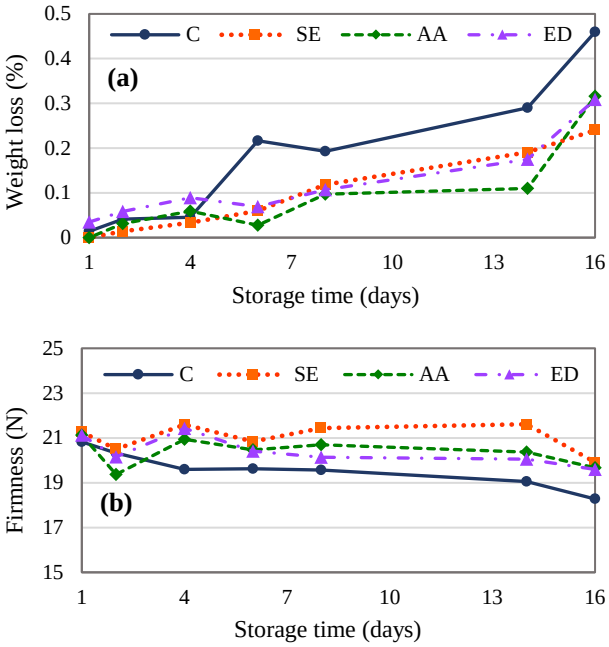


Fig. 1. Weight loss (a) and firmness (b) of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

Color is a crucial factor influencing the acceptability of fresh-cut fruits among consumers. It is reported that color changes are correlated with several biochemical activities such as degradation and synthesis of color pigments including chlorophylls and carotenoids in fruit. [14] In this study, the L^* value, indicating lightness, exhibited a significant increase (p<0.05) with storage time (Fig. 2a). Conversely, the a^* value, representing redness, decreased over time, indicating reduced browning. At the conclusion of the storage period, the uncoated sample displayed the highest a^* value (Fig. 2b). Changes in yellowness (b^*) of the samples during storage are attributed to enzymatic browning by polyphenol oxidase and the formation of pheophytin [15]. In this study, the sample coated with ED had lower b^* values compared to other coated and uncoated samples (Fig. 2c).

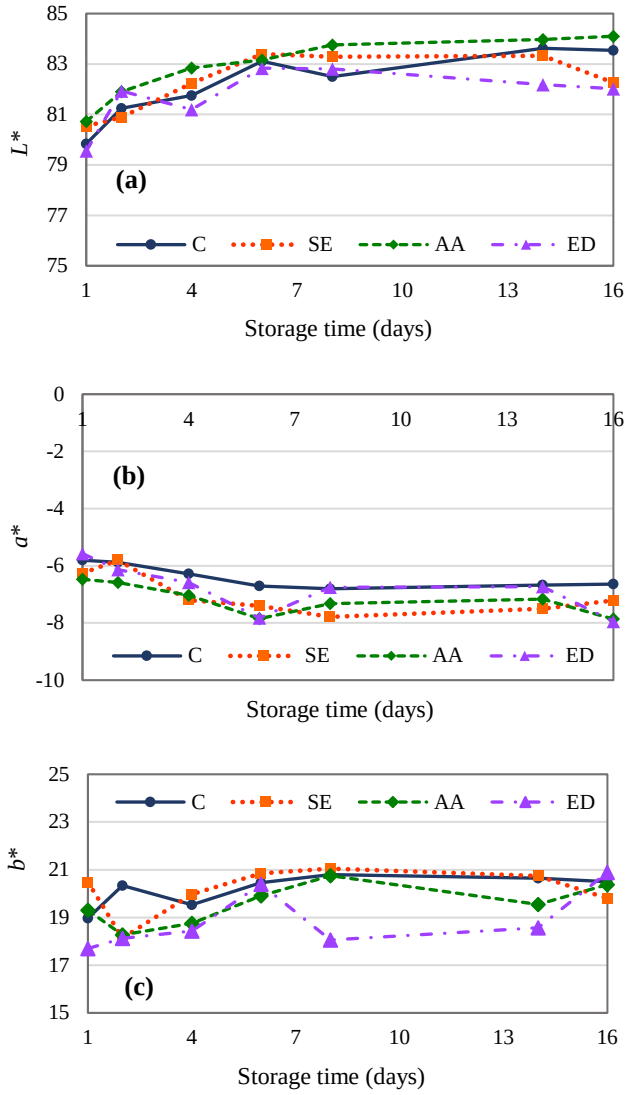


Fig. 2 Lightness (L^*) (a), redness (a^*) (b) and yellowness (b^*) (c) of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

Ascorbic acid, titratable acidity and soluble solids content

Vitamin C serves as a crucial indicator of guava quality [16]. By day 2nd of storage, the highest loss of ascorbic acid was observed in uncoated fresh-cut guava samples. ED coating exhibited a greater reduction in ascorbic acid loss compared to SE and AA coating, respectively. (Fig 3) These findings are consistent with the research conducted by Jung Chien et al. (2007) [17], who observed that low molecular weight chitosan-coated sliced red pitayas preserved their levels of ascorbic acid better than uncoated fruit.

The compositional change in terms of the TA is crucial to understand its metabolic processes such as ripening and senescence, which may affect the fruits quality. The TA of all samples significantly (p<0.05) decreased throughout storage period, with the uncoated sample exhibiting the most pronounced reduction. (Fig 4) These results align with the typical behavior observed in climacteric fruits, where the changes in the TA are influenced by its metabolic activities especially respiration. Respiration leads to the consumption of organic acids, resulting in a decrease in acidity [18].

The ripening of climacteric fruit such as guava may give rise to the soluble solid content due to polysaccharide degradation [19], whereby starch is broken down into soluble sugar, resulting in a flavor change, especially in terms of sweetness [20]. In this study, it was observed that there was no significant difference in total soluble solids content among the samples during the storage period (Fig 5).

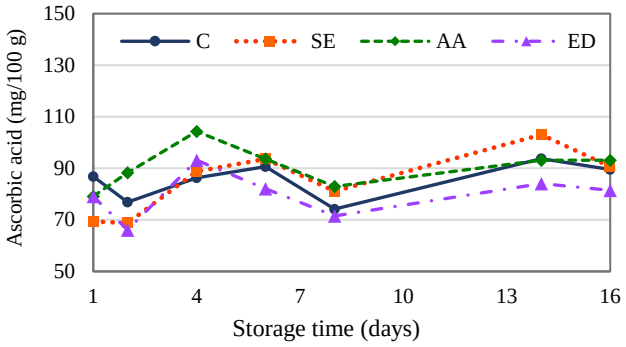


Fig. 3 Ascorbic acid of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

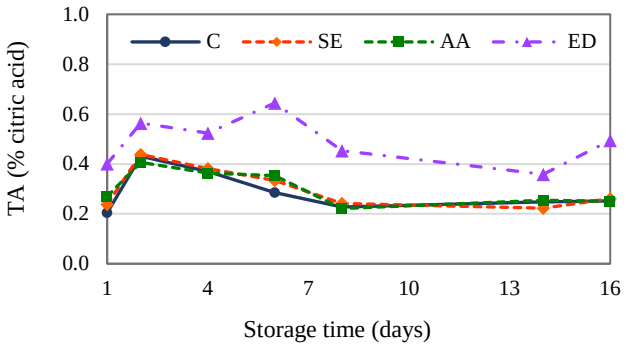


Fig. 4 Titratable acidity of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

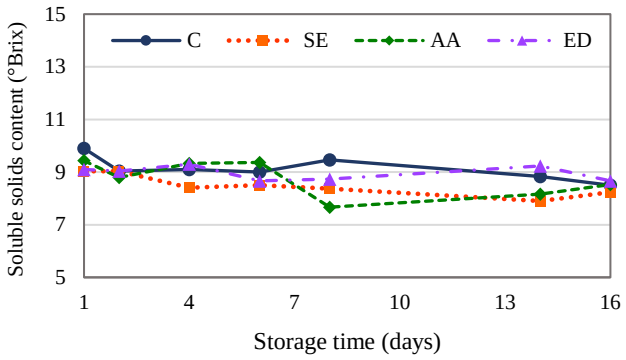


Fig. 5 Soluble solids content of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

Total phenolic content and antioxidant activity

The changes in the total phenolic content (TPC) in coated and control fresh-cut guavas during storage are shown in Fig. 6a. On the first day of storage, the control sample exhibited significantly lower TPC than the coated samples. This difference could be attributed to the incorporation of sericin and aloe vera gel in the coating solution, which served as a source of phenolic content. During storage period, TPC showed slight increased and decreased trend among the treatments which may be due to the metabolism during ripening that initially increase TPC followed by subsequent reductions due to oxidation reactions and senescence [21]. Armghan Khalid et al. (2019) [22] reported that the application of the edible coating on fruits could slow down the ripening process and contribute to higher retention of phenolic compounds. In this study, SE coating reduced the oxidation and effectively preserved higher phenolic content in fresh-cut guavas compared to control during storage. Moreover, fresh-cut guava coated with AA showed higher TPC throughout the storage period. This could be attributed to aloe vera gel containing various bioactive compounds, which serve as a rich source of phenolic content. [22]

The antioxidant activity, represented as %DPPH radical scavenging, of fresh-cut guavas is shown in Fig. 6b. The changes in antioxidant activity exhibited a trend similar to that observed in TPC. The lowest antioxidant activity was detected on 2nd day in SE, AA and ED coated samples, while the highest increase was detected on the 13th day in SE and AA coated samples, followed by a gradual decrease with the increase in the period of storage. These results are consistent with previous findings [21]. Moreover, fresh-cut guava coated with AA demonstrated higher antioxidant activity than the control throughout the entire storage period. Hasan & Nicolai (2014) [23] and Nair et al. (2018) [24] demonstrated that edible coatings modify the internal atmosphere of fruits, thereby maintaining metabolic activities and preventing the breakdown/oxidation of phenolics, flavonoids, carotenoids, and ascorbic acid [25], resulting in an increase in antioxidant activity in coated fruits during storage [26].

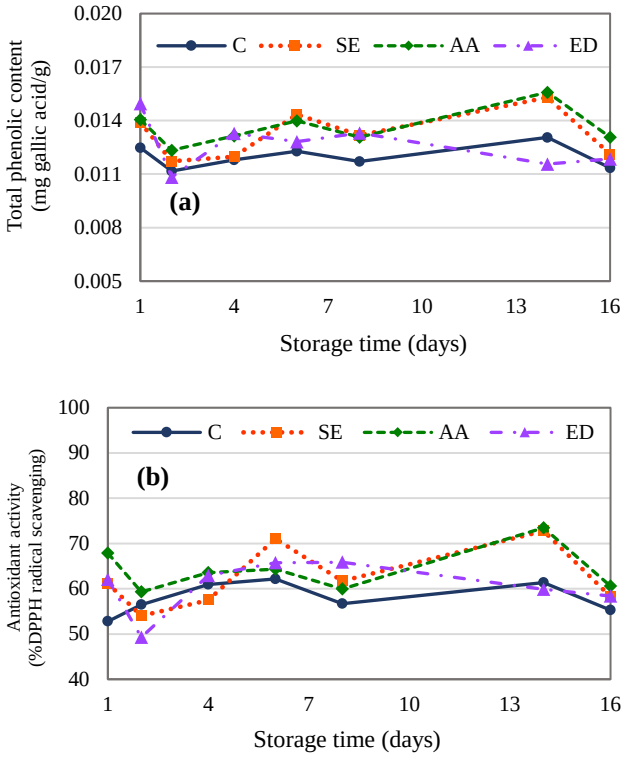


Fig. 6 Total phenolic content (a) and antioxidant activity (b) of fresh-cut guavas coated with SE, AA, ED and uncoated as a control during storage at 8±2 °C for 16 days

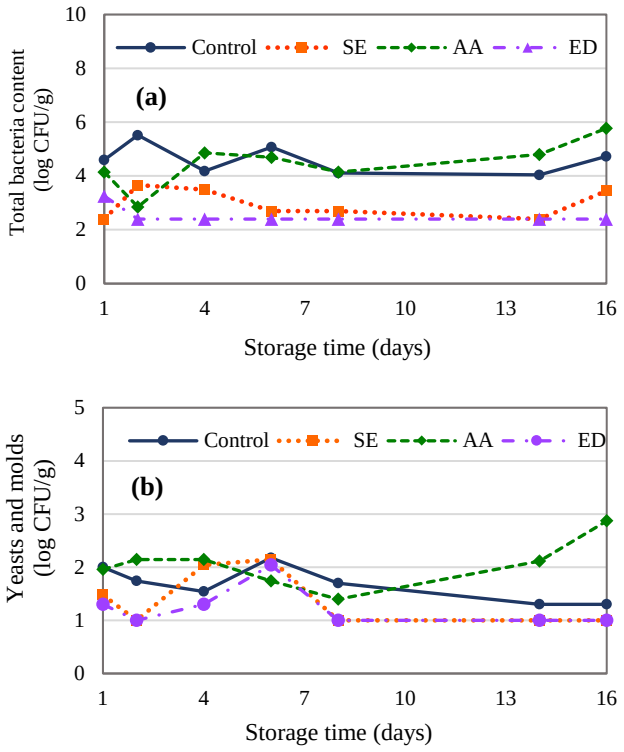


Fig. 7 Total bacteria content and yeasts and molds content in fresh-cut guava coated with various edible coatings and stored at 8±2 °C for 16 days

Microbiological properties

The microbial load in fresh-cut guavas is shown in Fig. 7a and 7b. Total bacterial content in both control and AA coated samples increased with storage time. The total bacterial content in control samples exceeded 10^4 CFU/g from the beginning of the study, which was significantly higher than in SE and ED coated samples (p<0.05). Coating samples with AA initially decreased total bacterial content of fresh-cut guava on the 2nd day, but then total bacterial content gradually increased (p<0.05). In contrast, both SE and AA coatings effectively inhibited bacteria growth in the samples. SE and AA coated samples maintained total bacterial content below 10^4 and 10^2 CFU/g, respectively. Yeasts and molds content exhibited a similar trend to total bacterial content, with SE and AA coated samples exhibiting yeasts and molds content lower than control and AA coated samples. These results may be attributed to the antimicrobial properties of sericin, which can delay the growth of microorganisms, thus extending the shelf life of fresh-cut guavas. The antimicrobial activity of sericin may stem from its role as a barrier in the coating solution, preventing easy access to substances such as water or other nutrients necessary for microbial growth naturally present in fruits. Additionally, it provides an extra barrier against microbial contamination during storage, thereby impeding their entry [12]. However, the antimicrobial activity of aloe vera gel was not demonstrated in this study.

IV. Conclusion

This study highlights the effectiveness of the developed edible coating derived from sericin and aloe vera gel compared to a commercial coating in preserving the physicochemical and microbiological properties of fresh-cut 'Kimju' guava. Coating fresh-cut guavas with SE, AA, and ED led to a reduction in weight loss percentage compared to non-coated guavas, with a significant reduction observed in the non-coated samples. The presence of SE coating was particularly effective in maintaining firmness, color, total soluble solids, and total acidity, while also delaying the degradation of ascorbic acid and increasing the total phenolic content and antioxidant activity in fresh-cut guavas during storage compared to the uncoated samples. Moreover, SE coating exhibited potential antimicrobial properties in fresh-cut guavas, although its effectiveness was less than commercial coating. Therefore, edible coatings derived from sericin hold promise for prolonging shelf life of fresh-cut guava. However, it is necessary to make efforts to improve their antibacterial properties to ensure competitiveness with commercial coatings.

Acknowledgment

This research project is supported by Thailand Science Research and Innovation (Fundamental Fund: fiscal year 2023-2024).

REFERENCES

[1] Baldwin, E. A., Nisperos-Carriedo, M. O. & Baker, R. A. Edible coatings for lightly processed fruits and vegetables. *HortScience* **30**, 35–38 (1995).
[2] Supapvanich, S. *et al.* Salicylic acid immersion maintains physiochemical quality and enhances bioactive compounds in 'kimju' guava fruit during cold storage. *Emirates Journal of Food and Agriculture* **29**, (2017).
[3] Dhall, R. K. Application of edible coatings on fruits and vegetables. *John Wiley & Sons, Inc. eBooks* 87–119 (2016)
[4] Raghav, P., Agarwal, N. & Saini, M. Edible Coating of Fruits and Vegetables: A Review. 2455-5630 (2016).
[5] Thongsook, T. & Tiyaboonchai, W. Inhibitory effect of sericin on polyphenol oxidase and its application as edible coating. *International Journal of Food Science & Technology* **46**, 2052–2061 (2011).
[6] Tripathi, P. & Dubey, N. K. Exploitation of natural products as an alternative strategy to control postharvest fungal rotting of fruit and vegetables. *Postharvest Biology and Technology* **32**, 235–245 (2004).
[7] AOAC. Official method 967.21. Vitamin C (ascorbic acid) in vitamin preparations and juices. *Official Methods of Analysis*. Washington DC: AOAC International; 1990.
[8] Maturin L, Peeler JT. Aerobic Plate Count. In: Food and Drug Administration (FDA), *Bacteriological Analytical Manual* Online. 2001.
[9] AOAC. Official method 997.02. Yeast and Mold Counts in Foods and Dried Cannabis Flower. *Official Methods of Analysis*. Washington DC: AOAC International; 2023.
[10] Pandey, S. K., Joshua, J. E., Bisen & Abhay. Influence of gamma-irradiation, growth retardants and coatings on the shelf life of winter guava fruits (Psidium guajava L.). *Journal of Food Science and Technology* **47**, 124–127 (2010).
[11] Vieira, J. M. *et al.* Effect of chitosan– Aloe vera coating on postharvest quality of blueberry (Vaccinium corymbosum) fruit. *Postharvest Biology and Technology* **116**, 88–97 (2016).
[12] Tarangini, K., Kavi, P. & Jagajjanani Rao, K. Application of sericin-based edible coating material for postharvest shelf-life extension and preservation of tomatoes. *eFood* **3**, (2022).
[13] Brishti, F. Effect of Biopreservatives on Storage Life of Papaya (Carica papaya L.). *International Journal of Food Studies* **2**, 126–136 (2013).
[14] Aina, J. O. & Oladunjoye, O. Respiration, pectolytic activity and textural changes in ripening African mango (Irvingia gabonensis) fruits. Journal of the Science of Food and Agriculture **63**, 451–454 (1993).
[15] Gil, M. I., Gorny, J. R. & Kader, A. A. Responses of `fuji' apple slices to ascorbic acid treatments and low-oxygen atmospheres. *Horticultural Science* **33**, 305–309 (1998).
[16] Shewfelt, R. L. Quality of minimally processed fruits and vegetables. *Journal of Food Quality* **10**, 143–156 (1987).
[17] Chien, P.-J., Sheu, F. & Yang, F.-H. Effects of edible chitosan coating on quality and shelf life of sliced mango fruit. *Journal of Food Engineering* **78**, 225–229 (2007).
[18] Khalifa, I., Barakat, H., El-Mansy, H. A. & Soliman, S. A. Enhancing the keeping quality of fresh strawberry using chitosan-incorporated olive processing wastes. *Food Bioscience* **13**, 69–75 (2016).
[19] Sharma, M. & Saini, C. S. Postharvest shelf-life extension of fresh-cut guavas (Psidium guajava) using flaxseed protein-based composite coatings. *Food Hydrocolloids for Health* **1**, 100015 (2021).
[20] Jongsri, P., Wangsomboondee, T., Rojsitthisak, P. & Seraypheap, K. Effect of molecular weights of chitosan coating on postharvest quality and physicochemical characteristics of mango fruit. *LWT* **73**, 28–36 (2016).
[21] Hasan, K., Islam, R., Hasan, M., Sarker, S.H. & Biswas, M.H. Effect of alginate edible coatings enriched with black cumin extract for improving postharvest quality characteristics of guava (Psidium guajava L.) Fruit. *Food and bioprocess technology* **15**, 2050–2064 (2022).
[22] Armghan Khalid, M. *et al.* Edible coatings for enhancing safety and quality attributes of fresh produce: A comprehensive review. *International Journal of Food Properties* **25**, 1817–1847 (2022).
[23] Hasan, S. M. K. & Nicolai, B. Quality of pears with permeability of Bio-FreshTM edible coatings. *African Journal of Food Science* **8**, 410–418 (2014).
20.
[24] Nair, M. S., Tomar, M., Punia, S., Kukula-Koch, W. & Kumar, M. Enhancing the functionality of chitosan- and alginate-based active edible coatings/films for the preservation of fruits and vegetables: A review. *International Journal of Biological Macromolecules* **164**, 304–320 (2020).
[25] Gonzalez-Aguilar, G. A., Villa-Rodriguez, J. A., Ayala-Zavala, J. F. & Yahia, E. M. Improvement of the antioxidant status of tropical fruits as a secondary response to some postharvest treatments. *Trends in Food Science & Technology* **21**, 475–482 (2010).
[26] Frusciante, L. *et al.* Antioxidant nutritional quality of tomato. *Molecular Nutrition & Food Research* **51**, 609–617 (2007).

A Comparative Study of a Preventive Maintenance Plan Using Historical Data: A Case Study of Skytrain in Thailand

Thirit Lertsumruai[1, a)], Kittchai Sojiphan [1, b)], Nuksit Noomwongs [2,c)] and Sompoap Talabgaew [3, d)]

Author Affiliations
[1] Railway Vehicles and Infrastructure Engineering Program, The Sirindhorn International Thai-German Graduate School of Engineering (TGGS), King Mongkut's University of Technology North Bangkok 1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok 10800 Thailand
[2] Department of Mechanical Engineering, Faculty of Engineering, Chulalongkorn University 254 Phayathai Road, Wangmai, Pathumwan, Bangkok 10330
[3] Department of Teacher Training in Mechanical Engineering, King Mongkut's University of Technology North Bangkok 1518 Pracharat 1 Rd., Wongsawang, Mangsue, Bangkok 10800 Thailand

Author Emails
[a)] s6409096860051@kmutnb.ac.th
[b)] Kittichai.s@tggs.kmutnb.ac.th
[c)] Nuksit.n@chula.ac.th
[d)] Corresponding author: Sompoap.t@fte.kmutnb.ac.th

Abstract. The number of people using the electric train was 22.97 percent, second only to buses. Ensuring high availability is crucial for providing reliable services. Therefore, applying the RAM (Reliability, Availability, and Maintainability) method to compare with current maintenance plans can help ensure system availability. Additionally, there are occasional shortages of spare parts for the brake system, posing challenges for maintenance. Utilizing reorder points and RAM to calculate the frequency of device replacements per year can help determine the appropriate storage levels. This research will be divided into two parts: comparing the results from the RAM method with the current maintenance plan and applying reorder point and RAM to determine annual usage.

***Index Terms*— RAM, Reliability, Availability, Rolling stock**

I. INTRODUCTION

The number of people using the electric train was 22.97 percent, second only to buses source from office of transport and traffic policy and planning. High availability is necessary to reduce service overload. Therefore, implementing RAM principles in maintenance planning, considering the real operational conditions of the rail system, is essential [1]. This strategic approach aims to enhance system reliability and availability by tailoring maintenance strategies to the actual usage patterns and demands of the system. It will help in adjusting the maintenance plan to be more efficient [2].

Due to the change of ownership, there has been some missing information passed on. This makes managing the spare parts difficult and causes problems causing the spare parts to be insufficient for maintenance. This problem often occurs with the brake system. The braking system directly affects the safety of train operations and passengers.

One way to increase efficiency in maintenance is to calculate the amount of maximum used per year so that spare parts are sufficient for maintenance.

The scope of this study encompasses data collection and the formulation of a research framework based on the EN 50126 [3]. Firstly, data utilized for analysis will be sourced from the skytrain in Thailand. Secondly, the study will employ a reliability model to calculate optimal maintenance timings, subsequently evaluating these against the current maintenance plan for all systems. Third, Utilization of RAM method will be central in determining the maximum level of spare parts required each year for effective management of the brake system. Fourth, Utilization of safety stock and reorder point formulas will be central in determining the appropriate level of spare parts required for effective management of the brake system.

II. METHODOLOGY

Gather information and study damage to parts. Information required for analyzing RAM and managing spare parts can be categorized into two main sections. The set of information, utilized for RAM analysis, involves fault data

necessary for applying RAM methodologies. This includes the notification time to calculate the duration until failure, the start and end timestamps of malfunctions to assess repair time, identification of the system encountering the failure, and the specific train number associated with the occurrence of the failure.

For analyzing the RAM of all system in this research involves a comprehensive methodology: Data Filtering, Train Categorization, Time to Failure (TTF) and Time to Repair (TTR) Calculations, Reliability Model Selection, Maintainability Assessment, Availability Computation, Time to Maintenance Determination.

In this research will separate the system into three groups, including the system that directly impacts train operation and passenger safety, the system that affects passenger comfort, and the system that does not affect train operation and passenger comfort from engineer at maintenance section. Group 1, Group 2, and Group 3 will have reliability set to 95%, 90%, and 85%, respectively. TABLE I provides the list of systems in each group.

TABLE I
GROUPING THE SYSTEM

Group1	Group2	Group3
Exterior Doors Systems	Heating Ventilating, Air Cond.	Driver Cab
Brake Systems	Interior Lighting	Interior Doors
Auxiliary Converter Unit	Passenger Information System	Passenger Compartment
Traction Converter		Exterior Lighting
Under Floor Equipment's		
Pantograph		

To analyze the RAM for all systems in this research, the initial step involves filtering the data. Subsequently, the trains will be categorized by their numbers, and the train with the highest number of faults will be selected. Following this, TTF will be calculated using the Notification date, and TTR will be derived using malfunction start/end dates. Then, a table depicting the reliability model will be created, allowing the choice between the Weibull and exponential models. The same methodology will be applied to assess maintainability. Then, availability will be computed, and the time to maintenance will be determined based on reliability calculations. Finally, compare the result with the actual plan and discuss.

Reliability Analysis

Reliability refers to the likelihood of an engineered component or system operating flawlessly as per its specifications within the operational conditions and timeframe [4]. The analysis of reliability models involves assessing the compatibility of statistical models using historical and current maintenance data.

The linearization method, reliant on Linear Regression, is a significant aspect of this process. The guiding principle revolves around comparing Weibull and Exponential models [5], aiming to derive values where the Coefficient of Determination (R^2) is at its highest. The operational approach involves crafting a Reliability Plotting Diagram using the regression technique.

The Failure and reliability functions of the Weibull model is:

$$In[-InR(\mathrm{t})] = \beta Int - \beta In\eta \quad (1)$$

The Failure and reliability functions of the Exponential model is:

$$-InR(\mathrm{t}) = \lambda t \quad (2)$$

Availability Calculation

Availability quantifies the operational preparedness of equipment, representing the ratio of the actual working duration to the total time [6], which encompasses both working time and downtime as shown in (8). It provides a comparative measure indicating the extent of readiness exhibited by the equipment for operational purposes.

Mean Time to Repair (MTTR) is the average duration taken to fix a failed system or component and return it to normal operation.

$$\mathrm{MTTR} = \frac{\Sigma TTR}{n} \quad (3)$$

Mean time to failure (MTTF) is the difference between Weibull function and exponential function are according to the equation below as follows and availability is the same.

$$\mathrm{MTTF}_{Weibull} = \eta \quad (4)$$

$$\mathrm{MTTF}_{Exp} = \frac{1}{\lambda} \quad (5)$$

In this research, the maximum spare parts will be calculated solely for the brake system. This will be achieved by utilizing reliability analysis with the spare parts [11] and the problem of lack of stock can be solved by using reorder point.

Safety Stock

Determining the appropriate amount of safety stock involves balancing the risk of stockouts against the carrying costs of excess inventory [7]. Companies need to consider various factors, such as demand patterns, lead times, service level goals, and the trade-offs between inventory costs and potential lost sales or customer dissatisfaction due to stockouts.

$$\mathrm{SS} = (\text{Max Lead Time} \times \text{Max Used}) - (\text{Avg. Lead time} \times \text{Avg. Used}) \quad (6)$$

Reorder Point

The reorder point is the inventory level at which a new order for replenishment is triggered. It is determined by considering the lead time to receive inventory and the expected demand during that lead time, along with the safety stock. When the actual inventory level reaches the reorder point, a new order is placed to restock and maintain adequate inventory levels [8, 9]

$$\text{ROP} = \text{SS} + \text{Avg. Used} \times \text{Lead time} \quad (7)$$

Figure 1 displays the components of the brake system ranked from the highest to the lowest failure rates. The top five components will be used to calculate the reorder point using equations (6) and (7). Subsequently, MTTF will be computed using reliability data, and the number of maximum uses per year will be obtained by dividing the number of days in a year by the MTTF.

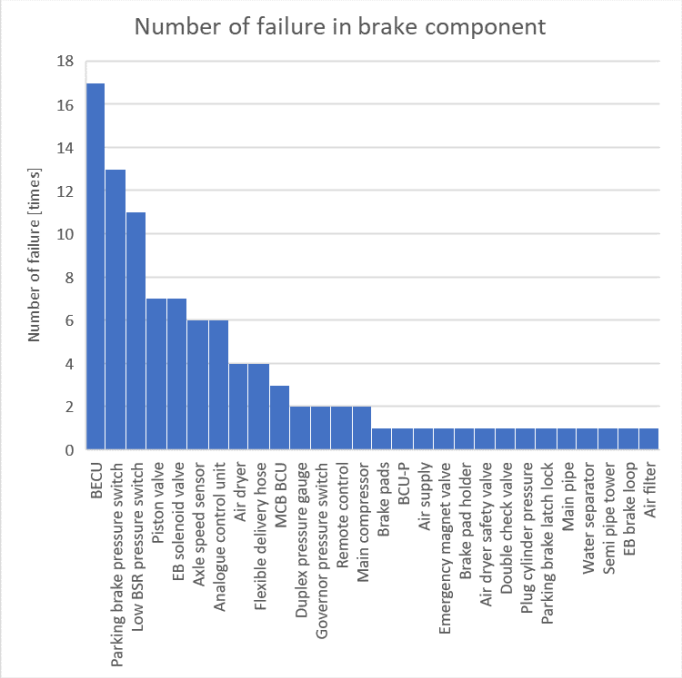


Figure 1 Failure count of spare parts in brake system

III. RESULTS AND DISCUSSION

There are 9 trains in this system, separated into 2 types: Model A, which has 4 carriages, and Model B, which has 3 carriages.

The analysis of the Exterior Door System as shown in TABLE II reveals that the most frequent failures occur in Model A-2, Model A-1, and Model B-1 , with an average time to failure of 28.5 days. Preventive maintenance is conducted monthly. The observed average time to failure closely aligns with the standard and the existing preventive maintenance plan, highlighting the effectiveness of implemented strategies. This comprehensive plan includes rigorous function tests, measurements, and calibration checks, ensuring the continued efficiency and safety of the door systems while enhancing the overall passenger experience. Such systematic evaluations mitigate potential disruptions and uphold the system's operational integrity.

TABLE II
MTTF OF EXTERIOR DOOR SYSTEM

Variable	A-2	A-1	B-1
R_Beta	1.587	1.408	1.344
R_Neta	435.7	531.7	546.0
R_Lamda	0.934	0.714	0.677
R target	95%	95%	95%
MTTF(d.)	24.65	30.08	30.89
Current plan(d.)	30	30	30

The examination of the Interior Lighting System as shown in TABLE III reveals that the predominant failures occur in Model B-1, Model A-1, and Model A-2, with an average duration to failure of 49.2 days. This system primarily consists of light bulbs installed within the gangway and saloon areas. Unlike other systems, it currently lacks a structured preventive maintenance regimen. Instead, maintenance interventions are triggered reactively, either upon detection by the driver or notification from a technician, ensuring prompt corrective actions before the train is deployed for service. However, this reactive approach underscores the need for further assessment to potentially introduce proactive measures that could enhance reliability and minimize disruptions within the system.

TABLE III
MTTF OF INTERIOR LIGHTING SYSTEM

Variable	B-1	A-1	A-2
R_Beta	1.098	1.246	1.517
R_Neta	504.7	434.8	460.0
R_Lamda	0.547	0.661	0.921
R target	90%	90%	90%
MTTF(d.)	53.19	45.82	48.47
Current plan(d.)	-	-	-

The three most common failures of the Driver Cab as shown in TABLE IV occur in Model B-3, Model B-2, and Model A-3, with an average time to failure of 56.1 days. This system encompasses the Driver console and Driver seat. Furthermore, it does not have a preventive maintenance plan, similar to the interior lighting system. Maintenance is conducted based on notifications from the Driver and Technician team.

TABLE IV
MTTF OF DRIVER CAB

Variable	B-3	B-2	A-3
R_Beta	1.305	0.936	1.205
R_Neta	423.1	311.2	301.2
R_Lamda	0.757	0.562	0.729
R target	85%	85%	85%
MTTF(d.)	68.77	50.58	48.95
Current plan(d.)	-	-	-

The analysis reveals that the most prevalent failures within the Air Conditioning System as shown in TABLE V are predominantly noted in Model A-2, Model B-3, and Model B-1, with an average duration to failure of 41.6 days. This system, critical for ensuring passenger comfort, encompasses 2 HVAC units situated within Saloon, as well as a single unit serving the Driver Cab. Scheduled maintenance for this system is currently set on a monthly basis, involving comprehensive tasks such as system cleaning, filter replacements, and thorough inspections to preempt potential issues. This can be extended to 40 days.

TABLE V
MTTF OF AIR CONDITIONING SYSTEM

Variable	A-2	B-3	B-1
R_Beta	1.051	0.890	1.195
R_Neta	330.1	538.9	316.4
R_Lamda	0.611	0.416	0.724
R target	90%	90%	90%
MTTF(d.)	34.78	56.78	33.34
Current plan(d.)	30	30	30

The analysis indicates that the three most prevalent failures within the Brake System as shown in TABLE VI predominantly occur in Model B-5, Model B-2, and Model A-3, with an average time to failure of 27.6 days. As a critical component ensuring the safety and operational integrity of the train. Maintenance for the brake system is meticulously arranged on a weekly, biweekly, and monthly basis, involving various inspections and tests. Overall, the value from RAM method is more than the existing preventive maintenance plan. So, the interval of preventive maintenance can be extended to about 25 days.

TABLE VI
MTTF OF BRAKE SYSTEM

Variable	B-5	B-2	A-3
R_Beta	3.195	2.505	1.565
R_Neta	535.5	506.1	420.0
R_Lamda	1.886	1.583	0.997
R target	95%	95%	95%
MTTF(d.)	30.30	28.63	23.76
Current plan(d.)	7	7	7

The three most frequent failures of the Auxiliary Converter System as shown in TABLE VII are observed in Model A-4, Model B-1, and Model A-2, with an average failure time of 13.4 days. The maintenance tasks of this system include visual inspection and reading the data log every week. If there is anything wrong in the data log, the maintenance will be addressing any failures on a case-by-case basis. From the result, it seems that the preventive maintenance can be extended to 2 weeks, but the data log is included in other system. So, it should be downloaded every time that train comes to depot.

TABLE VII
MTTF OF AUXILIARY CONVERTER SYSTEM

Variable	A-4	B-1	A-2
R_Beta	1.109	1.159	0.490
R_Neta	316.5	273.0	195.8
R_Lamda	0.748	0.820	0.287
R target	95%	95%	95%
MTTF(d.)	16.23	14.00	10.05
Current plan(d.)	7	7	7

The Traction Converter System experiences its three most common failures in Model A-3, Model B-5, and Model A-1, as shown in TABLE VIII with an average downtime of 16.3 days. The scheduled maintenance tasks of this system include visual inspection and reading the data log and done every week similar to auxiliary converter system. The interval of maintenance can be extended to 14 days.

TABLE VIII
MTTF OF TRACTION CONVERTOR SYSTEM

Variable	A-3	B-5	A-1
R_Beta	1.611	1.191	1.317
R_Neta	389.5	283.6	275.3
R_Lamda	0.956	0.697	0.734
R target	95%	95%	95%
MTTF(d.)	19.98	14.55	14.12
Current plan(d.)	7	7	7

The Interior doors system encounters its top three prevalent failures primarily in Model B-5, Model B-2, and Model A-1 as shown in TABLE IX, typically resulting in an average downtime of 45.6 days. It is noteworthy that this system operates as a secondary system in conjunction with the exterior door system, which encompasses functionalities such as open/close status lights and warning signals before door operation, a protocol that undergoes inspection before each service cycle. This system can extend the interval of maintenance, but this system does preventive maintenance along with the exterior door system. So, the interval should be based on the exterior door system.

TABLE IX
MTTF OF INTERIOR DOORS SYSTEM

Variable	B-5	B-2	A-1
R_Beta	0.746	1.336	0.860
R_Neta	117.407	271.326	453.882
R_Lamda	0.506	0.919	0.551
R target	85%	85%	85%
MTTF(d.)	19.08	44.10	73.76
Current plan(d.)	30	30	30

The passenger information system experiences its primary Three common failures mainly in Model B-1, Model A-1,

and Model A-2 as shown in TABLE X, with an average mean time to failure of 34.7 days. This system facilitates communication between the train driver and passengers. As part of the weekly preventive maintenance routine, the technician will test the function of PA handset, which serves as the device for the train driver to relay information to passengers. The passenger alarm unit, considered the most critical subsystem of the passenger information system, is subject to monthly maintenance and is utilized by passengers in emergency situations to alert the driver. Comparing the maintenance plan with the mean time to failure, the existing plan is aligned with the result from RAM method.

TABLE X
MTTF OF PASSENGER INFORMATION SYSTEM

Variable	B-1	A-1	A-2
R_Beta	1.067	0.783	0.531
R_Neta	289.7	580.9	226.2
R_Lamda	0.627	0.392	0.382
R target	90%	90%	90%
MTTF(d.)	30.53	61.21	23.84
Current plan(d.)	30	30	30

The top three failures of underfloor equipment are observed in Model A-3, Model B-1, and Model B-4 as shown in TABLE XI, with a mean time to failure equal to 15.76 days. The relatively low mean time to failure is attributed to the inclusion of all undercarriage equipment such as the air dryer, main transformer, and main compressor in this system. However, the preventive maintenance procedures for this system, conducted on a weekly basis, involve visual inspections and measurements of certain equipment, such as checking the level of coolant in the main transformer, conducting air dryer function tests, and draining water from the main air supply tank. The Mean Time to Failure value of this system already aligns with the current maintenance plan and thus does not necessitate any modifications.

TABLE XI
MTTF OF UNDER FLOOR EQUIPMENT'S

Variable	A3	B1	B4
R_Beta	0.843	0.624	0.805
R_Neta	434.6	267.8	219.3
R_Lamda	0.461	0.372	0.560
R target	95%	95%	95%
MTTF(d.)	22.29	13.74	11.25
Current plan(d.)	14	14	14

The top three failures of exterior lighting system observed in Model A-4, Model B-2, and Model A-2 as shown in TABLE XII, with an average time to failure of 25.9 days, highlight critical issues within the system. Ensuring the proper functioning of both the headlight and rear light before operating the train is imperative for adhering to the safety protocols established by the train operator, aimed at minimizing the risk of accidents and collisions. Technicians and train drivers consistently inspect these lights prior to operation, prioritizing the immediate resolution of any detected failures to maintain operational safety. This system does preventive maintenance every week, it can increase to 25 days once a time.

TABLE XII
MTTF OF EXTERIOR LIGHTING SYSTEM

Variable	A-4	B-2	A-2
R_Beta	0.809	2.347	0.711
R_Neta	433.5	531.6	530.8
R_Lamda	0.463	1.351	0.387
R target	95%	95%	95%
MTTF(d.)	22.24	27.27	27.23
Current plan(d.)	7	7	7

The top three failures of passenger compartment observed in Model A-1, Model A-3, and Model A-2 as shown in TABLE XIII underscore critical issues pertaining to essential equipment within the train's saloon, including handrails, passenger seats, and fire extinguishers. With a mean time to failure of 34.72 days for the passenger compartment. Visual inspections are incorporated into the weekly preventive maintenance plan. This plan can extend to once a month because of the relatively low probability of failures within this system.

TABLE XIII
MTTF OF PASSENGER COMPARTMENT

Variable	A-1	A-3	A-2
R_Beta	0.667	0.992	2.100
R_Neta	161.2	218.9	260.8
R_Lamda	0.364	0.593	1.249
R target	85%	85%	85%
MTTF(d.)	26.20	35.58	42.39
Current plan(d.)	7	7	7

Only Model A-3 and Model B-2 can be analyzed for pantograph reliability, as Weibull's method requires a minimum of four data points for calculation as shown in TABLE XIV. The mean time to failure, at 17.4 days, is notably low, primarily attributed to a misunderstanding between the supplier and the rolling stock department concerning the drain plug at the carbon current collector. Rectifying this issue necessitates replacing the rubber plug with a steel plug. The system undergoes bi-weekly preventive maintenance. It is anticipated that the mean time to failure value may see an increase in the future and preventive maintenance plan can be increased too.

TABLE XIV
MTTF OF PANTOGRAPH

Variable	A-3	B-2
R_Beta	1.957	0.812
R_Neta	339.2	339.8
R_Lamda	1.571	0.519
R target	95%	95%

MTTF(d.)	17.40	17.43
Current plan(d.)	14	14

In this research, it is revealed that the overall availability of each train surpasses 98 percent, exceeding the designated Key Performance Indicator (KPI) set at 95 percent by the train operator as shown in TABLE XV. This remarkable achievement underscores the effectiveness of the maintenance strategies implemented and their significant contribution to ensuring the reliability and operational efficiency of the train fleet. Such high availability rates not only meet but exceed the performance expectations established by the operator, thereby enhancing the overall quality of service provided to passengers.

TABLE XV
AVAILABILITY

Variable	B-1	B-2	B-3	B-4	B-5
MTTF	441.50	481.70	517.98	515.17	497.78
MTTR	6.68	4.49	6.05	4.54	3.80
A(t)	98.51%	99.07%	98.85%	99.13%	99.24%
Variable	A-1	A-2	A-3	A-4	
MTTF	564.49	486.07	475.60	438.32	
MTTR	6.92	4.98	9.21	5.03	
A(t)	98.79%	98.99%	98.10%	98.86%	

TABLE XVI shows the safety stock, reorder point and maximum used per year from reliability model. When compared to current stock per year. It seems that the problem of insufficient spare parts can be solved by increasing the stock per year to equal to result from reliability model.

TABLE XVI
MAXIMUM USED OF CRITICAL SPARE PARTS IN BRAKE SYSTEM

Component	SS (pcs.)	ROP (pcs.)	MTTF (d.)	Max. used per year (pcs.)	Current stock (pcs.)
BECU	4	5	26	14	10
Parking brake pressure switch	6	8	27	13	9
Low BSR pressure switch	6	8	30	12	9
Piston valve	1	2	29	13	10
Solenoid valve	14	16	20	18	18

IV. CONCLUSION

In the first part, RAM was employed to analyze each system. The system that current preventive maintenance plan is align with the result from RAM method including exterior door system, interior door system, passenger information system, under floor equipment, and pantograph. Next group is the system that can extend the interval of schedule maintenance including air conditioning system, brake system, auxiliary convertor system, traction convertor system, exterior lighting system, and passenger compartment. So, the cost of maintenance can be reduced. Lastly, there are zero system that is needed to cut the maintenance time. This mean the current maintenance is of a good standard.

The next part concerns spare parts levels. The use of reliability to determine the amount that should be stocked per year is more than the current stock plan each year. There will be some parts that are not enough for use. If the calculation results from RAM method are applied, it will help reduce the problem of not having enough spare parts and also help increase the readiness of the system.

ACKNOWLEDGMENT

The authors would like to thank teams from SIEMENS Mobility Thailand for their kind support of time and resources for this study project.

References

[1] M. G. Park, RAMS management of railway systems, University of Birmingham, 2013.

[2] P. Srisombat , Data-Driven Approach to Bogie Overhaul Planning, Bangkok: The Sirindhorn International Thai-German Graduate School of Engineering King Mongkut's University of Technology North Bangkok, 2022.

[3] EN 50126:2017 Railway Applications - The Specification and Demonstration of Reliability, Availability, Maintainability and Safety (RAMS), 2017.

[4] S. Rezvanizaniani, J. Barabady, M. Valibeigloo, M. Asghari and U. Kumar, "Reliability Analysis of the Rolling Stock Industry: A Case Study," *International Journal of Performability Engineering*, vol. 5, pp. 167-175, 2009.

[5] S. Talabgaew , Reliability, Availability, Maintainability Approach for Infrastructure and Logistics of Railway Operations,, Bangkok: King Mongkut's University of Technology North Bangkok, 2020.

[6] C. N. REALES, X. R. Ramon and D. M. Rodríguez, "Reliability, Availability and Maintainability Study of a Light Rail Transit System,," UNIVERSITAT POLITÈCNICA DE CATALUNYA.

[7] T. Stoilov and K. Stoilova, "Inventory and Safety Stock Optimization," *2023 XXXII International Scientific Conference Electronics (ET)*, pp. 1-4, 2023.

[8] R. R. Isnanto, D. A. Efrilianda and Mustafid, "Inventory control systems with safety stock and

reorder point approach," *2018 International Conference on Information and Communications Technology (ICOIACT)*, pp. 844-847, 2018.

[9] R. Rosihan, S. Kartika, J. Supratman, M. Mr and R. Kumalasari, "Analysis of raw material inventory control for hinge upper assembly products using the economic order quantity method," *Teknologi dan Informatika*, pp. 120-125, 2024.

[10] "A Computer Implementation of the Separate Maintenance Model for Complex-system Reliability," *Quality and Reliability Engineering International*, vol. 22, pp. 757-770, 2006.

[11] N. Regan, The RCM solution: reliability-centered maintenance,, New York: Industrial Press, Inc., 2012.

[12] F. Fang, "Application of Reliability-Centered Maintenance in Metro Door System," in *IEEE Access 12*, 2019.

[13] J. Olivieri, "Modeling multiple preventative maintenance actions in a RAMS analysis," *2015 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-5, 2015.

[14] Z. Li, J. Guo and R. Zhou, "Maintenance scheduling optimization based on reliability and prognostics information," *2016 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-5, 2016.

[15] Z. Fu, G. Wang, F. Gao, X. Tian, Y. Li and B. Lu, "Review of high-speed train maintenance," *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pp. 419-422, 2012.

[16] W. Y. Yun, Y. J. Han and G. Park, "Optimal preventive maintenance interval and spare parts number in a rolling stock system," *2012 International Conference on Quality, Reliability, Risk, Maintenance, and Safety Engineering*, pp. 380-384, 2012.

[17] J. Guo, Z. Li and J. Wolf, "Reliability centered preventive maintenance optimization for aircraft indicators," *2016 Annual Reliability and Maintainability Symposium (RAMS)*, pp. 1-6, 2016.

[18] K. M. Ghafour, "Optimising safety stocks and reorder points when the demand and the lead-time are probabilistic in cement manufacturing," *2018 International Journal of Procurement Management*, vol. 11, pp. 387-388, 2018.

[19] H. Yijun and W. Jingjing, "Study on safety stock control based on System Dynamics," *2014 11th International Conference on Service Systems and Service Management (ICSSSM)*, pp. 1-5, 2014.

Rolling Stock Maintenance Analysis by Using Data-Driven Approach: A Case Study of Green Line Skytrain HVAC System

Suchanet Jaroenjittakam[1, a)], Kittichai Sojiphan[1, b)], Chayut Ngamkhanong[2, c)] and Sompoap Talabgaew[3, d)]

Author Affiliations
[1]Railway Vehicles and Infrastructure Engineering Program, The Sirindhorn International Thai-German Graduate School of Engineering (TGGS), King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok 10800 Thailand
[2]Advanced Railway Infrastructure, Innovation and Systems Engineering (ARIISE) Research Unit, department of Civil Engineering, Faculty of Engineering, Chulalongkorn University, Phayathai Rd., Pathumwan, Bangkok 10330 Thailand
[3]Department of Teacher Training in Mechanical Engineering, King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok 10800 Thailand

Author Emails
[a)] s6409096860085@email.kmutnb.ac.th
[b)] Kittichai.s@tggs.kmutnb.ac.th
[c)] Chayut.ng@chula.ac.th
[d)] Corresponding author: Sompoap.t@fte.kmutnb.ac.th

Abstract. The expanding Bangkok railway network demands top-notch reliability and availability of rolling stock for safety and efficiency. Ensuring the optimal performance of machinery and equipment is crucial, especially for maintenance of the rolling stock's components. Thus, scheduling maintenance to minimize downtime and analyzing root causes are imperative. One of the major and problematic maintenance tasks is a Heating, Ventilation, and Air Conditioning (HVAC) corrective maintenance due to the hot weather and excessive loads in Bangkok, which requires meticulous attention to reliability, availability during operational hours, and maintainability during breakdown durations. The first part of the paper describes the implementation of Reliability, Availability, and Maintainability (RAM) to prioritize critical systems based on their impact on technical casualty and frequency of occurrence. Secondly, Failure Mode, Effects and Criticality Analysis (FMECA) methodology for the HVAC system in rolling stock applications identifies potential failure modes, their effects, and criticality levels in both qualitative and quantitative approaches, then facilitating proactive maintenance duration and spare parts management. The entire procedure employs the failure rate as a systematic connection between various elements. Finally, the findings contribute to rescheduling a reliability-centered maintenance (RCM) as an alternative preventive maintenance for HVAC system of rolling stock and improve the maintainability of HVAC system by inventory management according to reliability analysis. The spares are stored on an annual basis to ensure sufficient spare parts for critical items. This eliminates waiting times, which minimizes the maintenance duration of 22.04% leaving only practical maintenance that is influenced by the number of hours dedicated to on-site repairs. Additionally, this study can be utilized for guiding data-driven strategic decision-making for other rolling stock systems.

***Index Terms*— Reliability analysis, HVAC system, Failure analysis, FMECA, Spare parts management.**

I. INTRODUCTION

The light rail and heavy rail including underground railway projects in Bangkok have significantly expanded the network by leveraging its strengths in safety and convenience as an outcome with 18 hours of operational hours per day. It offers a compelling service model that motivates customers to use it even more. However, the machinery and equipment's availability and performance are critical to the overall system [1], particularly when it comes to the rolling stock system's maintenance of its numerous parts. It is regarded as a primary task that is critical to the effectiveness and lifespan of various railway system equipment. When a malfunction or interruption stops the service, it will directly affect the entire system. For instance, if there are not enough trains delivered to meet demand, there will be delays that do not follow the planned train operation schedule, which will result in the system not operating as fully efficiently as it could and being less reliable.

The most comprehensive general maintenance strategies are Corrective Maintenance (CM) and Preventive Maintenance (PM). These maintenance methods involve maintenance costs and operating costs [2]. Reliability Centered Maintenance (RCM) regarding the resolution of

these issues, RCM is beginning to gain recognition. However, rolling stock system encounters the multifaceted challenge of enhancing the Reliability, Availability, and Maintainability (RAM) leading to suboptimal reliability and performance [3].

Rolling stock comprises of various complex systems. One of the major and problematic maintenance tasks is a Heating, Ventilation, and Air Conditioning (HVAC) corrective maintenance [4] due to the high temperature and excessive loads in peak hours in Bangkok, Thailand. It is frequently challenging to quickly obtain an ideal schedule that takes maintenance requirements into account when dispatchers schedule manually [5], since HVAC system is an important system for delivering vehicles to service providers, including the convenience and satisfaction of using the railway transportation.

There is a few research that involve the RAM approach with rolling stock maintenance [6], [7] and no research related between RAM approach and HVAC system of rolling stock before. By addressing these gaps, the goal is to create a comprehensive framework that evaluates the current RAM metrics of HVAC system in rolling stock and strategic plans.

Certain risk assessment tools, such as Fault Tree Analysis (FTA) and Fishbone Diagram, neglect to take failure frequency or detectability into account. Failure Mode, Effects and Criticality Analysis (FMECA) is the appropriate risk assessment method to take into consideration that plausible gap and intended to evaluate every possible failure scenario in each component [8], [9], [10] of the rolling stock according to the international standard ISO 31010: Effective Risk Assessment Techniques. FMECA can determine the exact service task that needs to be performed by staff members in the event of an incoming failure [11]. This approach can be applied to evaluate frequently damaged parts and analyze the HVAC system in rolling stock to shorten the lead time for replacement parts via the aid of spare parts management [12].

II. METHODOLOGY

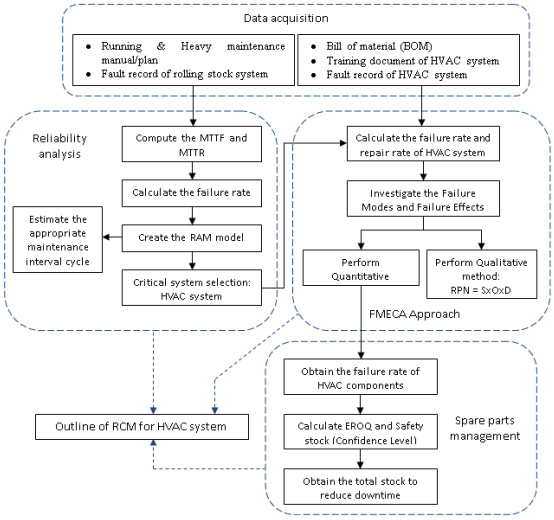


Figure 1 Overview of working process.

The methodology of this study is illustrated in Figure 1. The data of rolling stock maintenance from Siemens Mobility Thailand, Mochit depot, is extracted from the maintenance server between August 2021 and February 2024 (a total of 31 months). There are roughly 24,000 sets of data on the green line with 35 Siemens model first generation vehicles (EMU-A1). Data limitations result from modifications to the internal storage system, which only allow for the retrieval of data in the past three years.

The data analyzed in this section will only consider the maintenance record of corrective maintenance. This step is part of the data filtering process to exclude, rearrange, or allocate data according to reliability specification criteria [13]. The entire rolling stock system comprises a total of 21 systems as specified in the maintenance data. However, some fault notifications are merely related to minor dirtiness and minimal damage to decorative equipment both internal compartment and external of the vehicle, such as advertisement stickers on the windows. The beyond reason for not scrutinizing the Driver cab system or other systems as mentioned earlier is because it is lack of relevance to passenger-related considerations. Therefore, the assessed rolling stock configuration encompasses 12 systems, namely HVAC, Air supply, Auxiliary, Bogie, Brake, Communication, Destination, Door, Lighting, PIS and LCD, Propulsion, and Signaling systems. Every system under consideration has an impact on any technical casualties for passengers.

In this context, the failure rate is conventionally denoted in failures per hour. However, for this specific application, time-related data holds less significance compared to distance. Given that trains operate within specific timeframes, information regarding distance traveled becomes paramount and more meaningful than time. Additionally, certain subsystems may possess insufficient data sets for thorough analysis, potentially resulting in less reliable outcomes. Consequently, the failure rate is expressed in FPMK - failures per million kilometers. The calculation of the failure rate is as follows:

$$\lambda = FPMK\left[\frac{Failure}{10^6 km}\right] \times Operating\ speed\left[\frac{km}{h}\right] \quad (1)$$

Average running distance per train per month is attained from monthly report of E&M maintenance which is approximately 9,600 km. Hence, the total distance in service of 35 trains with the fault record of 31 months is equal to 10,416,000 km/35 trains in 31 months, while the operating speed is assumed to 60 km/h in this study.

The criteria to investigate to the critical system of the rolling stock system is defined by the frequency of failure occurrence or number of the failure notification (O) and the technical casualty (S) as in the following expression:

$$Criticality = O \times S \quad (2)$$

TABLE I CRITICALITY ASSESSMENT RELATED TO OCCURRENCE [FPMK] AND TECHNICAL CASUALTY [12]

Index number	Failure rate [FPMK]	Train service failure class	Failure classification	Failure definition
5	≥1.00E-02	A: Critical	Catastrophic	Potentially resulting in death or serious injury
4	5.00E-03	B: Significant	Hazardous	Reduces margin of safety, but excluding failure classified as catastrophic.
3	1.00E-03	C: Major	Significant	Causing termination, cancellation or requiring assistance, excluding those classified as hazardous and catastrophic.
2	1.00E-04	D: Major	Major	Lower train worthiness or increases the crew's workload but the journey can be complete safely, possibly with loss of time and possibly requiring rectification before the next day's diagram. Excluding those classified as significant, hazardous, and catastrophic.
1	≤1.00E-05	E: Minor	Minor	Causing an increase in maintenance workload. Excluding those classified as major, significant, hazardous, and catastrophic.

TABLE II CRITICALITY OF CONSIDERED OVERALL ROLLING STOCK SYSTEM

System	Failure rate [FPMK]	O	S	Criticality
HVAC system	4.718E-03	3	3	9
Air supply system	4.147E-04	2	4	8
Auxiliary system	2.673E-03	3	3	9
Bogie system	4.781E-04	2	5	10
Brake system	1.953E-03	3	5	15
Communication system	7.851E-03	4	1	4
Destination system	7.079E-03	4	1	4
Door system	5.628E-03	4	4	16
Lighting system	5.386E-03	4	2	8
PIS and LCD system	1.952E-02	5	1	5
Propulsion system	2.730E-03	3	4	12
Signaling system	4.090E-04	2	4	8

Classification criteria are used in order to consistently attribute the level of severity and occurrence to each analyzed system. The failure rate is a parameter to classify the range of failure occurrence in the means of appropriateness of the number of the considered trains. Failure severity or failure casualty is the failure of any parts of the train which causes a delay or a cancellation of a service train in excess of five minutes. In this study, the failure casualty criteria according to reliability specification [13] is adopted to relatively compare with the train service failure class of fault notifications.

TABLE I demonstrates the criteria for the choice of the occurrence assessment in the failure rate aspect, while the technical casualty applies to define the severity index compared to train service failure class which is mentioned in the Technical Train Service Failures document [14].

As a result, the criticality of the overall system regarding the collected data is summarized in TABLE II. The systems most significantly associated with both frequency and severity are the door system and brake system, registering criticality scores of 16 and 15, respectively. Conversely, the Communication and Destination systems exhibit the lowest levels of criticality, each garnering a score of 4. The HVAC system ranks fifth with a criticality score of 9. Notably, the criticality score attributed to the HVAC system consistently exceeds twice the magnitude of the least critical systems (Communication and Destination systems) and the railway system have strict regulations for reliability, availability, maintainability, and safety, as specified in the standard EN 50126-1 [15]. Thus, the maintainers of the HVAC system should replace critical components in early stages of degradation since HVAC system is considered as a critical system.

Reliability Analysis

After filtering and ranking the systems based on the frequency of malfunctions and the impact effects on passengers and vehicle operation, the top five systems in descending order of criticality are the Door, Brake, Propulsion, Bogie, HVAC, and Auxiliary systems. These systems will be subjected to a RAM model in the subsequent steps, and the datasets associated with these systems are deemed sufficient for analyzing failure rate and repair rate through the determination of Mean Time to Failure (MTTF) and Mean Time to Repair (MTTR) instead of employing the Failure per Million Kilometers (FPMK) approach.

Reliability, Availability, and Maintainability (RAM) relevant to the suitable maintenance plan can be investigated by analyzing the appropriate damage pattern that could be Weibull model or Exponential model [16]. The accuracy of the distribution model depends on the amount of data taken into account. The reliability model of the Weibull distribution and Exponential distribution are expressed in (3) and (4), respectively.

$$\ln(-\ln R(t)) = \beta \ln t - \beta \ln \eta \quad (3)$$

$$-\ln R(t) = \lambda t \quad (4)$$

Where β refers to a shape parameter which resembles the bathtub curve's phase, the model displays a constant failure rate when it equals to 1 in the random failure phase. A scale parameter is defined as η to represent the characteristic life or the typical Time to Failure (TTF) which is related to MTTF. λ represents the rate parameter in Exponential distribution which is the inverse proportion of MTTF. t is the TTF appearing in both expressions. In this analysis, the Weibull distribution is mostly implemented to describe the bathtub curve's failure distribution.

The probability plotting illustrates the graphical picture of linear equation patterns, which is correlated to (3) and (4), and quantitively indicates how closely distribution pattern fits the analyzed data. Linearization by linear regression is a tool to predict the value of a variable based on the value of another variable. The coefficient correlation, also known as the coefficient of determination, compares estimated and actual values and ranges a value of 0 to 1. In terms of hypothesis testing, any distribution that has a correlation coefficient close to 1 will be the best-fit distribution [17]. Its purpose is to find the perfect correlation between the Weibull distribution and the exponential distribution in the sample to determine the most appropriate functional parameters. Figure 2 represents the reliability of the Exponential distribution in orange and the reliability of the Weibull distribution in dark blue color. The method utilized for determining the reliability model is also applied to derive the maintainability function [3]. However, instead of analyzing the Time to Failure (TTF), the Time to Repair (TTR) is calculated as the variable of interest.

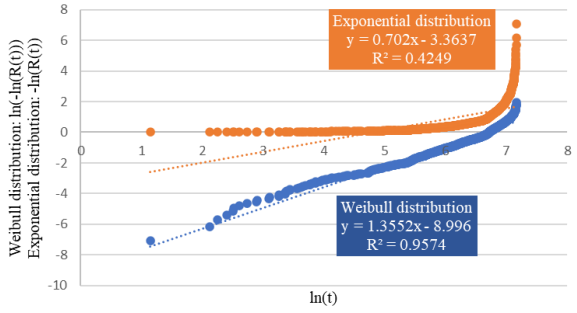


Figure 2 Reliability model of HVAC system.

To compute the availability of the interesting system, the relationship between MTTF and MTTR with availability under the constant failure rate condition is given in (5) [18]. Note that the rolling stock system has a limited-service period of 18 hours per day.

$$A(t) = \frac{MTTF}{MTTF+MTTR} \quad (5)$$

The consequences of the RAM model of the critical systems in rolling stock are tabulated in TABLE III with the interactive dynamic dashboard. The data visualization is created by the Microsoft Excel Software as illustrated in Figure 3 to serve as a friendly user's analytical interface and accessible for maintenance engineer's planning.

TABLE III RAM OF THE CRITICAL SYSTEM

System	MTTF [d]	MTTR [h]	R(t)	A(t)	M(t)
Door	45.89	19.20	99%	98%	100%
Brake	36.45	19.99	98%	97%	100%
Propulsion	39.40	14.27	99%	98%	100%
Bogie	30.76	28.02	98%	95%	99%
HVAC	42.42	34.80	99%	96%	100%
Auxiliary	44.15	10.86	99%	99%	100%

Figure 3 Rolling stock maintenance data analysis dashboard.

Two rationales justify the investigation further of failure root causes, alongside the formulation of a maintenance plan and an inventory management proposal for the critical components of the HVAC system. From TABLE III, it can be observed that the HVAC system requires the longest repair time compared to other systems. The second reason is that the Key Performance Indicator (KPI) of the HVAC system significantly impacts the delivery of vehicles to the service provider, as reported in the company's monthly report. This is because the HVAC system directly affects passenger satisfaction. Ratio of service car kilometers against failures amount related to Air Condition System's failures based on Mean Distance Between Failure (MDBF). The KPI target should be higher than 30,000 km for air conditioning system which calculate on 3-car train to achieve the target of each month. Therefore, developing a maintenance plan for the HVAC system using reliability analysis and failure analysis would be highly appropriate.

Heating Ventilation and Air Conditioning

This work deals with the root cause analysis of a Heating Ventilation and Air Conditioning (HVAC) system by Failure Mode, Effects and Criticality Analysis (FMECA) and the spare parts management of its components to ensure the availability of critical components which spare parts management directly impacts system reliability by reducing the likelihood and duration of failures.

HVAC systems play a critical role in enhancing indoor comfort, air quality, and energy efficiency, contributing to occupant well-being and productivity in diverse environments. The main HVAC unit is mounted on the roof of the carriage. Some installation details that distinguish the

C1-car apart from the A-car and C-car are the location of the emergency inverter (also known as an emergency switch) and the quantity of compressors that are in use. Nevertheless, both specifications have identical cooling capacity at 40kW. Therefore, both HVAC system's configurations can be considered similarly, and items can be categorized into ten subsystems that comprise to facilitate comprehension as follows:

- Compressor unit: Crankcase heater, Pressure switch, Solenoid valve-coil.
- Condenser unit: Condensing fan (Axial fan)
- Expansion unit: Expansion valve
- Evaporator unit: Radial fan (Supply fan), Emergency inverter, Air grill
- Mixed air compartment: Servo motor (Actuator), Mixed-air filter
- Junction box: Connector (Terminal), Contactor, MCB, Timer relay
- Sensor device: NTC-sensor
- Control unit: Controller FPC20/2, DC/DC Converter
- Conditioning driver's cab: Ventilation fan
- Unspecific unit: Refrigerant, Shims A/C Unit, Wiring

There are three element flows: Material element flows (Refrigerant and Air), Energy element flows (Electric and Thermal) and Signal element flows within a HVAC system as demonstrated in Figure 4. The refrigerant cycle is a thermodynamic process that enables the transfer of heat from one area to another. It involves the circulation of a refrigerant through various components to absorb heat from a low-temperature environment and release it to a higher-temperature environment. The refrigerant cycle of electrical multiple units (EMUs) is comparable to that of residences or buildings' refrigerant cycle. It consists of four main processes: compression, condensation, expansion, and evaporation, but with additional equipment such as sight glass, hand stop valves, and an attached crankcase heater with compressor.

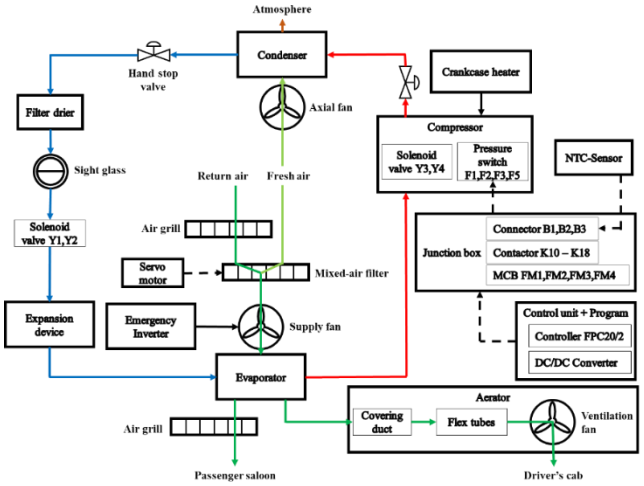


Figure 4 Functional block diagram of HVAC system.

Failure Analysis of HVAC

FMECA techniques, encompassing both quantitative and qualitative approaches, are employed to systematically analyze failures and assess their severity and criticality in the context of RAM enhancement for the HVAC system in rolling stock. This approach references MIL-STD-1629A Procedures for Performing a Failure Mode, Effects, and Criticality Analysis Standard. FMECA builds upon the Failure Mode and Effects Analysis (FMEA) methodology, incorporating Criticality Analysis (CA) to enhance comprehensiveness in failure analysis.

FMECA information of HVAC system is gathered from various sources that related to HVAC system of the rolling stock such as Fault records of HVAC system, HVAC system maintenance procedures, Training documentation of HVAC system, and Bill of materials (BOMs).

The data analyzed in this section pertains to corrective maintenance information for the HVAC system of 35 Electric Multiple Units (EMUs) spanning from January 2013 to December 2023, covering an 11-year period. The cumulative data of HVAC system is different from the rolling stock data because the HVAC system's data can be retrieved from the older version of the database. This analysis focuses on the components that underwent repair or replacement, the objective part codes providing brief information on the failure modes, and the damage codes indicating maintenance actions. These elements constitute part of the analysis of system indenture levels within the Failure Mode and Effects Analysis (FMEA) framework, specifically using the Hardware method or Bottom-up method [19].

A part of FMECA in both approaches is contained with Potential failure mode, Cause of the potential failure from local effects to the end of effects. On this basis the severity (S) of failure values can be evaluated which create the rank of the critical failure mode as compared to the TABLE IV including the occurrence and detectability ranking criteria [20].

TABLE IV CRITERIA CLASSIFICATION CATEGORY FOR HVAC FAILURE MODE [20]

Rating	Severity (S)	Occurrence (O)	Detection (D)
10	Potential failure mode affects the operation and/or involves noncompliance with government regulation without warning. (Hazardous)	Over 201 times	Control actions are not available for detection. (Not possible)
9	Potential failure mode affects the operation and/or involves noncompliance with government regulation without warning. (Very high)	from 151 to 200 times	Very remote likelihood that the current controls will detect the failure. (Very remote)
8	Inoperable with loss of primary function. (High)	from 101 to 150 times	Remote likelihood that the current controls will detect the failure. (Low)
7	Operable with loss of performance. The customer is dissatisfied. (Moderate High)	from 76 to 100 times	Low remote likelihood that the current controls will detect failure. (Low)
6	Operable but loss to comfort/convenience. Customer experience discomfort. (Moderate low)	from 51 to 75 times	Low remote likelihood that the current controls will detect failure. (Low)
5	Operable but loss to comfort/convenience. Customer has some discomfort. (Low)	from 31 to 50 times	Moderate remote likelihood that the current controls will detect failure. (Moderate)
4	Certain item characteristics do not conform to specification but noticed by most customers. (Very low)	from 21 to 30 times	Moderately high remote likelihood that the current controls will detect failure. (Moderate high)
3	Certain item characteristics do not conform to specification but noticed by average customers. (Very remote)	from 11 to 20 times	High remote likelihood that the current controls will detect failure. (High)
2	Certain item characteristics do not conform to specification but noticed by some discriminating customers. (Remote)	from 6 to 10 times	Very high remote likelihood that the current controls will detect failure. (Very high)
1	No effect (None)	from 1 to 5 times	Current controls certainly detect the failure.

According to the quantitative approach, the resulting consequence derived from the analysis is more closely associated with the level of available data compared to qualitative methodology. This is because it necessitates knowledge of variables such as failure rates, failure modes, failure mode ratios, and failure effect probabilities [21]. These variables are instrumental in computing a criticality number, which serves to prioritize items of concern. The criticality number assigned to an item is a relative indicator of both the consequences and frequency of its failure, as expressed in the following equation:

$$C_r = \sum_{n=1}^{j} (\beta \alpha \lambda_p t)_n \quad (6)$$

Item criticality number C_r, is described as the sum of the failure mode criticality number of each item. α is the probability of occurrence of the failure mode or failure mode ratio, while β refers to the conditional probability of the current failure mode's failure effect. β is generally set to 1 to solely account for the most severe potential outcome stemming from a failure mode. The duration of applicable mission phase t expresses in hours, 72,270 hours is the time of 11 years in operational service. The failure rate λ_p of an item is the ratio between the numbers of failures per unit of time and is typically expressed in failures per million hours or failures/10^6 hours. When evaluating the failure rates of a system employing redundant components to achieve its objectives, it is necessary to modify the failure rate to accurately represent the overall system failure rate.

$$\lambda_{(n-q)/n} = \frac{n!(\lambda)^{q+1}}{(n-q-1)!(\mu)^q} \quad (7)$$

For a redundant system where all units are active on-line with equal failure rates, which is based on an exponential distribution of failure (constant time between failures). Expression (7) takes repair time into consideration [18] where n and q represent a number of active on-line units and number of online units that can fail without system failure respectively, while μ stands for repair rate obtaining similarly to the process of failure rate derivation.

The outcomes of the FMECA quantitative methodology are presented in TABLE V, which selectively displays items with criticality numbers exceeding 100 from the entire pool of analyzed components. The failure modes depicted in the table constitute approximately 48% of the total. Primarily, these components pertain to equipment involved in regulating the refrigerant's circulation within the HVAC system. The items with the lowest criticality numbers include the Mixed-air filter, Ventilation fan, and Radial fan, with scores of 0.558, 2.275, and 3.837, respectively. Any equipment equipped with compensating provisions such as redundant items, safety devices, or automation self-compensating devices experience a substantial reduction in criticality, for instance, NTC-sensor, servo motor, and mixed-air filter.

In terms of comparison to the quantitative approach, the qualitative method is performed by subjectively classified ranking system in TABLE IV to assess the severity and occurrence of failures. Failure mode ratio and failure mode probability are not used in this analysis as well as the sake of redundancy. Instead, a Delphi technique is employed to solicit expert estimates of damage severity [7] by the assistance of group of overhaul engineers. Moreover, the Risk Priority Number (RPN) is utilized rather than the criticality number. The RPN ranking is determined by multiplying the Severity (1-10) ranking, the Occurrence (1-10) ranking, and the Detectability (1-10) ranking, as described by the following equation [22]:

$$RPN = S \times O \times D \quad (8)$$

TABLE V TOP CRITICALITY RANKING OF FMECA QUANTITATIVE METHOD OF HVAC SYSTEM

Item	Failure rate	Potential failure mode	Failure mode ratio	Modal criticality number	Item criticality number
Controller FPC20/2	2.918E-03	Notifications without a specific mode	0.51	108.383	210.854
		Electrical loads greater than board's specific tolerance	0.49	102.471	
Axial fan (Condenser fan)	2.447E-03	Uncontrol of the condenser fan	0.68	120.867	176.879
		Less efficient cooling of the condenser	0.25	44.220	
		Unbalance of the condenser fan	0.07	11.792	
Solenoid valve	2.291E-03	Less efficient refrigerant flow regulation	0.79	131.470	165.555
		Provide no refrigerant regulation	0.21	34.085	
Pressure switch	2.113E-03	Pressure greater than setpoint pressure	0.82	125.580	152.705
		Pressure less than setpoint pressure	0.18	27.125	
Connector (Terminal)	1.670E-03	Non continuous signal	0.82	99.289	120.705
		Terminal failure with no signal	0.18	21.415	
Expansion valve	1.462E-03	Pressure less than setpoint pressure or Droop	0.51	54.374	105.670
		Improper insulation	0.49	51.296	
Contactor	1.391E-03	Temperature greater than thermal stress of the material	0.52	52.715	100.549
		Misalignment and dirtiness of the contactor	0.48	47.834	

As a result, TABLE VI shows the level of RPN of top HVAC components' ranking by qualitative approach. The failure modes span a spectrum of scores ranging from 12 to 240, with 42% falling below 100 points. Failure mode in the range of 100-180 scores accounted for 39% of the total, which is worth following up and inspecting closely due to their considerable severity and frequency of occurrence not far from the top RPN ranking group.

TABLE VI TOP RPN RANKING OF FMECA QUALITATIVE METHOD OF HVAC SYSTEM

Item/Function	Potential Failure mode	Failure mechanism (Cause)	S	Number of occurrences	O	D	RPN
Solenoid valve	Less efficient refrigerant flow regulation	Damaged or worn membranes	5	135	8	6	240
Axial fan (Condenser fan)	Uncontrol of the condenser fan	Overheating or Faulty wiring	6	123	8	5	240
NTC-sensor	NTC-sensor failure	Mechanical or Electrical damages	5	115	8	5	200
Connector (Terminal)	Non continuous signal	Vibration or thermal stress	4	102	8	6	192
Controller FPC20/2	Electrical loads greater than board's specific tolerance	Power surge, voltage spike, or electrical fluctuations	6	104	8	4	192
NTC-sensor	Inaccurate temperature reading	Exposure to high temperature	4	141	8	6	192

The results of both FMECA approaches have the same direction of priority items, except NTC-sensor due to redundancy, converging to the strong evidence that occurrence is the most significant parameter because it is the main variable in the calculation of both methods. To prevent the recurrence of these failure modes, it is imperative to meticulously plan and execute preventive maintenance actions in a manner that is both cost-effective and timely [23]. Additionally, reducing lead time for spare parts contributes to enhancing the maintainability efficiency of the system, as downtime for a component includes not only the active maintenance duration but also delays incurred while awaiting spare parts, labor, facilities, transportation, and other related factors [12]. Among the equipment identified with the highest criticality number and Risk Priority Numbers (RPNs) in the results displayed in TABLE V and TABLE VI are the controller FPC20/2, solenoid valve, connector (Terminal), NTC-sensor, pressure switch, and condenser fan. These items require frequent maintenance and replacement. Therefore, employing spare parts management that utilizes failure rate estimates can help minimize the duration of repairs.

Spare parts management of HVAC components

Common challenges in spare parts management calculations often revolve around determining the appropriate quantity to stock in inventory. In case of unforeseen circumstances, a designated level of safety stock becomes imperative to ensure timely replacement of spare parts. Additionally, the Economic Order Quantity (EOQ) helps reduce overall holding and ordering costs in inventory management to the lowest possible level. Figure 5 elucidates the total spare parts inventory, enriching comprehension of the relationship between Safety Stock (SS) and Economic Order Quantity (EOQ) in inventory management.

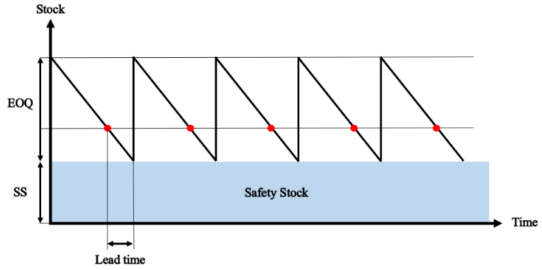


Figure 5 Theoretical total spare parts.

The fundamentals of EOQ comprise three key elements [24]. Firstly, the demand for a product (D), typically quantified in units per annum. This figure is pivotal in computing the most efficient order quantity that mitigates the cumulative expenses associated with inventory ordering and holding. Secondly, the ordering cost per order (S) or transaction cost constitutes another significant facet of the EOQ implementation. This expense encapsulates various costs incurred during the procurement process, encompassing administrative charges, transportation expenditures, and procedural costs. Lastly, the holding cost per unit per year (H) represents the expenses linked to inventory retention, including storage outlays, depreciation charges, and opportunity costs. It appears logical to infer that the quantity of spare parts needed corresponds to the number of failures [25]. Therefore, from the EOQ expression, the annual demand value is predicted based on the appropriate time for maintenance which is the reliability-based spare parts provision. This prediction stems from the failure rate of each component, which designates this order quantity as the Economic Reliability Order Quantity (EROQ), as represented by the following equation:

$$EROQ = \sqrt{\frac{2\times\frac{365}{t}\times S}{H}} \quad (9)$$

To determine the minimum quantity of spare parts or safety stock required to maintain stock levels considering failure rates, the quantity of installed parts, and the time required for resupply (order lead time), the "Spares analysis" theory will be employed [26]. The following expression is used to calculate the number of spare parts needed to reach a confidence level (CL) that is equal to or higher than CL:

$$CL \leq \sum_{k=0}^{S} \frac{(n\lambda R)^k e^{-(n\lambda R)}}{k!} \quad (10)$$

Where CL is the confidence level for the spare Line Replaceable Units (LRUs) adequacy which generally set to a minimum of 90%., while n and S are the number of the LRUs in operation and the minimum number of required spare LRUs, respectively. The resupply time interval in hours (the time to replace spare) is abbreviated as R and k is defined as a safety stock or minimum quantity of spare parts in the inventory store.

Thus, the cost of EROQ affects the inventory carrying cost, which in turn influences order quantity as determined in the inventory management system. In contrast, safety stock takes confidence levels into account rather than holding costs when calculating the minimum quantity of spares that must be kept in the warehouse in case of emergency. Hence, the total number of stocks in the inventory can be calculated as a following equation:

$$Total\ stock = EROQ + SS \quad (11)$$

Consequently, the required number of spare parts is calculated in the reliability-based statistical approaches according to the item reliability. The sample of the total stock calculation related to failure rate of the items is displayed in TABLE VII. Note that this spare parts provision procedure considers only corrective maintenance since the failure rate is collected from the unexpected situation, while the spare parts for preventive maintenance can be accurately predictable due to definite maintenance interval.

TABLE VII SAMPLE OF SPARE PARTS MANAGEMENT OF CRITICAL COMPONENTS

Material Units	Qty/Unit or n (pcs)	Re-supply time or R (day)	Item category	Item Failure rate (per hour)	Min. Required Spare/Unit or SS (pcs)	t (day)	Annual Demand (pcs)	EROQ (pcs)	Total stock (pcs)
Bearing 6007-2RS1	2	55	Axial fan	2.447E-03	10	2.39	152.62	119	129
O-Ring 65x3.5 mm.	1	44	Axial fan	2.447E-03	5	2.39	152.62	404	409
O-Ring 3x130mm	3	83	Connector	1.670E-03	14	3.50	104.15	290	304
NTC-SENSING DEVICE	3	125	NTC-sensor	3.456E-03	38	1.69	215.54	67	105
PRESSURE SWITCH PS3-AP4 HNK 23/20±0.3	1	97	Pressure switch	3.765E-04	2	15.55	23.48	26	28
Seal for EPDM80 Black	3	97	Pressure switch	3.765E-04	5	15.55	23.48	53	58
Hex cap screw for control box cover	10	83	Controller FPC20/2	1.463E-03	36	4.00	91.24	529	565
Solenoid Valve-coil 110VDC	2	101	Solenoid valve	2.264E-03	15	2.59	141.18	90	105
SOLENOID VALVE EVR10 (ODF-ODF 16 mm.)	2	101	Solenoid valve	2.264E-03	15	2.59	141.18	58	73

Applicable data to compute failure rates, particularly if employing detailed recordings such as Stock Keeping Units (SKUs) or Line Replaceable Units (LRUs), alongside accessing confidential factors like transaction costs, holding costs, and total resupply time (procurement duration, approved time, lead time), would lead to gain more accurate results. However, a proposed approach that utilizes the failure rate of various equipment to predict the rate of spare parts usage, ensuring appropriateness and adequacy for maintenance, can be considered as an alternative maintenance procedure for equipment requiring frequent maintenance and with high repair costs.

III. RESULTS AND DISCUSSION

The maintenance interval cycle is based on the mathematical models of reliable life between Weibull and Exponential functions as stated in (12) and (13) [27]. The reliability target R(t) is the probability that the component can perform the required function under the given conditions for a given time interval t. According to the weekly report maintenance, the critical systems should have a reliability target of over 95%. The HVAC system is one of the critical systems, thus the reliability target of each component is set to 95% and the result is summarized in TABLE VIII.

$$t = e^{\frac{\ln(-\ln R(t)) + \beta \ln \eta}{\beta}} \quad (12)$$

$$t = -\frac{\ln R(t)}{\lambda} \quad (13)$$

Finally, the maintenance actions are carried out at every estimated maintenance interval with the suitable maintenance strategy as follows:

- Malfunction test with service program SPP every 2 weeks with 1 h/EMU: NTC-sensor.
- Inspection every 3 weeks with 2 h/EMU: Axial fan, Solenoid valve, Pressure switch.
- Inspection every 1 month with 2.5 h/EMU: Controller FPC20/2, Connector, Expansion valve, Contactor.
- Inspection every 2 months with 6 h/EMU: Servo motor, Wiring, Refrigerant, Shims.
- Inspection every 6 months with 3 h/EMU: Timer relay, MCB, Emergency inverter, DC/DC Converter, Crankcase heater, Mixed-air filter.
- Inspection every 1 year with 2 h/EMU: Air grill, Radial fan, Ventilation fan.

TABLE VIII MAINTENANCE INTERVAL AND STRATEGY OF EACH SYSTEM

Item	Item criticality number	Maintenance strategy	Maintenance interval (day)
Controller FPC20/2	210.85	Preventive maintenance	35
Axial fan (Condenser fan)	176.88	Condition-based maintenance	21
Solenoid valve	165.56	Condition-based maintenance	23
Pressure switch	152.71	Preventive maintenance	24
Connector (Terminal)	120.70	Preventive maintenance	31
Expansion valve	105.67	Condition-based maintenance	35
Contactor	100.55	Preventive maintenance	37
Shims A/C Unit	58.15	Preventive maintenance	64
Refrigerant	52.26	Condition-based maintenance	71
Wiring	34.58	Preventive maintenance	107
Servo motor	30.91	Condition-based maintenance	56
Timer relay	23.62	Preventive maintenance	157
MCB	15.92	Preventive maintenance	233
Emergency inverter	14.25	Preventive maintenance	260
DC/DC Converter	6.02	Preventive maintenance	616
Crankcase heater	5.36	Preventive maintenance	691
NTC-sensor	4.43	Condition-based maintenance	15
Air grill	4.15	Preventive maintenance	894
Radial fan (Supply fan)	3.84	Condition-based maintenance	966
Ventilation fan	2.28	Preventive maintenance	1629
Mixed-air filter	0.56	Preventive maintenance	237

Preventive maintenance should be performed on all HVAC system components at the recommended intervals. Certain equipment can be closely monitored for the best practice alongside the suggested preventive maintenance through installing temperature sensors, pressure sensors, and integrating power consumption sensors.

The utilization of the detection sensors is one of the methods as a Condition-Based Maintenance (CBM). Instead of focusing solely on temperature sensors (NTC-sensors) and pressure sensors at the compressor's input and output, item related to controlling working fluid, such as solenoid valves, expansion valves, and equipment related to motors, has inspection techniques that are convenient and reliable based on failure modes of each item, so they are suitable for condition-based maintenance. The recommended sensors include pressure sensors between the condenser coil and expansion valve, as well as between the evaporator coil and expansion device. Additionally, integrating power consumption sensors in radial and axial fans can mitigate blower current fluctuation and enhance the detectability of air filter's degradation [28].

This procedure serves as a guideline for an alternative preventive maintenance of HVAC systems in rolling stock

by the principle of reliability-centered maintenance. The result of FMECA method, both quantitative and qualitative approaches, represents the items that are critical to the HVAC system and implicates at inspection and repair processes based on the priority indexes (criticality number and RPN index) and an appropriate maintenance interval. Furthermore, the EROQ and safety stock estimations are tools to improve the maintainability of the HVAC system in the spare parts management's manner.

Nevertheless, data collection is the most important state since the resolution of the data collected can be used to calculate the failure rate, which is the linkage parameter for the whole process. The group of components, number of SKUs in use, and failure modes should be stored in separated columns as a suggestion to achieve more accurate results of the overall proposed study.

IV. Conclusion

The alternative preventive maintenance based on reliability analysis and failure analysis is proposed in this paper. The connection between each step of the process is linked by the significant parameter called the failure rate. The RAM model is performed to assess the reliability, maintainability, and availability of the critical systems in the rolling stock. The FMECA approach is applied to the HVAC system as a case study to prioritize the components and emphasize the sequence of maintenance. The spare parts management estimates the number of spares in inventory for reasonable quantities for critical items. The spares are stocked on an annual basis to guarantee that there are enough spare parts. This eliminates the spare parts' waiting time and enhances the service duration. Therefore, the overall breakdown duration depends only on the hours of hands-on maintenance.

Following the recommended maintenance interval, the estimated maintenance duration of 35 EMUs totals 4,690-man hours per year. The on-duty shift plan for overhaul maintenance states that 11 staff members are needed to maintain the HVAC system at each time of day. The HVAC system requires 546.89 hours of preventive maintenance on average annually extracting from maintenance data. Hence, when the suggested preventive maintenance plan is executed, the total duration of preventive maintenance will be shortened by 22.04%.

Lastly, the availability and accessibility of maintenance data including data processing is a crucial step to attain the most precise maintenance guideline. The study's insights inform future data-driven strategic decisions for diverse rolling stock systems, enhancing operational efficiency and resource allocation strategies.

Acknowledgment

The author would like to acknowledge the rolling stock and overhaul departments of Siemens Mobility Thailand, Mochit depot, for the insight information of rolling stock maintenances and thank to TGGS, KMUTNB for giving the enriching and invaluable experience to do an internship at Siemens Mobility Thailand company. Moreover, this research is financially supported by RVIE master's degree programs.

References

[1] M. Bhebhe, N. Zincume and M. Philani, "Influence of different maintenance srategies on the availability of rolling stock," Stellenbosch, 2020.

[2] J. Raczyński, "Life cycle cost as a criterion in purchase of rolling stock," in *MATEC Web of Conferences*, 2018.

[3] M.-G. Park and Min-An, "Development of a framework for engineering RAMS into rolling stock through life cycle in the operator perspective," 2010.

[4] M. Catelani, L. Ciani, D. Galar, G. Guidi, S. Matucci and G. Patrizi, "FMECA Assessment for Railway Safety-Critical Systems Investigating a New Risk Threshold Method," *IEEE Access*, vol. 9, pp. 86243-86253, 2021.

[5] Q. Zhong, R. M. Lusby, J. Larsen, Y. Zhang and Q. Peng, "Rolling stock scheduling with maintenance requirements at the Chinese High-Speed Railway," *Transportation Research Part B: Methodological*, vol. 126, pp. 24-44, 1 8 2019.

[6] S. Muzakkir, H. Harish and K. Lijesh, "Failure mode and effect analysis of journal bearing," *International Journal of Applied Engineering Research*, vol. 10, no. 16, pp. 36843-36850, 2015.

[7] F. Dinmohammadi, B. Alkali, M. Shafiee, C. Bérenguer and A. Labib, "Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door System," *Urban Rail Transit*, vol. 2, no. 3-4, pp. 128-145, 1 12 2016.

[8] S. Rezvanizaniani, M. Valibeigloo, M. Asghari, J. Barabady and U. Kumar, "Reliability Centered Maintenance for Rolling Stock: A Case Study in Coaches' Wheel sets of Passenger Trains of Iranian Railway," *IEEE International Conference on Industrial Engineering and Engineering Management*, 2008.

[9] J. Kim, H. Jeong and J. Park, "Development of the FMECA process and analysis methodology for railroad systems," *International Journal of Automotive Technology*, vol. 10, no. 6, pp. 753-759, 20 3 2009.

[10] C. Srijuntra, "The condition-based maintenance module development by using FMECA technique and its application in rolling stock (Trains)," Bangkok, 2010.

[11] L. Jun and X. Huibin, "Reliability Analysis of Aircraft Equipment Based on FMECA Method," *Physics Procedia*, vol. 25, pp. 1816-1822, 2012.

[12] B. Rob and S. Gerard, "Spare Parts Management -

Successful MRO Spare Parts Management is more than just Software," SKF Reliability Systems, San Diego, 2003.

[13] R. Puntis and D. M. Walley, "The Use of Reliability Techniques on Traction and Rolling Stock," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 200, no. 4, 10 1986.

[14] SIEMENS MOBILITY, *Technical Train Service Failures*, 11 ed., Bangkok, 2022.

[15] B. E. 50126-1:2017, Railway Applications - The Specification and Demonstration of Reliability, Availability, Maintainability, and Safety (RAMS), BRITISH STANDARD INSTITUTION, 2017.

[16] S. Talabgaew, "Reliability, Availability, Maintainability Approach for Infrastructure and Logistics of Railway Operations," Bangkok, 2020.

[17] S. Rezvanizaniani, J. Barabady, M. Valibeigloo, M. Asghari and U. Kumar, "Reliability Analysis of the Rolling Stock Industry: A Case Study," *International Journal of Performability Engineering*, vol. 5, no. 2, pp. 167-175, 1 2009.

[18] A. Kumar, P. Sharma and A. Agarwal, "RAMS of HVAC for Rolling Stock Application," *International Journal of Engineering Research & Technology (IJERT)*, vol. 10, no. 8, 2021.

[19] U. Department of Defense, "Procedures for Performing a Failure Mode, Effects and Criticality Analysis," Department of Defense, Washington D.C., 1980.

[20] M. Sitarz, K. Chruzik and R. Wachnik, "Application of RAMS and FMEA Methods in Safety Management System of Railway Transport / ZASTOSOWANIE METOD RAMS I FMEA W SYSTEMACH ZARZĄDZANIA BEZPIECZEŃSTWEM W TRANSPORCIE KOLEJOWYM," *Journal of KONBiN*, 12 2012.

[21] A. J. Braaksma, A. J. Meesters, W. Klingenberg and C. Hicks, "A quantitative method for Failure Mode and Effects Analysis," *International Journal of Production Research*, vol. 50, no. 23, pp. 6904-6917, 1 12 2012.

[22] A. Walsh, R. Basile and A. Canhoto, "A Data-Derived Approach For Selecting Criticality Levels In FMECAs For Cleaning Process Risk Analysis," ASTM Pharmaceutical Cleaning Innovation, 2024.

[23] N. S. Arunraj and J. Maiti, "Risk-based Maintenance Techniques and Applications," *Journal of Hazardous Materials*, vol. 142, no. 3, pp. 653-661, 11 4 2007.

[24] B. Ghodrati, "Product Support and Spare Parts Planning Considering System Reliability and Operating Environment," Luleå University of Technology, Luleå, 2003.

[25] R. Barabadi, M. Ataei, R. Khalokakaie, A. Barabadi and A. Nouri Qarahasanlou, "Spare Part Management Considering Risk Factors," *International Congress and Workshop on Industrial AI*, pp. 24-39, 2021.

[26] ELLINIKO METRO, "International Tender Based on the Open Procedure for the Conclusion of PPP Contract concerning the Operation and Maintenance of the Thessaloniki Metro Network - Reliability, Availability, Maintainability and Safety (RAMS)," ELLINIKO METRO, Athens, 2022.

[27] I. De Sanctis, C. Paciarotti and O. Di Giovine, "Integration between RCM and RAM: a case study," *International Journal of Quality and Reliability Management*, vol. 33, no. 6, pp. 852-880, 6 6 2016.

[28] A. Gálvez, A. Diez-Olivan, D. Seneviratne and D. Galar, "Fault detection and RUL estimation for railway HVAC systems using a hybrid model-based approach," *Sustainability (Switzerland)*, vol. 13, no. 12, p. 6828, 2 6 2021.

Criticality Analysis of Automated People Mover Components By Applying Failure Mode And Effects Analysis

Thanodom Sirisongkram[1, a)] Kittchai Sojiphan[1, b),] Nuksit Noomwongs[2, c)] and Sompoap Talabgaew[3, d)]

Author Affiliations
[1]Railway Vehicles and Infrastructure Engineering Program, The Sirindhorn International Thai-German Graduate School of Engineering (TGGS), King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok 10800 Thailand
[2]Department of Mechanical Engineering, Faculty of Engineering, Chulalongkorn University 254 Phayathai Rd., Wangmai, Pathumwan, Bangkok 10330 Thailand
[3]Department of Teacher Training in Mechanical Engineering), King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Rd., Wongsawang, Bangsue, Bangkok 10800 Thailand

Author Emails
[a)] s6409096860042@email.kmutnb.ac.th
[b)] Kittichai.s@tggs.kmutnb.ac.th
[c)] Nuksit.n@chula.ac.th
[d)] Corresponding author: Sompoap.t@fte.kmutnb.ac.th

Abstract. Suvarnabhumi Airport, Thailand's biggest airport, has introduced Automated People Mover (APM) systems, necessitating precise maintenance of core components. However, lacking historical data, maintenance procedures rely solely on manual instructions. This study assesses whether maintenance strategies should be uniform or diversified, considering safety and service impacts. Utilizing Failure Mode and Effects Analysis (FMEA) theory to assess the criticality of various systems in APM trains. The Risk Priority Number (RPN) evaluates by severity, occurrence, and detectability. Ranking them into 3 levels are Significant, Major, and Minor from their RPN value. The application of FMEA principles to APM trains has resulted in the ranking of the reliability of each system. This assists in maintenance planning responses when multiple systems fail simultaneously, prioritizing based on their criticality.

***Index Terms*— APM trains, FMEA, Maintenance Planning, Risk analysis.**

I. INTRODUCTION

SUVARNABHUMI Airport is the first airport in Thailand to introduce an Automated People Mover (APM) system to facilitate passenger transportation between the North Main Terminal Building (NMTB) and Satellite Building (SAT-1). With approximately 1 kilometer between the two buildings, passengers often need to allocate more time for terminal connections, which could potentially lead to missed flights. Therefore, Suvarnabhumi Airport has implemented the APM system to reduce transit time between buildings. The APM can travel at speeds of up to 80 kilometers per hour and takes an average of 3 minutes to connect between terminals. Additionally, it can transport passengers approximately 3,590 passengers per hour, operating continuously 24 hours a day, 7 days a week.

The primary objective of providing continuous APM service is to ensure maximum availability, thereby enhancing passenger transit efficiency and maintaining safety, punctuality, and operational effectiveness standards. Commencing service from September 2023, the airport will operate two APM trains per route, with an additional two trains reserved for internal depot operations, featuring four running tracks, designated by red and green lines for passenger ingress and egress.

The cornerstone of APM service lies in the maintenance of the rolling stock, comprising 11 critical subsystems. However, due to the lack of historical maintenance, usage, and damage data, maintenance procedures have been limited to following manual instructions, primarily focusing on preventive and corrective maintenance protocols. Therefore, this research aims to assess the criticality of each subsystem to determine appropriate maintenance strategies. Utilizing the Failure Mode and Effects Analysis (FMEA) methodology, the study will evaluate severity, occurrence,

and detectability factors to derive risk assessments crucial for maintenance planning. Ultimately, the findings will serve as vital data for optimizing maintenance practices and enhancing operational reliability at Suvarnabhumi Airport.

II. METHODOLOGY

The FMEA method involves ranking the importance of equipment to address the root cause of problems and its failure modes. Next, to assess the Risk Priority Number (RPN) of the equipment includes the severity of abnormal symptoms, assign a severity rating: for each potential failure mode, assign a severity rating based on the impact it could have on the system and the passengers. Severity can be rated on a scale of 1-10, with 10 being the highest severity. Table I [5] illustrates the criteria for the choice of the severity index.

TABLE I
CRITERIA FOR SEVERITY S ASSESSMENT.

Severity	Criteria	Rating
Hazardous	Safety related failure modes causing non-compliance with government regulations without warning	10
Serious	Safety related failure modes causing non-compliance with government regulations with warning	9
Very high	Failure modes resulting in loss of primary system/component function	8
High	Failure modes resulting in a reduced level of system/component performance and customer dissatisfaction	7
Moderate	Failure modes resulting in a loss of function by comfort/convenience system/component	6
Low	Failure modes resulting in a reduced level of performance of comfort/convenience system/component	5
Very low	Failure modes resulting in a loss of fit and finish, squeak, and rattle functions	4
Minor	Failure modes resulting in partial loss of fit and finish, squeak, and rattle functions	3
Very minor	Failure modes resulting in minor loss of fit and finish, squeak, and rattle functions	2
None	No effect	1

the occurrence of abnormal symptoms occurring, assign an occurrence rating: For each potential failure mode, assign an occurrence rating based on the probability of the failure occurring. Occurrence can also be rated on a scale of 1-10, with 10 being the highest occurrence. Table II [12] gives the criteria for the assignment of occurrence values.

TABLE II
CRITERIA FOR OCCURRENCE O ASSESSMENT.

Occurrence	Criteria	Rating
Very high (failure is almost inevitable)	≥1 in 2	10
	1 in 3	9
	1 in 8	8
High (frequently repeated failures)	1 in 20	7
	1 in 80	6
Moderate (occasional failures)	1 in 400	5
	1 in 2000	4
Low (rare failures)	1 in 15,000	3
	1 in 150,000	2
Remote (failures are unlikely)	≤1 in 1,500,000	1

and whether abnormal symptoms can be detected, assign a detection rating: For each potential failure mode, assign a detection rating based on the ability to detect the failure before it causes harm. Detection can be rated on a scale of 1-10, with 10 being the highest detection. Table III [12] shows the assessment of the detection value.

TABLE III
CRITERIA FOR DETECTION D ASSESSMENT.

Detection	Criteria	Rating
Uncertain	Design control will not and/or cannot detect a potential cause/mechanism and subsequent failure mode.	10
Very remote	Very remote chance the design control will detect a potential cause/mechanism and subsequent failure mode.	9
Remote	Remote chance the design control will detect a potential cause/mechanism and subsequent failure mode.	8
Very low	Very low chance the design control will detect a potential cause/mechanism and subsequent failure mode.	7
Low	Low chance the design control will detect a potential cause/mechanism and subsequent failure mode.	6
Moderate	Moderate chance the design control will detect a potential cause/mechanism and subsequent failure mode.	5
Moderately high	Moderately high chance the design control will detect a potential cause/mechanism and subsequent failure mode.	4
High	High chance the design control will detect a potential cause/mechanism and subsequent failure mode.	3
Very high	Very high chance the design control will detect a potential cause/mechanism and subsequent failure mode.	2

Almost certain	The design control will almost certainly detect a potential cause/mechanism and subsequent failure mode.	1

Calculate the Risk Priority Number (RPN): The RPN is calculated by multiplying the severity, occurrence, and detection ratings:

$$RPN = S \cdot O \cdot D \quad (1)$$

This method helps prioritize the systems into failure category, determining whether a system or component should be maintained through condition monitoring or preventive maintenance. It can be summarized as follows:

- For RPN is over 200: the most critical damage to the transportation system is deemed "Significant" which means damage that disrupts scheduled operations or results in expenses exceeding the allocated budget.
- For RPN is 100-199: "Major" damage to the transportation system is repairable or improvable according to established standards and does not exceed the designated maintenance budget.
- For RPN is 1-99: "Minor" damage to the transportation system does not affect operations and will not lead to Major or Significant damage in the end.

This flowchart illustrates the process of Failure Mode and Effects Analysis (FMEA) process involves systematically identifying potential failure modes within the automated people mover (APM) system.

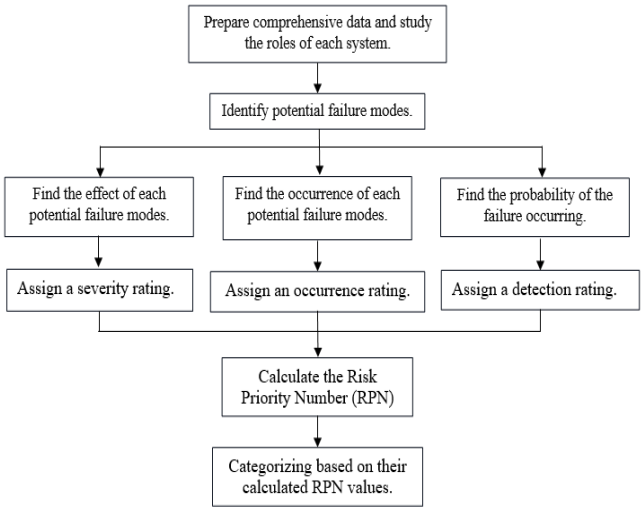


Fig. 1. Flow chart for the sequential procedure of FMEA analysis.

III. RESULTS AND DISCUSSION

The identification of failure modes for all 11 subsystems of the rolling stock system of the APM system and assign a severity rating based on its failure modes, aiming to determine the potential damages that may occur to the system.

TABLE IV
FAILURE MODE AND SEVERITY RATING FOR EACH SUBSYSTEM IN THE ROLLING STOCK SYSTEM.

Subsystem	Failure modes	S
Traction	Traction Motor Failure, Power Supply Interruptions, Control System Malfunctions, Wheel Slip or Wheel Slide, Overheating, Faulty Sensors or Feedback Devices, Mechanical Failures.	8
Axle	Axle Misalignment, Shock Absorber Leakage, Tire Puncture or Blowout, Tire Wear and Degradation, TPMS Sensor Malfunction, Speed Sensor Failure, Gearbox Lubrication Issues.	8
Guiding and Hydraulic system	Hydraulic Fluid Leakage, Hydraulic Pump Failure, Guideway Misalignment, Accumulator Lost Pressure, Pressure Limiter Failure, Guiding Roller coated wheel detachment, Derailment sensor malfunctions.	8
Pneumatic and Brakes System	Air Leakage, Brake Pad Wear, Pneumatic System Blockage, Brake System Overheating, Failure of Pressure Sensors, Valve Failure, Seal Deterioration	9
Coupler and Draw Bar	Coupler Misalignment, Draw Bar Fatigue, Coupler Jamming, Draw Bar Breakage, Coupler Release Failure, Draw Bar Wear, Coupler Locking Failure, Coupler Sensor Malfunction	6
Doors	Door Motor Failure, Door Sensor Malfunction, Door Interlock Failure, Door Seal Leakage, Door Control System Malfunction, Emergency Door Release Failure, Door Alignment Problems	7
MDCU, LVP, VCMS	MDCU Software Glitch, LVP Power Supply Failure, VCMS Sensor Malfunction, MDCU Hardware Failure, MDCU Control Interface Malfunction, LVP Circuit Short Circuit, VCMS Data Corruption	7
Carbody	Exterior and interior Panel Damage, Window Seal Failure, Exterior Coating Degradation, Interior Lighting Failure, Gangway System Malfunction.	3
HVAC	Compressor Failure, Refrigerant Leak, Condenser Blockage, Evaporator Coil Corrosion, Fan Motor Malfunction, Sensor Malfunction, Control System Failure, Drainage Issues, Power Supply Interruption.	5

High Voltage	High Voltage Cable Damage, Insulation Failure, High Voltage Component Failure, Overvoltage, Ground Faults, Overcurrent, Sensor Malfunction,	10
Auxiliary Power Supply	Power Supply Failure, Battery Degradation, Charging System Malfunction, Overload Protection Failure, Control System Malfunction, Emergency Power Failure, Inverter Failure, DC-DC Converter Failure	10

For determining occurrence, maintenance data collected during actual service hours for passengers since September 2023 will be utilized, and for detection, root cause of system failure recorded in the maintenance data management program and the failure modes will be used to analyze each instance of damage for assessing detectability capability.

TABLE V
ROOT CAUSES, OCCURRENCE, AND DETECTION RATING FOR EACH SUBSYSTEM IN THE ROLLING STOCK SYSTEM.

Subsystem	Root cause	O	D
Traction	Error from hardware	1	6
Axle	Installation, vibration, assembly process.	3	5
Guiding and Hydraulic system	Rail alignment, materials and conditions, heat, over and lost pressure.	7	6
Pneumatic and Brakes System	Vibration or not tighten enough, malfunction of sensors.	1	6
Coupler and Draw Bar	No failure with this subsystem	1	5
Doors	From operation and error from VCU, installation process	3	5
MDCU, LVP, VCMS	Vibration, defect from installation	1	3
Carbody	From sealing process	1	3
HVAC	From dirty by dust, error from software	3	6
High Voltage	No failure with this subsystem	1	6
Auxiliary Power Supply	Defect from control command board and software.	1	6

For each subsystem, the Risk Priority Number (RPN) will be obtained from equation (1) and then managed by categorizing into groups according to the criteria of Significant, Major, and Minor as outlined in Table VI.

TABLE VI
THE RISK PRIORITY NUMBER AND FAILURE CATEGORY FOR EACH SUBSYSTEM IN THE ROLLING STOCK SYSTEM.

Subsystem	RPN	Failure Category
Traction	48	Minor
Axle	120	Major
Guiding and Hydraulic system	336	Significant
Pneumatic and Brakes System	54	Minor
Coupler and Draw Bar	30	Minor
Doors	105	Major
MDCU, LVP, VCMS	21	Minor
Carbody	9	Minor
HVAC	90	Major
High Voltage	60	Minor
Auxiliary Power Supply	60	Minor

The results indicate that the Guiding and Hydraulic system falls into the Significant group, requiring adjustment of the maintenance plan according to the manufacturer's manual due to a higher-than-expected frequency of occurrences leading to potential damages. This will lead to considering Reliability, Availability, Maintainability, and Safety (RAMS) to determine the suitability for future maintenance plan adjustments.

On the other hand, the Axle and Doors system are categorized as Major. Although its occurrence frequency is not significant, it is still crucial for planning detection measures as these systems may encounter repeated damages like those experienced by the Guiding and Hydraulic system. This proactive approach aims to prevent frequent damage and ensure the continuous operational readiness of the APM railcars. And here are the details of failure to each system that may occur from consulting with expert engineers and maintenance advice from the APM train manufacturer's manual as shown in Table VII.

TABLE VII
FAILURE EFFECTS FOR EACH SUBSYSTEM IN THE ROLLING STOCK SYSTEM.

Subsystem	Failure effects
Traction	• Failures or malfunctions in the traction system can lead to significant disruptions in service, including delays, cancellations, and potential safety hazards. • Issues such as traction motor failures, power supply interruptions, or control system malfunctions can severely impact the APM's ability to operate effectively. • These failures can result in inconvenience to passengers, financial losses, and reputational damage to the operator. Therefore, any issues with the traction system must be promptly addressed to minimize disruptions and maintain passenger satisfaction and safety.

Axle	• Failures or malfunctions in the axle system can have severe consequences, including derailments, train delays, and potential safety risks to passengers and personnel. • Issues such as axle fractures, bearing failures, or wheel misalignments can lead to catastrophic accidents and significant damage to infrastructure. • Any abnormalities or irregularities in the axle system must be promptly addressed to prevent accidents and ensure the continued safe operation of the APM.
Guiding and Hydraulic system	Failure modes in the Guiding and Hydraulic system of an APM like AirVAL Siemens can range from critical issues like guiding roller detachment and hydraulic accumulator pressure loss, posing severe risks to passenger safety and system integrity, to less severe concerns like hydraulic system leaks and guiding misalignment, impacting operational efficiency and maintenance requirements.
Pneumatic and Brakes System	Critical for braking and suspension, failures in the Pneumatic System may not directly pose immediate safety risks as the Brakes System. However, malfunctions could still lead to operational disruptions and potential safety concerns. this system collisions or derailments.
Coupler and Draw Bar	Critical for the structural integrity and operational efficiency of the train, failures in the Coupler System may not directly pose immediate safety risks. malfunctions in the Draw Bar System could lead to operational disruptions and compromise the stability of the train. They are crucial for maintaining operational efficiency and passenger comfort.
Doors	Malfunctions or failures in the Door System can lead to operational disruptions and causing delays in service. Issues such as doors failing to open or close properly, door jams, or sensor failures can disrupt passenger flow and pose safety risks.
MDCU, LVP, VCMS	The MDCU is ensuring the safety of passengers and personnel in the event of automation malfunctions or emergencies. Without it, there could be disruptions in service. Issues with the LVP may not directly impact safety as much as other systems. However, failures could lead to disruptions in service and inconvenience for passengers. The VCMS is vital for monitoring and controlling critical vehicle functions, ensuring passenger safety and operational efficiency. Malfunctions or failures could lead to safety or disruptions in service.
Carbody	Failures or damage to any of these components are not significant implications for passenger safety, operational efficiency, and system reliability. While issues with gangways may impede passenger movement and evacuation procedures during emergencies, they may not pose immediate safety hazards in all scenarios.
HVAC	The HVAC System is essential for passenger comfort and safety. Any failure or malfunction in this system could lead to discomfort, but not direct impact in immediately after malfunction occurred.
High Voltage	The High Voltage System is indispensable for the functioning of the APM. Any failure or malfunction in this system can lead to significant disruptions in service, loss of propulsion, and safety hazards due to electrical faults or fires.
Auxiliary Power Supply	The Auxiliary Power Supply System is vital for maintaining the safety and functionality of the APM during power outages or emergencies. Any failure or malfunction in this system could lead to disruptions in service, compromised safety features, and inconvenience to passengers.

IV. CONCLUSION

This paper analyzes the rolling stock system, consisting of 11 subsystems of the APM trains, using the Failure Mode and Effects Analysis (FMEA) method. This method provides

accurate analysis as it comprehensively examines various types of potential damage, their causes, severity, occurrence frequency, and detectability. The analysis follows a flowchart, starting with the identification of all variables, including failure modes, causes of failure, Severity (S), Occurrence (O), and Detectability (D). Subsequently, the Risk Priority Number (RPN) is calculated, and failure categories are determined.

The results indicate that the Guiding and Hydraulic system falls into the "significant" group, while the Axle and Doors system falls into the "Major" group. The remaining systems are categorized as "Minor." This categorization prompts the need to adjust maintenance plans to ensure the completeness of APM railcar services.

ACKNOWLEDGMENT

The authors would like to express my sincere gratitude to New Railroad Technology company limited for providing valuable maintenance data for the APM trains. I also extend my appreciation to all the engineers and professors who generously offered their guidance and expertise in reviewing this research. Your contributions have been instrumental in enhancing the development of maintenance planning.

REFERENCES

[1] M. Catelani, L. Ciani, D. Galar, G. Guidi, S. Matucci, G. Patrizi, “FMECA assessment for railway safety-critical systems investigating a new risk threshold method,” IEEE Access 9, 86243–86253, 2021.
[2] V. Anes, T. Morgado, A. Abreu, J. Calado, L. Reis, “Updating the FMEA Approach with Mitigation Assessment Capabilities—A Case Study of Aircraft Maintenance Repairs,” Appl.Sci., 12, 11407. 2022.
[3] F. Dinmohammadi, B. Alkali, M. Shafiee, C. Bérenguer, & A. Labib, “Risk Evaluation of Railway Rolling Stock Failures Using FMECA Technique: A Case Study of Passenger Door Syste,” Urban Rail Transit, 2(3–4), 128–145, 2016.
[4] M. Singh, D. Sarkar, H. Raghavendra, “Comparative study of EVM and Fuzzy EVM Methods for Risk Analysis of Infrastructure Transportation Project”. International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 6, Issue 7, ISSN 2319-4847, 2017
[5] S.C. Panja, “Ray, P.K. Failure Mode and Effect Analysis of Indian Railway Signalling System,” Int. J. Perform. Eng. 5, 131–141, 2009.
[6] W. Huang, Y. Li, X. Kou, W. Wang, and Y. Xu, “Using a FMEA–TIFIAD approach to identify the risk of railway dangerous goods transportation system,” Group Decision and Negotiation, 30(1), 63-95, 2021.
[7] E. Kassa, “Analysis of Failures within Switches and Crossings using Failure Modes and Effects Analysis Methodology,” Intelliswitch Symposium. Copenhagen, Denmark, 2017.
[8] M. Szkoda, and G. Kaczor, "Application of FMEA analysis to assess the safety of rail vehicles," In Proceedings of the 23rd International Symposium EURO-ZEL, 2015.
[9] S. Ebrahimi, K. Vachal, and J. Szmerekovsky, “A Delphi-FMEA model to assess county-level speeding crash risk in North Dakota,” Transportation Research Interdisciplinary Perspectives, 16, 100688, 2022.
[10] C.-I. Pascu, R. Malciu, and I. Dumitru, “Perspective Chapter: Defining and Applying the FMEA Process Method in the Field of Industrial Engineering,” DOI: http://dx.doi.org/10.5772/intechopen.107401, 2022.
[11] F. Rahman, M. S. Said, A. Hassan, S. Yusoff, and S. Atmadyaya, “Using QFD and FMEA to Improve Maintenance Effectiveness in a Petroleum Refinery,” 2021.
[12] Talabgaew Sompoap, “Failure Mode & Effects Analysis (FMEA)” [PowerPoint slides], 2021.
[13] U, Y., Qin, Y. Wang, W. Liu, X. and L. Jia, “An Extended FMEA Model Based on Cumulative Prospect Theory and Type-2 Intuitionistic Fuzzy VIKOR for the Railway Train Risk Prioritization. Entropy,” 22, 1418, 2020.
[14] Srisombat Panin Data-Driven Approach to Bogie Overhaul Planning [Journal]. - The Sirindhorn International Thai-German Graduate School of Engineering King Mongkut's University of Technology North Bangkok, 2022.
[15] Talabgaew Sompoap Reliability, Availability, Maintainability Approach for Infrastructure and Logistics of Railway Operations [Journal]. - King Mongkut's University of Technology North Bangkok, 2020.

Effects of temperature and heat treatment time on the hardness of stainless steel

Kawin Jewsuwun[1,a] and Prayoon Surin[2,b]

[1]Department of Advanced Manufacturing Technology, Faculty of Engineering, Pathumwan Institute of Technology, Bangkok, Thailand
[2]Department of Industrial Engineering, Faculty of Engineering, Pathumwan Institute of Technology, Bangkok, Thailand

Corresponding Author:
[a)] Corresponding author: kawin.jew@gmail.com
[b)] another author: prayoon99@gmail.com

Abstract. This research studied the effect of factors affecting the hardness of super-alloy steel after heat treatment. In the experiment, an Analysis of Variance (ANOVA) experimental design was used 2 variables: hardening temperature and hardening time. There are two levels of temperature factors for hardening: 900 and 1000 ºC. The heat treatment times were 1, 2, 3, and 4 hours, respectively, and only the hardness value was considered as the dependent variable. The results of the experimental analysis showed that the factor that most affected the hardness value was the hardening temperature of 1000 ºC. In addition, using a longer heat treatment period will result in an increase in the hardness value, but it should not exceed 3 hours because heat treatment for too long will cause the hardness value to decrease. Examination of the microstructure and chemical composition of the super-alloy steel using scanning electron microscopy (SEM) revealed that the structure of the super-alloy steel that had undergone heat treatment displayed white crystals of the element chromium (Cr) and niobium (Nb) on the surface of the super-alloy steel results in a higher hardness.

Keywords: Design of Experiment, Hardness Testing, Heat treatment

I. INTRODUCTION

Nowadays, super alloy steel is one of the most popular materials used in industrial applications because it has the ability to increase hardness and strength. It can also withstand wear and tear and high temperatures [1]. The super-alloy steels used must have their mechanical properties improved to suit the application. The addition of various elements to super-alloy steel results in improved structure and improved mechanical properties [2]. Increasing hardness by heat treatment is another method to improve mechanical properties, which results in an increase in hardness and resistance [3]. Z.M. Shi et al. [4] found that the addition of erbium at 0.3% by weight had the best effect. The Fe-containing Al_3Er phase introduced by T6 treatment improves the hardness of A356 aluminum alloy. At the same time, research studies that add silicon to aluminum alloys will increase the grain size of aluminum alloys. Grain pretreatment experiments found that the grain size became finer when more than 3% silicon was added, which increased the mechanical properties of the metal [5]. Raghav Srivastava et al [6] investigated the influence of sulfur (S) and manganese (Mn) on the mechanical properties and microstructure of gray cast iron. It was found that manganese will react with sulfur to form MnS. Manganese also counteracts the formation of FeS, which causes the strength to decrease at high temperatures in M-alloys. Mohammadnezhad et al. [7] studied the effect of vanadium addition on the microstructure, mechanical properties, and wear resistance of Ni-Hard 4 white cast iron. The results showed that when increasing the amount of vanadium concentration, the microstructure of Ni-Hard 4 white cast iron became more refined. Hardness and wear resistance are improved without compromising fracture toughness.

The key to optimizing the performance of super-alloy steels is to control the level of factors that affect structural improvement and mechanical properties. Therefore, experimental design principles were applied to find appropriate parameters in order to reduce surface defects that affect the structure and mechanical properties of super-alloy steel. W. Chalermsmithchai et al. [8] studied the adjustment of the amount of elements that are chemical ingredients of iron

water used in the production of raw steel using an experimental design. Results of the experiment using Factorial Design in designing the experiment. Then, factors influencing the strength of steel were used to design the experiment using Box- Behnhen Design, resulting in a regression equation that was related to the strength of steel. Cristie Diego Pimenta et al. [9] studied the variables that affect the hardening process, namely temperature, time, and medium that affect the hardness value after quenching and designed a factorial experiment. The results of the experimental design indicate that finding the appropriate value of parameters (Response Optimization) involves finding the level of factors that cause the change in the response to be as desired. T. Varun Kumar et al [10] studied the increase in hardness in duplex stainless steel by suitable annealing process. The trial design used 3k Full Factorial Design to find the appropriate factors. The results of the experimental design showed that heating at a temperature of 1240 °C with a residence time of 4.5 hours and cooling at a temperature of 1043.90 C resulted in the duplex stainless steel having the highest hardness. Ramos-Azpeitia et al. [11] studied the effect of each parameter related to the heat treatment of CA- 6NM cast alloy with a multi- layer experimental design process. The experimental design can effectively select the appropriate level of experimentation to improve the production factors. It can also be used to find appropriate factors in other production processes. S. Udomboonyanupap et al. [12] determined the optimum level of the carbon steel hardening process. Factors in the study included quenching temperatures of 750, 800 and 830 °C, and quenching times of 2 and 5 minutes. Then, the hardness of the work-piece was checked and it was found that the best condition for hardening was at a temperature of 764 °C and quenching time of 5 minutes.

The objective of this research is to study the appropriate factors between temperature and heat treatment time that affect the hardness and distribution of chromium (Cr) and niobium (Nb) in super-alloy steel. The results obtained from this research can be used as an alternative to improve the microstructure and mechanical properties of super-alloy steel for further industrial applications.

II. METHODOLOGY

1.Experimental design

The experimental design used an Analysis of variance (ANOVA) experimental plan, with hardness being the main answer of the experiment. This is because the hardness value is an element that is taken into account in the mechanical properties of super-alloy steel after heat treatment. There are two factors used in the study: temperature used in heat treatment and time used in heat treatment as shown in Tables 1 and 2, respectively.

Table 1 Level of temperature factor for hardening

Factor	Temperature
Levels	2
Values	900 °C, 1000 °C

Table 2 Level of time factor for hardening

Factor	Time
Levels	4
Values	1, 2, 3, 4 Hour

2. Preparing the specimen

In this experiment, super-alloy steel was cut into trapezoidal sheets to prepare for heat treatment. The chemical composition of super-alloy steel was examined using an ARL-3460 spectrometer as shown in Table 3.

Table 3 Chemical composition of super-alloy steel

Chemical composition			
C	2.9242	Si	1.0813
Mn	0.3417	P	0.0443
S	0.0149	Ni	0.2611
Cr	21.1092	Mo	1.5024
Cu	0.6227	Nb	2.8856

3.Experiment

The experiment was carried out by subjecting the prepared super-alloy steel to a structural modification heat treatment at 1500 °C for 4 h and allowing it to cool in air. It is then subjected to heat treatment, where each heat treatment is determined by the conditions of each experimental sequence. After the work-piece has been heat treated according to the conditions, it is tested for overall hardness using a Vickers test. In this research, the hardness value used is in accordance with ASTM E-92-82 standards. Hardness testing involves testing specimens that have not been milled using a diamond indenter. The pressing weight was 30 kgf for 15 seconds with 9 pressure points per specimen and the average was taken. The test specimen was rough-polished with sandpaper no. 120, 200, 400, 700, and 1000, respectively. Then it was finely polished with diamond powder of sizes 8, 4.5 and 1 microns and the surface of the test piece was etched with a solution. The specimen was then studied for microstructure using a scanning electron microscope.

III. RESULTS AND DISCUSSION

Statistical analysis

Statistical analysis to find appropriate factors is essential to verify the accuracy of the data. From the results of ANOVA analysis, it was found that the P-value of the heat treatment temperature was equal to 0.00, which was less than 0.05, which is shown in Figure 1. While the time used for heat treatment has test results shown in Figure 2, it has a P-value of 0.00, which is also less than 0.05. Therefore, it can be concluded that the data have a normal probability distribution, are independent and have no difference in variance. Therefore, the data can be used for further analysis.

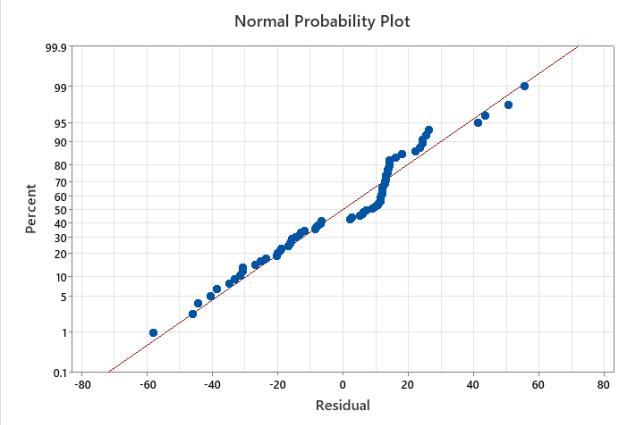


Figure 1 Probability plot of heat treatment temperature

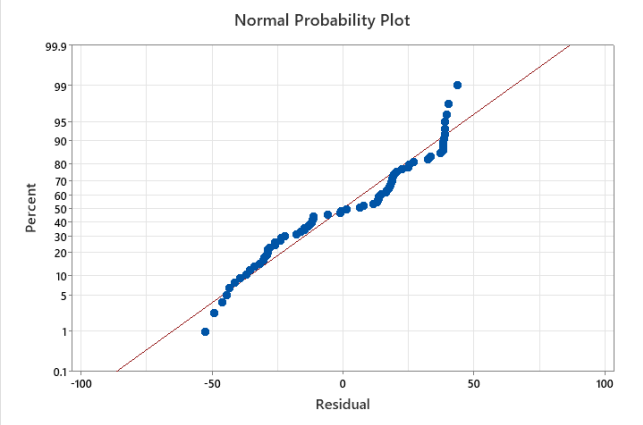


Figure 2 Probability plot of heat treatment time

Experiments to confirm analytical results

In order to verify the correctness of the appropriate factor levels obtained from the statistical values, a heat treatment experiment on the super-alloy steel was carried out according to the temperature and time conditions obtained from the statistical analysis. The microstructure was then examined using a scanning electron microscope (SEM). Figure 3 shows the microstructures examined using a scanning electron microscope at 1000X magnification of specimens heat treated at 1000 °C and with heat treatment times of 1, 2, 3, and 4 h, respectively.

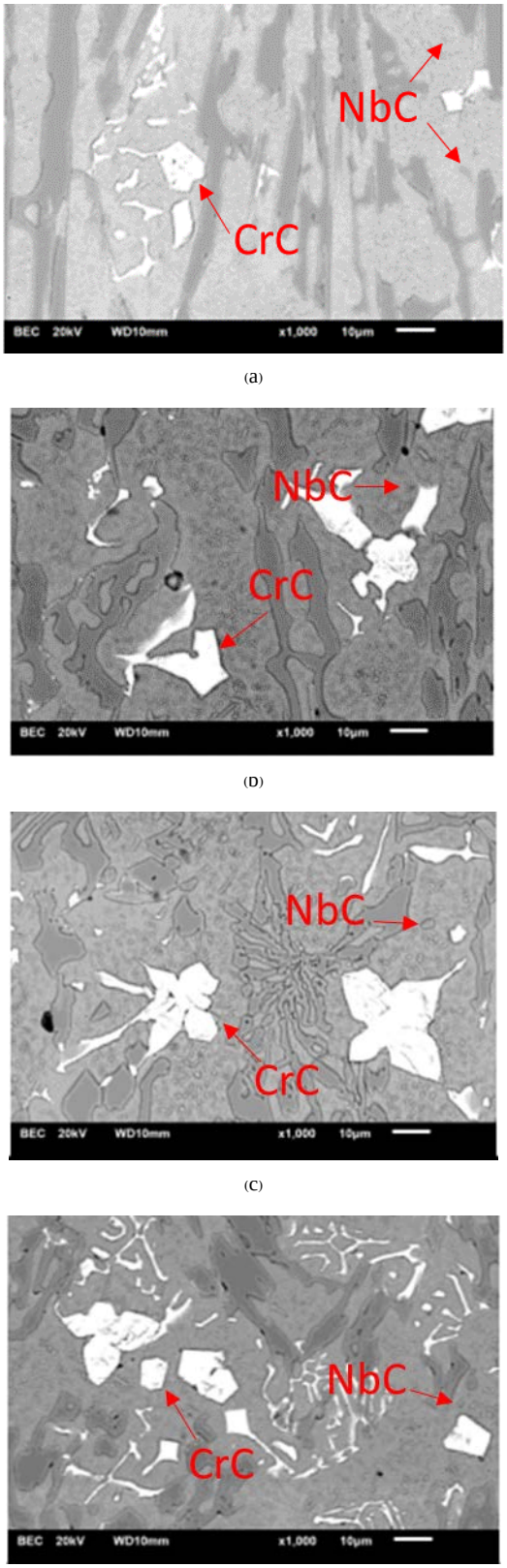


Figure 3. Microstructure of super-alloy steel after heat treatment at 1000 °C at (a) 1 h (b) 2 h (c) 3 h (d) 4 h.

IV. CONCLUSION

This research used an Analysis of variance (ANOVA) experimental design to analyze the appropriate factors in super-alloy steel plated frames. From the results of the statistical analysis, it can be concluded that temperature and heat treatment time have a significant effect on the average hardness value. It was found that high heat treatment temperature and long heat treatment time resulted in improved microstructure. It also results in the hardness value increasing, but it decreases when the heat treatment time is too long. In addition, scanning electron microscope (SEM) images show that the longer the heat treatment time, the larger the crystallization of chromium (Cr) and niobium (Nb) in the super-alloy steel, which will decrease after heat treatment time of more than 3 hours.

REFERENCES

[1.] KARL- HEINZ ZUM GAHR. (1987) . MICROSTRUCTURE AND WEAR OF MATERIALS, ELSEVIER SCIENCEP PUBLISHERS B.V. (Vol.10, p.1-562).

[2.] Z.R. Nie, J.B. Fu, J.X. Zou, T.N. Jin, J.J. Yang, G.F. Xu, H.Q. Ruan, T.Y. Zuo, "Advanced Aluminum Alloys Containing Rare- earth Erbium," Materials Forum Volume 28 pp.197-201 , 2004.

[3.] Gowri Shankar M C, Manjunath Shettar, Sharma SS, Achutha Kini and Jayashree, " Enhancement in Hardness and Influence of Artificial Aging on Stir Cast Al6061-B_4C and Al6061-SiC Composites" Materials Today: Proceeding 5 pp. 2435-2443, 2018

[4.] Z. M. Shi, Q. Wang, G. Zhao and R. Y. Zhang, "Effects of erbium modification on the microstructure and mechanical properties of A356 aluminum alloys" Materials Science & Engineering A 625 pp. 102-107, 2015

[5.] Y.C. Lee , A.K. Dahle, D.H. StJohn and J.E.C. Hutt, " The effect of grain refinement and silicon content on grain formation in hypoeutectic Al–Si alloys" Materials Science and Engineering A259 pp. 43–52, 1999

[6.] Raghav Srivastava, Bharat Singh and Kuldeep K. Saxena, " Influence of S and Mn on mechanical properties and microstructure of grey cast iron" Materials Today: Proceeding volume 26 part 2 pp. 2770-2775, 2020

[7.] M. Mohammadnezhad, V. Javaheri, M. Shamanian, M. Naseri and M. Bahrami, " Effects of vanadium addition on microstructure, mechanical properties and wear resistance of Ni-Hard4 white cast iron" Materials and Design 49 pp. 888–893, 2013

[8.] W. Chalermsmithchai and S. Jongprasitphon, "The study for element optimization in billet production by design of experiment" Operation Research Network Conference (OR-NET 2008), pp. 245-253, 2008

[9.] C. D. Pimenta, M. Borges Silva, R. L. de M. Campos and W. R. de C. Junior, "Desirability and Design of Experiments Applied to the Optimization of the Reduction of Decarburization of the Process Heat Treatment for Steel Wire Sae 51B35" American Journal of Theoretical and Applied Statistics, pp. 35-44, 2018

[10.] T.V. Kumar, K. Karthick, K. Dhanushkaran and D. Jagadeesh, "OPTIMISING THE PARAMETERS AND INCREASING THE HARDNESS OF DUPLEX STAINLESS STEEL BY ANNEALING PROCESS IN VALVE MANUFACTURING INDUSTRIES" Materials Science and Engineering 954, 2020

[11.] M. Ramos-Azpeitia and C. A. Perez-Alonso, "HEAT TREATMENT OPTIMIZATION OF CA-6NM CAST ALLOY USING A FULL FACTORIAL DESIGN OF EXPERIMENTS APPROACH" International Journal of Metalcasting, 2020

[12.] S Udomboonyanupap, and N Pianthong, "Optimization factors in hardening of AISI 5106 carbon steel" Journal of Science & Technology MSU, Vol. 42 Issue 2, pp. 65-72, 2023

Alteration of Relative Permeability and Investigation of Oil Recovery Mechanism during Low Salinity Waterflooding in Different Rock Formations

Tina THARN[1, a)], Falan SRISURIYACHAI[2, b)] and Chandoeun ENG [3, c)]

Author Affiliations
[1]Department of Mining and Petroleum Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand
[2]Department of Mining and Petroleum Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand.
[3]Faculty of Geo-resources and Geotechnical Engineering, Institute of Technology of Cambodia, Russian Federation Blvd., P.O. Box 86, Phnom Penh, Cambodia.

Author Emails
[a)] Corresponding author: 6570802121@student.chula.ac.th
[b)] falan.s@chula.ac.th
[c)] chandoeun.eng@itc.edu.kh

Abstract. This research investigates alteration of relative permeability curves during low salinity waterflooding (LSWF) in different rock formations including clean sandstone, shaly-sandstone, and dolostone (carbonate), utilizing the Johnson–Bossler–Naumann (JBN) method at unsteady-state condition. Coreflood experiments were conducted to gather various data such as oil recovery factor, fluid production rates, and differential pressure across the core sample. Through the application of the JBN method, the data was converted into relative permeability data. Then, Corey's correlation was employed to generate relative permeability curves that matched the calculated data. The Corey's exponent derived from the most suitable curves was scrutinized. Analysis of ion compositions in effluents from the coreflood test and observation of fluid turbidity aided in understanding the mechanism of oil recovery. The experiments revealed that the curvature of relative permeability curves undergoes minimal changes, suggesting that oil recovery mechanisms may occur without altering the rock's preference toward more water wetness. Nevertheless, changing amounts of calcium and magnesium ions in effluents from cases with high oil recover indicates the occurrence of multi-component ion exchange (MIE). Furthermore, the turbidity of effluent during the transition from high salinity water injection to low salinity waterflooding provides additional evidence for the oil recovery mechanism involving fine migration.
***Index Terms*—Low-salinity waterflooding, Un-steady state, relative permeability curve, JBN method, Corey's correlation.**

I. INTRODUCTION

In the petroleum production stage, there are three main consequences, including primary, secondary, and tertiary. To satisfy growing energy demand, oil and gas remain top priorities, providing sufficient energy to a variety of sectors. As a consequence, multiple Enhanced Oil Recovery (EOR) methods were developed, including low salinity water flooding (LSWF). Different researchers have revealed and explored the benefits of LSWF. Low salinity brine is more effective than fresh water for fluid injection, resulting in more oil recovery in sandstone formation, approximately 15% [1]. Moreover, fine migration is also observed during LSWF to enhance oil recovery and obtain oil recovery 21.3% better than using high-salinity water [2]. Another detail investigation revealed that sodium chloride solution at 1,500 ppm was suitable to gain oil recovery. The suggested optimal range of low-salinity water is between 1,000 and 2,000 ppm salinity, which has an adequate effect on enhanced oil recovery and the salinity threshold for LSWF is approximately in the range of 4,000–5,000 ppm [3-8] which is specially for sandstone. The incremental oil recovery from LSWF could be achieved by different mechanisms, including Multi-component Ion Exchange generation (MIE), wettability alteration, raising of pH value, generation of in-situ surfactant, and liberation of fine particles together with oil drop [4]. Whereas, the investigation by Yousef et al. (2011) [9] revealed that LSWF on carbonate rocks could enhance oil recovery by dilute seawater twice with 7% incremental oil recovery and reach maximum recovery with 20-time dilute seawater. Yet, pinpointing the specific function engaged by a particular set of mechanisms has become more complicated. Analyzing the petrophysical characteristics of rock before and after the operation might offer confirmation to support the outcomes. This study investigates the alteration of the relative permeability curves, which are relative permeability to water and to oil. The Johnson-Bossler-Nauman (JBN) method, which is an unsteady-state technique to generate relative permeability data, was employed to construct the whole relative permeability curve. The obtained data from the JBN method

is then matched with the relative permeability curves constructed by Corey's correlation to identify the best fit for Corey exponent, representing the curvature of the curve. The final discussion was assisted by measuring the concentration of key ions in the effluent from the coreflood experiment. Since sandstone and shaly-sandstone may contain some clays such as illite and kaolinite, Ca^{2+} and Mg^{2+} are detected and used as key indicators of the MIE mechanism. Both ions are able to make the interaction between the solid wall and organic acid in the oil phase. Additionally, shaly-sandstone is enriched with these clay minerals, and hence, this causes shaly-sand's surface to be more oil-wet.

New finding of this study should be a vital evaluation of the oil recovery mechanism from LSWF as an additional rational for future research, particularly for the design of water formulation to satisfy different reservoir characteristics, and the outcome from relative permeability curves can be a significant support for studies at the macroscopic scale by using a reservoir simulator, which is an appropriate step to scale up prior to field implementation.

II. METHODOLOGY

A. *Coreflood Experiment*

The Coreflood system was employed to conduct Coreflood experiments. In this study, reservoir rock were represented by Sirulian dolostone, Barea sandstone, representing a cleaner sandstone, and a shaly-sandstone from Sirikit oilfield, which is located in the north of Thailand. Three core samples were tested in this study. Barea sandstone labeled as SS has 19.33% porosity and 61.48 mDarcy in absolute permeability. Shaly-sandstone (SHS) has 19.49% porosity with an absolute permeability of 131 mDarcy. Sirulian dolostone (DL) has 8.15% porosity and 198.3 mDarcy in absolute permeability. Sirikit oilfield's crude oil was combined with dodecane at ratio 3:7 (crude oil: dodecane) to prevent formation of wax throughout the experiment. Formation water for SHS and SS was prepared using the water analysis obtained from the Sirikit oilfield with a total dissolution solid (TDS) of 14,098 ppm. For DL connate, water was designed with a TDS of 28,196 ppm, and the formulation of water was obtained from the averaging of formation water in limestone and dolostone around the globe. The selection of low-salinity water (LSW) was based on previous studies. For SS and SHS, LSW's were prepared at the TDS of 2,000 ppm and 1,000 ppm, respectively. The formulation of LSW for SS core sample was 75% of $CaCl_2$ and 25% of KCl, whereas for the SHS core sample, LSW was prepared using 85% of $CaCl_2$ and 15% of KCl. For DL core sample, LSW comprises $CaCl_2$, $MgCl_2$, and Na_2SO_4 with a ratio of 1:1:1, and TDS is 10,000 ppm. Even though the TDS is above the recommended values from the literatures, it was observed from the previous study that the difference between the TDS of formation water and injected water is also another key success for LSWF. The laboratory setup of this experiment is displayed in Figure 1. Coreflooding experiment was performed at reservoir temperature 50 Celsius with injection rate of 0.5 cc/min under confining of pressure of 2,500 psi. During the test, produced oil was collected to measure the oil recovery factor, and various

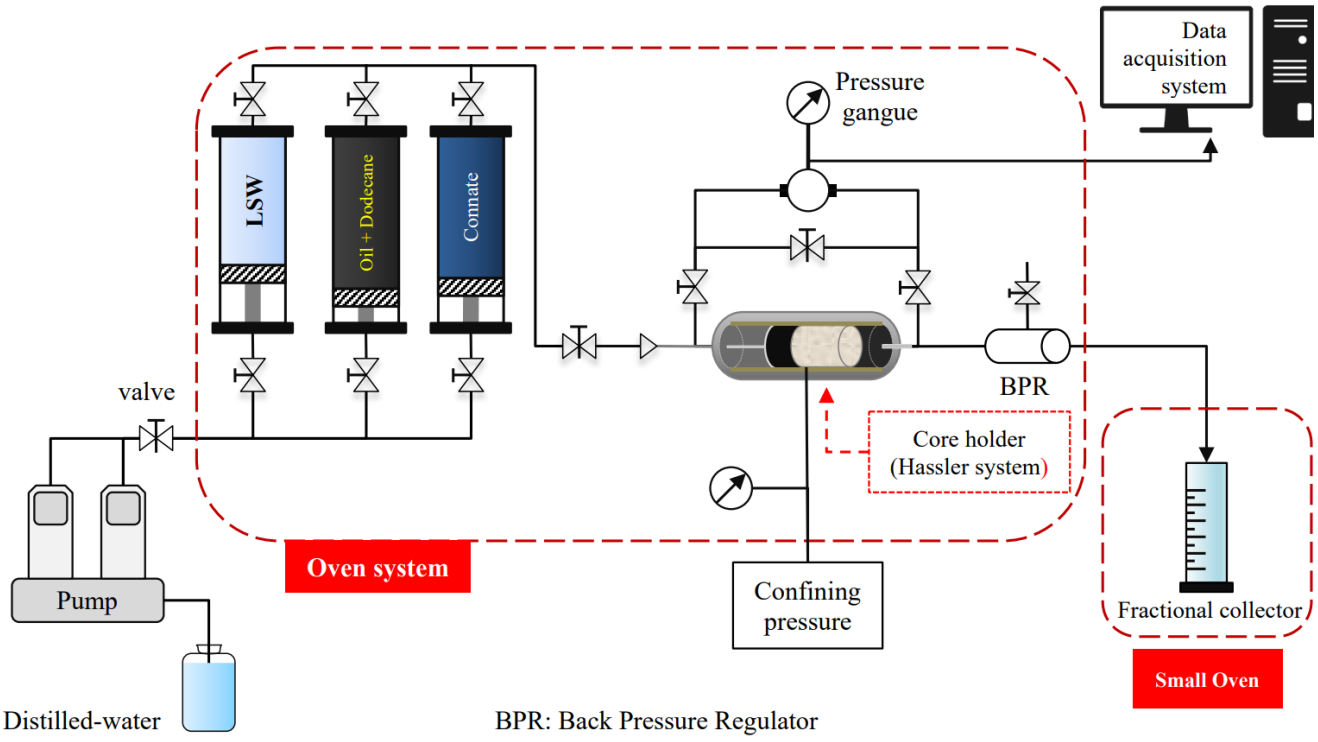


Figure 1 The coreflooding laboratory setup consists of three chambers containing low-salinity water, oil, and connate water. Core holders are commonly made utilizing the Hassler system. Back pressure regulators (BPR) should be employed to reduce pressure shock. The injection pressure must be 300 psi lower than the confining pressure.

differential pressures at specific Pore Volumes (PV) were recorded for the calculation of relative permeability curve. Moreover, produced oil and water effluent were collected in cylinders to observe the occurrence of fine migration and to further investigate the liberated ions during the process. To identify the concentration of calcium ion and magnesium ion, color titration was employed by using Ethylenediaminetetraacetic acid (EDTA) as a titrant and utilizing Eriochrome Black T (EBT) and Hydroxy Naphthol Blue (HNB) as color indicators. The end-point of color titration can be detected from the change from purple to blue without red hint.

B. *Johnson–Bossler–Naumann (JBN)*

The results of coreflood experiments are essential for a scientific examination of the relative permeability characteristics. In unsteady-state regimes, Johnson, Bossler, & Naumans (JBN) approach is the most beneficial for two-phase flow over porous medium [10]. It is particularly important for the oil and gas sector to comprehend fluid movement in reservoir rock. With the goal of executing JBN analysis for identifying the relative permeability of each phase at a specific water saturation, experimental data from the coreflood must be obtained for each time period, including:

- Q_i : Quantity of displacing phase injected

$$Q_i = \frac{Q_t}{V_P} \quad (1)$$

Where:

Q_t	: Volume of liquid produced	(ml)
V_P	: Volume of pore space	(ml)

- ΔP : Pressure differential (psi)
- ΔP_i : Pressure differential at initial conditions (psi)
- Q_o : Volume of oil produced (ml)
- Q_t : Volume of water produced (ml)

There are a few procedures in the computation process that are necessary to assess the relative permeability curve from the JBN study, including:

- The value of S_{wav} after breakthrough will be calculated from equation (2)
- The value of k_{ro} & k_{rw}

The following equation will be used to compute the fractional flow:

$$f_o = \frac{dS_{wav}}{dQ_i} = \frac{Q_o}{Q_t} = \frac{Q_o}{Q_{w,inj}} = \frac{1}{1+\left(\frac{k_{rw}\times\mu_o}{k_{ro}\times\mu_w}\right)} \quad (2)$$

where:

S_{wav}	: Average water saturation	(fraction)
μ_w	: water viscosity	(cp)
μ_o	: Oil viscosity	(cp)

The ratio of effective to absolute permeability (from the coreflooding experiment) will be used to compute the relative permeability value.

$$k_{r,n} = \frac{k_{eff,n}}{k_a} \quad (3)$$

Darcy's equation

$$k_{eff,n} = \frac{q\times\mu_n\times L}{A\times\Delta P} \quad (4)$$

Where:

k_a	: Absolute permeability	(md)
n	: fluid phase	(oil, water)
q	: fluid flow rate	(cc/min)
A	: Cross area of core	(cm^2)
ΔP	: Pressure drop	(psi)

C. *Data Extrapolation*

In this case, Corey's correlation must be the favored method, as displayed in equations (5) and (6) because the majority of the real data (k_{rw} and k_{ro}) were gathered at the end-point saturation. Nevertheless, the Corey exponent number must be sufficient and correspond to the actual data. Corey exponent represents the curvaceous of relative permeability curves, reflecting the interference between oil and water on rock surface. The closer value to unity indicates less interference, allowing oil to be produced without being impeded by water. Predicted values are Corey's correlation results (k_{rw} and k_{ro}) which are computed in a range that resembles the real saturation data. In numerical regression approaches, the RSME is also employed, as shown in equation (7). The RMSE is a frequently used metric for estimating a regression model's prediction error. A more optimal fit for the results is demonstrated by a lower RMSE. Up until the RSME reaches its lowest value, the Corey exponent with the best match will be iterated. The incremental of Corey exponent will develop the curvaceous of the relative permeability curve.

Corey's correlation

$$k_{rw}(S_w) = k_{rw,or}\left(\frac{S_w - S_{cw}}{1 - S_{cw} - S_{or}}\right)^{n_w} \quad (5)$$

$$k_{ro}(S_w) = k_{ro,cw}\left(\frac{1 - S_w - S_{or}}{1 - S_{cw} - S_{or}}\right)^{n_o} \quad (6)$$

Where:

$k_{rw,or}$	: relative water permeability at S_{or}	(fraction)
$k_{ro,cw}$	: relative oil permeability at S_{cw}	(fraction)
S_{cw}	: connate water saturation	(fraction)
S_w	: water saturation	(fraction)
S_{or}	: irreducible oil saturation	(fraction)
n_w, n_o	: Corey exponent for water, oil	(-)

Root means square error (RSME)

$$RSME = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2} \quad (7)$$

Where:

N : The total amount of data (-)
$\hat{y}_i$: Predicted $k_{r,n}$ from Corey's correlation (-)
y_i : Actual $k_{r,n}$ from JBN analysis (-)

III. RESULTS AND DISCUSSION

Coreflooding experiment has provided the initial information before acquiring the relative permeability curve, including differential pressure and recovery factor at each pore volume interval from different core samples and conditions. For SS, the pressure difference reaches approximately from 37 to 44 psi, as illustrated in Figure 2. For SHS, the pressure difference reaches approximately from 22 to 28 psi, as demonstrated in Figure 3. For DL, the pressure difference reaches approximately from 1.8 to 3.3 psi, as shown in Figure 4. As a close look at incremental oil production, the SS core sample obtained an increase in oil recovery of 6.62% and 10.37% for the SHS core after utilizing the selected LSWF, as shown in Figures 2 and 3, respectively. The incremental oil recovery leads endpoint of relative permeability to oil and water shifted the irreducible oil saturation from 72% to 75.56% (Figure 5) and the SHS case shifts from 74% to 82.14% (Figure 6). Additionally, the DL core sample possesses a smaller delta P compared to the SHS and SS cores. However, the increment in oil recovery was around 16.21% by LSWF. The experiment was performed by injecting formation water as high salinity water flooding (HSWF), and then the process was switched to LSWF, as summarized in Table 1. Relative permeability to water and to oil data could be obtained from the JBN technique, as there is a gradual change in oil recovery from 0PV to 3PV in Figures 2, 3, and 4, from 3.5 to 7PV in Figures 3 and 4, and from 9 to 12PV in Figure 2. The relative permeability curves were subsequently generated using Corey's correlation, and the best-fitting curves were chosen using RSME as explained in Section C.

From Figure 5, the endpoint relative permeability to water remained almost the same after changing from HSWF to LSWF, whereas in Figures 6 and 7, there is an obvious reduction of endpoint relative permeability to water after LSWF. Additionally, all cases show the increment of end-point water saturation. These could be explained by the low-salinity effect which triggers oil liberation to clear the flow pathway. The differential pressure was slightly dropped, which caused the relative permeability to water became less during LSWF. For SS, Corey exponent for oil is increased from 2.1 to 3 and Corey exponent for water is quite stable which are 1.0 and 0.9. For SHS, Corey exponent for oil is increased from 1.5 to 1.8, whereas Corey exponent for water is nearly constant (0.7 and 0.6). For DL, Corey exponent for both oil and water are constant at 1.6 and 1.0, respectively. From the summary of Corey exponent for each case as shown in Table I, it reveals that the incremental of oil production should be caused from oil recovery mechanism without alternation of wetting preference.

From the results summarized in Table II, the concentrations of calcium ions and magnesium ions changed during the injection of HSW and LSW, meaning that low-salinity water disturbs the equilibrium in porous media at the contact liquid-liquid and liquid-rock walls. During the Multi-component Ion Exchange (MIE), some particles are able to lose their attachment to the rock surface, leading to the appearance of turbidity in the effluent after LSWF.

TABLE I
SUMMARY OF RECOVERY FACTOR AND COREY EXPONENT FOR EACH CASE

	Core SS	Core SHS	Core DL
Formation water (ppm)	14,098	14,098	28,196
LSW (ppm)	2,000	1,000	10,000
Incremental RF (%)	6.62	10.37	16.21
Corey exponent oil (Formation water/LSW)	2.1/3	1.5/1.8	1.6/1.6
Corey exponent water (Formation water/LSW)	1/0.9	0.7/0.6	1/1

TABLE II
SUMMARY OF IMPORTANT PARAMETERS FROM COREFLOOD TESTS

Ion	Mode	Core SS (ppm)	Core Shs (ppm)	Core DL (ppm)
Ca^{2+}	WF	121.57	110.21	300.59
	LSW	152.30	80.16	320.62
Mg^{2+}	WF	42.94	54.69	1318.55
	LSWF	139.35	34.43	1251.71
Appearance of turbidity		YES	YES	NO

For SS and SHS, where magnesium ions were not included in the injected LSW, the appearance of magnesium confirms the MIE, especially in the case of SS. In the case of SHS, due to a complex system of clays, some of the liberated magnesium ions may be re-consumed to balance the abundance of negative changes in clay structures. The appearance of turbidity in effluents in both SS and SHS cores confirms the major oil recovery mechanisms, which are MIE and fine migrations. On the contrary, turbidity was not observed after LSWF. The input concentrations of calcium ions and magnesium ions in LSW in the DL case were 1,201 and 815 ppm, respectively. Comparing these data with the concentration obtained from the effluent, it can be observed that calcium ions were consumed, whereas magnesium ions were expelled from the structure. As dolostone is composed of both calcium and magnesium ions in its structure,

exposure of rock to low salinity water may result in a change in rock composition due to the dissolution of magnesium ions. Nevertheless, dissolution of calcium may also occur, but calcium ions may be spent in creating the calcium carboxylate complex, which is an initial substance to trigger the overall oil recovery mechanism. From the results for DL, the main oil recovery mechanism can be MIE together with the dissolution of rock.

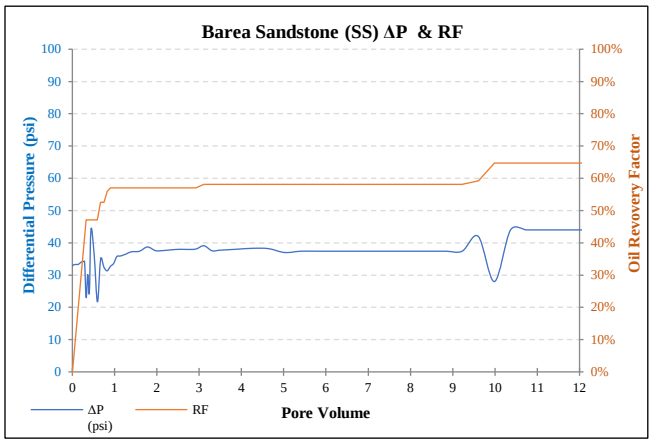


Figure 2 Oil recovery factor and delta P of smith sandstone (SS), injected from 0PV to 12PV

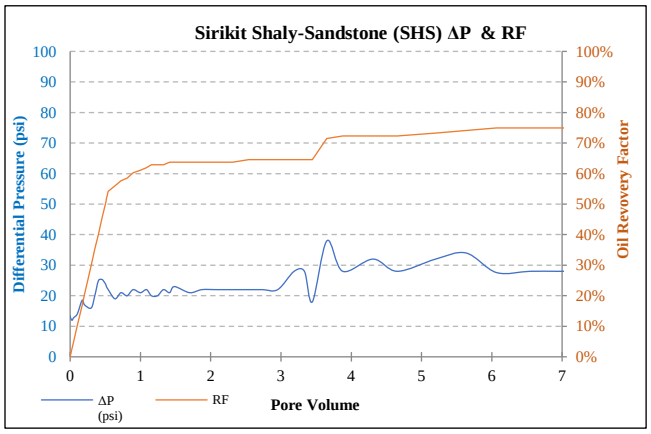


Figure 3 Oil recovery factor and delta P of core SHS, injected from 0PV to 7PV

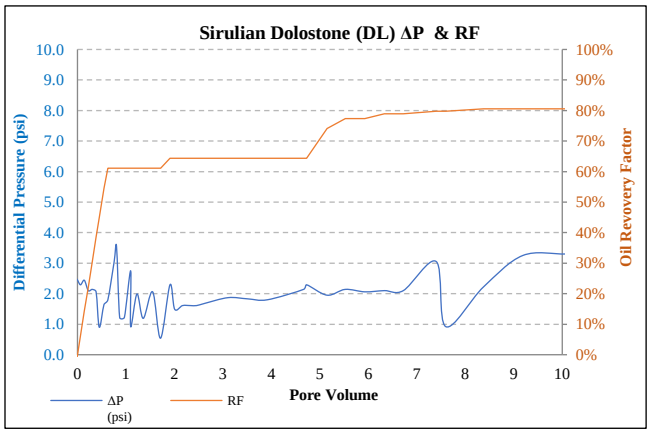


Figure 4 Oil recovery factor and delta P of Carbonate (DL), injected from 0PV to 10PV

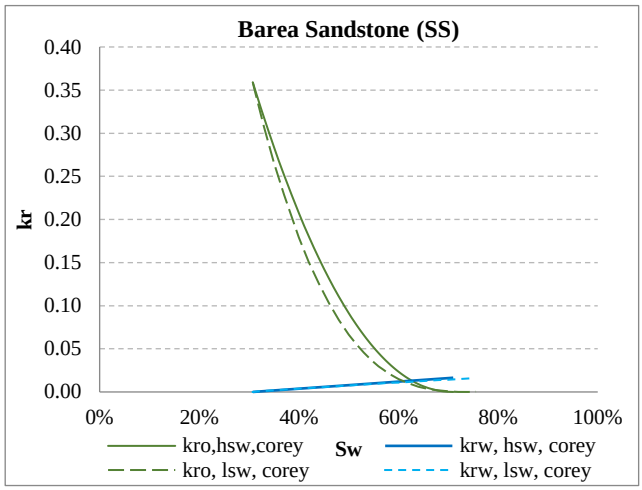


Figure 5 Comparison of relative permeability to oil and to water obtained from injection of formation water and low salinity waterflooding (Core SS)

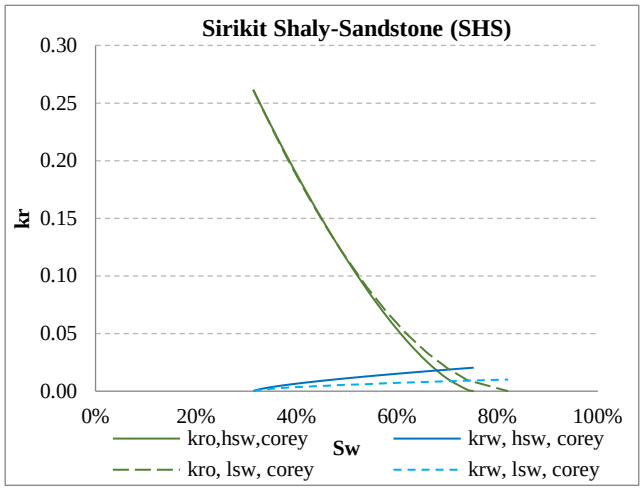


Figure 6 Comparison of relative permeability to oil and to water obtained from injection of formation water and low salinity waterflooding (Core SHS)

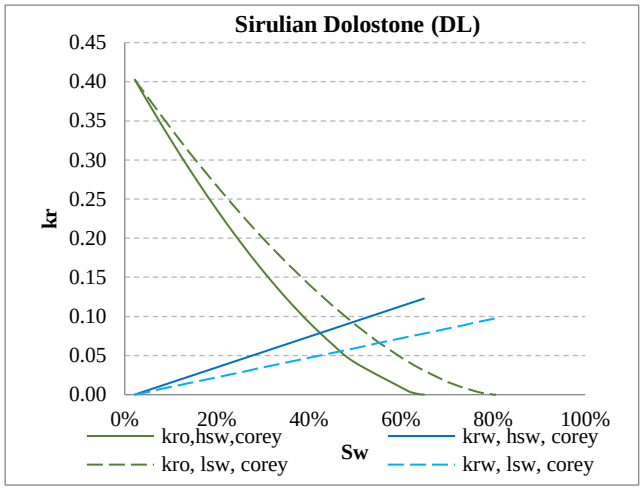


Figure 7 Comparison of relative permeability to oil and to water obtained from injection of formation water and low salinity waterflooding (Core DL)

I. CONCLUSION

From the study, different rocks respond in different ways to low salinity waterflooding. However, it can be observed that all the rocks do not show significant changes in curvaceous of relative permeability to oil and to water. The major change in relative permeability to water was the reduction of end-point relative permeability to water, which can be explained by the clearing of the flow pathway. From this step, a change in rock preference toward water wetness cannot be observed.

Regarding the observation of MIE, all cases react to LSW, as can be seen from the release of magnesium ions. Nevertheless, the appearance of turbidity in effluence was only observed in the cases of shaly-sandstone and sandstone. Therefore, the major mechanisms of oil recovery in shaly-sandstone and sandstone are multi-component ion exchange together with fine migration, whereas in the case of dolostone, dissolution of magnesium ions caused by MIE is the major oil recovery mechanism.

To complete these explanations, an analysis of the rock composition both before and after getting in contact with low salinity water should be performed. The mechanisms enabling oil recovery would be confirmed during each step by measuring elemental change and observing pore surface structure.

ACKNOWLEDGMENT

The authors would like to thank PTT Exploration and Production Public Company Limited for providing rock samples, oil samples, and information of formation water.

NOMENCLATURE

BPR	:	Back Pressure Regulator
DL	:	Dolostone
EBT	:	Eriochrome Black T
EDTA	:	Ethylenediaminetetraacetic Acid
EOR	:	Enhanced Oil Recovery
HNB	:	Hydroxy Naphthol Blue
HSW	:	High-salinity Water
HSWF	:	High-salinity Water Flooding
LSW	:	Low-salinity Water
LSWF	:	Low-salinity Water Flooding
MIE	:	Multi-component Ion Exchange
RSME	:	Root Mean Square Error
SHS	:	Shaly-sandstone
SS	:	Sandstone
TDS	:	Total Dissolution Solid

REFERENCES

[1.] Smith, K.W. *Brines as Flooding Liquids. Seventh Annual Technical Meeting*. in *Seventh Annual Technical Meeting*. 1942. Mineral Industries Experiment Station: Pennsylvania State College.

[2.] Reiter, P.K., *A water-sensitive sandstone flood using low salinity water*. 1961, University of Oklahoma.

[3.] Zhang , Y., X. Xie , and N.R. Morrow *Waterflood Performance by Injection of Brine with different Salinity for Reservoir Cores*, in *SPE Annual Technical Conference and Exhibition*. 2007: Anaheim, California, USA.

[4.] Al-Shalabi, E.W. and K. Sepehrnoori, *A comprehensive review of low salinity/engineered water injections and their applications in sandstone and carbonate rocks.* Journal of Petroleum Science and Engineering, 2016. **139**: p. 137-161.

[5.] Boussour, S., et al. *Oil recovery by low salinity brine injection: Laboratory results on outcrop and reservoir cores*. in *SPE Annual Technical Conference and Exhibition*. 2009. OnePetro.

[6.] Lager, A., et al. *Low salinity oil recovery-an experimental investigation*. in *International Symposium of the Society of Core Analysts*. 2006.

[7.] Ligthelm, D.J., et al. *Novel waterflooding strategy by manipulation of injection brine composition*. in *EUROPEC/EAGE conference and exhibition*. 2009. OnePetro.

[8.] McGuire, P., et al. *Low salinity oil recovery: An exciting new EOR opportunity for Alaska's North Slope*. in *SPE western regional meeting*. 2005. OnePetro.

[9.] Yousef, A.A., et al., *Laboratory Investigation of the Impact of Injection-Water Salinity and Ionic Content on Oil Recovery From Carbonate Reservoirs.* SPE Reservoir Evaluation & Engineering, 2011. **14**(05): p. 578-593.

[10.] Apostolos, K., J. Bryan, and S. Taheri, *Fundamentals of Fluid Flow in Porous Media.* PERM Inc TIPM laboratory, 2016.

Simple and Effective Approach for Fresh or Frozen Fish Classification Using Image Processing and Deep Learning

Taweepol Suesut [1, a)], Chanitnunt Sanbooranapunt[1, b)] and Ari Kuswantori[2, b)]

Author Affiliations

[1]School of Engineering King Mongkut's Institute of Technology Ladkrabang Bangkok, Thailand
[2] Politeknik Gajah Tunggal, Indonesia

Author Emails
[a)] Corresponding author: taweepol.su@kmitl.ac.th
[b)] 64601030@kmitl.ac.th
[c)] arikuswantori77@gmail.com

Abstract. Computer and machine vision is an emerging topic that developed massively this decade in many fields, including the food and fish industries. Image processing and deep learning as the core have been developed for various purposes in the fish industry. However, the authors could not find published work on classifying fresh and frozen fish based on deep learning and image processing, even though this work might be very useful for developing automation in that field. For this reason, this work is intended to fill this gap. The image dataset was made from tilapia fish obtained from traditional markets in Thailand. Several effective image pre-processing techniques were proposed. Resizing, cropping, and Lanczos resampling were utilized to minimize computation without eliminating (even focusing on) important features. Then a simple deep-learning algorithm was employed for classification. However, despite its simplicity, this approach could completely classify all fresh and frozen tilapia in the dataset. The resulting confidence score was also very high, reaching an average of 99.82% for all images in all classes.

***Index Terms*— fish classification, deep learning, image processing, Lanczos resampling**

I. INTRODUCTION

Computer and machine vision are also part of industrial automation and appear as a very interesting topic that continues to be developed today. It is an emerging topic being developed massively this decade in many fields, including the food and fish industries. Image processing and deep learning are the core of its technology and development. And in the fish or aquaculture industry, the development of automation processes using computer and machine vision based on image processing and deep learning is widely used for auto-detection and classification, detection of disease, detection of trajectory, evaluating feeding intensity, analyzing the freshness, and so on [1]. Several approaches have been proposed for fish auto detection and classification in the aquaculture industry based on image processing and deep learning. For example, with a deep convolutional autoencoder [2], CNN (Convolutional Neural Network) [3-9], Faster R-CNN [10, 11], Alex-Net [12, 13], SE-ResNet152 [14], VGGNet [15], SSD (Single Shot Detector) [16-18], and YOLO (You Only Look Once) [19-21].

The detection and classification of disease severity in fish are exemplified by the work [22]. They use SVM (Support Vector Machine) to classify healthy and diseased salmon. For trajectory detection of fish, [23] utilizes Faster R-CNN and time stamping to detect fish and obtain the trajectory record. Meanwhile, [24] proposes a modified YOLO (MSR-YOLO) for this task. To evaluate fish feeding intensity, [25] suggests a two-stage approach. An optical flow neural network is first applied to generate optical flow frames, which are then inputted into a 3D-CNN. To analyze the freshness of fish, [26] presents CNN-based deep learning, which is applied to common carp. Apart from the works that have been briefly discussed, there is still much work related to the application of image processing and deep learning in the fish industry [1]. Due to Thailand imports frozen tuna for many kind of food processing.

However, from all the works that have been published, the authors could not find a classification of fresh and frozen fish based on image processing and deep learning, even though this work might be very useful for the development of automation in the fish industry. While the classification of fresh and frozen fish based on visual (image) has some difficult challenges; visually similar of fresh and frozen fish, random positions and various sizes of fish, etc. And this paper aims to fill that research gap. This work aims to classify fresh and frozen fish based on visual (image) using

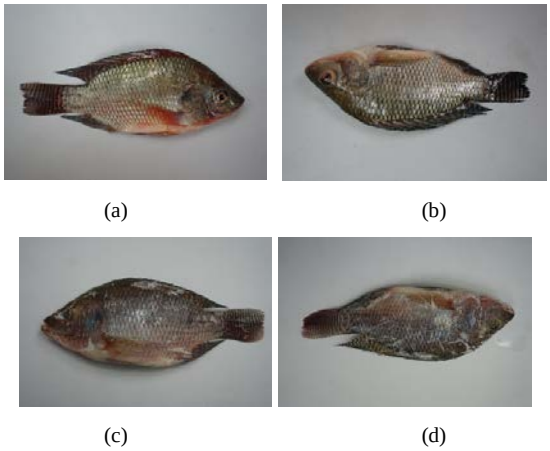
image processing and deep learning. Appropriate image processing techniques are proposed, and effective deep-learning models are employed. The results of this approach are tested on prepared test images, and the model is evaluated for accuracy.

II. Material and Methodology

A. Image Dataset and Augmentation

In this work, we composed our own dataset because there is no publicly available dataset containing images of fresh and frozen fish that the authors could find that could be utilized for this work. We took fresh tilapia *(Oreochromis niloticus)* from traditional markets in Thailand as samples. Tilapia was chosen because this type of fish is one of the most popular for being cultivated and consumed by the people of Thailand and its surroundings. The fish were harvested and sold on the same day we got them.

We brought the fish to the KMITL Food Engineering laboratory for the first photo shoot. Then after three days in the freezer, the fish became frozen, temperature of frozen fish is approximately -35 °C [31]. and the second photo shoot was carried out. Both shots were taken in the same location, conditions and camera. The lighting came from the room lighting with the help of an additional LED light. The measured light intensity was stable throughout the shootings at 846 lux/ 79 FC. The camera used was SONY Model ILCE-7, and the background captured was white and constant (same for every image generated). The fish images are taken under constant conditions (background and lighting) and by the same camera because this work is intended for automation applications in the fish industry, where the environment can be controlled. And finally, we got 50 images of fresh tilapia and 50 of frozen with image sizes of 4,000x6,000 pixels, respectively. Some examples of images of these fish can be seen in Fig. 1.



(a) (b)

(c) (d)

Fig. 1 Sample images of the dataset; (a,b) fresh fish, (c,d) frozen.

Then the augmentation was accomplished. This process aims to meet the minimum number of samples, enrich the data (especially for the training), and make the model more accurate [27]. The techniques used were vertical and horizontal flips, which are suitable for fish image data [28]. After the augmentation, the new images became 200 for fresh fish and 200 for frozen. This number has exceeded the minimum number of samples, which requires 100 images per class [13]. Then the images were randomly split into 70% for training and 30% for testing [13]. Details of this process are summarized in Table I.

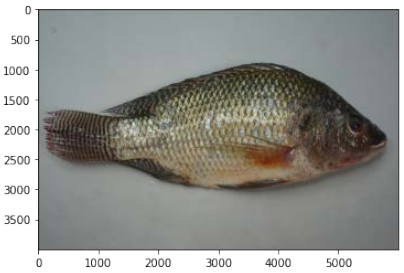
Table I DATASET AND IMAGE AUGMENTATION

Fish class	No. of original images	No. of augmented images (Flip Hz, Flip V, Flip Hz-V)	New Image dataset	For training (70%)	For testing (30%)
Fresh	50	150	200	140	60
Frozen	50	150	200	140	60
Total	100	300	400	280	120

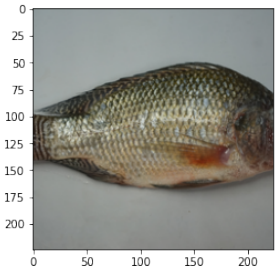
B. Image Pre-processing and Mode

The images were resized to 224x224 pixels to reduce the computations. Lighter computation makes the model work faster and requires low hardware specifications. Resizing also makes the model able to accept various sizes of input images. However, in this process, the important features in the image should not be degraded or lost [12]. So, before this process is carried out, we must first understand the important features in the fresh and frozen fish images that can distinguish them. Those features should not be degraded or lost during the resizing process so that the model can still extract them during training. And if we observe, the fresh fish have a brighter and shiny colour, while the frozen are darker. In addition, some parts are iced in frozen fish, especially the fish's body. These conditions are the most obvious to distinguish between fresh and frozen fish.

For that reason, a unique resizing technique is proposed. It is combined with cropping and Lanczos [29] resampling. This technique produces a new image with a size of 224x224 pixels, and the fish is cropped and focused on the body. The important features are maintained. So, the model could well extract the important distinguishing features between fresh and frozen fish from the fish body because these features are most abundant there and still be kept. In addition, the image also remains in RGB format so that the model can also extract the colour features. The results of this process can be observed in Fig. 2. Then, the images were arrayed and normalized before inputting into the model for the training process. The training process flow is illustrated in Fig. 3 as part of the overview of this entire work.



(a)



(b)

Fig. 2 (a) Input image (6,000x4,000) pixels.
(b) after resizing (224x224) pixels.

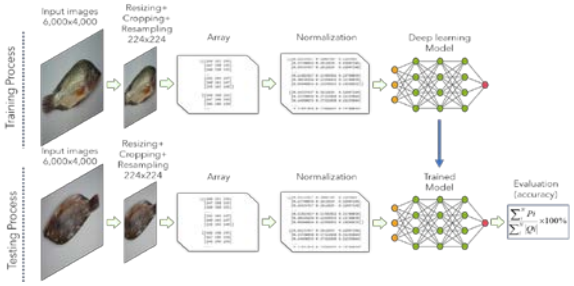


Fig. 3 Workflow (illustration of the entire work).

The model utilized in this work as a deep learning algorithm consists of only two sequential layers. The first layer has an output shape of 1,280 with 410,208 parameters. The second and last layer has an output shape of two because there are only two classes in this work (fresh and frozen fish). This layer has 128,300 parameters. So, the total parameters for this model are 538,508, with details of 524,428 trainable and 14,080 non-trainable. The details or architecture of this model are summarized in Fig. 4. This model was generated and trained involving the GTM (Google Teachable Machine) [30]. The learning rate during training was set at 0.001, and the training process could be achieved well with only 50 epochs with 16 batch sizes.

```
Model: "sequential_4"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 sequential_1 (Sequential)   (None, 1280)              410208

 sequential_3 (Sequential)   (None, 2)                 128300

=================================================================
Total params: 538,508
Trainable params: 524,428
Non-trainable params: 14,080
_________________________________________________________________
```

Fig. 4 Model summay.

$$Accuracy = \frac{\sum_i^N Pi}{\sum_i^N |Qi|} \times 100\% \quad (1)$$

III. RESULTS AND DISCUSSION

A. Experimental Results

The trained model was then utilized to classify the remaining 30% of the images in the dataset (the test data) to evaluate the model. The image was resized and cropped using the same method as the previous process (training), then arrayed and normalized before being inputted into the model. The results of this test were assessed by accuracy by comparing the correct classification with the total data. Equation (1) describes how the accuracy is obtained, where $\sum_i^N Pi$ is the number of correct predictions, and $\sum_i^N |Qi|$ is the total number of predictions [20, 21].

From the test results, 100% accuracy was obtained for both fresh and frozen fish classes. The confidence score was also very high, reaching an average of 99.99% for the fresh fish class and 99.65% for the frozen class. In other words, the proposed approach and the developed model can perfectly answer the challenge of classifying fresh and frozen tilapia in the dataset created. The test results are summarized in table 2, and the illustration of the testing process can be seen in Fig. 3 as part of the overview of this entire work.

Table II EXPERIMENTAL RESULTS

Fish class	Test Images	Correct Classification	Wrong Classification	Accuracy (%)	Averaged Confidence Score (%)
Fresh	60	60	0	100	99.99
Frozen	60	60	0	100	99.65
Total	120	120	0	-	-
Average	-	-	-	100	99.82

B. Limitations and Future Developments

Although this work produces very good results, several concerns have not been observed, and the variety of data is considered limited. For example, the fish object in each image is relatively similar in size and position. For that, this work can be developed in the future. The input data can be

more varied; for example, the fish in the image is simulated to zoom, rotate, and translate. The model performance and overall approach can be evaluated and even improved.

IV. Conclusion

This paper presents a simple but effective approach for classifying fresh or frozen fish images based on image processing and deep learning. We created our own dataset consisting of images of fresh and frozen tilapia. The images were then augmented to enrich and exceed the minimum data. The technique used in image pre-processing is considered very appropriate because it can still maintain the important features, impact on the proposed deep learning model is very simple. However, despite its simplicity, this approach can classify all fresh and frozen tilapia in the dataset with an accuracy score of 100%. The resulting confidence score is also very high, reaching an average of 99.82% for all images in all classes. However, further studies are needed in the future, especially for a more diverse variety of fish image inputs, and the overall model and approach can be further evaluated and improved.

Acknowledgment

This work was fully supported by King Mongkut's Institute of Technology Ladkrabang (KMITL).

References

[1] S. Zhao, S. Zhang, J. Liu, H. Wang, J. Zhu, D. Li, *et al.*, "Application of machine learning in intelligent fish aquaculture: A review," *Aquaculture*, vol. 540, p. 736724, 2021.

[2] A. Banerjee, A. Das, S. Behra, D. Bhattacharjee, N. T. Srinivasan, M. Nasipuri, *et al.*, "Carp-DCAE: Deep convolutional autoencoder for carp fish classification," *Computers and Electronics in Agriculture*, vol. 196, p. 106810, 2022/05/01/ 2022.

[3] S. A. Shammi, S. Das, M. Hasan, and S. R. H. Noori, "FishNet: Fish Classification using Convolutional Neural Network," in *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2021, pp. 1-5.

[4] H. T. Rauf, M. I. U. Lali, S. Zahoor, S. Z. H. Shah, A. U. Rehman, and S. A. C. Bukhari, "Visual features based automated identification of fish species using deep convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 167, p. 105075, 2019/12/01/ 2019.

[5] Z. Zheng, C. Guo, X. Zheng, Z. Yu, W. Wang, H. Zheng, *et al.*, "Fish Recognition from a Vessel Camera Using Deep Convolutional Neural Network and Data Augmentation," in *2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, 2018, pp. 1-5.

[6] M. A. Iqbal, Z. Wang, Z. A. Ali, and S. Riaz, "Automatic Fish Species Classification Using Deep Convolutional Neural Networks," *Wireless Personal Communications*, vol. 116, pp. 1043-1053, 2021/01/01 2021.

[7] R. B S, D. S. G N, S. Reddy, D. Kakwani, and N. Bhattad, "Fish detection and classification using convolutional neural networks," ed, 2020, pp. 1221-1231.

[8] A. A. Dos Santos and W. N. Gonçalves, "Improving Pantanal fish species recognition through taxonomic ranks in convolutional neural networks," *Ecological Informatics*, vol. 53, p. 100977, 2019/09/01/ 2019.

[9] T. Miyazono and T. Saitoh, "Fish Species Recognition Based on CNN Using Annotated Image," 2018.

[10] M. A. Rosales, M. G. B. Palconit, V. J. D. Almero, R. S. Concepcion, J.-A. V. Magsumbol, E. Sybingco, *et al.*, "Faster R-CNN based Fish Detector for Smart Aquaculture System," in *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, 2021, pp. 1-6.

[11] Y. Adiwinata, A. Sasaoka, I. A. Bayupati, and O. Sudana, "Fish species recognition with Faster R-CNN Inception-v2 using QUT FISH dataset," *Lontar Komputer: Jurnal Ilmiah Teknologi Informasi*, vol. 11, pp. 144-154, 2020.

[12] Z. Ju and Y. Xue, "Fish species recognition using an improved AlexNet model," *Optik*, vol. 223, p. 165499, 2020/12/01/ 2020.

[13] . N.S, S. D, and R. K. S, "Naive Bayesian fusion based deep learning networks for multisegmented classification of fishes in aquaculture industries," *Ecological Informatics*, vol. 61, p. 101248, 2021/03/01/ 2021.

[14] X. Xu, W. Li, and Q. Duan, "Transfer learning and SE-ResNet152 networks-based for small-scale unbalanced fish species identification," *Computers and Electronics in Agriculture*, vol. 180, p. 105878, 2021/01/01/ 2021.

[15] Prasetyo, N. Suciati, and C. Fatichah, "Multi-level residual network VGGNet for fish species classification," *Journal of King Saud University - Computer and Information Sciences*, 2021/06/05/ 2021.

[16] G. Tian, D. Li, W. Li, L. Zhang, H. Zhang, and Q. Duan, "A detection method of the turned white belly fish based on improved SSD," *Journal of Physics: Conference Series*, vol. 1856, p. 012035, 2021/04/01 2021.

[17] G. Yu, L. Wang, M. Hou, Y. Liang, and T. He, "An adaptive dead fish detection approach using SSD-MobileNet," in *2020 Chinese Automation Congress (CAC)*, 2020, pp. 1973-1979.

[18] P. D. Hung and N. N. Kien, "SSD-Mobilenet Implementation for Classifying Fish Species," in *Intelligent Computing and Optimization*, Cham, 2020, pp. 399-408.

[19] K. Cai, X. Miao, W. Wang, H. Pang, Y. Liu, and J. Song, "A modified YOLOv3 model for fish detection based on MobileNetv1 as backbone," *Aquacultural Engineering*, vol. 91, p. 102117, 2020/11/01/ 2020.

[20] . Kuswantori, T. Suesut, W. Tangsrirat, and N. Nunak, "Development of object detection and classification with YOLOv4 for similar and structural deformed fish," *EUREKA: Physics and Engineering*, vol. 2, pp. 154–165, 2022a.

[21] A. Kuswantori, T. Suesut, W. Tangsrirat, and S. Satthamsakul, "Fish Recognition Optimization in Various Backgrounds Using Landmarking Technique and YOLOv4,"

presented at the. The 37th International Technical Conference on Circuits/Systems, Computers, and Communications (ITC-CSCC), Phuket, Thailand, 2022b.
[22] M. S. Ahmed, T. T. Aurpa, and M. A. K. Azad, "Fish Disease Detection Using Image Based Machine Learning Technique in Aquaculture," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, pp. 5170-5182, 2021.
[23] W. Xu, Z. Zhu, F. Ge, Z. Han, and J. Li, "Analysis of Behavior Trajectory Based on Deep Learning in Ammonia Environment for Fish," *Sensors*, vol. 20, p. 4425, 2020.
[24] H. E.-D. Mohamed, A. Fadl, O. Anas, Y. Wageeh, N. ElMasry, A. Nabil, *et al.*, "MSR-YOLO: Method to Enhance Fish Detection and Tracking in Fish Farms," *Procedia Computer Science*, vol. 170, pp. 539-546, 2020.
[25] N. Ubina, S.-C. Cheng, C.-C. Chang, and H.-Y. Chen, "Evaluating fish feeding intensity in aquaculture with convolutional neural networks," *Aquacultural Engineering*, vol. 94, p. 102178, 2021/08/01/ 2021.
[26] A. Taheri-Garavand, A. Nasiri, A. Banan, and Y.-D. Zhang, "Smart deep learning-based approach for non-destructive freshness diagnosis of common carp fish," *Journal of Food Engineering*, vol. 278, pp. 109930, 2020/08/01/ 2020.
[27] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *arXiv*, 2017.
[28] Liu, X. Jia, and X. Xu, "Study of shrimp recognition methods using smart networks," *Computers and Electronics in Agriculture*, vol. 165, pp. 104926, 2019/10/01/ 2019.
[29] B. Madhukar and R. Narendra, "Lanczos resampling for the digital processing of remotely sensed images," in *Proceedings of International Conference on VLSI, Communication, Advanced Devices, Signals & Systems and Networking (VCASAN-2013)*, 2013, pp. 403-411.
[30] M. Carney, B. Webster, I. Alvarado, K. Phillips, N. Howell, J. Griffith, *et al.*, "Teachable machine: Approachable Web-based tool for exploring machine learning classification," in *Extended abstracts of the 2020 CHI conference on human factors in computing systems*, 2020, pp. 1-8.
[31] I. Tolstorebrov, T.M. Eikevik, M. Bantle,"*Effect of low and ultra-low temperature applications during freezing and frozen storage on quality parameters for fish"*, vol.63, 2016, pp.37-47.

Development of Hemagglutination Classification Using Machine Learning Techniques Based on Proposed Shadow Augmentation

Aneesah Satoh[1, a)], Somyot Chirasatitsin[1, b)] and Jermphiphut Jaruenpunyasak[1, c)]

[1]Department of Biomedical Science and Biomedical Engineering, Faculty of Medicine,
Prince of Songkla University, Hat Yai, Songkhla 90110, Thailand

[a)] anisah44919@gmail.com
[b)]somjot.c@psu.ac.th
[c)]Corresponding author: jjermphi@medicine.psu.ac.th

Abstract. The utilization of microfluidic technology in blood typing has shown great potential owing to its high sensitivity, rapidity, and minimal sample volume demands. Additionally, in an emergency situation, this system can be utilized for real-time blood typing measurements. However, when employing this method, it remains challenging for non-experts to observe blood agglutination or hemagglutination on the device with the naked eye. Therefore, computer vision with machine learning is required to assist in observing hemagglutination classification to overcome the difficulty associated with observing microfluidic devices. Nonetheless, the performance of hemagglutination classification is negatively impacted by human and environmental artifacts during image capture such as orientation, blur, and shadow from the capture device such a mobile phone. This research studies machine learning techniques to classify hemagglutination. It also proposes shadow augmentation and Grid HSV (Hue, Saturation, and Value) for feature extraction to simulate the human and environment artifact using a mobile phone capturing. Various percentages of the proposed shadow augmentation techniques are applied to the dataset in this experiment. The models are trained utilizing three main classifiers: SVM (Support Vector Machine), kNN (k-Nearest Neighbors), and MLP (Multi-Layer Perceptron). The results indicated that the Grid HSV features with the MLP classifier outperformed with a significant level of accuracy (≥ 0.999). Finally, for various human and environmental artifacts in dataset, this Grid HSV features based on the MLP classifier is feasible to maintain a high accuracy for hemagglutination classification. This system has the potential to be utilized for a mobile application-based online hemagglutination classification.

***Index Terms*— Microfluidic channel, Hemagglutination, Shadow augmentation, Grid HSV.**

I. INTRODUCTION

Road traffic injuries present not only immediate physical hazards but also intricate healthcare complications, including the requirement for blood transfusions to treat severe injuries. In addition to mitigating the effects of these injuries, this responsibility requires a comprehensive strategy to address the complications of blood transfusion, including but not limited to blood shortages, storage and transportation issues, and compatibility and matching problems [1], [2], [3]. In case of compatibility and matching problems. Typically, slide tests are employed in hospitals for blood type screening due to their cost-effectiveness and speed (e.g., 5-10 min/test) [4]. However, they have a low sensitivity to weak antigens, which diminishes their utility. As a result of their greater sensitivity than slide tests [5], tube and gel column tests are considered gold standard methods for blood typing. However, they are costly and time-consuming (e.g., 30-45 min/test). Moreover, this process might not be suitable for a large number of patients at once; furthermore, the device is barely installed in an ambulant car.

Presently, innovative blood typing techniques, including solid phase and microfluidic chip methods, have been developed to address the issues that plagued laboratory blood testing in the past. In the case of solid phase, lateral flow techniques are used to observe blood agglutination on the surface. This expedites the testing process, but it requires a buffer wash, which may not be adequate for observing blood agglutination [6]. In terms of microfluidic chip, it employs capillary action to determine the blood typing reaction, as opposed to washing with buffer in the solid phase process. It has additional benefits including compact size, low sample volume, low cost, and efficacy [7]. However, this technique still has the drawback of being difficult to observe with the naked eye or for non-experts to discern blood agglutination on this device.

To overcome the challenge associated with observing the microfluidic device, a computer vision is an interdisciplinary domain encompassing computer science and artificial intelligence, with the primary objective of empowering computers to comprehend and interpret the visual environment. It requires the development of algorithms and methods that enable the extraction of information from digital images and videos by computers [8]. Image classification is also a fundamental undertaking in the field of computer vision. It entails the allocation of a label or category to an image in accordance with its contents. Regarding hemagglutination classification, it is proposed to be a tool to monitor and assist the positive and negative hemagglutination. However, variations in light, shadow, and angle can diminish the accuracy of classification [9].

This paper studies the classification of hemagglutination

utilizing machine learning methods and the proposed shadow augmentation. Various percentages of the proposed shadow augmentation techniques are applied to the dataset in this experiment. Additionally, two types of features are utilized: HSV (Hue, Saturation, and Value) and Grid HSV feature extraction. Three classifiers are used to train the models: SVM (Support Vector Machine), kNN (k-Nearest Neighbors), and MLP (Multi-Layer Perceptron). The results are assessed with respect to their accuracy and standard deviation.

II. METHODOLOGY

This section provides an overview of the dataset employed in the research, including details on data preprocessing, CNN models, and our experimental evaluation. The training of a channel classification consists of four fundamental processes, as illustrated in Fig. 1. dataset, data preparation, proposed shadow augmentation, CNN model, and its evaluation.

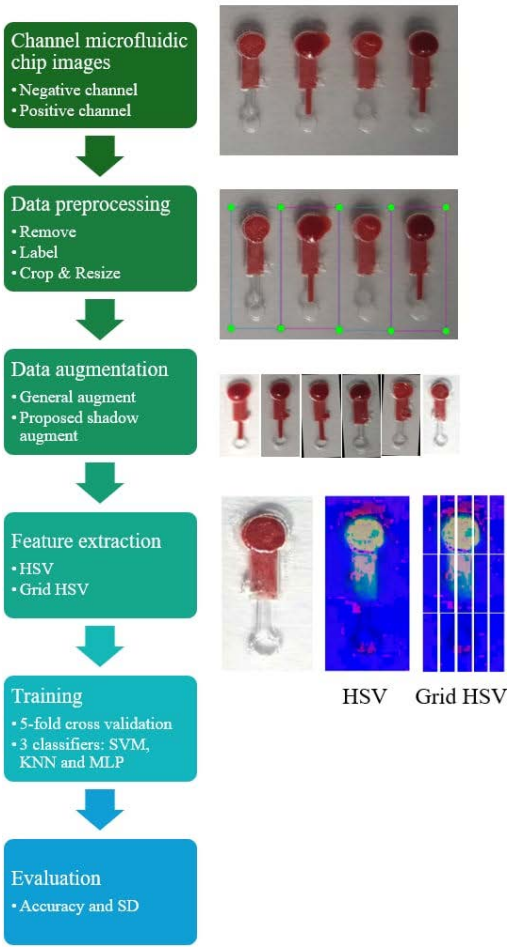


Fig. 1. Flowchart of this study.

A. Channel microfluidic chip images

To simulate the shadow artifact issues investigated in this study, we acquired channel microfluidic chip images from the microfluidic laboratory and assigned labels to one hundred and twelve of the acquired images to represent negative (N=56) and positive (N=56) channels, respectively. Fig. 2. provides an illustration of the configuration. With a high resolution of 3000x4000, the chip is captured in the laboratory under room-luminous conditions using a mobile phone.

Fig. 2. The configuration of capturing the microfluidic chip: (A) White light color, (B) Camera capture by researcher using a mobile phone with resolution of 3000x4000, (C) Tripod light with 60W 5400K (D) Microfluidic chip, (E) Standard red cells solution according to the Thai Red Cross Society.

B. Data preprocessing

In the process of data preprocessing for channel microfluidic chip image datasets, Experts carefully removed unclear images or bad-quality data such as shadow and blur. Then they labelled the dataset on positive and negative channel of microfluidic chip by using LabelImg [16]. Finally, all images were resized by 128x256.

C. Data augmentation

There are two main processes of data augmentation: general augmentation and proposed shadow augmentation. In the initial stages, the images undergo general augmentation techniques, including sharpening, brightness, and rotation, Gaussian blur, and vertical flipping modifications, as detailed in Table I. Then, for each dataset that had undergone general augmentation, shadow augmentation using (1) was applied in varying percentages in three datasets. In this experiment, negative (N=20,000) and positive (N=20,000) samples were selected, and shadow augmentation of 0%, 50%, and 100% was then proposed. Finally, this investigation comprised three main datasets. The examples of datasets are shown in Fig. 3. and Fig. 4.

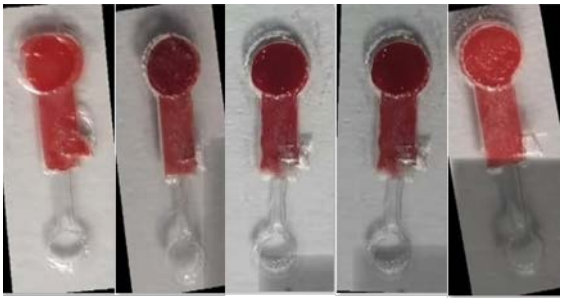


Fig. 3. Example of positive images of hemagglutination with proposed shadow augmentation.

Fig. 4. Example of negative images of hemagglutination with proposed shadow augmentation.

TABLE I GENERAL AUGMENTATION TECHNIQUES AND THEIR PARAMETERS

Augmentation	Parameters
Rotation	Degree: -10° and 10°, increased by 1°
Gaussian Blur	Radius: 1 to 5
Increasing Brightness	Pixel size: 1 to 5
Decreasing Brightness	Pixel size: 1 to 5
Increasing Sharpness	Pixel size: 1 to 5
Decreasing Sharpness	Pixel size: 1 to 5
Vertical Flip	Degree: 180°

$$SI(x,y,c) = OI(x,y,c) \times S(x,y) \quad (1)$$

Here, SI(x,y,c) is a shadowed image, x is the x-axis coordination, y is the y-axis coordination, c is the color channel index, OI(x,y,c) is an original image, and S(x,y) is the proposed shadow augmented function that consists of five main processes: generate shadow mask as rectangle shape to simulate a mobile phone shadow, apply Gaussian blur to create soft edges, normalize mask, generate random intensity, and apply shadow intensity with glow effect using Gaussian blur.

D. Feature extraction

Two feature extraction methods were contrasted in this experiment: HSV (Hue, Saturation, and Value) and Grid HSV. First of all, HSV feature is a widely employed color space in the field of digital image processing, providing a graphical depiction of colors that includes their hue, saturation, and brightness. Hue, as it pertains to the type of color, signifies the wavelength at which it is most prominent on the color wheel. Saturation measures the intensity or purity of a color and ranges from faint variations of gray to vibrant hues. Lastly, value corresponds to the brightness of the color, spanning from black to white. This experiment modified HSV as the Grid HSV feature by employing a grid pattern consisting of 3x5 windows, which corresponded to the feature of each grid in the image.

E. Training process

Following feature extraction, the dataset underwent training for hemagglutination classification using five-fold cross-validation. Additionally, a comparison was made between three distinct classifier types: SVM, kNN, and MLP. To begin with, the SVM classifier is a highly effective classification algorithm. SVM determines, given a labeled dataset, the hyperplane that maximizes the margin between classes to best separate them. Additionally, the kNN classifier is a machine learning algorithm that is both simple and efficient when applied to classification tasks. All training data is committed to memory by the model during the training phase. The kNN method selects the k-nearest neighbors in accordance with the distance metric computed between an unknown point and every other point in the training set when making a class classification for the unknown data point. In order to designate the class label to the unknown data point for classification tasks, a majority vote is conducted among the k-nearest neighbors. The last classifier is the MLP. It is a neural network architecture that consists of numerous layers of interconnected nodes and is specifically engineered to handle classification tasks. An input layer, one or more concealed layers, and an output layer comprise these networks. Different values of configured parameters were assigned to each of the three classifiers, as shown in Table II.

F. Evaluation

In order to assess the performance of each classifier in the hemagglutination classification result, the average classification accuracy was calculated utilizing a binary class confusion matrix. Furthermore, a comparison was made between the standard deviation (SD) and 5-fold cross-validation for the purpose of hemagglutination classification.

TABLE II CLASSIFICATION MODELS AND THEIR PARAMETERS

Classifier	Parameters	Value
SVM	Kernel function	linear, polynomial (degree=3), and radial basis function
kNN	Distance equation	Euclidean, Manhattan, and Minkowski distance
	number of k	3, 5, 7, and 9
MLP	number of hidden layers	1
	number of nodes in hidden layer	8, 16, 32, 64, and 128

III. Results and Discussion

This section describes the two main evaluation criteria utilized in the classification of hemagglutination: the averaged accuracy and standard deviation obtained from the extraction of two main features across three distinct datasets.

First of all, Table III-V show the performance of SVM, kNN, and MLP when the HSV feature is utilized for hemagglutination classification. This experiment classified three distinct datasets using various classifier parameter values. It is evident that the SVM and kNN classifiers achieved an average accuracy greater than 0.80, while the MLP, employing a number of node in a hidden layer exceeding 8, attained the maximum average accuracy than 0.87. However, when general augmentation and proposed shadow are applied to a variety of human and environmental artifacts, the performance of kNN classifiers decreases by approximately 0.05. In case of standard deviation, its values for these classifiers varied from 0.001 to 0.005, as shown in Table III-V.

In term of Grid HSV, Table III-V illustrate the performance of SVM, kNN, and MLP in the context of hemagglutination classification. It is obvious that all classifiers achieved an average accuracy greater than 0.99. In particular, the MLP classifier exhibits optimal performance across three distinct datasets when the number of nodes in the hidden layer is appropriately parameterized. Moreover, the SD values based on MLP classifiers for three datasets varied between 0.000 and 0.003, as shown in Table III-V.

The results presented in Table III-V indicate that the performance of each classifier improved from 0.81 to 0.94 for the HSV feature across all three datasets. It appears that this HSV feature is incapable of capturing image features associated with hemagglutination in the training set for various classifier parameters. Furthermore, a slight reduction of 0.05 is observed in the performance of the kNN classifier when 100% of the proposed shadow augmentation is added. Regarding the Grid HSV feature, the accuracy of all three classifiers in this study is above 0.999 across all datasets. The Grid HSV feature with three classifiers appears to be unaffected by a variety of environmental and human artifacts when general augmentation and proposed shadow augmentation are utilized. In particular, when only eight nodes are utilized in the hidden layer to classify hemagglutination, the MLP classifier outperforms other models. It appears that this small MLP classifier model could potentially be utilized for the classification of hemagglutination in real time.

IV. Conclusion

This study explores hemagglutination classification using machine learning. For mobile phone-captured human and environment artifact simulation, it proposed shadow augmentation and Grid HSV feature extraction. According to the results, the Grid HSV feature employing a conventional classifier demonstrates superior performance across all datasets when general and shadow augmentation is implemented. Moreover, using only 8 nodes of single hidden layer in MLP classifier could be suitable for monitoring in real time for hemagglutination classification. We intend to conduct further experiments involving additional human and environmental artifacts, including noise and flare light effects. In addition, the model will be customized to be compatible for mobile applications.

Acknowledgment

This work is supported by the graduate scholarship by Faculty of Medicine, and research grant for thesis by graduate school, Prince of Songkla University. In addition, the authors would like to thank to “AI and Robotics for All - Super AI Engineer” project organized by the Artificial Intelligence Association of Thailand.

TABLE III THE PERFORMANCE OF SVM, KNN, AND MLP
FOR HEMAGGLUTINATION CLASSIFICATION WITH 0% OF PROPOSED SHADOW

Classifier	Parameters		Feature extraction			
			HSV		Grid HSV	
			Accuracy	SD	Accuracy	SD
SVM	Kernel					
	Linear		0.80823	0.00237	1.00000	0.00000
	Polynomial (degree=3)		0.82285	0.00346	0.99985	0.00018
	Radial basis function		0.87200	0.00173	0.99999	0.00005
kNN	k	Distance				
	3	Euclidean	0.93985	0.00290	0.99972	0.00015
		Manhattan	0.93535	0.00427	0.99957	0.00020
		Minkowski	0.93683	0.00389	0.99962	0.00016
	5	Euclidean	0.92865	0.00462	0.99907	0.00028
		Manhattan	0.93855	0.00298	0.99972	0.00020
		Minkowski	0.93985	0.00290	0.99972	0.00015
	7	Euclidean	0.93228	0.00505	0.99942	0.00026
		Manhattan	0.94205	0.00272	0.99972	0.00015
		Minkowski	0.92865	0.00462	0.99962	0.00028
	9	Euclidean	0.93683	0.00389	0.99962	0.00016
		Manhattan	0.93157	0.00492	0.99960	0.00015
		Minkowski	0.93228	0.00505	0.99942	0.00026
MLP	Number of nodes in a hidden layer					
	8		0.87385	0.00417	1.00000	0.00000
	16		0.87980	0.00417	0.99997	0.00005
	32		0.88372	0.00438	1.00000	0.00000
	64		0.89033	0.00155	1.00000	0.00000
	128		0.89223	0.00355	1.00000	0.00000

TABLE IV THE PERFORMANCE OF SVM, KNN, AND MLP
FOR HEMAGGLUTINATION CLASSIFICATION WITH 50% OF PROPOSED SHADOW

Classifier	Parameters		Feature extraction			
			HSV		Grid HSV	
			Accuracy	SD	Accuracy	SD
SVM	Kernel					
	Linear		0.81335	0.00589	0.99983	0.00015
	Polynomial (degree=3)		0.82758	0.00183	0.99990	0.00009
	Radial basis function		0.87300	0.00305	0.99997	0.00005
kNN	k	Distance				
	3	Euclidean	0.91483	0.00196	0.99948	0.00015
		Manhattan	0.91358	0.00157	0.99950	0.00019
		Minkowski	0.91375	0.00203	0.99927	0.00009
	5	Euclidean	0.91280	0.00278	0.99905	0.00013
		Manhattan	0.91565	0.00109	0.99960	0.00015
		Minkowski	0.91483	0.00196	0.99948	0.00015
	7	Euclidean	0.91337	0.00204	0.99915	0.00030
		Manhattan	0.91585	0.00156	0.99965	0.00020
		Minkowski	0.91280	0.00278	0.99905	0.00013
	9	Euclidean	0.91375	0.00203	0.99927	0.00009
		Manhattan	0.91310	0.00174	0.99945	0.00024
		Minkowski	0.91337	0.00204	0.99915	0.00030
MLP	Number of nodes in a hidden layer					
	8		0.87203	0.00562	1.00000	0.00000
	16		0.87895	0.00235	0.99998	0.00005
	32		0.88175	0.00322	0.99997	0.00005
	64		0.88675	0.00259	1.00000	0.00000
	128		0.89015	0.00194	1.00000	0.00000

TABLE V THE PERFORMANCE OF SVM, KNN, AND MLP
FOR HEMAGGLUTINATION CLASSIFICATION WITH 100% OF PROPOSED SHADOW

Classifier	Parameters		Feature extraction			
			HSV		Grid HSV	
			Accuracy	SD	Accuracy	SD
SVM	Kernel					
	Linear		0.81158	0.00423	0.99953	0.00030
	Polynomial (degree=3)		0.82643	0.00279	0.99973	0.00018
	Radial basis function		0.86845	0.00130	0.99997	0.00005
kNN	k	Distance				
	3	Euclidean	0.87607	0.00343	0.99925	0.00033
		Manhattan	0.88110	0.00244	0.99948	0.00027
		Minkowski	0.87938	0.00231	0.99910	0.00031
	5	Euclidean	0.88503	0.00287	0.99878	0.00029
		Manhattan	0.87985	0.00254	0.99948	0.00017
		Minkowski	0.87607	0.00343	0.99925	0.00033
	7	Euclidean	0.88195	0.00157	0.99903	0.00018
		Manhattan	0.87703	0.00263	0.99960	0.00015
		Minkowski	0.88503	0.00287	0.99878	0.00029
	9	Euclidean	0.87938	0.00231	0.99910	0.00031
		Manhattan	0.88360	0.00243	0.99950	0.00026
		Minkowski	0.88195	0.00157	0.99903	0.00018
MLP	Number of nodes in a hidden layer					
	8		0.86708	0.00281	0.99980	0.00013
	16		0.87465	0.00374	0.99990	0.00015
	32		0.87833	0.00290	0.99993	0.00010
	64		0.88037	0.00338	0.99995	0.00006
	128		0.88265	0.00200	0.99997	0.00005

REFERENCES

[1] C.-Y. Huang et al., "Characteristics and Outcomes of Patients Injured in Road Traffic Crashes and Transported by Emergency Medical Services," Int. J. Environ. Res. Public. Health, vol. 13, no. 2, Art. no. 2, Feb. 2016, doi: 10.3390/ijerph13020236.

[2] A. A. Alqarni et al., "Injuries type and its relation with Glasgow Coma Scale, injury severity score and blood transfusion in road traffic accident Victims.," Medicni Perspekt., vol. 23, no. 2, Art. no. 2, 2018, doi: 10.26641/2307-0404.2018.2.133954.

[3] H. Saidi, B. K. Mutiso, and J. Ogengo, "Mortality after road traffic crashes in a system with limited trauma data capability," J. Trauma Manag. Outcomes, vol. 8, no. 1, p. 4, Feb. 2014, doi: 10.1186/1752-2897-8-4.

[4] W. Malomgré and B. Neumeister, "Recent and future trends in blood group typing," Anal. Bioanal. Chem., vol. 393, no. 5, pp. 1443–1451, Mar. 2009, doi: 10.1007/s00216-008-2411-3.

[5] E. Spada, R. Perego, L. Baggiani, and D. Proverbio, "Comparison of Conventional Tube and Gel-Based Agglutination Tests for AB System Blood Typing in Cat," Front. Vet. Sci., vol. 7, p. 312, Jun. 2020, doi: 10.3389/fvets.2020.00312.

[6] T. Songjaroen et al., "A simple and low-cost portable paper-based ABO blood typing device for point-of-care testing," J. Immunoassay Immunochem., vol. 39, no. 3, pp. 292–307, May 2018, doi: 10.1080/15321819.2018.1486856.

[7] Y. Xie, L. Dai, and Y. Yang, "Microfluidic technology and its application in the point-of-care testing field," Biosens. Bioelectron. X, vol. 10, p. 100109, May 2022, doi: 10.1016/j.biosx.2022.100109.

[8] S. Zhou, B. Chen, E. S. Fu, and H. Yan, "Computer vision meets microfluidics: a label-free method for high-throughput cell analysis," Microsyst. Nanoeng., vol. 9, no. 1, pp. 1–15, Sep. 2023, doi: 10.1038/s41378-023-00562-8.

[9] P. Palevičius, M. Pal, M. Landauskas, U. Orinaitė, I. Timofejeva, and M. Ragulskis, "Automatic Detection of Cracks on Concrete Surfaces in the Presence of Shadows," Sensors, vol. 22, no. 10, Art. no. 10, Jan. 2022, doi: 10.3390/s22103662.

[10] N. Su, Y. Zhang, S. Tian, Y. Yan, and X. Miao, "Shadow Detection and Removal for Occluded Object Information Recovery in Urban High-Resolution Panchromatic Satellite Images," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 9, no. 6, pp. 2568–2582, Jun. 2016, doi: 10.1109/JSTARS.2016.2570234.

[11] H. Fan, M. Han, and J. Li, "Image Shadow Removal Using End-To-End Deep Convolutional Neural Networks," Appl. Sci., vol. 9, no. 5, Art. no. 5, Jan. 2019, doi: 10.3390/app9051009.

[12] N. Inoue and T. Yamasaki, "Learning From Synthetic Shadows for Shadow Detection and Removal," IEEE Trans. Circuits Syst. Video Technol., vol. 31, no. 11, pp. 4187–4197, Nov. 2021, doi: 10.1109/TCSVT.2020.3047977.

[13] S. Bang, F. Baek, S. Park, W. Kim, and H. Kim, "Image augmentation to improve construction resource detection using generative adversarial networks, cut-and-paste, and image transformation techniques," Autom. Constr., vol. 115, p. 103198, Jul. 2020, doi: 10.1016/j.autcon.2020.103198.

[14] S. Park, Y. Choi, and H. Hwang, "SACuP: Sonar Image Augmentation with Cut and Paste Based DataBank for Semantic Segmentation," Remote Sens., vol. 15, no. 21, Art. no. 21, Jan. 2023, doi: 10.3390/rs15215185.

[15] G. Li, L. Wen, Z. Huang, R. Xia, and Y. Pang, "Data augmentation and shadow image classification for shadow detection," IET Image Process., vol. 16, no. 3, pp. 717–728, 2022, doi: 10.1049/ipr2.12377.

[16] Tzutalin. LabelImg. *Git code* (2015). Accessed: Feb. 14, 2024. [Online]. Available: https://github.com/tzutalin/labelImg

Hydroxyapatite mixed with polycaprolactone and carboxymethyl cellulose sodium salt crosslinked with acetic acid for wound dressing application

Orranat Eiangmee[1,a)], Jatupat Maikaew[2], Naruebodee Srisang[3] and Siriwan Srisang[3,b)]

[1] Student in Master degree of Mechanical Engineering, King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon, 86160, Thailand
[2] Student in Doctoral degree of Mechanical Engineering, King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon, 86160, Thailand
[3] Department of Engineering, King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon, 86160, Thailand

Author Emails
[1)] orranatt.0701@gmail.com
[b)] Corresponding author: siriwan.sr@kmitl.ac.th

Abstract: Wound dressings are used to prevent the germs from entering the wound. The appropriate properties of the wound dressing such as absorbing blood well and excellent mechanical strength are considered. This research aims to use the polycaprolactone (PCL) and carboxymethyl cellulose sodium salt (CMC) via the hydroxyapatite (HA) for forming the wound dressing sample. This work varied the optimum of HA concentration at 15, 20, and 25%wt and then investigation the properties of sample in part of the mechanical properties (tensile strength and elongation at break), and the water absorption in three liquid solution as deionized water (DI water), sodium chloride solution (NaCl), and phosphate-buffered saline (PBS). Results showed that the PCL/CMC can mix with HA which presented the optimal condition of HA at 20%wt. This condition provided the strength value as 6.05 ± 0.6 MPa, elongation at break at 34.3 ± 9.34% and the water absorption values in PBS, DI water, and NaCl solution were 22.47 ± 3.65, 77.16 ± 1.48, and 92.66 ± 3.85%, respectively. All above results indicated the excellent potential of biopolymer and HA to produce the biomedical device which can be further studied in term of biocompatibility and antibacterial activity.

***Keywords*—Wound dressing, Polycaprolactone (PCL), Carboxymethyl cellulose sodium salt (CMC), Hydroxyapatite (HA), Tensile strength, Water absorption**

I. INTRODUCTION

Wounds divide into two main types: closed wounds and open wounds. Closed wounds are the tissue with no tearing under the skin, the bruises, and hematoma. While, in the case of an open wound, It appears the separation of the skin then exposes the blood. Usually, when the body is injured, the body has a mechanism for healing itself naturally (Phase of wound healing) [1]. Self-healing depends on the severity of each wound. The uncared of wound dressing concerned due to result in serious infections ranging from mild to severe. Currently, the materials to produce the wound dressing made from biopolymers for biocompatibility. The advantages of wound dressing product are the increasing moisturization , antimicrobial, and biocompatibility [2]. The popular biopolymer use to produce the wound dressing was poly (lactic acid), polycaprolactone, and poly (lactic-co glycolic acid).

Polycaprolactone (PCL) is semicrystalline and has been widely used in many industries, especially in tissue engineering and biological materials, which approved by the Food and Drug Administration (FDA) [3, 4]. The PCL is an aliphatic polyester with hydrophobic properties, strength, and non-toxic properties [5]. In addition, various research reported the carboxymethyl cellulose sodium salt (CMC) which used in the variety of industries such as the pharmaceutical, food industries, paper, textile, explosive materials industries, and biomedical materials [6]. The good point of CMC has flexible, hydrophilic, and biocompatible properties including non cytotoxicity with cell [7]. Hence, the combination between PCL and CMC can formulate to the production of wound dressing via using the crosslink agent as acetic acid that reported by reference [8]. Moreover, the reinforcement with hydroxyapatite (HA) can also generate the strength of sample. HA is widely composed in the bone tissue engineering due to it is inorganic components, biodegradation, good cytotoxicity, biocompatitibility, and improving the mechanical properties of materials [9].

Hence, The aim of this work was to study the suitable of HA concentration into the combination between PCL and CMC for production the wound dressing sample, which investigation the physical and mechanical properties in part

of the absorption efficiency and tensile strength.

Methodology

A. Material

Polycaprolactone (PCL) (the number-average molecular weight of 80,000), hydroxyapatite (HA) with a powder particle size of 10 µm, acetone as a solvent agent, and acetic acid as a crosslink agent were purchased from Sigma-Aldrich (Damstadt, Germany). Carboxymethyl cellulose sodium salt (CMC) was purchased from Tokyo Chemical Industry co., ltd. (Toshima, Japan). Deionized water (DI water) was purchased from Manufactured by lee Huad ltd.,part. (Thailand).

B. Method

The concentration of PCL is 10%w/v, CMC is 4%w/v, and acetic acid is 17%v/v were fixed in the mixture [11-13]. In this work, the difference concentrations of HA at 15, 20, and 25%wt were varied. PCL was dissolved in acetone by the hot plate at temperature of a 60 ºC, stirrer at 200 rpm for 1 hr. Then added the CMC, HA and acetic acid were mixed into the homogenous for 45 min. Finally, the mixtures were poured into the chamber and were set for 24 hr.

A. Water absorption

The samples were tested the water absorption as following the standard ASTM D-570. The size of wound dressing was 2x2 mm for three set. The sample was weighed (dry weight; w_0) and was soaked in the liquid of the deionized water (DI water), sodium chloride solution (NaCl), and phosphate-buffered saline (PBS) for 1 day. After that, the sample was weighed (wet weight; w_1) then calculation the percentage of water absortion according to the equation (1) [15].

$$\%\,WA = \frac{(w_1 - w_0)}{w_0} \times 100 \qquad (1)$$

%WA is per cent water absorption
w_1 is Wet weight
w_0 is Dry weight

B. Tensile strength

The samples were tested the tensile strength as following the standard ASTM D-882. The samples were test by the Universal Testing Machine, Bravo. The size of wound dressing sample was 5x70 mm. The head speed was at 10 mm/min, and the gage length at 50 mm [16].

II. Results and Discussion

Fig. 1 shows the wound dressing sample made from PCL/CMC and combined with the HA at different concentrations. The appearance sample of PCL/CMC and HA demonstrated the color as milk-white that corresponed with the Meedecha, Srisang [16]. Meedecha, Srisang [16] reported the production of wound dressing from the PCL/CMC by crosslink method. In this work, the various of the HA concentration was studied. Result shows the homogenous of sample due to the white color of HA and become a milk-white color [17]. However, the sample was evaluated the properties in part of the physical and mechanical properties.

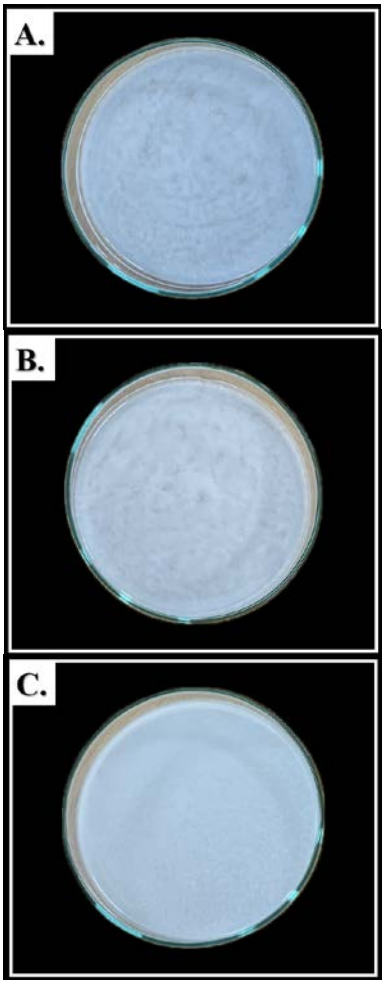


Fig. 1. Wound dressing products from PCL/CMC/and HA at A) HA15%wt, B) HA 20%wt, and C) HA 25%wt

A. water absorption

Fig. 2 shows the water absorption of sample in various of solution as DI water, NaCl, and PBS. The concentration of HA at 25%wt presented the highest the water absorption due to the HA composition of the calcium phosphate group. This increasing of HA can increase the pores in the matraix of the sample. Therefore, the efficiency of water absorption decreases [18]. While the concentration of HA at 20%wt presented the suitable the percentage of water absorption in every solution. This is different from the control, which has hydrophilicity caused by CMC. The result of water absorption in DI water showed the highest range at HA 25%wt, and the lowest range at HA 20%wt at values of 97.19 and 77.16%, respectively. It was caused by the increased amount of HA resulting to the increased pore space. The water absorption of NaCl is the same as the DI water. However, both are very different from PBS because NaCl contains minerals, which make the sample absorbtion the

water better [20]. Whereas the PBS has properties similar to human blood, it may be suitable for degradation testing [15].

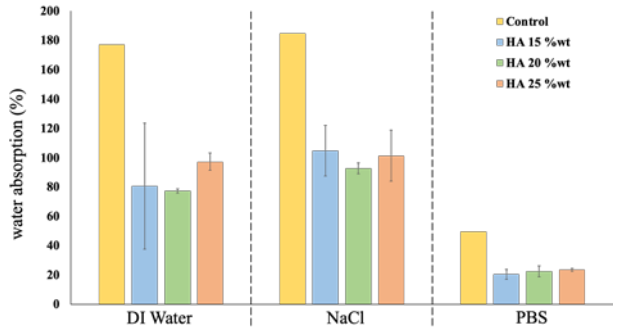


Fig. 2. Water absorption of the wound dressing sample in each condition.

B. Tensile strength

Fig. 3 shows the tensile strength and elongation at the break of PCL/CMC wound dressings when adding HA in each concentration. The mixture of HA at 20 %wt was slightly higher the strength than the mixture of HA at 15 %wt, while the HA concentration at 25% wt showed the reduction of tensile strength because the HA consists of calcium phosphate. The increasing of HA concentration related the brittleness of sample due to it can be increased the pores in the matrix of sampl [14]. Thus, the HA concentration at 25%wt provided the low tensile strength of the PCL/CMC wound dressing at 4.15 ± 0.55 MPa. Meedecha, Srisang [16] reported that the appropriate tensile strength of wound dressings product was about 1-10 MPa. Therefore, this research found that adding HA at a concentration of 20 %wt with a tensile of 6.05 ± 0.6 MPa, which can improve the tensile strength. Fan, Lei [21] demonstrated that adding the HA increases the tensile strength of the product. In addition, Meedecha, Srisang [16] demonstrated that the tensile strength of the wound dressing product from PLC/CMC that crosslinked with acetic acid was 4.5 MPa. For this reason, adding HA can enhance to increase the tensile strength of the sample.

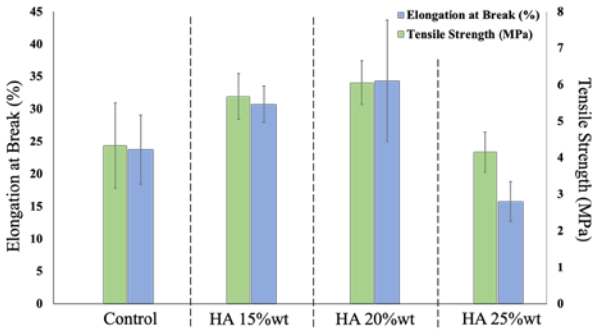


Fig. 3. Mechanical results: elongation at break and tensile strength.

I. Conclusion

The PCL/CMC wound dressing was successfully added the HA at 20% wt, which can be fabricated to the wound dressing. The wound dressing sample provided the tensile strength and the elongation at break at 6 ± 0.6 MPa and 34.3 ± 9.34%, respectively. Furthermore, the water absorption of DI water and NaCl were 77.16 and 92.66%, respectively. All results implied the excellent potential of biopolymer and HA to produce the biocomposite as a biomaterials which can be further studied in term of biocompatibility, antibacterial activity, and this formulation can be applied with 3D printing.

ACKNOWLEDGEMENTS

This work was financially supported by King Mongkut's Institute of Technology Ladkrabang [KREF016627].

REFERENCES

[1] Rezvani Ghomi E, Khalili S, Nouri Khorasani S, Esmaeely Neisiany R, Ramakrishna S. Wound dressings: Current advances and future directions. Journal of Applied Polymer Science. 2019;136(27):47738.
[2] Priya S, Choudhari M, Tomar Y, Desai VM, Innani S, Dubey SK, Singhvi G. Exploring polysaccharide-based bio-adhesive topical film as a potential platform for wound dressing application: A review. Carbohydrate Polymers. 2024;327:121655.
[3] Hu W-W, Lin Y-T. Alginate/polycaprolactone composite fibers as multifunctional wound dressings. Carbohydrate Polymers. 2022;289:119440.
[4] Woodruff MA, Hutmacher DW. The return of a forgotten polymer—Polycaprolactone in the 21st century. Progress in Polymer Science. 2010;35(10):1217-56.
[5] Kennedy SW, Roy Choudhury N, Parthasarathy R. 3D printing soft tissue scaffolds using Poly(caprolactone). Bioprinting. 2023;30:e00259.
[6] Zennifer A, Senthilvelan P, Sethuraman S, Sundaramurthi D. Key advances of carboxymethyl cellulose in tissue engineering & 3D bioprinting applications. Carbohydrate Polymers. 2021;256:117561.
[7] Kanikireddy V, Varaprasad K, Jayaramudu T, Karthikeyan C, Sadiku R. Carboxymethyl cellulose-based materials for infection control and wound healing: A review. International Journal of Biological Macromolecules. 2020;164:963-75.
[8] D'Asta F, Challoner T, Wilson YT, Thomas C, Moiemen NS. Acetic acid dressings converted skin graft donor sites into full thickness wounds in a burned infant, a case report. Burns Open. 2023;7(3):85-8.
[9] Qiu H, Wang J, Hu H, Song L, Liu Z, Xu Y, et al. Preparation of an injectable and photocurable carboxymethyl cellulose/hydroxyapatite composite and its application in cranial regeneration. Carbohydrate Polymers. 2024;333:121987.
[10] Amenaghawon AN, Anyalewechi CL, Darmokoesoemo H, Kusuma HS. Hydroxyapatite-based adsorbents: Applications in sequestering heavy metals and dyes. Journal of Environmental Management. 2022;302:113989.
[11] Teodoro KBR, Alvarenga AD, Rocha Oliveira LF, Marques Chagas PA, Lopes RG, Andre RdS, et al. Fast Fabrication of Multifunctional PCL/Curcumin Nanofibrous

Membranes for Wound Dressings. ACS Applied Bio Materials. 2023;6(6):2325-37.
[12] Wang Y, Shi D. In vitro and in vivo evaluations of nanofibrous nanocomposite based on carboxymethyl cellulose/polycaprolactone/cobalt-doped hydroxyapatite as the wound dressing materials. Arabian Journal of Chemistry. 2022;15(11):104270.
[13] Fakhrali A, Semnani D, Salehi H, Ghane M. Electrospun PGS/PCL nanofibers: From straight to sponge and spring-like morphology. Polymers for Advanced Technologies. 2020;31(12):3134-49.
[14] Li T, Peng Z, Lv Q, Li L, Zhang C, Pang L, et al. SLS 3D Printing To Fabricate Poly(vinyl alcohol)/Hydroxyapatite Bioactive Composite Porous Scaffolds and Their Bone Defect Repair Property. ACS Biomaterials Science & Engineering. 2023;9(12):6734-44.
[15] Zhu Y, Zhou W, Xiang J, Wu M, Chen Z, Yang Z, et al. Deferoxamine-loaded Janus electrospun nanofiber dressing with spatially designed structure for diabetic wound healing. Materials & Design. 2023;233:112166.
[16] Meedecha P, Srisang N, Eawsakul K, Ongtanasup T, Tambunlertchai S, Sokjabok S, et al. Preparation and evaluation of blend polymer films for wound dressing using vancomycin-loaded polycaprolactone and carboxymethyl cellulose via crosslinking methods: Effect of mechanical strength, antibacterial activity, and cytotoxicity. Journal of the Mechanical Behavior of Biomedical Materials. 2024;151:106339.
[17] Sun T-W, Zhu Y-J, Chen F. Highly Flexible Multifunctional Biopaper Comprising Chitosan Reinforced by Ultralong Hydroxyapatite Nanowires. Chemistry – A European Journal. 2017;23(16):3850-62.
[18] Al-Mofty SE-D, Karaly AH, Sarhan WA, Azzazy HME. Multifunctional Hemostatic PVA/Chitosan Sponges Loaded with Hydroxyapatite and Ciprofloxacin. ACS Omega. 2022;7(15):13210-20.
[19] Qianqian O, Songzhi K, Yongmei H, Xianghong J, Sidong L, Puwang L, Hui L. Preparation of nano-hydroxyapatite/chitosan/tilapia skin peptides hydrogels and its burn wound treatment. International Journal of Biological Macromolecules. 2021;181:369-77.
[20] Parikh DV, Fink T, DeLucca AJ, Parikh AD. Absorption and swelling characteristics of silver (I) antimicrobial wound dressings. Textile Research Journal. 2010;81(5):494-503.
[21] Fan J, Lei T, Yu M, Wang Y, Cao F, Yang Q, et al. Keratin/PEO/hydroxyapatite Nanofiber Membrane with Improved Mechanical Property for Potential Burn Dressing Application. Fibers and Polymers. 2020;21(2):366-75.

Exploring Underground Railway Structures Using Ground Penetrating Radar: Testing Sites in Bang Klam District, Songkhla Province, Southern Thailand

Nisarat Prangsuwan[1, a)], Kamhaeng Wattanasen[1, b)], Sawasdee Yordkayhun[1] and Thanan Chub-uppakarn[2]

[1]Division of Physical Science, Faculty of Science, Prince of Songkla University,
HatYai, Songkhla, Thailand
[2]Division of Civil Engineering and Environment, Faculty of Engineering, Prince of Songkla University,
HatYai, Songkhla, Thailand

[a)] Corresponding author: phy.kate@gmail.com
[b)] kawageophysics@gmail.com

Abstract. Testing an application of GPR for detecting railway subsurface structure was conducted in Bang Klam District, Songkhla Province, southern Thailand. A GPR system was placed on a specially designed wooden plank for collecting data. It can save time and easy to perform. Two GPR frequencies of 100 and 200 MHz were used with the aims to characterize the railway subsurface structures and to localize the defect zone or zone of potential damage within the railway substructures. The results found that the 200 MHz antenna can provide the subsurface railway structures and some potential defected zones in the ballast and sub-ballast layers. This testing results will be a useful information for the maintenance team of the State Railway of Thailand (SRT) and demonstrate that GPR is a powerful tool and can help in surveying and maintaining the railway in the future.

Keywords: Geophysical survey, Ground penetrating radar (GPR), Railway structures

I. INTRODUCTION

Rail transport is used for passenger and freight transport globally. It can carry passengers and cargo with heavy loads, great energy efficiency, potentially high speed and safety. Thus, the railway infrastructures maintenance is essential for ensuring safe and efficient transportation systems. Typical railway track structuresare shown in Fig.1. The substructures of railway tracks, the ballast or sub-ballast layer plays a crucial role in supporting the tracks, dispersing loads, and facilitating drainage [1]. However, as the time passed, deterioration and settlement of the supporting layer can occur due to various factors such as weathering, traffic loads, and subgrade conditions. The defected zone or soft spot within the supporting layers relates to an occurrence of mud pumping when the train passing. This phenomenon deteriorates in load track performance and in train ride quality. Thus, the location of underground defected zone needs to be investigated for railway track maintenance.

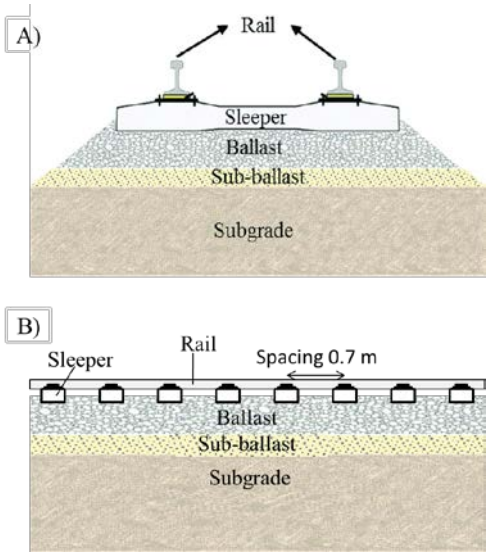


Fig. 1. Typical railway track structures, A) cross-section and B) longitudinal section, (Modified from [2])

Traditional method for assessing subsurface railway condition has been done by visual or manual inspection, whichis often involve time-consuming and labor-intensive and the comprehensive insights into subsurface conditions may not be provided. Therefore, thenon-destructive testing (NDT) method like GPR (Ground Penetrating Radar) has

been tried to investigate the distribution of mud pumping along a railways line [3], [4] and to inspect the railway track [5]. In Thailand, a mud pumping in railway embankment in Pitsanuloke province has been studied [6]. However, there is still relatively little use of GPR for detecting subsurface railway structures in Thailand. This test is therefore the first of its kind for GPR at the Geophysical Research Centre, Faculty of Science Prince of Songkla University that are used to investigate the subsurface structure of railway tracks.The test was carried out on a section of railway track located between Ban Din Lan Railway Station and HatYai Junction Railway Station, Bang Klam District, Songkhla Province.

The railway track at Bang Klam District is the main track for the train from Bangkok to southern Thailand. Therefore, it is an important railway route for transporting goods and traveling both within the country and connecting to foreign countries. The condition of railway ballast in this region is a cause for concern. This is because some positions of the rails that are jointed have moved up-down when a train passes and some sections of the railway line have been used for 10 years after the latest maintenance [7]. Therefore, the railway track here is appropriate to be a testing site by GPR measurement with the aims:

1. To characterize the underground railway structures such as ballast, sub-ballast, subgrade, and any additional layers, contributingto a comprehensive understanding of the railway's composition.

2. To localize the defect zone or zone of potential damage or deterioration within the railway substructure such as voids, moisture ingress, and changes in material composition beneath the railway track which will be recommended for localize maintenance.

Study area

Two areas in Bang Klam District, Songkhla Province were selected as the testing areas for GPR measurement on the single-track railway (Fig 2). Area 1 is close to the railway bridge across the Rama 1 canal. The distance of survey line is approximately 52 meters. In this area the railway track has been maintained latest in 2022 [7]. However, there is a point of concern on the route where the rails jointed have moved up-down when a train passing which may affect the ballast and deeper layers. Area 2 is close to Ban Din Lan railway station. The survey line here is approximately 58.2 meters long. This site has an older route sectionthan the route in area 1 due to there was the latest maintenance of the railway track in 2014 [7]. Therefore, it may be deterioration in some parts of underground railway structure.

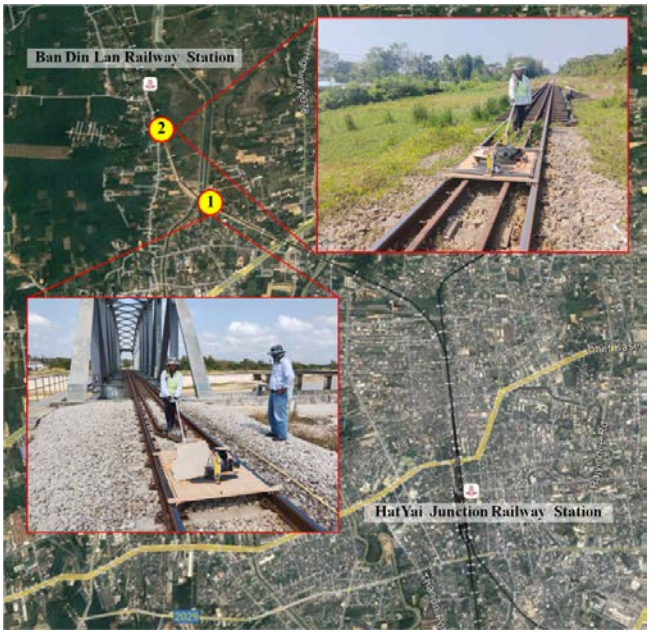


Fig. 2. Two study areas at Bang Klam District, Songkhla Province. Number 1 and 2 are represent the area 1 and area 2, respectively.

II. METHODOLOGY

Groundpenetratingradar (GPR) is a geophysical technique used to investigate subsurface structures and materials by transmitting and receiving electromagnetic (EM) waves. GPR waves will be reflected at the lithologic boundaries where there is a contrast in the dielectric constants and to create images of subsurface features such as voids, moisture content, and changes in material properties [8]. Thus, GPR may map the continuity of ballast, sub-ballast, sub-grade layers and the defect zone within the railway substructure.

1. GPR measurement

The GPR system (RAMAC pulse GPR) consists of antennas (Transmitter, Tx) and Receiver, Rx), control unit and display unit (Fig.3). A system was mounted on a wooden plank with wheels which can easily move on the railway track by man power. A common-offset GPR survey with antennas frequencies of 200 and 100 MHz, and an antenna offset of 0.6 m and 1.0 m respectively were conducted along the survey line in area1 and 2. The height of antennas is 0.18 m above the railway sleepers. A measuring interval of 0.05 m was used for all survey lines. The reflected signals from both frequencies were clearly seen down to a depth of > 250 ns of the time scale, which corresponds to a depth of > 5 m, when an average normal ground GPR wave velocity of 0.0401 m/μs obtained from WARR (Wide Angle Refraction and Reflection) measurement was used.

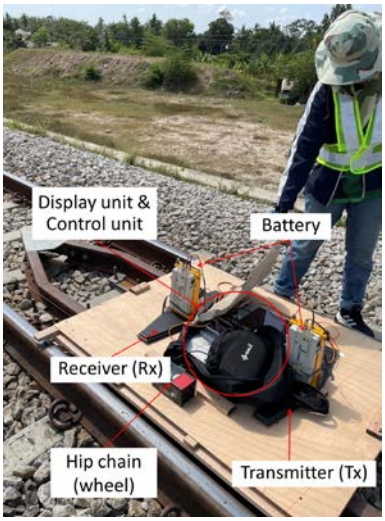


Fig. 3 GPR system is adapted for conducting on the railway track by mounting it on the wooden plank.

WARR measurement was performed with the aim to estimate the average velocity of EM wave that propagate in the subsurfaceby using 100 MHz antennas on the same survey lines of common-offset GPR survey in both areas. This type of measurement, theTx antenna will be fixed and the Rx antenna will move to the next station for recording the reflected signal with the station interval of 0.05 m until the reflected signal is weakened or the hyperbolic shape of the reflected signal is appeared on the radargram.

2. Depth determination

In GPR measurement, the EM waves have travelled from the Tx antennaand have then been reflected at the interface of the media. The two-way travel times and the amplitude of the returning signals are recorded by the Rx antenna (Fig.4). The depth to the reflectors (h) can be calculated if the velocity of EM wave in the upper medium (v) is known [9].

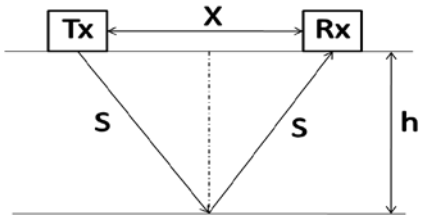


Fig. 4. The symbol S denotes the travel path of a reflected wave the depth to the reflector (h) can be calculated from the equation related to the travel time (t) of the reflected wave and antenna offset (x) [9].

$$(S)^2 = \left(\frac{x}{2}\right)^2 + h^2$$
$$2s = vt$$
$$(vt)^2 = x^2 + 4h^2$$
$$h = \sqrt{\frac{(vt)^2 - x^2}{4}} \quad (1)$$

3. Data Processing

The GPR data were processed by using software ReflexW version 8.0.2 [10]. Data processing includes: (1) move start time for determining where the surface is, (2) apply Subtract-mean (de-wow) option for removing low frequency noise, (3) apply Background removal, this filter performs a subtracting of an averaged trace (tracerange) which is built up from the chosen time/distance range of the current section, (4) apply Running average, this filter performs a running average over a selectable number of traces for each time step, (5) apply Bandpass frequency to cut off only the desired frequencies, (6) apply manual gain (y) for enhancing the reflected signal from the deeper layer, and (7)apply depth conversion option for converting GPR arrival-times to depth, which requires the GPR propagation velocity in the ground of each survey area.

III. RESULTS AND DISCUSSION

Signal characteristics of different antennas frequencies

The characteristics of reflected signal from a common-offset GPR survey conducted in an area 2 with the 100 and 200 MHz antennas are shown in Fig. 5-6. The amplitude of the reflected signal of both frequencies is almost the same (see Fig. 5A1 and Fig. 6A1). The difference is found in the resolution of data that the 200 MHz gives the higher resolution than the 100 MHz (see Fig. 5A2 and Fig. 6B2). This is due to the sampling interval of the higher frequency is shorter than the lower frequency. Therefore, the 200 MHz antenna were found to be suitable for the aims of this testing.

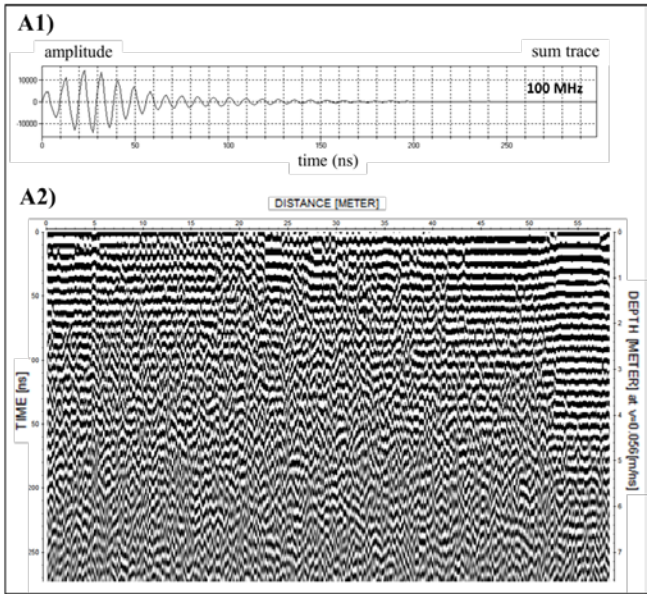


Fig.5. A1) A signal amplitude of the 100 MHz antenna and A2) GPR section was conducted by the 100 MHz antenna in area 2.

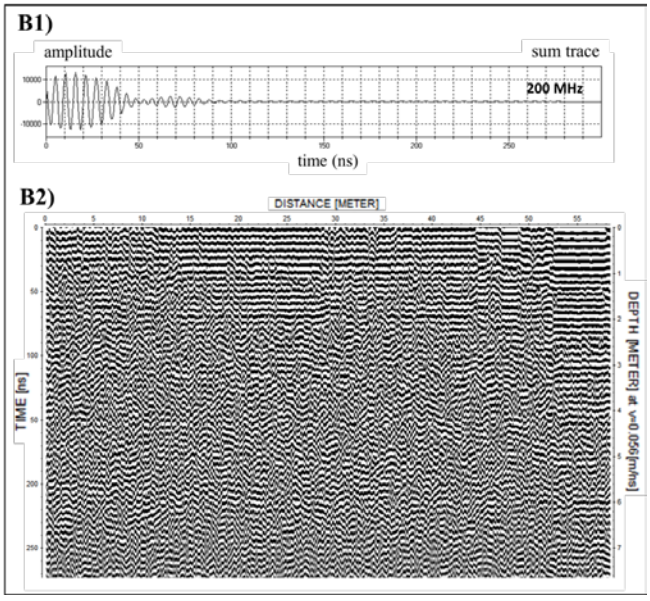


Fig. 6. B1) A signal amplitude of the 200 MHz antenna and B2) GPR section was conducted by the 200 MHz antenna in area 2.

WARR measurement result

The result of WARR measurement in both areas is showed in Fig. 7. The velocity of EM waves which propagetes in the subsurface can be obtained by fitting the hyperbolic pattern of reflected signal in the radargram. The average propagation velocities of 0.040 m/nsand 0.056 m/ns in survey area 1 and 2 are found, respectively. These values of EM wave velovity will be used for determining the depth to the reflector in the subsurface.

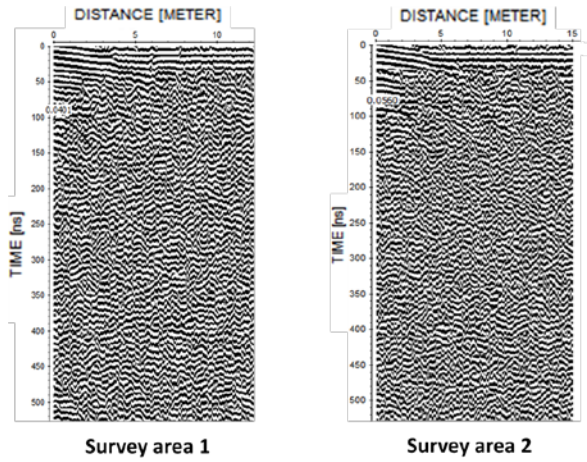


Fig. 7. GPR section from WARR measurement determining an EM wave propagation velocity in the subsurface of survey area 1 and 2.

As we found the difference in an average EM wave propagation velocity in survey area 1 and 2. It means that the subsurface property that related to the velocity of EM wave such as dielectric constant in both areas are different. This is because the EM velocity in the medium has a negative relationship to the dielectric constant and the dielectric constant has a positive relationship to the moisture contentof the medium. Thus, it can be suggested that the drainage efficiency of ballast layer in the area 1 might be less than in the area 2.

200 MHz antenna - GPR survey results

The GPR section in the area 1 is shown in Fig. 8.

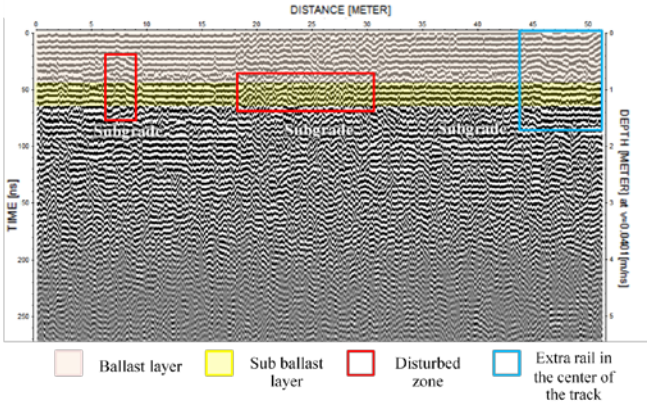


Fig. 8.GPR section of survey area 1.

The reflected signals were clearly seen down to a depth of >250 ns and the subsurface railway structure could be divided into 3 distinct layers. The first layer is a ballast layer (pink transparent layer) which corresponds to the layering of low frequency reflected signal. It has a thickness about 0.90 m and is underlain by a layer of prominent amplitudes of the higher frequency of reflected signal (yellow transparent layer), which may represent the layer of sub-ballast. This layer is somewhat thin about 0.4 m thick. Below the base of this second layer, as defined by strong irregular-reflected signal, which may correspond to the subgrade layer. For the depth conversion of the GPR section, a constant average propagation velocity of 0.040 m/ns was used. The anomalous in this GPR section that found the irregular and lateral un-continuity of the reflected signal are seen at a distance about 7.5 m and at a distance from about 18 – 31 m (red-square block). This first zone is coincided with the location of railroad joints. It should be the cause of potential damage to the ballast and deeper layers when the trains passing. At the second zone in the middle of GPR section, irregularly reflected signal in the ballast and sub-ballast layer is observed. This zone here is expected that the ballast and sub-ballast layers may be damaged.

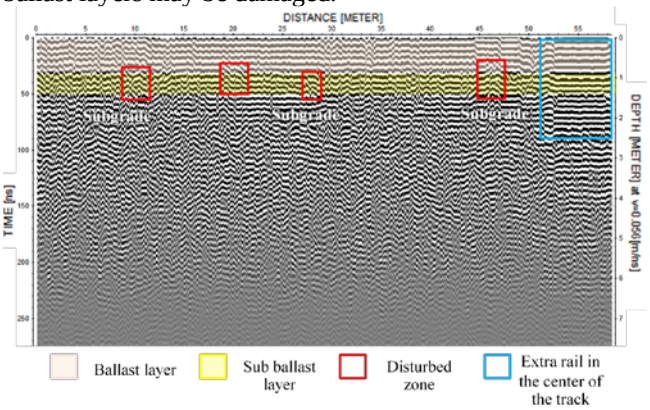


Fig. 9.GPR section of survey area 2.

The GPR section in the area 2 is shown in Fig. 9. The radargram is also clearly seen down to a depth of >250 ns and the subsurface railway structure could be divided into 3 distinct layers as well. The first layer which corresponds to the layering of low frequency reflected signal represents a ballast layer (pink transparent layer). It has a thickness about 0.95 m and is underlain by a layer of prominent amplitudes of the reflected signal (yellow transparent layer), which may represent the layer of sub-ballast. This second layer is somewhat thin about 0.45 m thick.The third layer is defined by strong irregular-reflected signal, which may correspond to the subgrade layer.In this area, a constant average propagation velocity of 0.056 m/ns was used for the depth conversion. Four anomalous found the irregular and lateral un-continuity of the reflected signal are seen at a distance about 10.0, 20.0, 29.0 and 46.0 m (red-square block). These positions (disturbed zone) should be the zone of damage or deform in the ballast and sub-ballast layer. For the zones of blue-square block in Fig. 8 and Fig. 9, the GPR data cannot be interpreted because at these distances of GPR measurements, there is a special rail (Guard Rail) installed at the center of the railroad connecting to the bridge. The reflected signals here are less confident to be interpreted.

The drainage property is a main function of the supporting layer (ballast and sub-ballast), the drainage efficiency of the subsurface railway tracks in the area 2 should be greater than that in the area 1. This suggestion relies on the relationship of EM wave velocity, dielectric constant and the moisture content, in which the first 2 factors can be obtained from GPR measurement. However, the corrected positions/zones of deformation in the ballast and sub-ballast layers which found from this study should be approved by the maintenance team of the State Railway of Thailand (SRT).

IV. Conclusion

The testing application of GPR for detecting underground railway structure has been carried out in Bang Klam District, Songkhla Province, southern Thailand. By putting a GPR system on a specially designed wooden plank for this work. It saves a time and easy for collecting data. The testing results found that the 200 MHz can provide the subsurface railway structure and some potential defected zones in the ballast and sub-ballast layers. These finding results will inform the maintenance team of the State Railway of Thailand (SRT) fortheir decision-making regarding the maintenance and rehabilitation strategies. The correlation between the signal pattern reflected from GPR and the mapped subsurface railway structure suggests that GPR will be a powerful tool for future railway surveying and maintenance.

Acknowledgment

The authors would like to acknowledge Thitinan Indhanu, a Ph.D. candidate, Faculty of Engineering, Prince of Songkla University for providing a useful information and supporting during field work. We thank to Jamras Na Suwan, Division of Physical Science, Faculty of Science, Prince of Songkla University for building a wooden plank cart. Thank you very much to Panumas Promchot, a staff of the State Railway of Thailand (SRT) in HatYai Junction Railway Station for facilitating of this work.

References

[1] G. Liu, Z. Peng, G. Jing, S. Wang, Y. Li, and Y. Guo, “Railway ballast layer inspection with different GPR antennas and frequencies,” Transportation Geotechnics 36, 2022.
[2] P. Barkhordari, R. Galeazzi, A. M. Tejada and I. F. Santos, “Identification of behavioural models for railway turnouts monitoring”. arXiv preprint arXiv:1910.06582, 2019.
[3] A. Benedetto, F. Tosti, L. B. Ciampoli, A. Calvi, M. G. Brancadoro, and A. M. Alani b, “Railway ballast condition assessment using ground-penetrating radar–An experimental, numerical simulation and modelling development,” Construction and Building Materials,vol. 140, 508–520, 2017.
[4] N. N. Yahaya, A. Ibrahim, J. Ahmad, A. Ahmad, M. I.F. Rozli, and Z. Ramli “Assessment of Subgrade Degradation Induced Mud Pumping at Railway Track: A Review,” Civil Engineering and Architecture,vol. 10, no.1, 345-355, 2022.
[5] C. Kuo, “Ground penetrating radar to investigate mud pumping distribution along a railway line,” Construction and Building Materials, vol. 290, 123104, 2021.
[6] C.H. Hsu, “Using characterize grids resolving method for GPR on investigating railway mud pumping,” Master’s thesis, Minghsin University of Science and Technology, 2017.
[7] P. Maneechote, “Railway engineer, personal communication,” 2024.
[8] J. Hugenschmidt, “Railway track inspection using GPR,” J. Appl. Geophys, vol.43, no.2-4, 147–155, 2000.
[9] K. Wattanasen, “A Geophysical Study of an Arsenic Contaminated Area in Ron Phibun District, Southern Thailand”. Licentiate thesis, Lulea University of Technology, Sweden, 2004.
[10] K.J. Sandmeier, “REFLEXW manual,”1998-2024.

Detection of Coconut Black Headed Caterpillar Using Aerial Imagery Combined with Artificial Intelligence

Orapadee Joochim[1, a)] Kridtat Satharanond [1, b)] and Wirachat Kumkun [1, c)]

[1]Institute of Field Robotics, King Mongkut's University of Technology Thonburi, Bangkok, Thailand

[a)] Corresponding author: orapadee.joo@kmutt.ac.th
[b)] Kridtat_mac@hotmail.com
[c)] wirachat.kamkun@gmail.com

Abstract. In this paper, the methods to detect coconut black headed caterpillars are investigated, which are presently difficult for locating and correcting in a timely manner. The infestation of coconut black headed caterpillars causes severe damage to the coconut trees resulting in reduction of yield significantly. The coconut trees, which are most facing this problem are mostly in the southern region, in Prachuap Khiri Khan Province and Surat Thani Province of Thailand that has the coconut trees, which are more than 20 meters high, making it impossible to observe and take care of it thoroughly. The developed drone can be used to inspect and find the abnormal area that may be caused by the infestation of coconut black headed caterpillars. The images from the drone will be processed by artificial intelligence system, when the flight is completed. These images are approved for the characteristic check by working with an entomologist, Department of Agriculture of Thailand, and additionally done for the ground check by placing the colour papers at the bottom of the coconut trees that are destroyed by coconut black headed caterpillars for labeling and for processing using the artificial intelligence system easily. RetinaNet model is used in this research for accuracy and precision. The result has 93% accuracy for rounds 3 of modeling.

***Index Terms*—Unmanned aerial vehicle (drone), coconut black headed caterpillar, artificial intelligence system.**

I. INTRODUCTION

THAILAND is primarily an agricultural country with the majority population engaged in agriculture. Coconut is one of the significant economic crops in Thailand. In the current situation, coconut production in Thailand is facing significant challenges due to the widespread infestation and destruction by plant pests. This has caused severe damage, leading the government to allocate billions of baht in budget to provide the assistance. The significant factors contributing to the spread of the infestation include continuous year-round cultivation, high seed rates, the use of non-specific chemical pesticides, and the lack of surveillance and monitoring of the pest situation. Farmers are aware that when plant pests infest severely, it becomes challenging to control them effectively. Generally, farmers rely solely on chemical pesticides for pest control, and often use them incorrectly. Apart from the inability to control plant pests, this practice can increase production costs. Plant pests can also spread widely. The coconut black headed caterpillar is the most damaging coconut pest compared to other types. It can cause severe damage to coconuts, sometimes leading to the death of coconut trees [1].

Drones, or Unmanned Aerial Vehicles (UAVs), are aircraft that operate without a human pilot on board. The initial use of drones in Thailand was primarily in the military sector, for reconnaissance missions or even armed with weapons to engage enemies from a safe distance without risking human lives in dangerous areas. Subsequently, there has been a continuous development, and drones have been utilized in various forms. For example, attaching a camera to a drone allows for high-angle photography, traffic monitoring, and geographic data collection. With numerous advantages, drones have also been developed and utilized in agriculture. In modern agriculture, drones play a significant role in improving efficiency and have become an essential part of the country's economic development. Drones are increasingly being used in agriculture to enhance productivity. Drones can oversee vast areas, and they can be used to spray chemicals, fertilizers, and plant hormones. Currently, drones are expensive, making it difficult for farmers to access them. Moreover, in the development of models to detect abnormal tree conditions, diseases, or damages caused by pests, there is a problem with data, particularly the scarcity of data on trees exhibiting abnormalities. Research on drones in agriculture is increasing continuously. However, the majority of the research focuses on drones for agricultural surveying or chemical spraying to control pests. Research on drones for disease or pest detection remains minimal, especially regarding the development of drones for detecting diseases

or pest infestations in coconut plantations. No one has conducted significant research in this area [2, 3].

Therefore, this research aims to develop drones for agricultural use for detecting abnormalities, whether it is related to fertility or the presence of coconut black headed caterpillars. This research uses the artificial intelligence technique along with analyzing and improving color images through the image processing. The study utilizes the RetinaNet model, which is the state-of-the-art object detection algorithm from 2018. Since coconuts are a significant economic crop, the drone will be used to inspect and locate trees affected by pests, particularly coconut black headed caterpillars. The drone can fly to inspect and identify areas with pest infestations in coconut trees.

II. METHODOLOGY

The research team has received the consultation from a group of experts from the Ministry of Agriculture and Cooperatives, Thailand for additional information regarding the project plan. From the suggestion of the experts, it can be concluded that drones will be beneficial in terms of assessing the affected areas. We will be able to see the overall picture of how much area is affected, the severity level, the recovery after chemical spraying, and be able to disseminate information to farmers. If we can identify the problem from its early or moderate stage, it will be more beneficial to farmers rather than in severe stages. Table I demonstrates the severity level of damage from coconut black headed caterpillars (Opisina arenosella). The optimal number of leaves for a coconut tree without affecting its productivity is between 13 to 24 leaves. The characteristic damage pattern of the coconut black headed caterpillar involves destruction primarily from the underside of the canopy, with leaves exhibiting dryness and signs of being chewed or eaten. However, the petioles still retain their green color.

TABLE I
SERVERITY LEVEL OF DAMAGE

Low	Moderate	High
The leaves are healthy with no signs of damage from 13 leaves or more.	The leaves are healthy with no signs of damage of 6-12 leaves.	The leaves are healthy with no signs of damage of 5-0 leaves.

After that, the research team plans to collect survey flight data to capture additional images in Prachuap Khiri Khan Province and Bang Khun Thian District and Nong Chok District, Bangkok, Thailand for the infestation at the early, moderate, and severe stages. The research team will mark the trees affected by the infestation by placing colored papers near the infestation points. Different colors will be used to indicate the severity of the infestation, for example, red indicates severe infestation (Fig. 2), while yellow indicates moderate infestation (Fig. 3).

Fig. 1. Example of Abnormal Coconut Tree Symptoms Caused by Coconut Black Headed Caterpillar.

Fig. 2. Placing colored papers for trees with a severe infestation.

Fig. 3. Placing colored papers for trees with a moderate infestation.

2.1 Specification of Drone and Camera

The drone is capable of performing according to the following specifications: it can fly and capture images automatically, carry a payload of up to 1,000 grams, cover an area of 16,000 square meters or approximately 12 rai per battery cycle, and operate for at least 25 minutes per battery cycle. The camera attached to the drone has a resolution of 16 million pixels. Based on AI processing tests, using an RGB lens with 16 million pixels resolution is sufficient. The drone flies 15 meters above the top of the coconut tree. The flight duration is estimated to be around 4 minutes per rai.

Fig. 4. An example of how to partition an image into 9 sub-images.

2.2 Object Detection

Detecting the position of abnormal coconut tree leaves through AI training will be divided into two stages: data preparation for coconut tree leaf detection, AI creation and creating an AI to detect the position of abnormal coconut tree leaves.

Image partitioning is crucial in preparing the dataset to train the AI to learn about abnormal coconut trees. Since the images taken by the drone are too large and do not allow for a clear view of the coconut tree details, it is necessary to divide the images into 9 sub-images before creating the dataset (Fig. 4). The split images will have filenames that represent the sequence of parts from the original image. The file naming format is "XXXX_YY," where XXXX refers to the sequence number of the image taken by the drone, and YY indicates the number of the part from the original image. The divided images will have dimensions of 628 x 512 pixels.

To determine the location to teach AI to learn what abnormal coconut leaves look like. Visual Geometry Group Image Annotator (VIA) program is used, which is a web platform for creating location information within the image data for AI. The program can create a file specifying locations from drawing bounding boxes to use as Sample data for developing AI Object Detection. The nature of specifying the location to cover the overall nature of the destruction of coconut trees by coconut black headed caterpillars must be in the form of teaching the computer to recognize the characteristics of the entire tree. The classification example is divided into two types: trees that are damaged by moderate level of infestation of coconut black headed caterpillars and trees that are damaged by low level of infestation of coconut black headed caterpillars.

To create an AI to detect coconut trees with abnormalities caused by the coconut mite, the team has chosen to use the RetinaNet AI algorithm [4] for object detection and localization. The developers have already trained the AI to recognize objects from a large dataset (Big Data). Hence, the task at hand is to provide specific and targeted data. For example, the positions of coconut trees with abnormalities, to further teach the AI.

The number of data used to teach AI is adjusted by adding data per version (Round 0, 1, 2 and 3) to study the trend of AI ability in detecting abnormal coconut tree positions. This results in having a total of four versions of datasets used to train AI. The dataset can be categorized into the number of coconut trees affected by black headed caterpillars in the early and intermediate stages.

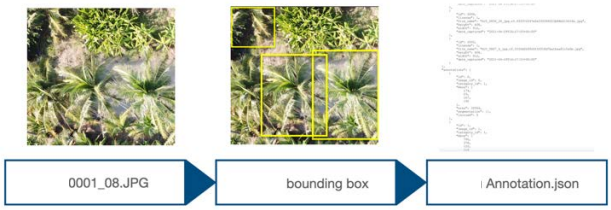


Fig. 5. Example of creating the dataset.

III. Results and Discussion

Fig. 6 shows the evaluation of the distribution of infestation by coconut black headed caterpillars at a low level. The red color represents the bounding boxes generated by the AI, and the yellow color represents the Ground Truth of the coconut trees affected by the coconut black headed caterpillar infestation at the initial stage. The Ground Truth is assisted by experts in the field to pinpoint and identify coconut trees affected by black headed caterpillars. Fig. 7 shows the evaluation of the distribution of infestation by coconut black headed caterpillars at a moderate level. The red color represents the bounding boxes generated by the AI, and the blue color represents the Ground Truth of the coconut trees affected by the coconut black headed caterpillar infestation at the moderate stage. Figure 8 illustrates the evaluation of the distribution of infestation by coconut black headed caterpillars at a low level. The red color represents the bounding boxes generated by the AI, and the yellow color represents the Ground Truth of the coconut trees affected by the coconut black headed caterpillar infestation at the initial stage. Figure 9 illustrates the evaluation of the distribution of infestation by coconut black headed caterpillars at a moderate level. The red color represents the bounding boxes generated by the AI, and the blue color represents the Ground Truth of the coconut trees affected by the coconut black headed caterpillar infestation at the moderate stage.

The evaluation of Precision and Recall values based on the assessment from the Precision and Recall score in each model is shown in Table II. The models with the highest Precision and Recall scores are Round 2 Arch 101 and Round 3 Arch 50, respectively.

Based on research and experimentation, we create an AI for detecting abnormal coconut trees with a mathematical model using the RetinaNet technique. The model using the RetinaNet-50 Round 3 technique give the best performance results compared to Precision, and Recall. It had precision and recall values of 0.87 and 0.93, respectively.

TABLE II
PRECISION AND RECALL SCORES FOR EACH MODEL

Round	Arch	Precision	Recall
0	50	0.743	0.701
0	101	0.654	0.591
0	152	0.756	0.617
1	50	0.705	0.681
1	101	0.717	0.694
1	152	0.707	0.646
2	50	0.848	0.882
2	101	0.888	0.906
2	152	-	-
3	50	0.871	0.929
3	101	0.813	0.898
3	152	-	-

Fig. 6. Example of the distribution of infestation of coconut black headed caterpillars at a low level.

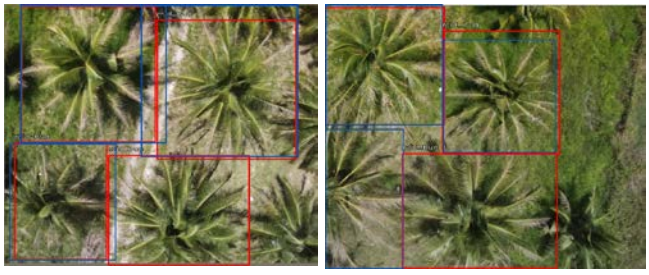


Fig.7. Example of the distribution of infestation of coconut black headed caterpillars at a moderate level.

Fig.8. Example of the distribution of infestation of coconut black headed caterpillars at a low level.

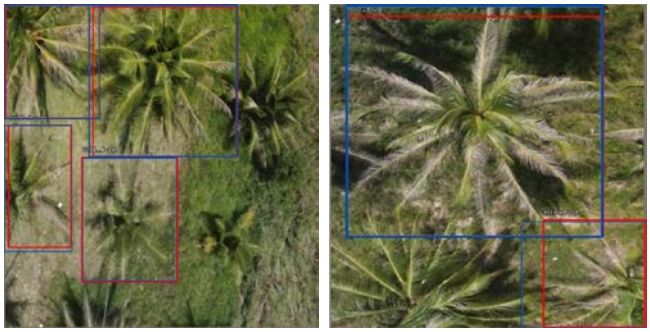


Fig.9. Example of the distribution of infestation of coconut black headed caterpillars at a moderate level.

IV. CONCLUSION

The research begins with data collection from farmers and experts specializing in the infestation caused by coconut black headed caterpillars. The study involves a review of related research, documents, and reports. Subsequently, it assesses and evaluates the insect's destructive behavior through its initial, moderate, and severe infestation stages, according to the recommendations of experts. Following this, the drone is designed to meet the requirements, and plans for bringing drone images into the processing process with artificial intelligence. For the detection of abnormalities caused by the infestation of coconut black headed caterpillars, once the drone flight is completed, the images are processed using an artificial intelligence system. The image characteristics are examined through collaboration with agricultural extension experts and ground checks. Ground checks involve placing paper at the base of coconut trees affected by the coconut black headed caterpillar infestation to facilitate labeling and ensure accurate and precise processing by the AI system. The RetinaNet model is used, achieving an accuracy rate of 93%. Consequently, the drone survey results allow for the detection of areas of coconut trees damaged by coconut black headed caterpillars, with an accuracy rate of 93%.

ACKNOWLEDGMENT

This work would not have been possible without the financial support of National Research Council of Thailand (NRCT).

REFERENCES

[1] The Coconut Situation in Thailand, The Department of Foreign Trade, Office of Commercial Affairs, Thailand, 2021.
[2] Li, W. Q., Han, X. X., Lin, Z. B., & Rahman, A., Enhanced Pest and Disease Detection in Agriculture Using Deep Learning-Enabled Drones. Acadlore Trans. Mach. Learn., 3(1), 1-10, 2024.
[3] Jere Kaivosoja, Ari Ronkainen, Juho Hautsalo, Juha Backman, Raimo Linkolehto, Miguel San Emeterio & Juha-Pekka Soininen, Automatized Rapeseed Pest Detection and Management with Drones. Lecture Notes in Networks and Systems, 2022.
[4] Hong Tian, Yufu Zheng, and Zhaozhao Jin, Improved RetinaNet model for the application of small target detection in the aerial images, 6th International Conference on Energy, Environment and Materials Science. 2020.

A Demonstration of Solar Pyrolysis of Agricultural Waste Biomass for Green Chemical Production

Tawanrung Artphan, Tanapat Sartsiri, Surakiat Yeamamphan, Kanyarat Jitsuwan, Rangsima Plapngam, Sirapob Boonyawanthang, Ratikorn Kulporm and Srirat Chuayboon[a)]

Department of Mechanical Engineering, King Mongkut's Institute of Technology Ladkrabang, Prince of Chumphon Campus, Chumphon 86160, Thailand

[a)] Corresponding author: srirat.ch@kmitl.ac.th

Abstract. Solar pyrolysis offers a promising carbon-neutral pathway to thermochemically convert waste biomass and solar energy into high-quality syngas and green chemicals. In this study, solar pyrolysis of agricultural waste biomass from palm oil empty fruit bunch (EFB) was performed in a one kW_{th} prototype chemical solar reactor to produce liquid bio-oil, syngas, and biochar. A study was experimentally conducted with a maximum temperature reaching 600 °C under heating rate of 12.2°C/min. As a result, syngas increased with increasing temperature. H_2 and CO were mainly produced, followed by CO_2, CH_4, and C_2H_m. Within 600 °C, the liquid tar was mainly produced, followed by char and syngas. The process demonstrated the concept of converting both renewable biomass and sunlight into chemicals in a single process.

Keywords: Biofuel, Concentrated solar power, solar thermochemistry

I. INTRODUCTION

The depletion of fossil fuel resources and the escalating concerns over environmental degradation have propelled the quest for sustainable and eco-friendly alternatives in chemical production. Biomass and solar energy are promising candidates due to their abundance, renewability, and potential to mitigate carbon emissions [1]. Notably, Thailand has great potential in renewable solar energy [2] and agricultural solid biomass supply, particularly bio-waste from oil palm plantations [3]. The oil palm is an agro-industrial commodity that produces cooking oil and fuels (biodiesel). The residues of cultivated palm oil plantations are plentiful and consist of tree branches (31%), fiber material (18%), fruit shells (10%), fronds and trunks (38%), and palm kernels (4%) [4, 5]. Thailand's average high solar radiation is estimated at 1800 kWh/m^2 annually [2]. Concerning solar energy, in Thailand, solar power is mainly used for generating electricity, primarily using non-concentrating solar power systems. Alternatively, it is worth noting that concentrating solar power (CSP) is a promising method of harnessing solar energy, which is more efficient than non-concentrating solar power systems [6]. CSP technologies are the parabolic trough, linear Fresnel, dish/engine, and central receiver. They can provide very high-temperature heat, which can be used for solar thermochemical conversion [6].

Solar biomass pyrolysis offers a promising avenue to integrate two renewable energies considering solid biomass and solar power in a single process [7-10]. With this concept, concentrated solar energy provides an external heat source to drive the pyrolysis reaction [11]. Therefore, the pyrolysis reaction can occur without any combustion in the process [12]. In addition, solar pyrolysis involves the thermal decomposition of biomass in the absence of oxygen, leading to the generation of bio-oil, biochar, and syngas, which can serve as precursors for a wide range of chemicals and fuels.

Therefore, the present study aims to experimentally investigate solar pyrolysis of oil palm agricultural bio-wastes. Empty fruit bunch (EFB) from oil palm was selected due to its high availability and suitable properties [13, 14]. A 1 kW_{th} indirectly irradiated solar reactor was utilized and operated. A representative study of solar pyrolysis was experimentally conducted with the maximum temperature reaching 600 °C under heating rate of 12.2°C/min. Syngas production rate, syngas yield, liquid bio-oil, and biochar was assessed and compared thoroughly.

II. MATERIALS AND METHODS

EFB biomass was acquired from the palm oil milling industries in Chumphon province, Thailand. Prior to its utilization, EFB was dried to reduce moisture content below ten wt% and then cut into tiny parts by a crusher machine to obtain a size distribution between 1-2 mm with a bulk density of around 0.17 g/cm^3. The physical and chemical properties of EFB biomass are presented in a previous study [15].

Fig. 1 shows the schematic diagram of the solar reactor setup. The design of this solar reactor was inspired by an

indirect irradiated solar chemical reactor, which comprises a tube inside the reactor where the chemical reaction takes place. The tube inside the reactor is made of alumina thanks to its excellent thermal and chemical stability, high mechanical strength, and good corrosion resistance. In addition, two 500W electrical heaters were supplied to heat the reactor.

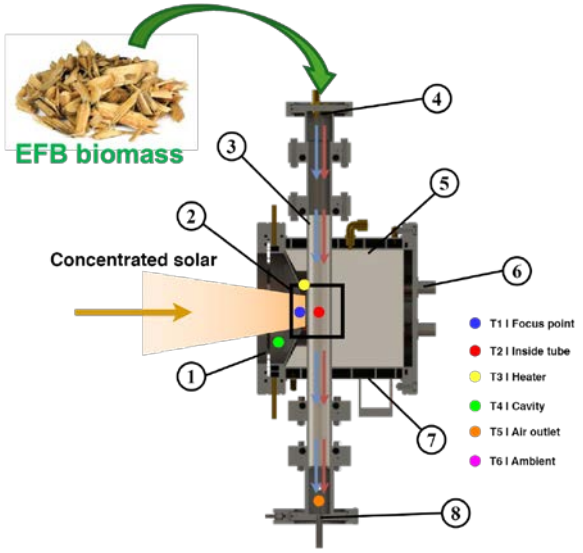


Fig.1. Schematic diagram of the cross-section of a solar reactor: (1) glass window, (2) reaction zone, (3) alumina tube, (4) reactants inlet, (5) insulation, (6) heater, (7) reactor's shell, (8) products outlet.

Prior to experiments, EFB biomass with 1000 mg mass was loaded into the reaction zone. Then, the reactor was gradually heated by electrical heaters under an N_2 carrier gas flow rate of 0.2 NL/min. As seen in **Fig. 2**, the pyrolysis reaction occurred at temperatures above 150 °C, as evidenced by a growth in the syngas production rate. Syngas increased, reached a peak, and decreased with increasing temperature. The reaction finished after 40 min, as evidenced by the zero-syngas production rate. The maximum temperature reached 600 °C at 37 min. Then, the temperature suddenly dropped due to a shutdown of the reactor heating process.

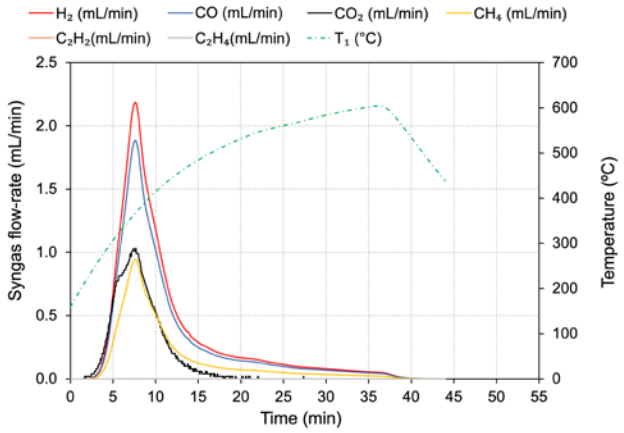


Fig. 2. Representative syngas production rate of solar EFB pyrolysis (heating rate 12.2 °C/min)

Syngas left the solar reactor through the product outlet port. Syngas flowed through a microfilter (pore diameter of 0.1 µm) before gas products analysis. The concentration of each species in evolved syngas was measured continuously by an online gas analyzer (Testo 350, uncertainty ≤ 2% FS). All the measured data were recorded by an automatic data collector (GRAPHTEC GL240). Based on the standard deviation of gas analysis measurements, the relative uncertainty on the syngas yields and reactor performance metrics was estimated to be ±2.%.

III. RESULTS AND DISCUSSION

Fig. 3 shows a representative global mass balance, represented by the ratio of products' mass output to reactants' mass input. Syngas products were quantified by integrating their flow rates with time, while char was quantified by weighing them after each experiment. Note that the liquid bio-char was quantified by the mass difference (1000 mg - (biochar mass + syngas mass)). It was found that products mass output was mainly composed of liquid bio-oil (67.3 w%), followed by biochar (29.4%), and syngas (3.4 w%), respectively, as a result of low heating rate, in agreement with a previous study [16], thereby demonstrating slow pyrolysis process.

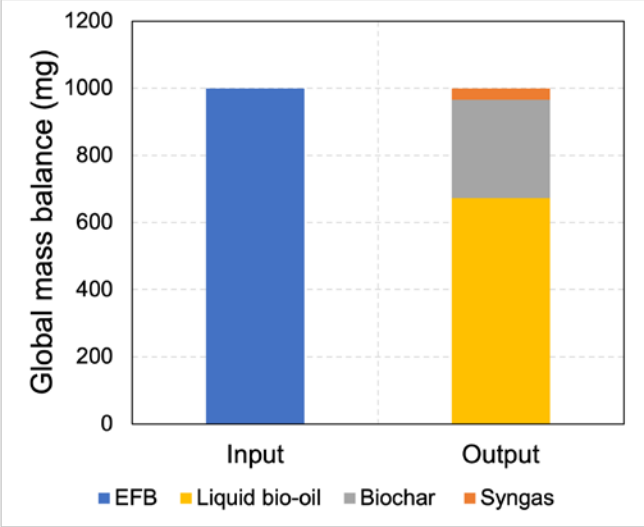


Fig. 3 Comparison of input reactants and output products (heating rate 12.2 °C/min).

Fig.4 shows that syngas yields produced from slow pyrolysis reaction consisting mainly of H_2 (0.71 mmol/ $g_{dry\text{-}EFB}$), followed by CO (0.61 mmol/ $g_{dry\text{-}EFB}$), CO_2 (0.28 mmol/ $g_{dry\text{-}EFB}$), CH_4 (0.30 mmol/ $g_{dry\text{-}EFB}$), C_2H_m (0.08 mmol/ $g_{dry\text{-}EFB}$), respectively, resulting in a total syngas yield of 1.98 mmol/ $g_{dry\text{-}EFB}$. The H_2/CO ratio was 1.16, which is in agreement with the previous study [16, 17].

Fig.5 shows the representative photograph of biochar collected after an experiment. Its mass was found to be 293.60 mg. It appeared in fine black particles. Its composition will be further characterized.

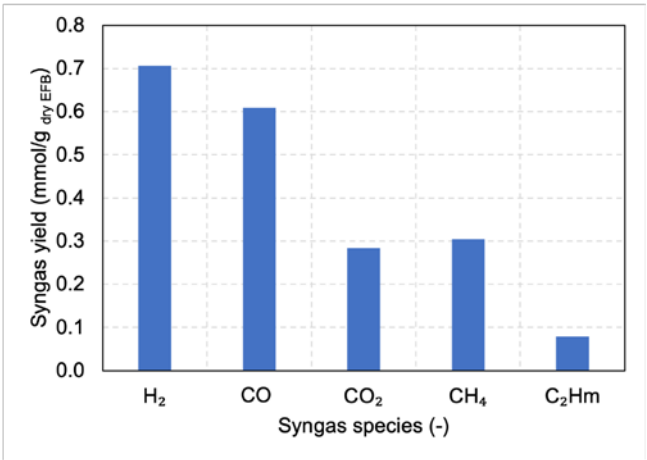


Fig. 4. Syngas yield (heating rate 12.2 °C/min)

Fig.5. Photograph of biochar collected in the rector after an experiment.

IV. CONCLUSION

Solar pyrolysis of EFB for green chemical production was demonstrated. It was efficiently performed under a low heating rate, demonstrating the feasibility and reliability of the solar-driven process in valorizing EFB waste biomass. Thanks to low heating rate, chemical production yields mainly composed of liquid bio-oil (67.3 w%), followed by biochar (29.4%), and syngas (3.4 w%), while syngas was produced in the form of mainly H_2 (0.71 mmol/ $g_{dry\text{-}EFB}$), followed by CO (0.61 mmol/ $g_{dry\text{-}EFB}$), CO_2 (0.28 mmol/ $g_{dry\text{-}EFB}$), CH_4 (0.30 mmol/ $g_{dry\text{-}EFB}$), C_2H_m (0.08 mmol/ $g_{dry\text{-}EFB}$), respectively. Biochar collected after an experiment was found in fine black particles. The solar pyrolysis EFB proved to be a promising route for agricultural biomass valorization and solar energy storage into sustainable carbon-neutral chemical products. Future work will be focused on the effect of heating rate and temperature to identify optimal conditions and evaluate the process large-scale feasibility.

ACKNOWLEDGMENT

This research is a result of the project entitled "A prototype of high-quality solar fuels from renewable oil palm waste biomass and biogas from palm oil mill effluent" by King Mongkut's Institute of Technology Ladkrabang (KMITL), which has been received funding support from the NSRF (grant number: RE-KRIS/FF67/004).

REFERENCES

[1] C. Bonechi, M. Consumi, A. Donati, G. Leone, A. Magnani, G. Tamasi, C. Rossi, 1 - Biomass: An overview, in: F. Dalena, A. Basile, C. Rossi (Eds.), Bioenergy Systems for the Future, Woodhead Publishing2017, pp. 3-42.
[2] N. Chimres, S. Wongwises, Critical review of the current status of solar energy in Thailand, Renew. Sust. Energ. Rev. 58 (2016) 198-207.
[3] S. Prasertsan, P. Prasertsan, Biomass residues from palm oil mills in Thailand: An overview on quantity and potential usage, Biomass Bioenergy 11(5) (1996) 387-395.
[4] S.H. Shuit, K.T. Tan, K.T. Lee, A.H. Kamaruddin, Oil palm biomass as a sustainable energy source: A Malaysian case study, Energy 34(9) (2009) 1225-1235.
[5] S. Prasertsan, B. Sajjakulnukit, Biomass and biogas energy in Thailand: Potential, opportunity and barriers, Renew. Energy 31(5) (2006) 599-610.
[6] R. Bader, W. Lipiński, Solar Thermochemical Processes, Sol Energy, pp. 345-394.
[7] S. Chuayboon, S. Abanades, S. Rodat, Insights into the influence of biomass feedstock type, particle size and feeding rate on thermochemical performances of a continuous solar gasification reactor, Renew. Energy 130 (2019) 360-370.
[8] H. Boujjat, S. Rodat, S. Chuayboon, S. Abanades, Experimental and numerical study of a directly irradiated hybrid solar/combustion spouted bed reactor for continuous steam gasification of biomass, Energy 189 (2019) 116118.
[9] S. Chuayboon, S. Abanades, S. Rodat, Experimental analysis of continuous steam gasification of wood biomass for syngas production in a high-temperature particle-fed solar reactor, Chem. Eng. Process. 125 (2018) 253-265.
[10] Q. Bellouard, S. Abanades, S. Rodat, N. Dupassieux, Solar thermochemical gasification of wood biomass for syngas production in a high-temperature continuously-fed tubular reactor, Int. J. Hydrogen Energy 42(19) (2017) 13486-13497.
[11] J.L.J. Ling, E.S. Go, Y.-K. Park, S.H. Lee, Recent advances of hybrid solar-Biomass thermo-chemical conversion systems, Chemosphere 290 (2022) 133245.
[12] P. Lichty, C. Perkins, B. Woodruff, C. Bingham, A. Weimer, Rapid High Temperature Solar Thermal Biomass Gasification in a Prototype Cavity Reactor, J. Sol. Energy Eng. 132(1) (2010).
[13] M.A.A. Mohammed, A. Salmiaton, W.A.K.G. Wan Azlina, M.S. Mohamad Amran, Gasification of oil palm empty fruit bunches: A characterization and kinetic study, Bioresour. Technol. 110 (2012) 628-636.

[14] R. Omar, A. Idris, R. Yunus, K. Khalid, M.I. Aida Isma, Characterization of empty fruit bunch for microwave-assisted pyrolysis, Fuel 90(4) (2011) 1536-1544.
[15] S. Chuayboon, S. Abanades, Continuous solar-driven gasification of oil palm agricultural bio waste for high-quality syngas production, Waste Manage. (Oxford) 154 (2022) 303-311.
[16] K. Zeng, D. Gauthier, J. Soria, G. Mazza, G. Flamant, Solar pyrolysis of carbonaceous feedstocks: A review, Sol Energy 156 (2017) 73-92.
[17] A.H. Rony, D. Mosiman, Z. Sun, D. Qin, Y. Zheng, J.H. Boman, M. Fan, A novel solar powered biomass pyrolysis reactor for producing fuels and chemicals, J. Anal. Appl. Pyrolysis 132 (2018) 19-32.

Estimation Methods for the Difference between Variances of Birnbaum-Saunders Distributions Containing Zero Values with Application to Wind Speed Data in Thailand.

Natchaya Ratasukharom[1, a)], Sa-Aat Niwitpong[1, b)] and Suparat Niwitpong[1, c)]

[1]Department of Applied Statistics, Faculty of Applied Sciences,
King Mongkut's University of Technology North Bangkok, Thailand.

[a)] natchaya.r@email.kmutnb.ac.th
[b)] sa-aat.n@sci.kmutnb.ac.th
[c)] Corresponding author: suparat.n@sci.kmutnb.ac.th

Abstract. Wind power is a renewable energy source that occurs naturally and is classified as clean energy, as it does not cause environmental pollution. One significant benefit of wind energy is its capacity to generate electricity through wind speed. Since wind speed data contain both zero and positive values, they are suitable for fitting a delta-Birnbaum-Saunders distribution. However, wind speed dispersion may vary in different climate conditions, leading to differences in variances. Therefore, the comparison of variances can be used to evaluate wind speed dispersion in different environments. To estimate this, credible intervals are constructed using the generalized confidence interval and bootstrap confidence interval methods. These approaches rely on variance stabilizing transformations and the Wilson method. The performance of the proposed methods was evaluated using coverage probabilities and average widths through Monte Carlo simulations. The results indicate that the generalized confidence interval based on the Wilson method outperformed the other methods. Wind speed datasets from Pattaya City and Laem Chabang, Chonburi Province, Thailand, are utilized to demonstrate the practical application of the proposed credible intervals.

***Index Terms*— confidence interval, Birnbaum–Saunders (BirSau) distribution, variance, wind speed.**

I. INTRODUCTION

Wind is a form of renewable energy that can be transformed from its original state into other forms such as thermal energy, electrical energy, and mechanical energy [1]. This makes wind an essential alternative to environmentally harmful fossil fuels, with the advantage of being inexhaustible. The conversion of wind energy into electricity is effectively accomplished through wind turbines, which require steady wind speeds to operate efficiently. Thailand boasts areas with high potential for wind energy production, especially those near the coast or in regions with consistently strong winds, like the area around Laem Chabang port and the wind turbine fields on Koh Larn Island in Chonburi Province. Research by Leiva et al. [2] and Mohammadi et al. [3] has shown that the wind speed data fit well with the Birnbaum-Saunders (BirSau) distribution. However, due to the natural variability of wind speeds, inconsistencies arise, leading to instances when the wind speed may drop to zero. In such cases, wind speed data are more accurately described by the delta-Birnbaum-Saunders (delta-BirSau) distribution. Therefore, in this research, we are interested in studying the delta-BirSau distribution.

A delta- BirSau distribution can be applied to data having a mixture of non-negative and zero observations: the zero value are binomially distributed with proportion of zeros (δ), while the positive values follow BirSau distribution with shape (α) and scale (β) parameters. The Birnbaum-Saunders (BirSau) distribution has found applications across various fields, particularly in environmental studies including rainfall patterns [4], air quality assessments [5], and wind energy analysis [6]. Many researchers are interested in studying how to estimate the parameters of the BirSau distribution. For instance, Birnbaum and Saunders [7] and Engelhardt et al. [8] presented the maximum likelihood estimators (MLEs) for α and β. Ng et al. [9] introduced modified moment estimators (MMEs) for α and and β utilized the Jackknife resampling method to reduce biases in both the MLEs and MMEs. Wang [10] investigated the application of the generalized confidence interval (GCI) for estimating parameter α, while Sun [11] proposed a method for parameter β. Several researchers developed various methods for estimating the δ parameter when dealing with a binomial distribution. For instance, Wu and Hsieh [12] used the generalized pivotal quantity (GPQ) based on the variance stabilized transformation (VST) to create intervals of δ. Wilson [13] proposed an interval technique that makes use of δ the Wilson interval. In statistics, numerous parameters of interest exist. However, since we cannot find their values from the entire population, we conclude population characteristics. These conclusions stem from calculations based on a random sample, aiding in the estimation of population parameters.

In probability and statistical inference, two main methods

are used for estimating parameters: point estimation and confidence interval (CI). Among these, the CI method is widely preferred for its ability to provide a range of values rather than a single point estimate. There are several parameters of statistical interest, and one of them is variance. Variance is a crucial parameter in probability and statistical inference. It is significant because it measures the spread or variability of a distribution, serves as the second central moment, and is essential for understanding distribution variability. This importance is underscored by its positive square root, known as the standard deviation [14]. In addition to variance, another interesting parameter to consider is the difference between variances. This parameter is regarded as the focal point for comparing the difference between the variances is the result of subtracting or comparing two variances. To compare the variance between the two population groups and assess their differences. If they exhibit significant disparity, it implies that the population from group one is considerably more dispersed than that of group two. There are numerous studies that investigate the estimation of CIs for the difference between variances, as follows: Mathew and Webb [15] examined CIs for variance components using the concepts of p-value and GCI. Cojbasic and Tomovic [16] proposed CIs for the variance and difference between variances of exponential distributions based on the ordinary t-statistics combined with the bootstrap method. Niwitpong [17] developed CI for the difference of the variances of two normal populations based on the closed form method of variance estimation. Maneerat et al. [18] estimated the CIs for the difference between delta-lognormal distributions using the highest posterior density based reference and probability-matching priors. Puggard et al. [19] presented CIs for the variance and difference of variances of BirSau distributions by using the bootstrap CI, the GCI, the Bayesian confidence interval, and the highest posterior density interval. Kooriphan et al. [20] studied CI for the difference between variances of delta-gamma distributions by using various Bayesian and highest posterior density methods based on the Jeffrey's, uniform, or normal-gamma-beta priors and compared with the fiducial quantity approach.

Although there have been several advancements in theoretical and methodological approaches to constructing CIs, the disparities in variances of delta-BirSau distributions have not yet been explored in statistical inference studies. To fill this research gap, we introduce CIs for the difference of variances of delta-BirSau distributions. Therefore, the objective of this study is to compare the efficiency of methods for estimating the CIs of the difference between the variances when the population follows a delta-BirSau distributions. The methods utilized include the GCI and bootstrap confidence interval (BCI) methodologies. These are based on δ estimation using the VST and Wilson techniques. We utilized these methods to compute 95% CIs for difference of variances in delta-BirSau distributions. The effectiveness of our proposed CIs was assessed through coverage probabilities (CPs) and average widths (AWs). Moreover, we applied the proposed methods to the wind speed data collected in Chonburi Province, Thailand, specifically in Pattaya City and Laem Chabang, during the period from September to November 2023 to illustrate their effectiveness.

II. METHODOLOGY

Let $X_{ij}=(X_{i1},X_{i2},\ldots,X_{in_i})$, $i=1,2$ and $j=1,2,\ldots,n_i$ be a random sample from the delta-BirSau distribution, denoted as $X_{ij} \sim delta-BirSau(\alpha_i,\beta_i,\delta_i)$, with shape parameter α_i, scale parameter β_i, and the proportion of zero values δ_i. The distribution function X_{ij} is introduced by Atichison [21] and Birnbaum-Saunders [22] as

$$F(x_{ij};\alpha_i,\beta_i,\delta_i)=\begin{cases}\delta_i & ;x_{ij}=0\\ \delta_i+(1-\delta_i)H(x_{ij};\alpha_i,\beta_i) & ;x_{ij}>0\end{cases}, \quad (1)$$

where $\delta_i=P(X_{ij}=0)$, $H(X_{ij};\alpha_i,\beta_i)$ is a BirSau distribution function, denoted as H_{ij} when $x_{ij}\neq 0$. The zero observations follow a binomial distribution denoted as $n_{i(0)} \sim bi(n_i,\delta_i)$; $n_i=n_{i(0)}+n_{i(1)}$, where $n_{i(0)}$ is the number of zero values and $n_{i(1)}$ is a number of non-zero values. The maximum likelihood estimator of δ_i is $\hat{\delta}_i=n_{i(0)}/n_i$. The population variance of delta-BirSau distribution is

$$\eta_i=(1-\delta_i)(\beta_i^2)\left[\alpha_i^2\left(1+\frac{5}{4}\alpha_i^2\right)+\delta_i\left(1+\frac{1}{2}\alpha_i^2\right)^2\right]. \quad (2)$$

For focusing on the parameter of interest, the difference of variances is defined as

$$\begin{aligned}\nu&=\eta_1-\eta_2\\ &=(1-\delta_1)(\beta_1^2)\left[\alpha_1^2\left(1+\frac{5}{4}\alpha_1^2\right)+\delta_1\left(1+\frac{1}{2}\alpha_1^2\right)^2\right]\\ &\quad-(1-\delta_2)(\beta_2^2)\left[\alpha_2^2\left(1+\frac{5}{4}\alpha_2^2\right)+\delta_2\left(1+\frac{1}{2}\alpha_2^2\right)^2\right].\end{aligned} \quad (3)$$

To construct the CIs for difference between variances, the following concepts are elaborate.

A. Generalized confidence interval

A GPQ is defined by Weerahandi [23] and has a distribution that is independent of any unknown parameters. Additionally, this GPQ's observed value does not depend on nuisance parameters. In the overarching theory of this method, it has been established that GPQ are allowed to be a function of nuisance parameters. In contrast, conventional

pivotal quantities can only be a function of the sample and the parameter of interest.

Let $X_{ij}=(X_{i1},X_{i2},...,X_{in_i})$ be an independent random sample from the BirSau distributions. Sun [11] and Wang [10] first presented the idea of GPQ for β_i and α_i :

$$G_{\beta_i}:=G_{\beta_i}(x_{ij};G_i)=\begin{cases}\max(\beta_{i1},\beta_{i2}), & T_i\le 0;\\ \min(\beta_{i1},\beta_{i2}), & T_i>0,\end{cases} \tag{4}$$

and

$$G_{\alpha_i}:=G_{\alpha_i}(x_{ij};v_i,G_i)=\left[\frac{S_{i2}{G_{\beta_i}}^2-2n_iG_{\beta_i}+S_{i1}}{G_{\beta_i}\tau_i}\right]^{1/2}, \tag{5}$$

where $x_{ij}=(x_{i1},x_{i2},...,x_{in_i})$ are the observed values of X_{ij}, $G_{\beta_i}(x_{ij};G_i)$ is defined in (4) and $T_i \square t(n_i-1)$ refers to t-distribution with a degree of freedom, $S_{i1}=\sum_{j=1}^{n_i}x_{ij}$, $S_{i2}=\sum_{j=1}^{n_i}1/x_{ij}$, $\tau_i \square \chi^2_{(n_i)}$. By utilizing (4), we can obtain two solutions, β_{i1} and β_{i2} for

$$[(n_i-1)K_i{}^2-\frac{1}{n_i}M_iG_i{}^2]\beta_i{}^2-2[(n_i-1)N_iK_i-(1-N_iK_i)G_i{}^2]\beta_i + (n_i-1)N_i{}^2-\frac{1}{n_i}O_iG_i{}^2=0, \tag{6}$$

where $K_i=n_i^{-1}\sum_{j=1}^{n_i}1/\sqrt{X_{ij}}$, $M_i=\sum_{j=1}^{n_i}(1/\sqrt{X_{ij}}-K_i)^2$, $N_i=n_i^{-1}\sum_{j=1}^{n_i}\sqrt{X_{ij}}$, and $O_i=\sum_{j=1}^{n_i}(\sqrt{X_{ij}}-N_i)^2$.

The generalized confidence interval based on VST (GCI-VST)

The CPs of VST were substantially better than the Wald interval, as demonstrated by Dasgupta [24]. Wu and Hsieh [12] subsequently used the GPQ based on VST to create intervals, as

$$G_{\delta_i.VST}=\sin^2(\arcsin\sqrt{\hat{\delta}_i}-\frac{T_i}{2\sqrt{n_i}}), \tag{7}$$

where $T_i=2\sqrt{n_i}(\arcsin\sqrt{\hat{\delta}_i}-\arcsin\sqrt{\delta_i}) \square N(0,1)$. By utilizing the three pivots from (4), (5), and (7), the GPQs ν based on the VST method are defined as

$$G_{\nu.GCI.VST}=(1-G_{\delta_1.VST})(G_{\beta_1}{}^2)\left[G_{\alpha_1}^2\left(1+\frac{5}{4}G_{\alpha_1}^2\right)+G_{\delta_1.VST}\left(1+\frac{1}{2}G_{\alpha_1}^2\right)^2\right] -(1-G_{\delta_2.VST})(G_{\beta_2}{}^2)\left[G_{\alpha_2}^2\left(1+\frac{5}{4}G_{\alpha_2}^2\right)+G_{\delta_2.VST}\left(1+\frac{1}{2}G_{\alpha_2}^2\right)^2\right]. \tag{8}$$

Therefore, the $100(1-\gamma)\%$ two-sided CI for ν based on GCI-VST is

$$CI_{\nu.GCI.VST}=[L_{\nu.GCI.VST},U_{\nu.GCI.VST}] =[G_{\nu.GCI.VST}(\gamma/2),G_{\nu.GCI.VST}(1-\gamma/2)], \tag{9}$$

where $G_{\nu.GCI.VST}(\gamma/2)$ denote the $(\gamma/2)100$th percentile of $G_{\nu.GCI.VST}$.

The generalized confidence interval based on Wilson (GCI-W)

The interval that Wilson [13] suggested for the binomial is commonly referred to as the Wilson interval. The following procedure results in an approximation GPQ for δ_i when using the approximate result.

$$G_{\delta_i.W}=\frac{n_{i(0)}+Z_{i.W}^2/2}{n_i+Z_{i.W}^2}-\frac{Z_{i.W}}{n_i-Z_{i.W}^2}\left[n_{i(0)}\left(1-\frac{n_{i(0)}}{n_i}\right)+\frac{Z_{i.W}^2}{4}\right]^{1/2}, \tag{10}$$

where $Z_{i.W}=\left(N_{i(0)}-n_i\delta_i\right)/\left(\sqrt{n_i\delta_i(1-\delta_i)}\right) \square N(0,1)$ and $N_{i(0)}$ be a binomial variable following $bi(n_i,\delta_i)$, when $N_0=n_0$ the observed value $G_{\delta.W}$ is equal to δ . By (4), (5), and (10), the GPQ of ν between the variances based on the Wilson method are given by

$$G_{\nu.GCI.W}=(1-G_{\delta_1.W})(G_{\beta_1}{}^2)\left[G_{\alpha_1}^2\left(1+\frac{5}{4}G_{\alpha_1}^2\right)+G_{\delta_1.W}\left(1+\frac{1}{2}G_{\alpha_1}^2\right)^2\right] -(1-G_{\delta_2.W})(G_{\beta_2}{}^2)\left[G_{\alpha_2}^2\left(1+\frac{5}{4}G_{\alpha_2}^2\right)+G_{\delta_2.W}\left(1+\frac{1}{2}G_{\alpha_2}^2\right)^2\right]. \tag{11}$$

Therefore, the $100(1-\gamma)\%$ two-sided CI for ν based on GCI-W is

$$CI_{\nu.GCI.W}=[L_{\nu.GCI.W},U_{\nu.GCI.W}] =[G_{\nu.GCI.W}(\gamma/2),G_{\nu.GCI.W}(1-\gamma/2)]. \tag{12}$$

where $G_{\nu.GCI.W}(\gamma/2)$ denote the $(\gamma/2)100$th percentile of $G_{\nu.GCI.W}$.

B. Bootstrap confidence interval

Efron and Tibshirani [25] introduced the fundamental principles of bootstrapping. Ng et al. [9] showed that when employing Monte Carlo simulation, the MLEs of shape and scale parameters tend to be biased. Later, Lemonte et al. [26] discovered that the constant-bias-correcting (CBC) bootstrap method, initially proposed by Mackinnon and Smith [27], provided the most effective means of reducing bias. The process of constructing a CI for the difference between variances using the CBC method can be explained as follows.

Consider $\mathbf{x}_{ij}=(x_{i1},x_{i2},...,x_{in_i})^T$ as a random sample of size n_i from the $BirSau(\alpha_i,\beta_i)$ distribution, where $\hat{\alpha}_i$, and $\hat{\beta}_i$ represent the MLEs of parameters α_i and β_i. Subsequently, a bootstrap sample is independently generated from the original sample $\mathbf{x}_{ij}$. The Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton nonlinear optimization algorithm is utilized to compute the corresponding bootstrapped estimations $\hat{\alpha}_i$ and $\hat{\beta}_i$ (represented as $\hat{\alpha}_i^* = (\hat{\alpha}_{i,1}^*,\hat{\alpha}_{i,2}^*,...,\hat{\alpha}_{i,B_i}^*)$) and $\hat{\beta}_i^* = (\hat{\beta}_{i,1}^*,\hat{\beta}_{i,2}^*,...,\hat{\beta}_{i,B_i}^*)$) based on B_i replications.

The approximate bootstrap estimators are calculated by

$$\hat{\alpha}_i^{*(\cdot)} = (1/B_i)\sum_{b_i=1}^{B_i}\hat{\alpha}_{i,b_i}^* \text{ and } \hat{\beta}_i^{*(\cdot)} = (1/B_i)\sum_{b_i=1}^{B_i}\hat{\beta}_{i,b_i}^*. \quad (13)$$

Given the estimators $\hat{\alpha}_i$ and $\hat{\beta}_i$, their bias is given by

$$B(\hat{\alpha}_i,\alpha_i) = E(\hat{\alpha}_i) - \alpha_i \text{ and } B(\hat{\beta}_i,\beta_i) = E(\hat{\beta}_i) - \beta_i. \quad (14)$$

The bootstrap bias estimates are based on the bootstrap sample of $\hat{\alpha}_i$ and $\hat{\beta}_i$ can be written as

$$\hat{B}(\hat{\alpha}_i,\alpha_i) = \hat{\alpha}_i^{*(\cdot)} - \hat{\alpha}_i \text{ and } \hat{B}(\hat{\beta}_i,\beta_i) = \hat{\beta}_i^{*(\cdot)} - \hat{\beta}_i. \quad (15)$$

Based on the concept of CBC estimates first proposed by Mackinnon and Smith [27], the bias-corrected estimator that follows is computed. The method used to calculate these estimates is as follows:

$$\bar{\alpha}_{i,b}^* = \hat{\alpha}_{i,b}^* - 2\hat{B}(\hat{\alpha}_i,\alpha_i) = 2\hat{\alpha}_i - \hat{\alpha}_i^{*(\cdot)}, \quad b=1,2,...,B, \quad (16)$$

and

$$\bar{\beta}_{i,b}^* = \hat{\beta}_{i,b}^* - 2\hat{B}(\hat{\beta}_i,\beta_i) = 2\hat{\beta}_i - \hat{\beta}_i^{*(\cdot)}, \quad b=1,2,...,B. \quad (17)$$

The bootstrap confidence interval based on VST (BCI-VST)

By applying equations (7), (16) and (17) to the VST method, the bootstrap estimator of ν is defined as

$$\begin{aligned} G_{\nu.BCI.VST} &= (1-G_{\delta_1.VST})(\bar{\beta}_{1,b}^{*\,2})\left[\bar{\alpha}_{1,b}^{*\,2}\left(1+\frac{5}{4}\bar{\alpha}_{1,b}^{*\,2}\right)+G_{\delta_1.VST}\left(1+\frac{1}{2}\bar{\alpha}_{1,b}^{*\,2}\right)^2\right] \\ &\quad -(1-G_{\delta_2.VST})(\bar{\beta}_{2,b}^{*\,2})\left[\bar{\alpha}_{2,b}^{*\,2}\left(1+\frac{5}{4}\bar{\alpha}_{2,b}^{*\,2}\right)+G_{\delta_2.VST}\left(1+\frac{1}{2}\bar{\alpha}_{2,b}^{*\,2}\right)^2\right]. \end{aligned} \quad (18)$$

Therefore, the $100(1-\gamma)\%$ two-sided CI for ν based on BCI-VST is

$$\begin{aligned} CI_{\nu.BCI.VST} &= [L_{\nu.BCI.VST}, U_{\nu.BCI.VST}] \\ &= [G_{\nu.BCI.VST}(\gamma/2), G_{\nu.BCI.VST}(1-\gamma/2)], \end{aligned} \quad (19)$$

where $G_{\nu.BCI.VST}(\gamma/2)$ denote the $(\gamma/2)100$th percentile of $G_{\nu.BCI.VST}$.

The bootstrap confidence interval based on Wilson (BCI-Wilson)

The Wilson [13] method based bootstrap estimator in equations (10), (16), and (17), can be expressed as follows.

$$\begin{aligned} G_{\nu.BCI.W} &= (1-G_{\delta_1.W})(\bar{\beta}_{1,b}^{*\,2})\left[\bar{\alpha}_{1,b}^{*\,2}\left(1+\frac{5}{4}\bar{\alpha}_{1,b}^{*\,2}\right)+G_{\delta_1 W}\left(1+\frac{1}{2}\bar{\alpha}_{1,b}^{*\,2}\right)^2\right] \\ &\quad -(1-G_{\delta_2.W})(\bar{\beta}_{2,b}^{*\,2})\left[\bar{\alpha}_{2,b}^{*\,2}\left(1+\frac{5}{4}\bar{\alpha}_{2,b}^{*\,2}\right)+G_{\delta_2.W}\left(1+\frac{1}{2}\bar{\alpha}_{2,b}^{*\,2}\right)^2\right]. \end{aligned} \quad (20)$$

Therefore, the $100(1-\gamma)\%$ two-sided CI for ν based on BCI-W is

$$\begin{aligned} CI_{\nu.BCI.W} &= [L_{\nu.BCI.W}, U_{\nu.BCI.W}] \\ &= [G_{\nu.BCI.W}(\gamma/2), G_{\nu.BCI.W}(1-\gamma/2)], \end{aligned} \quad (21)$$

where $G_{\nu.BCI.W}(\gamma/2)$ denote the $(\gamma/2)100$th percentile of $G_{\nu.BCI.W}$.

III. Simulation Studies and Results

This study employed R statistical software to assess four estimation methods: GCI-VST, GCI-W, BCI-VST, and BCI-W, aimed to constructed CIs for the variance differences between two delta-BirSau distributions. The evaluation criteria for these methods included their CPs and AWs. Generally, the preferred CIs are those with CPs greater than or closest to the nominal confidence level of 0.95 and by possessing the shortest AWs. The data were generated from two independent delta-BirSau distributions, $x_1 \sim delta-BirSau(\alpha_1,\beta_1,\delta_1)$ and $x_2 \sim delta-BirSau(\alpha_2,\beta_2,\delta_2)$, with sample sizes of $(n_1,n_2)=$ (30,30), (50,50), (70,70) or (100,100). The probability of zeros was set to $(\delta_1,\delta_2)=$ (0.1,0.1), (0.1,0.3), (0.1,0.5), (0.3,0.3), (0.3,0.5) or (0.5,0.5), and the shape parameters (α_1,α_2) were configured as (0.5,0.5), (0.5,1) or (1,1), respectively. Lastly, the scale parameters (β_1,β_2) were fixed at (1,1) for all cases without loss of generality was applied in this simulation study. For all simulations, the number of replications was set at 2,000, with 5,000 pivotal quantities used for GCI-VST, GCI-W and $B=$ 500 for BCI-VST and BCI-W. This determination aligns with the research conducted by Puggard et al. [28].

The results in Table I reveal that the GCI-VST and GCI-W methods for the equivalent CI outperformed the others because the CPs were consistently greater than or close to the

nominal confidence level of 0.95 in most cases. Contrastingly, the BCI method exhibited CPs lower than 0.95 in all cases. Next, an examination of the AWs is warranted. It is evident that both the BCI-VST and BCI-W methods boast the shortest AW values. However, given the consistent CPs of these methods falling below the nominal confidence level of 0.95, they are deemed unsuitable for consideration. Upon comparing the AW values between the GCI-ST and GCI-W methods, it is discerned that the GCI-W method exhibits a shorter AW value than GCI-VST.

The CPs and AWs summary is also shown in Fig. 1 and 2, respectively. For both CIs for the difference of two variances, the AWs decreased as sample sizes increased. Conversely, an increase in the shape parameters led to an expansion of the AWs across all CIs. Overall, GCI-W emerged as the top performer in the simulation study, meeting the criteria of both requirements effectively.

This study aligns with the research of Puggard et al. [19, 28], who investigated the CIs for the variance, difference between the variances, and the coefficient of variation. From their study, it was observed that the CPs of GCI exceeded or closely approached the nominal confidence level of 0.95. The conclusions of this research are consistent with previous studies by Thangjai et al. [29-31] indicating that the GCI method is effective in estimating CIs.

TABLE I
THE COVERAGE PROBABILITIES AND AVERAGE WIDTHS OF NORMAL 95% TWO-SIDED CONFIDENCE INTERVAL FOR DIFFERENCE BETWEEN TWO VARIANCES OF DELTA-BIRSAU DISTRIBUTIONS.

n_1	n_2	δ_1	δ_2	α_1	α_2	Coverage probability (Average width)			
						GCI-VST	GCI-W	BCI-VST	BCI-W
30	30	0.1	0.1	0.5	0.5	**0.9520** *(1.1400)*	0.9330 (1.1320)	0.9250 (0.8032)	0.9365 (0.8239)
				0.5	1	**0.9540** (8.5900)	**0.9535** *(8.4720)*	0.9050 (5.3930)	0.9035 (5.3860)
				1	1	**0.9490** (14.7300)	**0.9475** *(14.5100)*	0.9320 (8.6680)	0.9320 (8.6340)
		0.1	0.3	0.5	0.5	**0.9500** *(1.2740)*	0.9380 (1.2500)	0.9180 (0.8389)	0.9135 (0.8475)
				0.5	1	**0.9460** *(10.8400)*	0.9430 (10.4500)	0.8970 (5.5310)	0.8920 (5.4850)
				1	1	**0.9450** (16.6800)	**0.9450** *(16.3300)*	0.9315 (8.7720)	0.9265 (8.7610)
		0.1	0.5	0.5	0.5	**0.9525** (1.5350)	**0.9465** *(1.4690)*	0.9140 (0.8797)	0.9160 (0.8934)
				0.5	1	**0.9515** (16.2500)	**0.9475** *(15.0000)*	0.8775 (5.4050)	0.8800 (5.4320)
				1	1	**0.9535** (22.6900)	**0.9490** *(21.2200)*	0.9375 (8.7600)	0.9345 (8.7450)
		0.3	0.3	0.5	0.5	**0.9545** (1.4290)	**0.9500** *(1.3860)*	0.9180 (0.8872)	0.9280 (0.8831)
				0.5	1	**0.9505** (11.1800)	**0.9500** *(10.7100)*	0.8970 (5.5900)	0.8990 (5.5530)
				1	1	0.9430 (18.2000)	0.9395 (17.6400)	0.9160 (8.6780)	0.9215 (8.6690)
		0.3	0.5	0.5	0.5	**0.9550** (1.6510)	**0.9550** *(1.5730)*	0.9155 (0.9085)	0.9225 (0.9121)
				0.5	1	**0.9550** (17.9100)	**0.9510** *(16.0800)*	0.8865 (5.4760)	0.8860 (5.4480)
				1	1	**0.9585** (25.9500)	**0.9545** *(23.3300)*	0.9290 (8.6820)	0.9275 (8.6990)
		0.5	0.5	0.5	0.5	**0.9595** (1.9250)	**0.9510** *(1.8000)*	0.9160 (0.9607)	0.9195 (0.9674)
				0.5	1	**0.9490** (18.0300)	**0.9485** *(16.4000)*	0.8810 (5.4610)	0.8850 (5.4260)
				1	1	**0.9545** (29.0200)	**0.9470** *(26.5100)*	0.9190 (8.8800)	0.9140 (8.8880)
50	50	0.1	0.1	0.5	0.5	**0.9550** *(0.7305)*	0.9415 (0.7281)	0.9315 (0.5839)	0.9410 (0.6040)
				0.5	1	**0.9445** (5.1470)	**0.9475** *(5.1050)*	0.9170 (3.9020)	0.9155 (3.9110)
				1	1	**0.9550** (8.0540)	**0.9545** *(8.0010)*	0.9415 (5.8710)	0.9420 (5.8810)
		0.1	0.3	0.5	0.5	**0.9510** *(0.8064)*	0.9395 (0.7981)	0.9260 (0.6128)	0.9240 (0.6233)
				0.5	1	**0.9555** (5.6160)	**0.9535** *(5.5320)*	0.9095 (3.8070)	0.9120 (3.8230)

TABLE I
(CONTINUE)

n_1	n_2	δ_1	δ_2	α_1	α_2	Coverage probability (Average width)			
						GCI-VST	GCI-W	BCI-VST	BCI-W
50	50	0.1	0.3	1	1	**0.9470** (8.6740)	**0.9445** *(8.5610)*	0.9330 (5.9180)	0.9325 (5.9340)
		0.1	0.5	0.5	0.5	**0.9535** *(0.8835)*	0.9405 (0.8654)	0.9200 (0.6324)	0.9230 (0.6478)
				0.5	1	**0.9570** (6.8270)	**0.9480** *(6.6450)*	0.9100 (3.8010)	0.9125 (3.8500)
				1	1	**0.9495** (9.4310)	**0.9500** *(9.2140)*	0.9320 (5.7430)	0.9330 (5.7860)
		0.3	0.3	0.5	0.5	**0.9575** (0.8797)	**0.9575** *(0.8656)*	0.9210 (0.6459)	0.9285 (0.6480)
				0.5	1	**0.9540** (5.7060)	**0.9520** *(5.5960)*	0.9175 (3.8500)	0.9165 (3.8570)
				1	1	**0.9545** (9.2400)	**0.9510** *(9.1080)*	0.9365 (5.9110)	0.9370 (5.9430)
		0.3	0.5	0.5	0.5	**0.9545** (0.9529)	**0.9480** *(0.9276)*	0.9140 (0.6624)	0.9155 (0.6681)
				0.5	1	0.9405 (6.8290)	0.9400 (6.5790)	0.9015 (3.7540)	0.9115 (3.7890)
				1	1	**0.9560** (10.0980)	**0.9505** *(9.7990)*	0.9340 (5.8340)	0.9345 (5.8760)
		0.5	0.5	0.5	0.5	**0.9510** (1.0266)	**0.9480** *(0.9936)*	0.9085 (0.6799)	0.9110 (0.6918)
				0.5	1	**0.9600** (6.6190)	**0.9525** *(6.4810)*	0.9090 (3.7100)	0.9125 (3.7760)
				1	1	**0.9515** *(11.2700)*	0.9440 (10.8100)	0.9210 (5.8780)	0.9215 (5.9430)
70	70	0.1	0.1	0.5	0.5	**0.9480** *(0.5764)*	0.9380 (0.5752)	0.9355 (0.4847)	0.9415 (0.5027)
				0.5	1	**0.9490** (3.9050)	**0.9465** *(3.8850)*	0.9185 (3.1740)	0.9210 (3.1820)
				1	1	**0.9485** (6.0190)	**0.9470** *(5.9830)*	0.9340 (4.7370)	0.9355 (4.7500)
		0.1	0.3	0.5	0.5	**0.9500** *(0.6316)*	0.9425 (0.6276)	0.9235 (0.5103)	0.9285 (0.5208)
				0.5	1	**0.9455** *(4.0770)*	0.9440 (4.0340)	0.9080 (3.0340)	0.9130 (3.0610)
				1	1	0.9395 (6.3310)	0.9395 (6.2750)	0.9255 (4.7590)	0.9290 (4.7800)
		0.1	0.5	0.5	0.5	**0.9550** *(0.6731)*	0.9410 (0.6647)	0.9265 (0.5234)	0.9305 (0.5382)
				0.5	1	**0.9475** *(4.5700)*	0.9430 (4.4730)	0.9125 (2.9880)	0.9100 (3.0290)
				1	1	**0.9465** (6.6670)	**0.9450** *(6.5490)*	0.9305 (4.6460)	0.9325 (4.6840)
		0.3	0.3	0.5	0.5	**0.9570** (0.6686)	**0.9585** *(0.6611)*	0.9325 (0.5253)	0.9320 (0.5281)
				0.5	1	**0.9545** (4.1430)	**0.9500** *(4.1040)*	0.9290 (3.0940)	0.9280 (3.1150)
				1	1	**0.9540** (6.5690)	**0.9555** *(6.5110)*	0.9290 (4.6880)	0.9325 (4.7260)
		0.3	0.5	0.5	0.5	**0.9545** (0.7102)	**0.9465** *(0.7000)*	0.9290 (0.5362)	0.9290 (0.5439)
				0.5	1	**0.9525** (4.4670)	**0.9490** *(4.4000)*	0.9145 (2.9270)	0.9165 (2.9820)
				1	1	**0.9465** (6.9650)	**0.9480** *(6.8580)*	0.9170 (4.6380)	0.9245 (4.6870)
		0.5	0.5	0.5	0.5	**0.9500** *(0.7631)*	0.9430 (0.7487)	0.9215 (0.5579)	0.9275 (0.5691)
				0.5	1	**0.9500** *(4.4820)*	0.9375 (4.4120)	0.9010 (2.9410)	0.9025 (3.0040)
				1	1	**0.9530** *(7.2260)*	0.9435 (7.1010)	0.9245 (4.5150)	0.9240 (4.5930)
100	100	0.1	0.1	0.5	0.5	**0.9550** *(0.4516)*	0.9410 (0.4509)	0.9405 (0.3946)	0.9415 (0.4109)
				0.5	1	**0.9525** (3.0470)	**0.9525** *(3.0330)*	0.9340 (2.6070)	0.9340 (2.6190)
				1	1	0.9425 (4.5610)	**0.9530** *(4.5480)*	0.9355 (3.8060)	0.9360 (3.8220)
		0.1	0.3	0.5	0.5	**0.9475** *(0.4917)*	0.9355 (0.4894)	0.9280 (0.4152)	0.9250 (0.4249)
				0.5	1	**0.9515** (3.1930)	**0.9510** *(3.1730)*	0.9280 (2.5390)	0.9295 (2.5610)
				1	1	**0.9545** (4.6860)	**0.9500** *(4.6570)*	0.9360 (3.7690)	0.9370 (3.7910)
		0.1	0.5	0.5	0.5	0.9425 (0.5125)	0.9335 (0.5082)	0.9175 (0.4199)	0.9215 (0.4334)
				0.5	1	**0.9525** (3.2440)	**0.9445** *(3.2140)*	0.9180 (2.3540)	0.9210 (2.4150)
				1	1	**0.9480** (4.8400)	**0.9465** *(4.8090)*	0.9325 (3.7090)	0.9330 (3.7600)
		0.3	0.3	0.5	0.5	**0.9475** *(0.5258)*	0.9440 (0.5218)	0.9205 (0.4314)	0.9190 (0.4344)
				0.5	1	**0.9550** (3.1580)	**0.9540** *(3.1280)*	0.9200 (2.5100)	0.9215 (2.5240)
				1	1	0.9440 (4.8180)	0.9440 (4.7850)	0.9215 (3.7240)	0.9240 (3.7530)
		0.3	0.5	0.5	0.5	**0.9540** (0.5493)	**0.9485** *(0.5448)*	0.9225 (0.4394)	0.9275 (0.4468)
				0.5	1	**0.9480** *(3.2750)*	0.9435 (3.2380)	0.9195 (2.3940)	0.9220 (2.4520)
				1	1	**0.9490** (4.9470)	**0.9450** *(4.9040)*	0.9275 (3.6490)	0.9310 (3.7140)
		0.5	0.5	0.5	0.5	**0.9525** *(0.5776)*	0.9395 (0.5702)	0.9165 (0.4512)	0.9205 (0.4616)
				0.5	1	**0.9505** *(3.3040)*	0.9405 (3.2690)	0.9150 (2.3970)	0.9210 (2.4530)
				1	1	**0.9505** *(5.0910)*	0.9405 (5.0160)	0.9230 (3.5700)	0.9240 (3.6440)

Notes: The CPs greater than the nominal confidence level of 0.95 are in bold, and the shortest AWs are in italics.

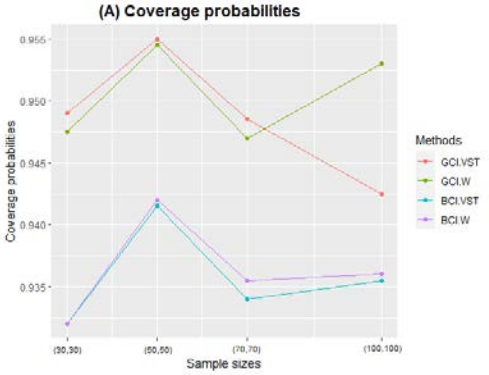


Fig. 1. Comparison of the CPs for estimation of the 95% CI for the difference between two variances of delta-BirSau distributions for various sample sizes: $(\delta_1, \delta_2) = (0.1, 0.1)$, $(\alpha_1, \alpha_2) = (1,1)$, and $(\beta_1, \beta_2) = (1,1)$.

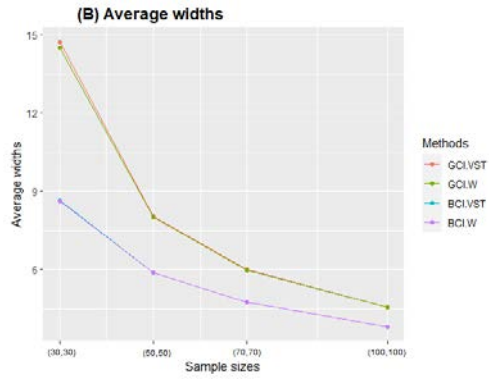


Fig. 2. Comparison of the AWs for estimation of the 95% CI for the difference between two variances of delta-BirSau distributions for various sample sizes: $(\delta_1, \delta_2) = (0.1, 0.1)$, $(\alpha_1, \alpha_2) = (1,1)$, and $(\beta_1, \beta_2) = (1,1)$.

IV. AN EMPIRICAL APPLICATION

Wind is a naturally occurring renewable energy that can be harnessed indefinitely, offering an alternative to fossil fuels, the current primary energy source. The combustion of fossil fuels leads to the emission of pollutants. Consequently, this study is centered on wind energy, with a specific focus on wind speed, as the wind speed can be utilized to generate electricity. Wind turbines convert wind energy into electrical energy, making wind speed crucial. The research examines wind speed data from Pattaya City and Laem Chabang in Chonburi Province, Thailand, regions known for their wind energy potential. These coastal areas experience solid and consistent winds, making them ideal locations for wind farms to generate electricity. Studies by Leiva et al. [2] and Mohammadi et al. [3] found that wind speed data are suitable for the BirSau distribution. This study analyzed average wind speed data (in knots) from Pattaya City and Laem Chabang, Chonburi Province, Thailand, from September to November 2023.

The wind speed data are displayed in Table II, and the density and Q-Q plots from Pattaya City and Laem Chabang are presented in Fig. 3 and 4,respectively. This analysis aimed to evaluate the effectiveness of CIs for estimating the difference between two variances of the delta-BirSau distribution. These CIs were derived using four methods: GCI-VST, GCI-W, BCI-VST and BCI-W. These wind speed data sets were obtained from the Thai meteorological department automatic weather system [32]. Given that the dataset includes zero and positive values, we can fit the positive values to distributions such as normal, exponential, logistic, Cauchy, uniform, or BirSau distributions. To determine the best fit, we utilized the Akaike information criterion (AIC) and Bayesian information criterion (BIC) criterions. The results presented in Table III indicate that the AIC and BIC values for the BirSau distribution were the smallest among the various distributions considered. This suggests that the BirSau distribution is the most suitable choice for modeling these datasets. Summary statistics were derived from the wind speed dataset from Pattaya City, which included values such as $n_1 = 91$, $n_{1(0)} = 8$, $n_{1(1)} = 83$, $\hat{\delta}_1 = 0.0879$, $\hat{\alpha}_1 = 0.7260$, $\hat{\beta}_1 = 2.8898$, and $\hat{\eta}_1 = 7.7830$. Similarly, for Laem Chabang, the values were $n_2 = 91$, $n_{2(0)} = 3$, $n_{2(1)} = 88$, $\hat{\delta}_2 = 0.0330$, $\hat{\alpha}_2 = 0.7005$, $\hat{\beta}_2 = 2.2492$, and $\hat{\eta}_2 = 4.1232$. Therefore, in this study, the parameter of interest, the difference between variances, is denoted as $\hat{\nu} = \hat{\eta}_1 - \hat{\eta}_2 = 3.66$.

The 95% CIs for the difference between two variances of the delta-BirSau distribution, calculated using both the GCI-VST, GCI-W, BCI-VST, and BCI-W methods, are presented in the Table IV. These intervals provide an estimate that closely matches the results obtained from simulations. When examining sample sizes of $n_1 = 100$ and $n_2 = 100$, with parameters $\delta_1 = 0.1$, $\delta_2 = 0.1$, $\alpha_1 = 1$, and $\alpha_2 = 1$, It was found that the BCI-VST, and BCI-W method provided the shortest AWs. However, from the simulation results, it was discovered that both BCI-VST, and BCI-W methods had CP values lower than the nominal confidence level of 0.95, excluding these two methods from consideration. Consequently, the GCI-VST, and GCI-W method was chosen. Subsequent analysis revealed that the GCI-W method yielded shorter AW values. Hence, the GCI-W method is considered appropriate for estimating the CI of the variance difference between the two wind speed datasets. It aims to elucidate the difference in the wind speed between Pattaya City station and Laem Chabang station. Understanding these differences is crucial for the industry involved in electricity production using wind energy, as consistent wind speed is essential for achieving optimal efficiency in electricity generation through wind turbines.

TABLE II
THE DAILY WIND SPEED DATA FROM LAEM CHABANG AND PATTAYA CITY, CHONBURI PROVINCE, THAILAND.

Wind Speed (Knots)																
Pattaya City									Laem Chabang							
2	3	0	1	2	0	3	1	7	6	5	2	3	1	1	2	2	5
3	3	2	1	1	1	3	2	5	5	4	1	2	1	1	5	1	2
2	3	1	1	1	2	2	3	4	6	2	1	3	1	1	4	0	2
3	2	1	2	2	1	1	3	4	4	3	2	4	0	0	2	1	2
4	2	2	1	2	0	0	4	6	5	4	1	4	0	0	2	2	3
2	2	3	2	2	1	0	5	5	5	7	5	3	3	3	1	1	4
2	2	2	3	1	1	0	6	5	8	7	5	5	6	6	2	1	5
1	1	2	2	1	1	1	8	5	6	9	5	3	3	3	3	7	3
2	2	2	1	0	1	2	5	7	1	4	3	3	2	2	3	6	5
3	1	2	0	2	1	2	8	7	2	2	1	1	4	4	3	5	4
10									3								

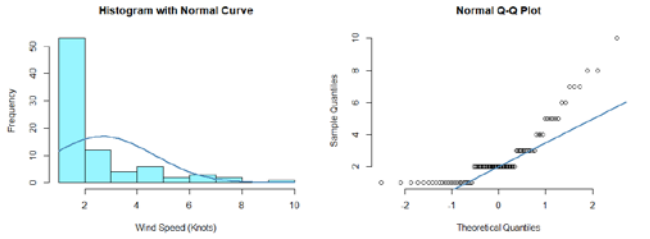


Fig. 3. The Histogram and Normal Q-Q plot of wind speed data from September to November in Pattaya City, Chonburi Province, Thailand.

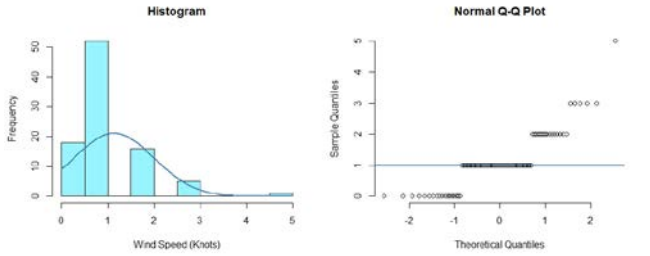


Fig. 4. The Histogram and Normal Q-Q plot of wind speed data from September to November in Laem Chabang, Chonburi Province, Thailand.

TABLE III
AIC AND BIC VALUES FOR FITTING THE POSITIVE WIND SPEED DATA FROM LAEM CHABANG AND PATTAYA CITY, CHONBURI PROVINCE, THAILAND.

Models	Laem Chabang		Pattaya City	
	AIC	BIC	AIC	BIC
Normal	369.50	374.50	248.88	353.71
Exponential	399.10	401.50	332.06	334.48
Logistic	327.80	377.80	339.03	343.86
Cauchy	406.70	411.60	332.61	337.44
Uniform	370.00	364.90	368.74	373.58
BirSau	**358.80**	**322.30**	**292.24**	**299.50**

TABLE IV
THE 95% CONFIDENCE INTERVALS FOR THE DIFFERENCE BETWEEN TWO VARIANCES OF WIND SPEED DATA FOR PATTAYA CITY AND LEAM CHABANG, CHONBURI PROVINCE, THAILAND.

Methods	Confidence Intervals for V		Width
	Lower	Upper	
GCI-VST	-0.9403	5.9050	6.8450
GCI-W	-1.0500	5.7900	6.8390
BCI-VST	-0.7154	5.1500	5.8650
BCI-W	-0.8296	5.0480	5.8770

V. CONCLUSIONS

We constructed CIs for the difference between the variances of two delta-BirSau distributions using GCI-VST, GCI-W, BCI-VST, and BCI-W methods. We evaluated the performance of these CIs based on their CPs and AWs. The simulation study results indicate that the CPs of GCI-VST, and GCI-W were consistently greater than or close to the nominal confidence level of 0.95 in nearly all cases. Additionally, the AWs of GCI-W were shorter than those of other methods for estimating the difference between the variances of two delta-BirSau distributions. When applying these methods to actual wind speed data to illustrate their efficacy, The wind speed data must include the value zero, and the wind speed data follow a delta-BirSau distribution. The result of the actual wind speed shows that the GCI-W method outperformed other methods in terms of AWs for estimating the difference between the variances of wind speed between Pattaya City station and Leam Chabang station.

ACKNOWLEDGMENT

This work was partially financially supported by the Science Achievement Scholarship of Thailand (SAST) and the Graduate College, King Mongkut's University of Technology North Bangkok.

REFERENCES

[1] M. Douak, Z. Aouachria, R. Rabehi, N. Allam, "Wind energy systems: Analysis of the self-starting physics of vertical axis wind turbine". Renewable and Sustainable Energy Reviews, 2017.
[2] V. Leiva, E. Athayde, C. Azevedo, "Marchant, Modeli ng wind energy flux by a Birnbaum–Saunders distribution with an unknown shift parameter, Journal of Applied Statistics". 38(12); 2819-2838, 2011.
[3] K. Mohammadi, O. Alavi, J.G. McGowan, "Use of Birnbaum-Saunders distribution for estimating wind speed and wind power probability distributions: A review, Energy Conversion and Management". 143:109-122, 2017.
[4] J. Mills, "Robust estimation of the Birnbaum-Saunders distribution. Master thesis". Technical University of Nova Scotia, Canada, 1997.
[5] V. Leiva, C. Marchant, F. Ruggeri, H. Saulo, "A criterion for environmental assessment using Birnbaum-Saunders attribute control charts". Environmetrics. 26(7):463-476, 2015b.
[6] V. Leive, M. Ferreira, M. Gomes, C. Lillo, "Extreme value Birnbaum-Saunders regression models applied to environmental data". Stoch. Environ. Res. Assess. 30:1048-1058, 2016.
[7] Z.W. Birnbaum, S.C. Saunders, "Estimation for a family of life distributions with applications to fatigue". J. Appl. Probab. 6:328-347, 1969.
[8] M. Engelhardt, L.J. Bain, F.T. Wright, "Inference on the parameters of Birnbaum-Saunders fatigue life

distribution based on maximum likelihood estimation". Technomertrics. 23:251-255, 1981.
[9] H.K.T. Ng, D. Kundu, N. Balakrishnan, "Modified moment estimation for the two-parameter Birnbaum-Saunders distribution". Comput. Stat. Data Anal. 43: 283-298, 2003.
[10] B.X. Wang, "Generalized interval estimation for the Birnbaum-Saunders distribution via the generalized ratio-of-uniforms method". Comput. Stat. 31:207-225, 2016.
[11] Z.L. Sun, "The confidence intervals for the scale parameter of the Birnbaum-Saunders fatigue life distribution". Acta Armamentarii. 30: 1558-1561, 2009.
[12] W.H. Wu, H.N. Hsieh, "Generalized confidence interval estimation for the mean of delta lognormal distribution: an application to New Zealand trawl survey data". Journal of Applied Statistics. 41(7): 1471-1485, 2014.
[13] E.B. Wilson, "Probable inference the law of succession and statistical inference". Journal of the American Statistical Association. 22(158): 209-212, 1927.
[14] G. Caswlla, R.L. Berger, Statistical inference. Second Edition. Boston: Cengage Learning. 2001
[15] T. Mathew, D.W. Webb, "Generalized p values and confidence intervals for variance components: applications to army test and evaluation". Technometrics. 47: 312-322, 2005.
[16] V. Cojbasic, A. Tomovic, "Nonparametric confidence intervals for population variance of one sample and the difference of variances of two samples". Computational Statistics & Data Analysis, 51:5562-5578, 2007.
[17] S.A. Niwitpong, "A note on coverage probability of confidence interval for the difference between two normal variances". Applied Mathematical Sciences. 6: 3313-3320, 2012.
[18] P. Maneerat, S.A. Niwitpong, S. Niwitpong, "Bayesian confidence intervals for the difference between variances of delta-lognormal distributions". Biometrical Journal. 62(7):1769-1790, 2020.
[19] W. Puggard, S.A.Niwitpong, S. Niwitpong,"Confidenc e intervals for the variance and difference of variances of Birnbaum-Saunders distributions", Journal of Statistical Computation and Simulation. 92(13): 2829-2845, 2022.
[20] W. Khooriphan, S.A. Niwitpong, S. Niwitpong, "Confidence Interval for the Difference Between Variances of Delta-Gamma Distribution", Applied Science and Engineering Progress, 16(4), 2023.
[21] J. Aitchison, "On the distribution of a positive random variable having a discrete probability mass at the origin," J. Am. Stat. Assoc. 50:901-908, 1955.
[22] Z.W. Birnbaum, S.C. Saunder, "A new family of life distribution", J.Appl.Probab. 1:6319-327, 1969a.
[23] S. Weerahandi, "Generalized confidence intervals," J. Am. Stat. Assoc. 88(423):899-905, 1993.
[24] A. DasGupta, "Asymptotic theory of statistics and probability", Springer. New York, 2008.
[25] B. Efron, R. Tibishirani, "An introduction to the bootstrap", Chapman and Hall, London, UK, 1993.
[26] A.J. Lemonte, A.B. Simas, F. Cribari Neto, "Bootstrap based improved estimators for the two parameter Birnbaum Saunders distribution", J. Stat. Comput. 78:37-49, 2008.
[27] J.G. Mackinnon, J.A.A Smith, "Approximate bias correction in econometrics", J. econometrics. 85:205-230, 1998.
[28] W. Puggard, S.A. Niwitpong, S. Niwitpong. "Bayesian Estimation for the Coefficients of Variation of Birnbaum-Saunders Distributions". Symmetry 13(11): 2130, 2021.
[29] W. Thangjai, S.A. Niwitpong, S. Niwitpong, "Confidence intervals for variance and difference between variances of one-parameter exponential distributions. Advances and Applications in Statistics" 53(3): 259-283, 2018.
[30] W. Thangjai, S.A. Niwitpong, "The relative potency of two drugs using the confidence interval for ratio of means of two normal populations with unknown coefficients of variation. Journal of Statistics Applications and Probability" 11(1): 1-14, 2022.
[31] W. Thangjai, S.A. Niwitpong, S. Niwitpong, "Estimation of the percentile of Birnbaum-Saunders distribution and its application to PM2.5 in NorthernThailand. PeerJ 12: e17019, 2024.
[32] Thai Meteorological Department Automatic Weather System, https://www.tmd.go.th/service/tmdData.

Confidence Intervals for the Coefficient of Variation of Zero-Adjusted Inverse Gaussian Distribution with Application to Rainfall Dispersion Data

Wasurat Khumpasee[1, a)], Sa-Aat Niwitpong[1, b)] and Suparat Niwitpong[1, c)]

[1]Department of Applied Statistics, Faculty of Applied Science,
King Mongkut's University of Technology North Bangkok, Thailand.

[a)] k.wasurat@email.kmutnb.ac.th
[b)] sa-aat.n@sci.kmutnb.ac.th
[c)] Corresponding author: suparat.n@sci.kmutnb.ac.th

Abstract. In September was the middle of the rainy season in Thailand always had a heavy rain which can lead to unforeseen circumstances and natural disasters like floods in many parts of the country, especially in northern and northeastern. The zero-inflated inverse gaussian distribution, which combines the inverse gaussian and binomial distributions, is the distribution that applies to rainfall dispersion data that have both positive and zero values. To estimate the precipitation of rainfall in different parts of Thailand, we desired to construct the confidence intervals (CIs) for the coefficient of variation (CV) of zero-inflated inverse Gaussian populations. We proposed five methods to construct CIs using the Generalized Confidence Interval (GCI), Adjusted-Generalized Confidence Interval (AGCI), Bootstrap Percentile Confidence Interval (BPCI), Fiducial Confidence Interval (FCI), and Fiducial Highest Posterior Density Confidence Interval (F-HPDCI). Using a Monte Carlo simulation, the suggested CI method performances were assessed in terms of average widths (AWs) and coverage probabilities (CPs). The results from a simulation study show that F-HPDCI and AGCI had the best performance in term of CPs in almost all cases. But when considered in term of AWs, BPCI was the best method and followed by AGCI and F-HPDCI. When we considered both of CPs and AWs, the outstanding methods were F-HPDCI and AGCI. We applied all methods to construct the CIs for the CV of natural rainfall data from northern and northeastern in Thailand and was found that AGCI and F-HPDCI still the best methods for constructing Cis for the CV of ZAIG distribution.

***Index Terms*—coefficient of variation, confidence interval, rainfall dispersion, zero-inflated inverse gaussian distribution (ZAIG).**

I. INTRODUCTION

In various fields like economics, biomedical sciences, life science, and meteorology, it is common to encounter observations that are non-negative but have an excess of zeros. These observations often have a right-skewed and heteroscedastic distribution, with a large amount of zero values and positive values. For instance, in a clinical trial where a specific drug is administered to patients, the duration of activities, as deficient blood pressure, over sweating, and fast heartbeat, is evaluated in minutes. In this case, some patients may have a measurement of zero, and non-zero measurements may be right-skewed.

Liu et al. [1] were studied the model of data that were suitable in case of including the zero-inflated non-negative observations in a continuous distribution. They used the data from microbiome, medical expenses, and alcohol consumption. Subsequently, several researchers suggested using the Inverse Gaussian (IG) distribution to analyze these types of data. For example, Heller et al. [2] researched insurance data characterized by zero-inflated information, finding that the Inverse Gaussian distribution was more suitable for analysis compared to the Gamma distribution. Ma and Yan [3] analyzed the frequency of traffic accidents in Colorado, USA, and found that the Inverse Gaussian distribution could be employed for analyzing data above zero values. They used regression models to forecast accident occurrences and estimate parameters using Maximum Likelihood (ML). Additionally, Punzo [4] were used the Inverse Gaussian distribution in analyzing the economic data on household earnings in Italy and insurance data regarding claims for physical harm.

In analyzing to propose the confidence intervals for zero-inflated data, researchers have explored various distributions adapted to accommodate with these data characteristics. For instance, Zhou and Tu [5] investigated CIs for the ratio of means in log-normal distributions using ML-based methods and Bootstrap CI. Many other researchers have studied methods for estimating CIs for different parameters of the lognormal distribution and proposed diverse approaches.

However, it was later found that the Inverse Gaussian Distribution was not complicated than other distributions, such as the lognormal distribution, and offers diverse prefernces researched by Chhikara and Folks [6] and suggested that the Inverse Gaussian Distribution is more versatile. Subsequently, Chatzidiamantis et al. [7] further recommended that the lognormal distribution lacks certain properties because there is not a definite Moment Generating Function (MGF). Additionally, the distribution of the mean of the log-normal distribution differs from the population

distribution, making it potentially more difficult to analyze compared to other distributions. Conversely, the Inverse Gaussian Distribution is more suitable for wind energy data and allows for easier parameter estimation. It shares similarities with the Three-parameter Weibull distribution but less complexity [8].

The Zero-adjusted Inverse Gaussian (ZAIG) distribution can be used for estimation purposes in diverse scenarios. This includes estimating the average replacement cost of insurance, the number of accidents in a day, the amount of rainfall in a specific area, and the number of deaths from COVID-19 within a certain period in each country. However, there are many research study on regression models for zero-inflated inverse Gaussian distributed data and constructing CIs for the mean, differences in means and ratios of means for the zero-adjusted inverse Gaussian distribution [9]. But there is currently no research investigating the estimation of CIs for the CV of the zero-adjusted inverse Gaussian distribution which represents the ratio between the standard deviation and the mean indicating the spread of data with different units.

In this research will focus on studying about the estimation of Cis for the CV of the zero-adjusted inverse Gaussian distribution using various methods. Five methods are considered: Generalized Confidence Interval (GCI), Adjusted-Generalized Confidence Interval (AGCI), Bootstrap Percentile Confidence Interval (BPCI), Fiducial Confidence Interval (FCI), and Fiducial Highest Posterior Density Confidence Interval (F-HPDCI). The obtained confidence interval values are then compared under different conditions, considering the limitations and characteristics of each estimation method.

II. METHODOLOGY

Assume that all nonzero observations in a population are fitted by an appropriate continuous distribution, and that there are both zero and nonzero positive observations in the population, with a probability of p with $(0 < p < 1)$. The number of zero and nonzero observations in the sample is n_0 and n_1, respectively, where $n_0 + n_1 = n$ where n is a random sample. For the nonzero observations, we use an Inverse Gaussian distribution as an appropriate model. A random variable X according to the population is said to follow the zero-adjusted Inverse Gaussian distribution with mean (μ), shape parameter (λ), $(\mu, \lambda > 0)$, and a probability of p denoted as $X \sim ZAIG(p, \mu, \lambda)$ with the density

$$f_{ZAIG}(x; p, \mu, \lambda) = pI_{\{x=0\}} + (1-p)I_{\{x>0\}}\sqrt{\frac{\lambda}{2\pi x^3}}exp\left[\frac{-\lambda(x-\mu)^2}{2\mu^2 x}\right],$$
$$, \mu, \lambda > 0, \text{ where } P(X=0) = p. \quad (1)$$

The maximum likelihood estimates (MLEs) of parameters p, μ, λ are $\hat{p} = n_0/n_1$ $\hat{\mu} = \bar{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i$ and $n_1\hat{\lambda}^{-1} = V = \sum_{i=1}^{n_1}(X_i^{-1} - \bar{X}^{-1})$. The mean and the variance of the distribution are $E(X) = q\mu$ and $Var(X) = \mu^2 q\left(\frac{\mu}{\lambda} + p\right)$, respectively. Such that the CV of X is $\theta = \sqrt{(\mu + \lambda p)/q\lambda}$. Assume that a random sample n is drawn from the ZAIG population is $X = (X_1, X_2, \ldots, X_n)$. Given that n_0 represents the number of successful in n Bernoulli trials, and n_0 represents the occurrence of zero observations as a success, $n_0 \sim Bin(n, p)$.

A. Generalized Confidence Interval (GCI)

Weerahandi [10] proposed the GCI method, which is based on the Generalized Pivotal Quantity (GPQ) approach. The GPQ method is a generalization of the Pivotal Quantity (PQ) but it differs from GPQ in that GPQ is a function of nuisance parameters, whereas PQ can be considered from the perspective of the sample group and the parameters of interest. The observed value of the GPQ approach is independent of the nuisance parameter, and the method's parameters have unknown values that are not reliant on any distribution.

Ye, Ma, and Wang [11] introduced a method for constructing CIs for the common mean of the inverse Gaussian distribution in several populations using the GPQ method. Subsequently, Chankham, Niwitpong, and Niwitpong [12] suggested that the method for estimating CIs for the CV of the inverse Gaussian distribution should have coverage probability close to 0.95 and the shortest average width, which is achieved by estimating confidence intervals using the GCI method. Therefore, the GCI method was chosen as the basis method for comparing the performance of estimating confidence intervals with other methods.

Generalized Pivotal Quantity (GPQs) of parameters μ and λ proposed by Ye, Ma, and Wang are

$$R_\lambda = \frac{n_1\lambda V_i}{n_1 v_i} \sim \frac{\chi^2_{n_1-1}}{n_1 v_i}, \; i = 1,2,\ldots,k \quad (2)$$

where $\chi^2_{n_1-1}$ is Chi-square distribution with degree of freedom $n_1 - 1$ and

$$R_\mu = \frac{\bar{X}_i}{\left|1+\frac{\sqrt{n_1\lambda}(\bar{X}_i-\mu)}{\mu\sqrt{\bar{X}_i}}\sqrt{\frac{\bar{X}_i}{n_1 R_{\lambda_i}}}\right|} \sim \frac{\bar{X}_i}{\left|1+Z_i\sqrt{\frac{\bar{X}_i}{n_1 R_{\lambda_i}}}\right|} \quad (3)$$

where $Z_i \sim N(0,1)$.

Generalized Pivotal Quantity (GPQs) of parameters p proposed by Wu and Hsieh [13] is

$$R_p = sin^2\left[arcsin\sqrt{\hat{p}} - \frac{T}{2\sqrt{n}}\right] \quad (4)$$

where $T = 2\sqrt{n}(arcsin\sqrt{\hat{p}} - arcsin\sqrt{p}) \sim N(0,1)$

From the GPQs of the three parameters, the GPQ for the CV of a ZAIG distribution using GCI method is given by

$$R_\theta = \sqrt{(R_\mu + R_\lambda R_p)/(1-R_p)R_\lambda} \quad (5)$$

Then, the 100(1-α)% two-sided CI of CV for ZAIG distribution based on the GCI method is given by

$$CI_{GCI} = [L_{GCI}, U_{GCI}] = [R_\theta(\alpha/2), R_\theta(1-\alpha/2)] \quad (6)$$

where $R_\theta(\alpha/2)$ and $R_\theta(1-\alpha/2)$ are $100_{\alpha/2}$-th and $100_{(1-\alpha)/2}$-th percentiles of the distribution of R_θ, respectively.

Algorithm 1: The GCI method

1. Generate $X_1, X_2, \ldots, X_n$ from ZAIG distribution and compute for $\hat{\mu}_i$ and $\hat{\lambda}_i$.
2. Generate $\chi^2_{n_i-1}$ and Z_i from Chi-square and standard normal distributions, respectively.
3. Compute R_{μ_i} and R_{λ_i} using (2),(3) and compute R_{p_i} using (4).
4. Calculate R_{θ_i} using (5).
5. Repeat steps 2-4 for 5,000 times.
6. Obtain CI_{GCI} from $R_\theta(\alpha/2)$ and $R_\theta(1-\alpha/2)$.

B. Adjusted Generalized Confidence Interval (AGCI)

Upon proposing the GCI method to estimate CIs of CV for the ZAIG distribution, it was discovered that utilizing the GPQ of parameter μ suggested by Krishnamoorthy and Tian [14] as follows

$$R_{\tilde{\mu}} = \frac{\bar{X}_i}{max\left\{0, 1+t_{n_1-1}\sqrt{\frac{\bar{X}_i v_i}{n_1-1}}\right\}} \quad (7)$$

where t_{n_1-1} is students' t distribution with degree of freedom $n_1 - 1$. While the GPQs of parameters λ and p was similar to GCI method.

Thus, the GPQ for the CV of a ZAIG distribution using AGCI method is given by

$$R_{\tilde{\theta}} = \sqrt{\left(R_{\tilde{\mu}} + R_\lambda R_p\right)/(1-R_p)R_\lambda} \quad (8)$$

The 100(1-α)% two-sided CI of CV for ZAIG distribution estimated by the AGCI method is given by

$$CI_{AGCI} = [L_{AGCI}, U_{AGCI}] = [R_{\tilde{\theta}}(\alpha/2), R_{\tilde{\theta}}(1-\alpha/2)] \quad (9)$$

Algorithm of AGCI method was like GCI method but using (7) instead of (2) for computing GPQ of μ.

C. Bootstrap Percentile Confidence Interval (BPCI)

The BPCI method was established by Efron and Tibshirani [15] using resampling techniques, where the resampled sample group is selected from the original sample group. The specific characteristics of the estimated statistics are examined for each distribution. In this research, 1,000 resamples are used to estimate confidence intervals using the BPCI method.

However, Chankham, Niwitpong, and Niwitpong [11] discovered that the BPCI method is less effective in estimating confidence intervals than the GCI method. Nevertheless, to illustrate the diversity of estimation methods, this method will be compared with others.

Therefore, the 100(1-α)% two-sided CI of CV for ZAIG distribution from the BPCI method is given by

$$CI_{BPCI} = [L_{BPCI}, U_{BPCI}] = [\theta^*(\alpha/2), \theta^*(1-\alpha/2)] \quad (10)$$

where $\theta^*(\alpha/2)$ and $\theta^*(1-\alpha/2)$ are $100_{\alpha/2}$-th and $100_{(1-\alpha)/2}$-th percentiles of the distribution of θ^*, respectively.

Algorithm 2: The BPCI method

1. Generate $X_1, X_2, \ldots, X_n$ from ZAIG distribution.
2. Obtain bootstrap samples $X_1^*, X_2^*, \ldots, X_n^*$ from step 1.
3. Compute θ^*.
4. Repeat steps 2-3 for 1,000 times.
5. Obtain CI_{BPCI} using (10).

D. Fiducial Confidence Interval (FCI)

The frequentist inference kind of interpretation that makes advantage of the conditionality of the data is called fiducial inference. The Gibbs sampling method by Geman and Geman (1984) [16] can be used in Fiducial inference. In Bayesian inference, Gibbs sampling is a popular Markov Chain Monte Carlo (MCMC) technique for data generation from the posterior distribution. The method involves sampling by sweeping through variables from their conditional distributions while the values of other variables was the same as their present values. The sample distributions of the MLEs of parameters p, μ, λ are used in the ZAIG distribution and the Fiducial distribution of parameters μ, λ can be obtained by substituting the parameters with the MLEs. When both parameters appear in the sample distribution, they are treated as follows $\mu \sim IG(\hat{\mu}, n_1\hat{\lambda})$ and $\lambda \sim (\hat{\lambda}/n_1)\chi^2_{n_1-1}$ where $\hat{\mu}, \hat{\lambda}$ are the MLEs of μ and λ, respectively.

The 100(1-α)% two-sided CI of CV for ZAIG distribution from the FCI method is given by

$$CI_{FCI} = [L_{FCI}, U_{FCI}] = [\theta^t(\alpha/2), \theta^t(1-\alpha/2)] \quad (11)$$

Algorithm 3: The FCI method

1. Given the initial values (MLEs) of parameters $(\mu^{(0)}, \lambda^{(0)})$
2. Generate $\mu^{(t)} \sim IG(\hat{\mu}^{(t-1)}, n\hat{\lambda}^{(t-1)})$.
3. Generate $\lambda^{(t)} \sim (\hat{\lambda}^{(t-1)}/n)\chi^2_{n-1}$.
4. Repeat steps 2-3 for t = 20,000 times.
5. Burn-in 1,000 samples and calculate for the interest parameter.
6. Obtain CI_{FCI} using (11).

E. Fiducial Highest Posterior Density Confidence Interval (F-HPDCI)

The MCMC approach be used to derive the confidence interval created from the highest posterior density function of parameters [17], if all values inside the interval have higher posterior densities than values outside the interval

[18]. We computed $L_{F-HPDCI}$ and $U_{F-HPDCI}$ using *HDInterval* in the R software package to calculate the 100(1-α)% HPD intervals.

The 100(1-α)% two-sided CI of CV for ZAIG distribution based on the F-HPDCI method is given by

$$CI_{F-HPDCI} = [L_{F-HPDCI}, U_{F-HPDCI}] \quad (12)$$

III. RESULTS AND DISCUSSION

The study considered five methods (GCI, AGCI, BPCI, FCI, and F-HPDCI) to construct CIs for the coefficient of variation within ZAIG distribution. This analysis was conducted through a Monte Carlo simulation performed in the R programing. The performance of these five methods was assessed based on their coverage probabilities (CPs) and average widths (AWs). The selection of a preferred approach depends on two important factors: the CP is equal to or near the nominal confidence level of 0.95 and achieving the narrowest AW. The simulation setting consisting of the number of replications was 5,000, with 5,000 pivotal quantities using in GCI, B = 1,000 for BPCI, and t = 20,000 for FCI, and F-HPDCI. For generating sample data for coefficient of variation of ZAIG, the data were generated for $X \sim ZAIG(p, \mu, \lambda)$ with sample size n = 20, 50, 100 or 200, the mean μ = 0.5 or 1.0, the shape parameter λ = 1, 5 or 10 and probability p = 0.1, 0.3 or 0.5.

The results from Table 1 about CPs and AWs, and was illustrated as Fig. 1 for CPs and Fig. 2 for AWs show that in almost all cases, the CPs from AGCI, FCI, and F-HPDCI were close to the nominal coverage level of 0.95. AGCI method was better than others when the sample sizes were small. But when sample sizes were larger, FCI and F-HPDCI were better in CPs than the others. However, considering in the average widths, BPCI had narrowest among the others, following by AGCI and F-HPDCI. Meanwhile, when two criteria be considered together, the methods that had outstanding performance were AGCI and F-HPDCI. Following by FCI, GCI, and BPCI, respectively. Furthermore, AGCI had shorter duration to compute the CIs than F-HPDCI. However, F-HPDCI was less complex in estimating CIs than AGCI because there was package *HDInterval* in R programming can be used.

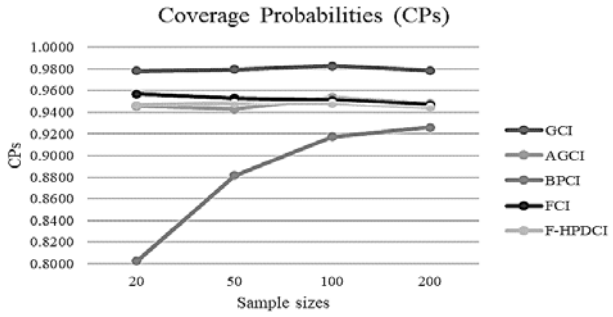


Fig. 1. Comparison of coverage probabilities of the five proposed methods with various sample size: $(p, \mu, \lambda) = 0.1, 0.5, 1$.

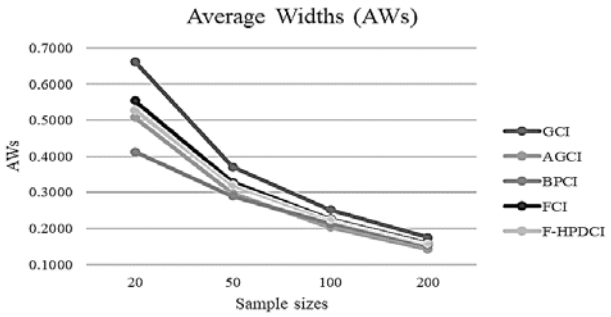


Fig. 2. Comparison of average widths of the five proposed methods with various sample size: $(p, \mu, \lambda) = 0.1, 0.5, 1$.

IV. AN EMPIRICAL APPLICATION

Using several kinds of mathematical models, Duangdai & Likasiri [19] forecasted the global temperature as well as the quantity of seasonal and forest rainfall in northern Thailand. Despite the fact that there was greater forest cover in the summer than in the rainy season, their research indicates that there was less variance in rainfall during the rainy season from 1973 to 2008 when compared to the summer seasons. Rice farming occupies over 60% of the arable land in the northeast [20]. Significantly, rain-fed planted rice, which is farmed in both the northeast and northern parts of Thailand, is the most highly prized product in the Lower Mekong Basin [21]. Our attention was drawn to the rainfall quantities in Thailand's northern and northeastern as a result of these findings, particularly regarding variations in rainfall during the rainy season. Crucially, we investigated and evaluated the ratio of variations for rainfall in the northern and northeastern areas using our suggested confidence intervals. The Thai Meteorological Center recorded rainfall data from a total of 272 substations across both areas. The 272 recorded rainfall values from July 2–8, 2018, in the northeast (210 substations) and north (62 substations) correspondingly, including 40 of 62 (64.52%) and 145 of 210 (69.05%) positive records. For both sets, the remaining observations were zero.

Since the data contain both of zero and positive values, they can be fitted to ZAIG distributions. Thus, we used the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) to assess the distributions of these rainfall datasets. From Table 2 findings indicate that the rainfall data in Thailand's northeast and north both fit the ZAIG distribution, which is supported by the AIC and BIC values, which were the lowest for the ZAIG distribution relative to other distributions.

The essential statistic estimates computed for the rainfall data are displayed in Table 3. The 95% CI for this CV, utilizing GCI, AGCI, BPCI, FCI, and F-HPDCI, is outlined in Table 4. According to the simulation outcomes, AGCI exhibited the narrowest AWs, followed by F-HPDCI, FCI, BPCI, and GCI, respectively. The results were also equivalent with the simulation study. In summary, AGCI and F-HPDCI were the most appropiate method for constructing a CI for the CV of ZAIG distribution.

TABLE I
THE COVERAGE PROBABILITIES AND AVERAGE WIDTHS OF 95% TWO-SIDED CONFIDENCE INTERVAL FOR COEFFICIENT OF VARIATION OF ZERO-ADJUSTED INVERSE GAUSSIAN DISTRIBUTIONS.

				Coverage Probabilities					Average Widths				
n	μ	λ	p	GCI	AGCI	BPCI	FCI	F-HPDCI	GCI	AGCI	BPCI	FCI	F-HPDCI
20	0.5	1	0.1	0.9782	0.9456	0.8024	0.9568	**0.9462**	0.6623	0.5082	**0.4121**	0.5544	0.5279
			0.3	0.9792	**0.9462**	0.7614	0.9568	0.9426	0.8292	0.6067	**0.4247**	0.6601	0.6102
			0.5	0.9768	**0.9474**	0.6978	0.9592	0.9366	1.3916	0.9000	**0.4726**	0.9558	0.8361
		5	0.1	0.9568	9.9456	9.8146	**0.9508**	0.9484	0.1928	0.1817	**0.1298**	0.1852	0.1733
			0.3	0.9590	0.9468	0.7724	**0.9530**	0.9450	0.1974	0.1841	**0.1104**	0.1879	0.1690
			0.5	0.9590	**0.9470**	0.7146	0.9548	0.9442	0.2910	0.2623	**0.1134**	0.2657	0.2231
		10	0.1	0.9534	0.9456	0.8136	**0.9508**	0.9476	0.1108	0.1074	**0.0738**	0.1085	0.1005
			0.3	0.9538	0.9480	0.7768	**0.9514**	0.9438	0.1053	0.1015	**0.0583**	0.1026	0.0913
			0.5	**0.9510**	0.9452	0.7170	0.9514	0.9442	0.1521	0.1436	**0.0587**	0.1444	0.1195
	1	1	0.1	0.9886	**0.9458**	0.7908	0.9552	0.9338	1.2271	0.7468	**0.6532**	0.8734	0.8345
			0.3	0.9878	**0.9474**	0.7460	**0.9526**	0.9312	1.7037	0.9369	**0.7159**	1.0859	1.0116
			0.5	0.9854	**0.9480**	0.6834	0.9560	0.9292	2.9583	1.4186	**0.8247**	1.5844	1.4041
		5	0.1	0.9642	0.9450	0.8122	0.9542	**0.9498**	0.3264	0.2913	**0.2181**	0.3021	0.2852
			0.3	0.9650	**0.9478**	0.7732	0.9548	0.9450	0.3639	0.3185	**0.2021**	0.3309	0.3011
			0.5	0.9662	**0.9484**	0.7094	0.9556	0.9416	0.5604	0.4615	**0.2145**	0.4730	0.4042
		10	0.1	0.9568	0.9456	0.8146	**0.9508**	0.9484	0.1928	0.1817	**0.1298**	0.1852	0.1733
			0.3	0.9590	0.9468	0.7724	**0.9530**	0.9450	0.1974	0.1841	**0.1104**	0.1879	0.1690
			0.5	0.9590	**0.9470**	0.7146	0.9548	0.9442	0.2910	0.2623	**0.1134**	0.2657	0.2231
50	0.5	1	0.1	0.9794	0.9428	0.8816	0.9530	**0.9484**	0.3702	0.2965	**0.2894**	0.3282	0.3216
			0.3	0.9822	0.9454	0.8712	0.9540	**0.9484**	0.4260	0.3380	**0.3129**	0.3732	0.3616
			0.5	0.9792	0.9466	0.8422	**0.9528**	0.9414	0.5717	0.4444	**0.3737**	0.4880	0.4642
		5	0.1	**0.9528**	0.9394	0.8898	0.9440	0.9444	0.1096	0.1041	**0.0921**	0.1064	0.1035
			0.3	0.9570	0.9472	0.8796	**0.9488**	0.9480	0.1037	0.0982	**0.0815**	0.1004	0.0961
			0.5	0.9592	0.9470	0.8464	**0.9488**	0.9444	0.1266	0.1193	**0.0892**	0.1217	0.1140
		10	0.1	**0.9476**	0.9420	0.8886	0.9416	0.9446	0.0623	0.0607	**0.0525**	0.0614	0.0594
			0.3	0.9540	0.9484	0.8794	0.9492	**0.9494**	0.0545	0.0530	**0.0431**	0.0536	0.0511
			0.5	0.9544	0.9478	0.8476	**0.9482**	0.9456	0.0650	0.0631	**0.0461**	0.0637	0.0594
	1	1	0.1	0.9932	0.9424	0.8764	**0.9512**	0.9452	0.6448	**0.4372**	0.4626	0.5265	0.5167
			0.3	0.9940	0.9462	0.8628	**0.9526**	0.9398	0.7912	**0.5268**	0.5303	0.6303	0.6128
			0.5	0.9906	**0.9462**	0.8392	0.9542	0.9364	1.1264	0.7166	**0.6603**	0.8495	0.8129
		5	0.1	0.9624	0.9410	0.8876	**0.9476**	0.9468	0.1862	0.1687	**0.1541**	0.1762	0.1720
			0.3	0.9654	0.9482	0.8758	**0.9506**	0.9486	0.1923	0.1734	**0.1490**	0.1810	0.1742
			0.5	0.9670	0.9474	0.8456	**0.9514**	0.9438	0.2432	0.2174	**0.1685**	0.2261	0.2131
		10	0.1	**0.9528**	0.9394	0.8898	0.9440	0.9444	0.1096	0.1041	**0.0921**	0.1064	0.1035
			0.3	0.9590	0.9468	0.7724	**0.9530**	0.9450	0.1037	0.0982	**0.0815**	0.1004	0.0961
			0.5	0.9592	0.9470	0.8464	**0.9488**	0.9444	0.1266	0.1193	**0.0892**	0.1217	0.1140

V. CONCLUSION

In this research, five methods (GCI, AGCI, BPCI, FCI, and F-HPDCI) to construct the CIs for the CV of ZAIG distributions were investigated to find the best method for estimating Cis for CV. The results of the simulation study show that in term of the CPs, AGCI, F-HPDCI, and FCI were close to the nominal confidence level of 0.95 for nearly all cases. But when we considered about the AWs, BPCI, AGCI, and F-HPDCI were better. However, when we considered both CPs and AWs, the methods that had outstanding performance were AGCI and F-HPDCI. In an empirical application about rainfall dispersion to demonstrate the effectiveness of these methods, the AGCI had the narrowest AWs and was better than the others. Such that, from the simulation study and an empirical application suggest that AGCI and F-HPDCI were the most suitable method. Since ZAIG was the widely used in economics, biomedical sciences, life science, and meteorology, appropriate methods selection for accurate estimation CIs for CV is important. This study was extended from estimation CIs for CV of IG distribution to ZAIG distribution which is more suitable for the real-world scenarios. Therefore, the most appropriate method for construct CIs for CV of ZAIG from the simulation study and empirical application were AGCI and F-HPDCI methods.

TABLE I (CONTINUED)

				Coverage Probability					Average Width				
n	μ	λ	p	GCI	AGCI	BPCI	FCI	F-HPDCI	GCI	AGCI	BPCI	FCI	F-HPDCI
100	0.5	1	0.1	0.9826	0.9538	0.9172	**0.9514**	0.9484	0.2518	**0.2039**	0.2129	0.2269	0.2244
			0.3	0.9828	0.9514	0.9086	0.9538	**0.9500**	0.2849	**0.2297**	0.2323	0.2552	0.2509
			0.5	0.9846	0.9554	0.9024	0.9590	**0.9518**	0.3679	0.2944	**0.2848**	0.3262	0.3179
		5	0.1	0.9622	0.9520	0.9172	0.9514	**0.9494**	0.0747	0.0711	**0.0677**	0.0727	0.0717
			0.3	0.9632	0.9522	0.9132	0.9504	**0.9496**	0.0692	0.0658	**0.0603**	0.0673	0.0658
			0.5	0.9656	0.9540	0.9008	0.9556	**0.9508**	0.0816	0.0774	**0.0678**	0.0792	0.0766
		10	0.1	0.9570	0.9518	0.9150	0.9514	**0.9486**	0.0423	0.0412	**0.0386**	0.0417	0.0410
			0.3	0.9580	0.9536	0.9166	0.9514	**0.9500**	0.0363	0.0354	**0.0319**	0.0358	0.0349
			0.5	0.9590	0.9538	0.8980	0.9540	**0.9508**	0.0418	0.0407	**0.0350**	0.0411	0.0397
	1	1	0.1	0.9930	**0.9514**	0.9114	0.9536	0.9450	0.4334	**0.3009**	0.3419	0.3653	0.3615
			0.3	0.9912	0.9534	0.9042	0.9546	**0.9490**	0.5216	**0.3594**	0.3973	0.4353	0.4288
			0.5	0.9926	0.9550	0.8948	**0.9538**	0.9450	0.7041	**0.4777**	0.5055	0.5789	0.5629
		5	0.1	0.9702	0.9540	0.9188	**0.9504**	0.9476	0.1271	0.1156	**0.1131**	0.1210	0.1194
			0.3	0.9692	0.9530	0.9120	**0.9500**	0.9480	0.1289	0.1170	**0.1105**	0.1225	0.1200
			0.5	0.9724	0.9540	0.8998	0.9546	**0.9500**	0.1572	0.1423	**0.1281**	0.1486	0.1441
		10	0.1	0.9622	0.9520	0.9172	0.9514	**0.9494**	0.0747	0.0711	**0.0677**	0.0727	0.0717
			0.3	0.9632	0.9522	0.9132	**0.9504**	**0.9496**	0.0692	0.0658	**0.0603**	0.0673	0.0658
			0.5	0.9656	0.9540	0.9008	0.9556	**0.9508**	0.0816	0.0774	**0.0678**	0.0792	0.0766
200	0.5	1	0.1	0.9786	**0.9482**	0.9262	0.9470	0.9446	0.1751	**0.1423**	0.1536	0.1586	0.1576
			0.3	0.9814	0.9476	0.9302	0.9536	**0.9518**	0.1968	**0.1598**	0.1697	0.1778	0.1761
			0.5	0.9826	0.9528	0.9246	0.9536	**0.9494**	0.2498	**0.2021**	0.2103	0.2247	0.2217
		5	0.1	0.9554	**0.9480**	0.9298	0.9448	0.9452	0.0519	0.0495	**0.0488**	0.0506	0.0502
			0.3	0.9542	0.9464	0.9350	0.9510	**0.9502**	0.0478	0.0455	**0.0442**	0.0465	0.0460
			0.5	0.9614	0.9516	0.9278	**0.9502**	0.9496	0.0554	0.0527	**0.0499**	0.0539	0.0529
		10	0.1	**0.9520**	0.9476	0.9302	0.9450	0.9454	0.0293	0.0286	**0.0279**	0.0290	0.0287
			0.3	0.9506	0.9466	0.9356	0.9482	**0.9500**	0.0250	0.0244	**0.0233**	0.0247	0.0243
			0.5	0.9560	0.9516	0.9256	**0.9498**	0.9488	0.0283	0.0276	**0.0258**	0.0279	0.0274
	1	1	0.1	0.9922	**0.9488**	0.9266	0.9476	0.9432	0.2999	**0.2102**	0.2475	0.2559	0.2543
			0.3	0.9942	0.9470	0.9272	0.9532	**0.9476**	0.3582	**0.2502**	0.2901	0.3043	0.3017
			0.5	0.9934	0.9538	0.9214	**0.9516**	0.9478	0.4739	**0.3289**	0.3735	0.3990	0.3941
		5	0.1	0.9602	**0.9482**	0.9282	0.9454	0.9436	0.0885	**0.0806**	0.0817	0.0844	0.0838
			0.3	0.9630	0.9464	0.9336	0.9518	**0.9504**	0.0891	0.0811	**0.0808**	0.0849	0.0839
			0.5	0.9690	0.9530	0.9284	0.9530	**0.9490**	0.1069	0.0972	**0.0946**	0.1017	0.1000
		10	0.1	0.9554	**0.9480**	0.9298	0.9448	0.9452	0.0519	0.0495	**0.0488**	0.0506	0.0502
			0.3	0.9542	0.9464	0.9350	0.9510	**0.9502**	0.0478	0.0455	**0.0442**	0.0465	0.0460
			0.5	0.9614	0.9516	0.9278	**0.9502**	0.9496	0.0554	0.0527	**0.0499**	0.0539	0.0529

TABLE 2
AIC AND BIC VALUES FOR FITTING THE POSITIVE RAINFALL DATA FROM NORTHERN AND NORTHEASTERN OF THAILAND.

	North		Northeast	
Models	AIC	BIC	AIC	BIC
Normal	348.812	352.19	1411.599	1417.552
Logistic	337.678	341.056	1376.586	1382.54
Cauchy	336.093	339.471	1331.557	1337.511
ZAIG	317.155	320.533	1253.438	1259.392

TABLE 3
ESTIMATES FROM RAINFALL DATA

Region	n	median	mean	variance	CV
North	62	6.65	12.12	294.99	1.6263
Northeast	210	6.60	17.70	809.04	2.3950

TABLE 4
THE 95% CONFIDENCE INTERVALS FOR THE COEFFICIENT OF VARIATION OF RAINFALL DATA FROM NORTHERN AND NORTHEASTERN OF THAILAND.

	Northern			Northeastern		
Method	Lower	Upper	Length	Lower	Upper	Length
GCI	1.3160	2.2976	0.9816	1.9594	3.1125	1.1531
AGCI	1.4144	2.0272	**0.6128**	2.1784	2.6885	**0.5101**
BPCI	1.2515	2.0168	0.7653	1.9757	2.8448	0.8692
FCI	1.3361	2.1172	0.7811	2.0147	2.8791	0.8644
F-HPDCI	1.3173	2.0845	0.7672	1.9869	2.8421	0.8552

ACKNOWLEDGMENT

This work was partially financially supported by the Science Achievement Scholarship of Thailand (SAST) and Graduate College, King Mongkut's University of Technology North Bangkok.

REFERENCES

[1] L. Liu, Y. T. Shih, R. L. Strawderman, D. Zhang, B. A. Johnson, H. Chai. "Statistical Analysis of Zero-Inflated Nonnegative Continuous Data: A Review." Statistical Science, 34(2) 253-279, 2019.

[2] G. Heller, M. Stasinopoulos, and B. Rigby. "The zero-adjusted inverse Gaussian distribution as a model for insurance claims". In Proceedings of the 21st International Workshop on Statistical Modelling, Galway, pp. 226–233, 2006.

[3] L. Ma, X. Yan, "Modeling zonal traffic accident counts with the regression under zero-adjusted inverse Gaussian distribution". Procedia-Social and Behavioral Sciences, 138:452–9, 2014.

[4] A. Punzo, "A new look at the inverse Gaussian distribution with applications to insurance and economic data". Journal of Applied Statistics, 46:7, 1260-1287, 2019.

[5] X. H. Zhou, W. Tu, "Interval estimation for the ratio in means of log-normally distributed medical costs with zero values". Computational Statistics & Data Analysis 35 (2):201–10, 2000.

[6] R. S. Chhikara, J. L. Folks, "The inverse Gaussian distribution as a lifetime model". Technometrics 19 (4): 461–8, 1977.

[7] N. D. Chatzidiamantis, H. G. Sandalidis, G. K. Karagiannidis, S. A. Kotsopoulos, "On the inverse-Gaussian shadowing". In 2011 International Conference on Communications and Information Technology (ICCIT), pp. 142–146, 2011.

[8] W. E. Bardsley, "Note on the use of the inverse Gaussian distribution for wind energy applications". Journal of Applied Meteorology 19 (9):1126–30, 1980.

[9] N. Jana, M. Gautam, "Confidence intervals of difference and ratio of means for zero-adjusted inverse Gaussian distributions", Communications in statistics - Simulation and Computation, 2022.

[10] S. Weerahandi, "Generalized confidence intervals". Journal of the American Statistical Association 88(423):899-905, 1993.

[11] R. D. Ye, T. F. Ma, S. G. Wang, "Inferences on the common mean of several inverse Gaussian populations". Computational Statistics & Data Analysis, 54:906-915, 2010.

[12] W. Chankham, S. A. Niwitpong, S. Niwitpong, "Confidence intervals for coefficient of variation of inverse Gaussian distribution". In: Proceedings of the 3rd international conference on vision, image and signal Processing, 1-6, 2019.

[13] W. H. Wu, H. N. Hsieh, "Generalized confidence interval estimation for the mean of delta-lognormal distribution: an application to New Zealand trawl survey data". Journal of Applied Statistics, 41(7):1471–1485, 2014.

[14] K. Krishnamoorthy, L. Tian, "Inferences on the difference and ratio of the means of two inverse Gaussian distributions". Journal of Statistical Planning and Inference, 138(7):2082-2089, 2008.

[15] B. Efron, R. Tibshirani, "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy". Statistical Science 1(1):54-77, 1986.

[16] S. Geman, D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images". IEEE Transactions on Pattern Analysis and Machine Intelligence, 6:721-741, 1984.

[17] M. H. Chen, Q. M. Shao, "Monte Carlo estimation of Bayesian credible and HPD intervals". Journal of Computational and Graphical Statistics, 8(1):69-92, 1999.

[18] G. E. P. Box, G. C. Tiao, "Bayesian inference in statistical analysis". Volume 40. New York: John Wiley and Sons, 2011.

[19] E. Duangdai, C. Likasiri, "Rainfall model investigation and scenario analyses of the effect of government reforestation policy on seasonal rainfalls: a case study from Northern Thailand". Atmospheric Research, 185:1-12, 2017.

[20] R. Nicely, A. Counselor, "Grain and feed update May 2018". Technical report, 2018.

[21] B. Zhang, L. Zhang, H. Guo, P. Leinenkugel, Y. Zhou, L. Li, Q. Shen, "Drought impact on vegetation productivity in the Lower Mekong Basin". International Journal of Remote Sensing, 35(8):2835-2856, 2014.

AI-generated animation for improving knowledge and comprehension of triptans for acute migraine treatment in hospital pharmacists

Budsarin Sivathitikul[1], Kemjira Chalermsuk[1]
Siranan Klanbut[1], Sumet Klinhom[2], Suteera Boonyapitaks[3],
Pattarawit Rukthong[1,4,*,a)], and Tulaya Potaros[1,*,b)]

Author Affiliations
[1]Faculty of Pharmacy, Srinakharinwirot University, Nakhon Nayok, 26120, Thailand
[2]Division of audiovisual education, Srinakharinwirot University, Bangkok, 10110, Thailand
[3]Department of Climate Change and Environment, Bangkok, 10400, Thailand
[4]Center for Excellence in Plant and Herbal Innovation Research, Strategic Wisdom and Research Institute, Srinakharinwirot University, Nakorn Nayok, 26120, Thailand
* Equally contributed corresponding author

[a)] Corresponding author: pattarawit@g.swu.ac.th
[b)] tulaya@g.swu.ac.th

Abstract. Artificially intelligent (AI) are extensively used in medical sciences, including for improving medical science education. In this study, the education about migraine management will be developed. Our quasi-experimental study aimed to create and evaluate the quality and satisfaction of AI-generated animation for enhancing medical science education, knowledge and comprehension of triptans for acute migraine treatment in hospital pharmacists. The study was conducted among 21 pharmacists at HRH Princess Maha Chakri Sirindhorn Medical Center (MSMC), Nakhon Nayok, between January 1 and January 31, 2024. The data collection tools included the 10-minute AI generated animation about triptans for acute migraine treatment, pretest and posttest, satisfaction forms, test's quality forms and animation's quality forms, which were approved by experts. The research findings indicated that the animation and tests were content validity and consistent with objectives (IOC 0.66 to 1.00). From the knowledge and comprehension evaluation, the 12-point pretest and posttest revealed that average score of the posttest was higher than the pretest, which was 5.05 ± 1.80 points (p-value < .001). In addition, it was found that age, period of employment and pharmacist positions in MSMC of the samples didn't affect the pre-post difference score. Regarding the satisfaction, the participants considered high satisfaction with the animation in terms of content, titivation, design and usability. In conclusion, the study found that this AI-generated animation could be an alternative to enhance knowledge and comprehension of triptans in the treatment of migraines for hospital pharmacists, community pharmacists, medical students and healthcare workers.

Index Terms—AI-generated animation, Triptans, Acute Migraine, Hospital pharmacists, Medical science education

I. INTRODUCTION

Artificially intelligent (AI) are used extensively in medical sciences including for improving medical science education. In this study, the education about migraine management will be developed. Nowadays, 90% of the migraineurs have moderate to severe pain. Migraine diagnosis and treatment guideline 2022 had recommended to use triptans for moderate to severe migraine treatment.[1] In recent studies, there were non-pharmacological and pharmacological migraine management in community pharmacies which still had some defects. It had no study about migraine management of hospital pharmacists in Thailand. The pharmacists in hospitals where had many migraineurs, should comprehend the migraine medications for treating them.[2, 3, 4, 5, 6] Nevertheless, none of study about improving knowledge of triptans for migraine in hospital pharmacists had been available. Therefore, critical information about triptans for migraine treatment was so important for the hospital pharmacists.

Our quasi-experimental study aimed to create and evaluate the quality and satisfaction of AI-generated animation for improving medical science education, knowledge and comprehension of triptans for acute migraine treatment in hospital pharmacists.

II. METHODOLOGY

The trial was approved by Srinakharinwirot ethics committee, SWUEC/E-358/2565. The study design was quasi-experimental and one group pretest posttest design - Experimental procedure or intervention. We used AI Botnoi voice for audio description and animaker program for AI-

generated animation creation.

Participants would join google meet and follow the research procedures; doing pretest, watching AI-generated animation, doing posttest and satisfaction forms. Populations were 28 pharmacists at HRH Princess Maha Chakri Sirindhorn Medical Center (MSMC), Nakhon Nayok, between January 1 and January 31, 2024.

$$n_0 = \frac{Z^2 pq}{e^2} \quad (1)$$

Here, n_0 is sample size, z is level of confidence according to the standard normal distribution (for a level of confidence of 95%, z = 1.96), p is estimated proportion of the population that presents the characteristic (p = 0.5), d = tolerated margin of error (within 5%). Equation (1) is used to determine sample size of the study. For a 95% confidence interval, the sample size should be more than 27 participants.

$$IOC = \frac{\sum R}{N} \quad (2)$$

The data collection tools included the 10-minute AI generated animation about triptans for acute migraine treatment, the pretest and posttest, the satisfaction forms, test's quality forms and the AI-generated animation's quality forms. These tools were approved by three experts to evaluate the content validity of the tests and three experts to evaluate the content validity and quality of the animation by IOC (index of item objective congruence) from equation (2). In equation (2), IOC is index of item objective congruence, R is total score from experts, N is the number of experts. If IOC was between 0.50 and 1.00, it means the animation or tests were content validity and consistent with objectives. Conversely, if the IOC was lower than 0.50 would mean that the animation or tests were not valid and consistent with objectives.

From the knowledge and comprehension evaluation, 12-point pretest and posttest, the pretest posttest scores were interval data, continuous and independent variables. For the distribution, we used Shapiro-Wilk test. Then pair t-test or Wilcoxon signed rank test was used if the scores were or were not normal distribution, respectively. For the proportion of the samples who had correct to incorrect responses ratio which was nominal data, we used McNemar test. In addition, for the factors which would affect the pre-post difference score and descriptive data, we used correlation coefficient. All statistical assessments were used at the significance level of 0.05.

III. Results and Discussion

A 10-minute AI-generated animation about triptans for acute migraine management was created by using AI Botnoi voice for audio description and animaker program for AI-generated animation creation. The contents of AI-generated animation included types and symptoms of migraine, diagnostic criteria, migraine medications, Migraine Diagnosis and Treatment Guideline 2022, triptans in Thailand, mechanism of action, administration, effectiveness, adverse reactions, drug interactions, dosage adjustments in patients with liver disease, contraindications and precautions. This animation was involved with the Migraine Diagnosis and Treatment Guideline 2022. For IOC (index of item objective congruence), the animation and tests were content validity and consistent with the objectives (IOC 0.66 to 1.00). Thus, the AI-generated animation and tests could be used for improving knowledge and comprehension of triptans for acute migraine treatment in hospital pharmacists and also in community pharmacists, medical students and healthcare workers.

Fig. 1. Examples of AI-generated animation using AI Botnoi voice for audio description and animaker program for AI-generated animation creation

The study was conducted among 21 of 28 pharmacists at HRH Princess Maha Chakri Sirindhorn Medical Center (MSMC), Nakhon Nayok, between January 1 and January 31, 2024 due to their hospital workload. Although the sample size was 21 participants, the 91% confidence interval which lower than the 95% expected in methodology, it was also an acceptable confidence interval.

TABLE 1
BASELINE CHARACTERISTICS OF THE PARTICIPANTS

Characteristics	Number of participants	Percentage
Age, years		
15-29	4	19.0
30-44	17	81.0
45-59	0	0
$\bar{x}$ ± S.D.	33.57 ± 4.29	

TABLE 1 (CONTINUE)
BASELINE CHARACTERISTICS OF THE PARTICIPANTS

Characteristics	Number of participants	Percentage
Period of employment in MSMC, years		
< 1	0	0
1-2	2	9.5
3-5	8	38.1
6-9	3	14.3
> 10	8	38.1
$\bar{x}$ ± S.D.	7.48 ± 4.76	
Knowledge about triptans		
Had	21	100
Had not	0	0
Pharmacist positions in MSMC		
In-patient department (IPD)	2	9.5
Out-patient department (OPD)	11	52.4
Clinical	2	9.5
Drug information service (DIS)	2	9.5
Inventory control & supply	1	4.8
Pharmaceutical preparations	3	14.3

Baseline characteristics of the participants showed that the mean age was 33.57 ± 4.29 years and almost all participants were aged between 30-44 years (81.0%). All participants had ever known about triptans (100%) in TABLE 1.

TABLE 2
THE MEAN SCORE OF PRETEST AND POSTTEST

	Paired differences				t	df	p-value
	Mean	Std.Dev	95% CI lower	95% CI Upper			
Pretest and Posttest scores	5.05	1.80	4.23	5.87	12.84	20	.000

From the knowledge and comprehension evaluation, 12-point pretest and posttest, the scores were continuous and continuous independent variables. For distribution, used the Shapiro-Wilk test revealed that the pre-post difference score was normal distribution (p-value = .065). It revealed that this AI-generated animation could enhance knowledge and comprehension of triptans in the treatment of migraines among participants. For the mean pre-post difference score, used the pair t-test revealed that average score of the posttest; 11.52 ± 1.81 points, was higher than the pretest; 6.48 ± 0.68 points, which was 5.05 ± 1.80 points (p-value < .001) in TABLE 2.

TABLE 3
THE FACTORS WHICH WOULD AFFECT THE PRE-POST DIFFERENCE SCORE

Factors	Statistical analysis	p-value
Aged	r = .307[1]	.176
Period of employment in MSMC	r = .259[1]	.256
Pharmacist positions in MSMC	df = 5, 15[2] F = 1.465	.273

1 Pearson's correlation, 2 Kruskal-Wallis H test

About the factors which would affect the pre-post difference score used the correlation coefficient, it was found that age, period of employment and pharmacist positions in MSMC of the samples didn't affect the pre-post difference score in TABLE 3. The knowledge about triptans factor could not be analyzed in this study. That's because 100% of participants had ever known about triptans thus it could not compare the results of the other group.

TABLE 4
THE PROPORTION OF THE PARTICIPANTS WHO HAD CORRECT RESPONSES INCREASED

Questions	Pretest (N)		Posttest (N)		P-value
	✓	X	✓	X	
The mechanism of action of triptans	4	17	21	0	<.001
Management of drug interaction between eletriptan and CYP 3A4 inhibitors	7	14	21	0	<.001

Note: ✓ defined as correct; X defined as incorrect

TABLE 4 (CONTINUE)
THE PROPORTION OF THE PARTICIPANTS WHO HAD CORRECT RESPONSES INCREASED

Questions	Pretest (N)		Posttest (N)		P-value
	✓	X	✓	X	
Resulting of triptans' drug interaction	8	13	19	2	<.001

Note: ✓ defined as correct; X defined as incorrect

The questions which had the most incorrect answers were about the mechanism of action and drug interaction of triptans using McNemar test in TABLE 4. However, participants watched AI-generated animation, they could choose the correct answer. It showed AI-generated animation was effective.

TABLE 5
THE SATISFACTION SCORE OF AI-GENERATED ANIMATION

Section	Mean	%	S.D.	Satisfaction
Design	4.16	83.24	0.71	High
Content	4.23	84.57	0.73	High
Titivation	4.17	83.43	0.70	High
Usability	4.16	83.14	0.71	High

Note: the satisfaction score; a scale of 1 to 5

Regarding the satisfaction, the participants considered high satisfaction with the animation; content, titivation, design and usability test respectively, in TABLE 5. Moreover, participants suggested that the communication of the animation was easy to understand. The content contained the main point which, was concise. It was considered that participants were very satisfied with the AI-generated animation.

AI in medicine involved in using machine learning models to process medical data and provide important insights for medical professionals, improving health outcomes and patient experiences. It is extensively applied in medical sciences, including for improving medical education, with examples such as SOFA, Chest 4 All, Agnos, and B-Hive1. Similarly, AI in pharmaceutical science is used for convenience and enhancement. The advantages of AI included analyzing data, improving diagnosis, enhancing patient care, reducing care costs, improving workload, and reducing staff stress. However, AI still required human input and might have inaccuracies. The AI-generated animation about migraine management in this study served as an alternative for enhancing knowledge and comprehension of triptans in migraine treatment among hospital pharmacists, community pharmacists, medical students, and healthcare workers.

Several studies had demonstrated the effectiveness of computer-generated animations in medical education. Prinz et al. (2005) found that three-dimensional animated teaching was advantageous over traditional surgical videos for teaching ophthalmic surgery, resulting in improved learning outcomes in a randomized study.[15] Similarly, Hoang et al. (2021) evaluated the use of computer-generated three-dimensional animations for resident education of hepatobiliary anatomy and found that these animations were effective in enhancing learning.[16]

Ruiz et al. (2009) conducted a critical literature review on the use of computer animations in medical education. They were concluded that animations could be a valuable tool for enhancing learning, particularly when it was combined with other instructional strategies.[17] Cleeren et al. (2014) investigated the role of 3D animation in periodontal patient education through a randomized controlled trial. It was found that the use of 3D animations led to improve patient knowledge and understanding.[18] Dahodwala et al. (2018) performed a scoping review to assess the impact of video-based educational interventions on patient outcomes in hospital settings. The review found that video-based interventions, including animations, could positively influence patient outcomes, such as increasing knowledge, reducing anxiety, and improving satisfaction.[19] While these publications were not specifically addressing AI-generated animations or triptans for acute migraine management, they provided insights into the use of computer-generated animations in medical and pharmacy education. The studies examined the effectiveness of 3D animations and video-based educational interventions in improving learning outcomes, patient education, and clinical skills across various medical specialties. These findings suggested that computer-generated animations could be a valuable tool in medical and pharmacy education, but more research was needed to evaluate their specific application in the context of triptans education for hospital pharmacists.

The study's limitations included small sample size and potential bias from the data collection method, such as the inability to regulate frequency that participants watched the animation. Future studies could integrate all tools, including the 10-minute AI-generated animation, pretest and posttest questionnaires, satisfaction forms, and quality assessment forms, into a single platform for better accessibility. Moreover, the study could be expanded to a multi-center

study or conducted with a larger sample size to allow for inferences about the population.

IV. Conclusion

In conclusion, the mean pre-post difference score revealed that the average post-test score was significantly higher than the pre-test score (p-value < .001). This finding, along with the evidence from the literature, suggested that AI-generated animation could be an effective alternative for enhancing knowledge and comprehension of triptans in migraine treatment among hospital pharmacists, community pharmacists, medical students, and healthcare workers. The use of computer-generated animations in medical education had been shown to improve learning outcomes, patient knowledge, and understanding, and had the potential to positively impact patient outcomes in various healthcare settings.

Acknowledgment

The authors would like to acknowledge and thank to pharmacists at HRH Princess Maha Chakri Sirindhorn Medical Center for the study, Faculty of Pharmacy, Srinakharinwirot University, Nakhon Nayok and Srinakharinwirot University, Bangkok for the guidance and support.

References

[1] Migraine Diagnosis and Treatment Guideline 2022 [cited 2022 Aug 20th]. Available from: http://www.neurothai.org/content.php?id=464.

[2] Hu Y, Guan X, Fan L, Jin L. Triptans in prevention of menstrual migraine: a systematic review with meta-analysis. The Journal of Headache and Pain. 2013; 14(1):1-9.

[3] Saengcharoen W, Lerkiatbundit S. Migraine Management in Community Pharmacies: Practice Patterns and Knowledge of Pharmacy Personnel in T hailand. Headache: The Journal of Head and Face Pain. 2013;53(9):1451-63.

[4] Viticchi G, Falsetti L, Buratti L, Plutino A, Provinciali L, Silvestrini M, et al. Triptan use among hospital workers affected by migraine. Neurological Sciences. 2015; 36(1):157-9.

[5] Höffler TN, Leutner D. Instructional animation versus static pictures: A meta-analysis. Learning and instruction. 2007;17(6):722-38.

[6] Thakur E, Recober A, Turvey C, Dindo LN. Benefits of an on-line migraine education video for patients with co-occurring migraine and depression. Journal of Psychosomatic Research. 2018;112:47-52.

[7] Dictionary O. Animation [cited 2022 Sep 3rd]. Available from: https:// en.oxforddic tionaries .com/definition/animation

[8] Micromedex I. Sumatriptan 2022 [cited 2022 Sep 13th]. Available from: https://www.micromedexsolutions.com/micromedex2/librarian/CS/DBD8DC/ND_PR/evidencexpert/ND_P/evidencexpert/DUPLICATIONSHIELDSYNC/A17151/ND_PG/evidencexpert/ND_B/evidencexpert/ND_AppProduct/evidencexpert/ND_T/evidencexpert/PFActionId/evidencexpert.DoIntegratedSearch?SearchTerm=Sumatriptan&UserSearchTerm=Sumatriptan&SearchFilter=filterNone&navitem=searchGlobal#.

[9] Micromedex I. Eletriptan 2022 [cited 2022 Sep 13th]. Available from: https://www. micromedexsolutions.com/micromedex2/librarian/CS/DBD8DC/ND_PR/evidencexpert/ND_P/evidencexpert/DUPLICATIONSHIELDSYNC/A17151/ND_PG/evidencexpert/ND_B/evidencexpert/ND_AppProduct/evidencexpert/ND_T/evidencexpert/PFActionId/evidencexpert.DoIntegratedSearch?SearchTerm=Eletriptan&UserSearchTerm=Eletriptan&SearchFilter=filterNone&navitem=searchGlobal#.

[10] MIMs. Sumatriptan 2022 [cited 2022 Sep 13th]. Available from: https://www.mims.com/thailand /drug/info/sumatriptan?mtype=generic

[11] MIMs. Eletriptan 2022 [cited 2022 Sep 13th]. Available from: https://www.mims.com/thailand/drug/info/eletriptan?mtype=generic

[12] Research MFfMEa. Migraine [cited 2022 Oct 3th]. Available from: https://www.mayoclinic.org/diseases-conditions/migraine-headache/symptoms-causes/syc20360201?p=1.

[13] Diener H-C, Jansen J-P, Reches A, Pascual J, Pitei D, Steiner TJ. Efficacy, tolerability and safety of oral eletriptan and ergotamine plus caffeine (Cafergot®) in the acute treatment of migraine: a multicentre, randomised, double-blind, placebo-controlled comparison. European neurology. 2002;47(2):99-107.

[14] Budiu Raluca. Confidence Intervals, Margins of Error, and Confidence Levels in UX. 2021 [cited 2024 Feb 20th]. available from: https://www.nngroup.com/articles /confidence-interval/

[15] Prinz, A., Bolz, M., & Findl, O. (2005). Advantage of three dimensional animated teaching over traditional surgical videos for teaching ophthalmic surgery: a randomised study. British Journal of Ophthalmology, 89(11), 1495-1499. DOI: 10.1136/bjo.2005.075077

[16] Hoang, D., Woodfield, C. A., Ahn, S., Lauffer, M., Corbin, Z. A., & Schapiro, R. (2021).

[17] Ruiz, J. G., Cook, D. A., & Levinson, A. J. (2009). Computer animations in medical education: a critical literature review. Medical Education, 43(9), 838-846. DOI: 10.1111/j.1365-2923.2009.03429.x

[18] Cleeren, G., Quirynen, M., Ozcelik, O., & Teughels, W. (2014). Role of 3D animation in periodontal patient education: a randomized controlled trial. Journal of Clinical Periodontology, 41(1), 38-45. DOI: 10.1111/jcpe.12170

[19] Dahodwala, M., Geransar, R., Babion, J., de Grood, J., & Sargious, P. (2018). The impact of the use of video-based educational interventions on patient outcomes in hospital settings: A scoping review. Patient Education and Counseling, 101(12), 2116-2124. DOI: 10.1016/j.pec.2018.06.018.

Techno-economic assessment of domestic technology pico-hydro power turbine

Nitikorn Nanthawirojsiri[1, a)], Netithorn Ditnin[1, b)], Attakarn Jansasithorn[1, c)] and Usa Boonbumroong[1, d)]

[1]Pilot Plant Development and Training Institute, King Mongkut's University of Technology Thonburi, Bangkok, 10150, Thailand.

[a)] Corresponding author: nitikorn.nan@kmutt.ac.th
[b)] netithorn.dit@kmutt.ac.th
[c)] attakarn.jan@kmutt.ac.th
[d)] usa.boo@kmutt.ac.th

Abstract. Currently, despite pico hydroelectric turbine technology has the lowest cost per unit of energy among renewable energy sources, in Thailand, the cost reaches as high as 80 baht/W due to the exclusive need to import this technology. This research aims to bridge these gaps through a techno-economic analysis of a pico-hydro power turbine utilizing domestic technology. The turbine, developed by the KIRIWONG Community Enterprise, has a 3 kW capacity, producing electricity at 220 V 50 Hz. The study involved installing and testing a pico hydroelectric turbine system in a laboratory to assess its electricity production potential and to evaluate related economic indicators. The performance analysis of the hydroelectric turbine set revealed that the maximum electrical power and efficiency are linearly related to the flow rate. It was possible to generate 3 kW at a head of 100 m and a flow rate of 5.05 l/s, achieving an energy conversion efficiency of 60.5%, which is close to the design specifications. The investment return analysis for the hydroelectric turbine set indicated that the turbine offers a cost-effective investment, with benefits surpassing costs by the end of the project. The payback period was estimated at 1 year and 4 months when compared with operating a generator, and 4 years and 4 months when compared with purchasing electricity from the PEA, respectively. This results in an electricity cost of 0.26 baht/kWh, which can serve as valuable data for the future development and promotion of pico hydroelectric turbine systems in the country.

Keywords **—Pico hydro turbine, Techno-economic assessment, Renewable energy, KIRIWONG hydro turbine.**

I. INTRODUCTION

Background and Statement of the problem, in an effort to ensure equitable access to electricity, especially in remote rural areas lacking power grid connections, the Thai government has significantly promoted the use of renewable energy sources, including solar, hydro, wind, biogas, and biomass, over the past thirty years [1]. This focus has led to the development of numerous projects aimed at improving the living standards of people in remote areas through renewable energy technologies. However, the strategy has primarily involved importing these technologies, notably pico turbine technology. This reliance on imported technology presents a considerable challenge to the cost-effectiveness of renewable energy initiatives within the country [2].

Large-scale hydropower projects, which typically require substantial reservoirs and dam structures, entail significant capital investments and have sparked concerns regarding their impact on aquatic ecosystems and local communities. In contrast, pico-scale hydropower projects, often operating on a 'run of river' basis, minimize ecological disruption by avoiding the need for large dam infrastructure, instead utilizing simpler barriers or weirs to manage water flow [3]. Defined as hydropower systems with a capacity of less than 5 kW, pico-hydropower has emerged over the past three decades as a viable, sustainable, and cost-effective solution for off-grid electricity and mechanical power generation, particularly in rural electrification efforts. The Pelton turbine, a specific type of water impulse turbine, has garnered increased attention from researchers for its suitability in these settings [4]. Over recent decades, extensive research, including both numerical and experimental studies, has been conducted to improve the Pelton turbine's efficiency and design [5][6].

Although the installation and use of pico hydropower turbines have been successful in several developing countries [7], in Thailand, the cost of importing hydropower turbine technology for electricity production can reach up to 80 THB/W. This high cost is due to the absence of domestic manufacturers of pico hydropower turbines, necessitating exclusive technology importation, thus hindering widespread installation and use. Recognizing this need, the KIRIWONG community enterprise has initiated a project to develop a 3 kW hydropower turbine prototype suitable for rural household use. However, past research assessing the cost-effectiveness of renewable energy installations has mainly focused on solar cells, with very few studies on pico hydropower turbine technology, especially in Thailand.

Objective, this research is undertaken to address this gap, focusing on assessing the cost-effectiveness of installing and operating the KRW-3000 model KIRIWONG turbine, a domestically engineered technology. It incorporates findings from technical evaluations alongside an examination of pertinent economic indicators, offering insights crucial for the future development and promotion of pico hydropower systems within the country.

Expected benefits, the study provides quantitative data on the performance and investment returns associated with the installation and use of hydroelectric turbines employing domestic technology. This information can serve as a valuable resource for all sectors involved in the production and promotion of renewable energy. It supports the efficient promotion and installation of small hydroelectric turbines within the country. Moreover, it offers rural communities an alternative technology for generating electricity for both domestic and agricultural purposes. This technology has the potential to improve the living standards of villagers, increase agricultural income, reduce expenses related to fuel procurement, and ensure 24-hour electricity supply to agricultural areas. Consequently, this contributes to enhanced stability for families and local communities.

Conceptual Framework, A conceptual framework for a pico hydropower system begins with a water source and a penstock channeling water to a turbine. This turbine converts the water's kinetic energy into mechanical energy, which is then transformed into electrical energy by a generator. A control system manages the flow and output, ensuring that the electricity is distributed to meet local needs with minimal environmental impact. The interactions between performance testing results and the assessment of relevant economic indicators are depicted in Figure 1.

II. METHODOLOGY

As shown in the figure, the performance analysis module is utilized to predict the electrical energy generated from specified flow rates and head levels. This prediction relies on equations developed from various experimental results obtained in a research laboratory setting. Meanwhile, the economic analysis module incorporates estimates of cost and performance into a discounted cash flow analysis. This analysis is instrumental in evaluating various economic metrics, such as Net Present Value (NPV), Internal Rate of Return (IRR), and the Levelized Cost of Energy (LCOE).

1. Performance analysis

The performance assessment of electricity generation using hydro turbines is primarily dependent on two main variables: flow rate and head level. The efficiency of the Pelton hydro turbine is calculated based on the input power supplied to the turbine. This input power is contingent upon the water's flow rate and head level, as well as the electrical power generated. This relationship is illustrated in Equation (1):

$$\eta_{turbine} = \frac{E_{power}}{\rho g Q H} \qquad (1)$$

Where, $\eta_{turbine}$ is the efficiency of the hydro turbine system, E_{power} is the electrical power output (W), ρ is the density of water (kg/m3), g is the acceleration due to Earth's gravity (m/s^2)

2. Economic analysis

The three principal economic indicators used in the economic evaluation of the pico hydro system, as considered in this article, are as follows:

(a) The total net present cost (NPC).

The Net Present Cost (NPC) of a system is calculated as the present value of all incurred costs over its lifetime, subtracted by the present value of all revenues it generates within the same period. Incurred costs encompass capital, replacement, operation and maintenance, fuel, emissions penalties, and costs associated with purchasing power from the grid. Conversely, revenues consist of salvage value and revenue from sales to the grid [8]. The formula for the total net present cost is presented in Equation (2):

$$C_{NPC} = \frac{TAC}{CRF(i,N)} \qquad (2)$$

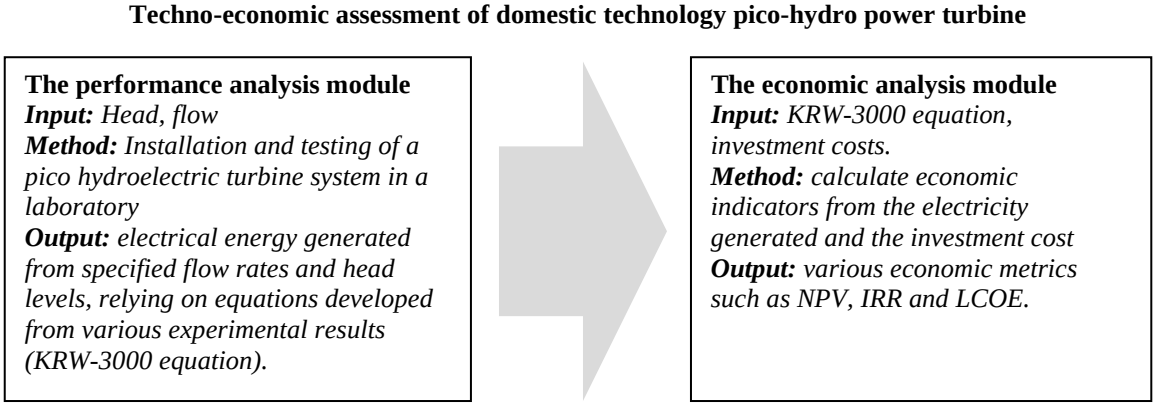


Fig. 1. A conceptual framework

Where, TAC is the total annualized cost (USD/year) and CRF, the capital recovery factor is a function of the annual real interest rate (i), and the project lifetime (N) in years. CRF is given in equation (3) [9]:

$$CRF(i, N) = \frac{i(1+i)^N}{(1+i)^N - 1} \quad (3)$$

The NPC criteria are as follows: If the NPC is positive, the investment is economical; a negative "NPC" denotes that the returns are worth less than the initial investment. The financial viability of a project is uncertain with a zero NPC. For independent projects, the higher the NPC value, the better.

(b) Levelized cost of energy (LCOE).
The LCOE is the ratio of lifetime costs to lifetime electric power generation, both discounted back to a base year reflecting the average capital cost. It is defined as the average cost per kWh of useful electrical energy produced by the system. The LCOE is given as follows:

$$LCOE = \frac{TAC}{E_{anloadserved}} \quad (4)$$

where: $E_{anloadserved}$ is the total annual load served by the system (kWh).

(c) Simple payback period
The simple payback period (SPB) is another crucial financial tool used to estimate the number of years required to recover the project cost of an investment. It is also important to note that the SPB does not consider the time value of money and is calculated with a zero-discount rate.

$$\sum_m \Delta IIC_m \leq \sum_m \Delta S_m \quad (5)$$

where ΔIIC_m is the incremental investment cost at zero discount rate in period m and ΔS_m is sum of the annual cash flows net annual costs at zero discount rate in period m.

3. Sample
In this study, the KIRIWONG KRW-3000 hydro turbine set, a 3kW hydroelectric turbine system produced using domestic technology, was used as the sample group. It comprises a Pelton turbine with a 200 mm PCD, chosen for its high-strength stainless steel construction designed to withstand water speed impacts at 36 m/s. It features a direct-drive 3-phase, 3 kW, 6-pole induction motor; a machine base with a protective cover made from cast aluminum for affordability and strength; adjustable nozzles for water flow control; and an induction generator control unit with capacitors for magnetic field stimulation. Additionally, it includes an automatic voltage control system to maintain stable electrical output, accommodating both main and dummy electrical loads to ensure consistent power delivery at 3 kW. This hydroelectric turbine, designed for optimal efficiency and durability, incorporates a Pelton turbine set (KRW-3000) tailored for a 100 m head and a 5 l/s flow rate. It produces single-phase AC electricity at 230 V and 50 Hz. Essential to its design are stainless steel Pelton blades, chosen for their high strength (750 – 850 N/mm^2) and cost-effectiveness, connected directly to a 3 kW, 3-phase, 6-pole induction motor that serves as an efficient generator. The structure is supported by an aluminum cast base, ensuring strength (90-250 N/mm^2) at a reduced cost, and features adjustable nozzles for precise water flow. The inclusion of a control unit for the induction generator, complete with capacitors and a voltage meter, alongside an automatic voltage control system that stabilizes power output through load division, highlights the prototype's advanced engineering. This system, capable of maintaining a consistent power level amid load changes, underscores the KIRIWONG turbine's potential as a forward-looking solution for hydroelectric power in appropriately selected locations, as shown in Figure 2.

4. Data collection
The data used in this study includes information on the amount of electricity generated, which is based on test installations in a laboratory setting, and data on production costs, which relies on actual data gathered from inquiries made to the KIRIWONG hydro turbine community enterprises. The details of the data collection are as follows:

4.1 Technical data collection
Hydro turbine testing in the laboratory is conducted in accordance with relevant standards, including (1) IEC TC 4: Hydraulic Turbines, (2) IEC 61116: Electromechanical Equipment Guide for Small Hydroelectric Installations, (3) IEC 620-06: 2011 Hydraulic Machines - Acceptance Tests of Small Hydroelectric Installations, and (4) IEC 60041:1995 Field Acceptance Tests to Determine the Hydraulic Performance of Hydraulic Turbines, Storage Pumps, and Pump-Turbines. Measurement and data collection tools for the hydro turbine electricity generation tests include a pressure gauge, brand SFC, model 300, for measuring head level; an ultrasonic flow meter, brand Portaflow, for measuring flow rate; electric bulbs, brand Philips, model E27, for electrical load; a tachometer, brand Digicon, model DT235T, for measuring shaft speed; a power meter, brand Fluke, model 179, for measuring electrical data; and an oscilloscope, brand GW Instek, model GDS-2102A, for measuring electrical signals, as shown in Fig. 1.

(a)

(b)

Fig. 2. The KRW-3000 KIRIWONG hydro turbine (a) and Test rig (b)

4.2 Economic data collection

The study aims to assess the economic viability of deploying KRW-3000 hydro turbine systems by comparing their costs, returns, and profitability with those of conventional electricity sources from regional power utilities and generator usage. It assumes a 10-year project lifespan, an interest rate of 6.244%, turbine efficiency of 60%, and specific maintenance costs, including capacitor replacement every three years. The analysis seeks to provide a clear financial picture, considering a capacity factor of 0.5 and comparing electricity rates for grid connections (4.42 THB/kWh) and generator replacements (20 THB/kWh based on 0.5 l/kWh fuel consumption), referenced from the supplier of a 3kW generator. thus, offering a comprehensive evaluation of the economic benefits of hydro turbine installation in comparison to traditional energy sources.

4.2.1) Investment Costs

The investment costs were obtained through inquiries to the KIRIWONG hydro turbine community enterprises and include material costs, equipment, and labor for constructing the hydro turbine system. The unit price data for the equipment used in the project are shown in Table 1.

TABLE I
THE INSTALLATION COSTS FOR THE RRW-3000 HYDRO TURBINE SET.

Item	Price (THB)
KRW-3000 hydro turbine set	52,000
- Machine base	8,200
- 200 mm Pelton turbine blade	8,000
- 2 Adjustable jet nozzles	7,000
- 3 kW 3 HP 6 Pole motor	16,000
- Stainless steel flexible pipe, Y-junction, and nipples	2,500
- Generator control unit	2,500
- Automatic voltage control unit	3,000
- Miscellaneous (e.g., nuts, machining, handling fees)	4,800
Installation accessories:	15,000
- Civil work (e.g., foundations, buildings)	10,000
- Electrical work (e.g., electrical cables, poles)	5,000
Water conveyance work:	182,400
- Water pipes and fittings	182,400
Total	249,400

4.2.2) Return on Investment

The return on investing in small hydroelectric turbine systems lies in providing villagers with an alternative technology for electricity production tailored to their area and meeting their consumption needs. The direct benefit realized is villagers having access to electricity 24 hours a day, a result of using the hydroelectric turbine system, which generates 13,140 kWh per year (with a capacity factor of 0.5).

III. RESULTS AND DISCUSSION

1.System performance

1.1) Testing at various head levels and flow rates

The results of the Pelton hydro turbine testing at a head level of 40 m and various flow rates are presented in Fig. 3. It was found that without connecting to an electric bulb, the turbine's speed is directly proportional to the flow rate (the opening extent of the Spear Valve). When the Spear Valve is turned by one rotation, the turbine reaches a maximum speed of 800 rpm at a flow rate of 1.2 l/s. Upon connecting to an electric bulb with a generator, the turbine's speed decreases linearly with electrical power, and when the electric bulb is fully connected, the system can deliver a maximum power of 230 W at a speed of 714 rpm. With an increased flow rate, it was found that the maximum power is linearly related to the flow rate, as shown in Fig. 4., where the system can deliver a maximum power of 1,385 W at a flow rate of 5.1 l/s, resulting in an efficiency of 69%, consistent with the tests at other head levels and flow rates.

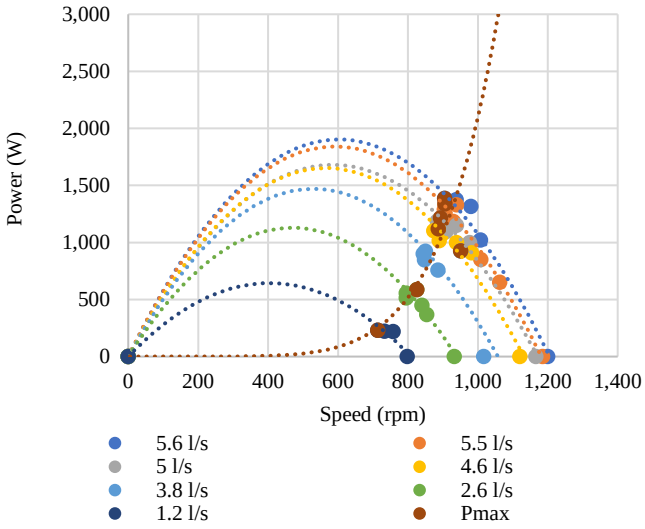


Fig. 3. Results of the hydro turbine testing at a head of 40 m.

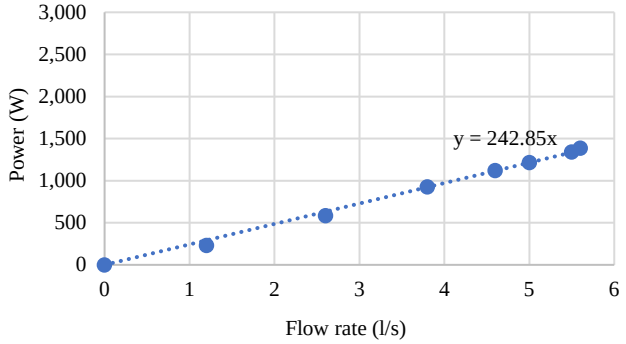


Fig. 4. Relationship between maximum power and flow rate at a head of 40 m.

1.2) Performance of the hydro turbine set at designed head levels and flow rates Due to limitations of the motor/pump set used in simulating head levels in the laboratory, a maximum of 60 m at a flow rate of 5 l/s could be achieved, thus unable to test the performance of the hydro turbine set at the designed head level (100 m) and flow rate (5 l/s). Therefore, the performance study of the hydro turbine set at designed head levels and flow rates was based on the relationship between electrical power and flow rate at different head levels, as shown in Fig. 5.

From the equations representing the relationship between electrical power and flow rate at different head levels in Fig. 5., it is found that:

P (W) at 90 m = 450.95 x Q (l/s) (6)
P (W) at 80 m = 408.77 x Q (l/s) (7)
P (W) at 70 m = 375.99 x Q (l/s) (8)
P (W) at 60 m = 345.42 x Q (l/s) (9)
P (W) at 50 m = 289.30 x Q (l/s) (10)
P (W) at 40 m = 242.37 x Q (l/s) (11)

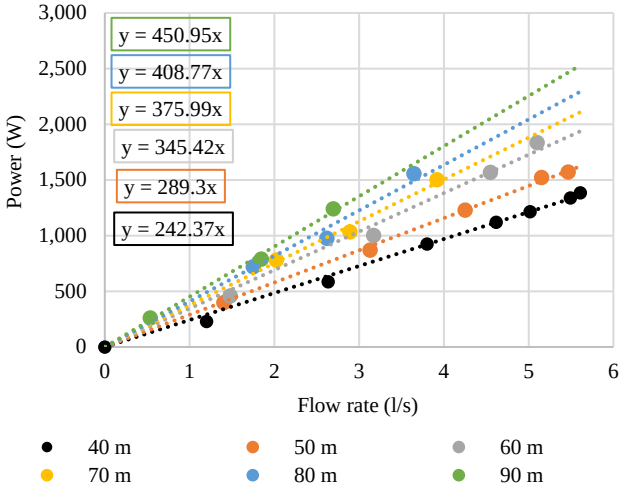


Fig. 5. Relationship between electrical power and flow rate at different head.

From the above equations, a graph showing the relationship between head level and flow rate at different electrical powers can be constructed, as presented in Fig. 6., and equations (12) – (13).

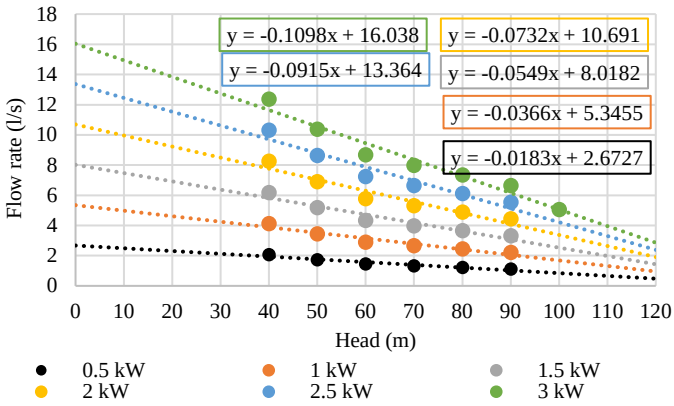


Fig. 6. Relationship between head level and flow rate at different electrical powers

Q (l/s) at P 3.0 kW = (-0.109 x H (m)) + 16.03 (12)
Q (l/s) at P 2.5 kW = (-0.091 x H (m)) + 13.36 (13)
Q (l/s) at P 2.0 kW = (-0.073 x H (m)) + 10.69 (14)
Q (l/s) at P 1.5 kW = (-0.054 x H (m)) + 80.18 (15)
Q (l/s) at P 1.0 kW = (-0.036 x H (m)) + 5.34 (16)
Q (l/s) at P 0.5 kW = (-0.018 x H (m)) + 2.67 (17)

From equation (12), it is found that the hydro turbine set can generate electricity of 3 kW at a head level of 100 m when operating at a flow rate of 5.05 l/s, achieving an efficiency of 60.5% in converting water energy into electrical energy, which is close to the design requirements.

2. Return on investment

The financial analysis of installing a hydroelectric turbine system amounts to 249,400 baht, with maintenance costs for capacitor replacement at 600 baht every three years over a 10-year lifespan. This includes the benefits of reducing oil usage for electricity production at 30 baht/kWh, maintenance costs for a generator at 3,000 baht per year, and the comparison of installing and using a generator or reducing electricity purchases at 4.42 baht/kWh compared to buying electricity from the Provincial Electricity Authority (PEA), as shown in Table 2.

TABLE II
THE RESULTS OF ECONOMIC ANALYSIS

Indicator	Compared to Generator	Compared to PEA
Net Present Value (Baht)	1,683,592	171,981
Internal Rate of Return (%)	106%	19%
Benefit/Cost Ratio (B/C Ratio)	7.72	1.7
Payback Period (Years)	1 Year	4 Years 4 Months
Electricity Rate (Baht/kWh)	0.26	0.26

The financial viability analysis, when compared to using a generator by evaluating the net present value of benefits over

the project's lifetime, found that the developed prototype Pelton hydro turbine set within the project is a worthwhile investment. At the project's end, the benefits exceeded the costs by 1,683,592 baht, with a payback period of 1 year, resulting in an electricity cost of 0.26 baht/kWh. This rate is 115 times lower than the cost of electricity from a generator (30 baht/kWh), aligning with the financial viability analysis of purchasing electricity from the PEA. It was also found that the developed prototype Pelton hydro turbine set is a cost-effective investment, with benefits exceeding the costs by 171,981 baht at the project's end. The payback period for installing the hydro turbine set is 4 years and 4 months, resulting in an electricity cost of 0.26 baht/kWh, which is about 17 times lower than the current electricity rate from the PEA for residential use exceeding 150 kWh/month (4.42 baht/kWh).

3. Sensitivity Analysis

The objective of this study is to explore the returns on investment when various related variables change. The analysis encompasses the following variables: electricity usage (load factor) at levels of 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, and 10%, as well as the lifespan of the hydroelectric turbine set, which ranges from 5 years, 10 years, to 15 years. This investigation aims to understand the impact on investment returns under different conditions, thereby aiding in more effective planning and decision-making regarding investments.

3.1) Electricity Usage (Load factor)

The study of cases where actual electricity usage is lower than the system's capacity to supply energy at 26,280 kWh/year (3 kW 24 hours 365 days/year), in terms of the percentage of the actual electricity energy produced (Capacity factor) at 90%, 80%, 70%, 60%, 50%, 40%, 30%, 20%, and 10%, as shown in Figure 6, found that the prototype hydro turbine set developed in the project is a worthwhile investment. The electricity cost produced from the hydro turbine set is lower than using a generator to produce electricity in all cases, even when the actual electricity usage drops below 90% of the actual electricity produced. This aligns with the comparison case with the electricity rate of the Provincial Electricity Authority for residential use, where at actual electricity usage lower than 10-80% of the actual electricity produced, the electricity cost from the system is lower than the rate of the Provincial Electricity Authority and will approximate the rate of the PEA when the estimated electricity usage decreases below 90% of the actual electricity produced.

3.2) Lifespan of the Hydroelectric Turbine Set

The study of the lifespan of the hydroelectric turbine set at 5 years, 10 years, 15 years, 20 years, as shown in Figure 7, found that the prototype hydro turbine set developed in the project is a worthwhile investment. The electricity cost produced from the hydro turbine set is lower than both using a generator and purchasing electricity from the PEA in all cases, even if the lifespan decreases to 5 years.

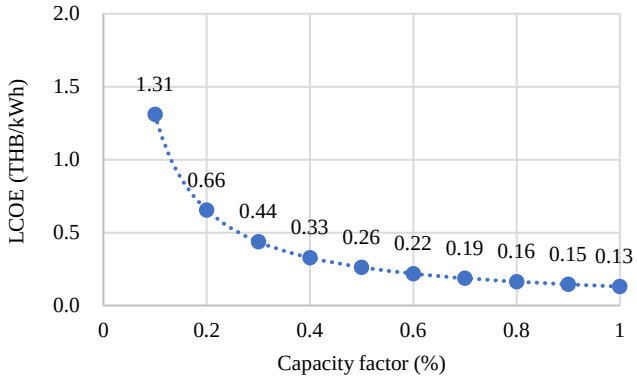


Fig. 6. Relationship between the electricity cost produced from the system and the electricity usage.

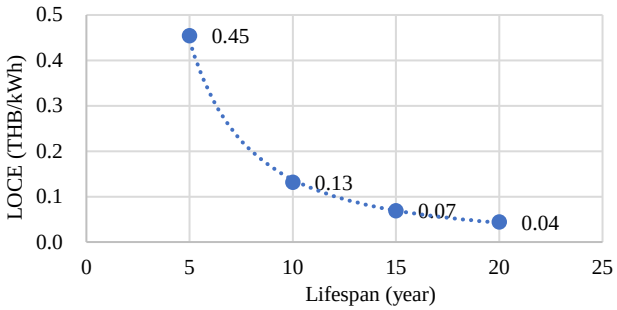


Fig. 7. Relationship between the electricity cost produced from the system and the lifespan.

Summary of the Study

Currently, small-scale hydroelectric turbines can be categorized into two groups based on design and manufacturing quality: simple designs and complex designs. The complex group requires high expertise in production, focusing on details such as blade geometry and twist angles, resulting in higher performance but at a greater cost. In contrast, the simple design group is characterized by lower initial costs, with many products manufactured in China and other Asian countries.

The performance testing of the hydro turbine set showed that the maximum electrical power and efficiency are exponentially related to the rotational speed, while they are linearly related to the flow rate. The prototype hydro turbine set can generate 3 kW of electricity at a head level of 100 m and a flow rate of 5.05 l/s, achieving an efficiency of 60.5% in converting water energy into electrical energy. Compared to research reports both domestically and internationally, the performance of the hydro turbine set in this research project surpasses that reported by Bryan R. Cobb and Kendra V. Sharp [4], which found the Turgo hydro turbine had an efficiency greater than 80% at a speed ratio of approximately 0.46, and Ibadullah Safdar et al., who studied the performance of impulse turbines and PMG in pico hydro systems, finding the maximum efficiency of hydro turbines to be 20.75% at a flow rate of 19.7 gpm (1.24 l/s) [3].

Compared with the performance of Pelton hydro turbines available internationally, the performance of the hydro turbine set in this research project is higher than that of the Pelton turbine set by Rainbow Power Company Ltd, which reported an efficiency of about 40%, but is similar the performance of the Pelton turbine set by Lingenhöle Technologie, which reported an efficiency of about 60% [12]. Meanwhile, the capability to maintain voltage and frequency levels could be achieved similarly, due to the use of the same voltage control techniques.

A study on the return on investment for installing a prototype Pelton hydro turbine, including water conveyance infrastructure, indicates an investment of 249,400 Baht. With benefits assessed in terms of Net Present Value over the project's lifespan, the analysis shows the investment to be worthwhile. The payback period for the prototype turbine is one year, with an electricity cost of 0.26 Baht/kWh, significantly lower than the cost from a conventional generator by 115 times. This aligns with the financial analysis comparing the option to purchasing electricity from the regional utility, where the investment also proves to be sound. The payback period for installing the hydro turbine set is 4 years and 4 months, with the produced electricity cost at 0.26 Baht/kWh, approximately 17 times lower than the regional electricity rates and less than the cost of small-scale commercial hydroelectric turbines available internationally, with a maximum power rating of 5 kW [13].

Discussions

The techno-economic study of the 3 kW KIRIWONG hydro turbine, utilizing domestic technology, was conducted through laboratory testing and investment indicator analysis. It was found that the maximum electrical power and efficiency are linearly related to the flow rate, with the system capable of generating 3 kW at a head level of 100 m and a flow rate of 5.05 l/s. This translates to an energy conversion efficiency of 60.5%, which is close to the design specifications. The study on the return on investment for installing and operating the hydroelectric turbine set found that the developed prototype turbine set offers a worthwhile investment, with benefits exceeding costs at the end of the project. The payback period was 1 year, and 4 years and 4 months when compared with installing and operating a generator and purchasing electricity from the Provincial Electricity Authority, respectively. This results in an electricity cost of 0.26 baht/kWh, which is lower than the cost of commercially available pico hydroelectric turbines abroad. Additionally, the sensitivity analysis revealed that the prototype hydroelectric turbine set remains a cost-effective investment even if its lifespan decreases to 5 years or if the electricity usage drops below 90% of the system's energy production capability. However, to maximize the benefits of installing the prototype hydro turbine set, the geographical conditions of the installation site must be considered, as the prototype in this research project is suitable for installations at head levels greater than 40 m.

IV. Conclusion

Although the hydroelectric turbine set studied generates electricity at a rate lower than that of conventional power generators and below the regional electricity tariff rates, it approaches the regional electricity tariffs when the estimated electricity usage falls below 90% of the actual power generated, even with a lifespan of less than 5 years. This indicates a high potential for installation, particularly in various mountainous regions of the country. However, to maximize the benefits of deploying these prototype hydroelectric turbines, the geographical conditions of the installation sites must be considered. The prototypes in this research are particularly suited for installations at head levels above 40 meters.

Acknowledgment

This research was funded by the National Science Research and Innovation Fund (NSRF)), managed by Program Management Unit for Competitiveness (PMU-C) in collaboration with the KIRIWONG Community Enterprise and King Mongkut's University of Technology Thonburi.

References

[1] AEDP, “Alternative Energy Development Plan (2012-2021),” International Journal of Renewable Energy, vol. 7, no. 1. pp. 1–10, 2012.
[2] K. Aroonrat and S. Wongwises, “Current status and potential of hydro energy in Thailand: A review,” Renewable and Sustainable Energy Reviews, vol. 46, pp. 70–78, 2015, doi: 10.1016/j.rser.2015.02.010.
[3] I. Safdar, S. Sultan, H. A. Raza, M. Umer, and M. Ali, “Empirical analysis of turbine and generator efficiency of a pico hydro system,” Sustainable Energy Technologies and Assessments, vol. 37, no. August 2019, p. 100605, 2020, doi: 10.1016/j.seta.2019.100605.
[4] B. R. Cobb and K. V. Sharp, “Impulse (Turgo and Pelton) turbine performance characteristics and their impact on pico-hydro installations,” Renew Energy, vol. 50, pp. 959–964, 2013, doi: 10.1016/j.renene.2012.08.010.
[5] S. Chitrakar, B. W. Solemslie, H. P. Neopane, and O. G. Dahlhaug, “Review on numerical techniques applied in impulse hydro turbines,” Renew Energy, vol. 159, pp. 843–859, 2020, doi: 10.1016/j.renene.2020.06.058.
[6] X. Liu, Y. Luo, B. W. Karney, and W. Wang, “A selected literature review of efficiency improvements in hydraulic turbines,” Renewable and Sustainable Energy Reviews, vol. 51, pp. 18–28, 2015, doi: 10.1016/j.rser.2015.06.023.
[7] E. Gallego, A. Rubio-Clemente, J. Pineda, L. Velásquez, and E. Chica, “Experimental analysis on the performance of a pico-hydro Turgo turbine,” Journal of King Saud University - Engineering Sciences, vol. 33, no. 4, pp. 266–275, 2021, doi: 10.1016/j.jksues.2020.04.011.

[8] W. M. Amutha and V. Rajini, "Cost benefit and technical analysis of rural electrification alternatives in southern India using HOMER," Renewable and Sustainable Energy Reviews, vol. 62, pp. 236–246, 2016, doi: 10.1016/j.rser.2016.04.042.
[9] L. Olatomiwa, S. Mekhilef, A. S. N. Huda, and O. S. Ohunakin, "Economic evaluation of hybrid energy systems for rural electrification in six geo-political zones of Nigeria," Renew Energy, vol. 83, pp. 435–446, 2015, doi: 10.1016/j.renene.2015.04.057.
[10] M. Emmanuel, D. Akinyele, and R. Rayudu, "Techno-economic analysis of a 10 kWp utility interactive photovoltaic system at Maungaraki school, Wellington, New Zealand," Energy, vol. 120, pp. 573–583, 2017, doi: 10.1016/j.energy.2016.11.107.
[11] "Rainbow Power 300W Hydro Generator," http://www.reuk.co.uk/wordpress/hydro/rainbow-power-300w-hydro-generator/.
[12] "Pico Hydroelectric Plants." Accessed: Apr. 30, 2023. [Online]. Available: https://www.lingenhoele.at/en/turbine-construction/pico-hydroelectric-plants/
[13] M. F. Basar, R. Sapiee, S. Rahman, Z. Hamdan, S. Borhan, and K. Sopian, "Cost analysis of pico hydro turbine for power production," Adv Environ Biol, vol. 8, no. 14, pp. 147–151, 2014.

Application of Remote Sensing and Machine Learning for Quantifying Ocean Acidification Parameters in a Marine Protected Area

Joel Colobong[1*, a)], Kanchana Nakhapakorn[1*, b)], Jirataya Roemmontri[1, c)] Phaisarn Jeefoo[2, d)] and Naruemon Pratanwanich[3, e)]

[1]Faculty of Environment and Resource Studies, Mahidol University, Salaya, Nakhon Pathom, Thailand
[2]School of Information and Communication Technology, University of Phayao, Phayao, Thailand
[3]Department of Mathematics and Computer Science, Chulalongkorn University, Bangkok, Thailand

[a)] joelbaysac.col@student.mahidol.ac.th
[b)] kanchana.nak@mahidol.ac.th; kanchana.nak@mahidol.edu
[c)] jirataya.pan@mahidol.ac.th; jirataya.pan@mahidol.edu
[d)] p.jeefoo@up.ac.th
[d)] naruemon.p@chula.ac.th

Abstract. Ocean Acidification is a slow-onset disaster. Current climate models estimate that the pH levels of the oceans are expected to become more acidic with the increased CO_2 emissions. Profound effects will be felt by different marine organisms particularly shellfish, crustaceans, and corals. Marine protected areas are centers of marine biodiversity and have a big role in the marine ecosystem. This study focuses on building a database of different ocean acidification parameters such as temperature, salinity, total alkalinity, pH, dissolved inorganic carbon, partial pressure of CO_2, and aragonite saturation rate in a marine protected area. MODIS-AQUA data was used as data source coupled with in-situ records on salinity, temperature, and pH in order to build machine learning models to calculate the salinity and pH parameters. This is then used as data inputs to the CO2Calc marine carbonate system calculator to obtain all other parameters. The study was able to build a time series data using support vector machine for salinity estimation and multilayer perceptron for pH. Looking at the different parameters, we see that both study areas are still within manageable range of its parameters in relation to ocean acidification. However, it was observed that there is a gradual reduction on the pH level suggesting increased acidity as well as reduction in the aragonite which is crucial for shell and coral formation.

***Index Terms* — CO2CALC, Marine Carbonate System, MODIS-AQUA, Ocean Acidification**

I. INTRODUCTION

Increased anthropological activities have accelerated climate change in our planet. Greenhouse gas (GHG) emissions have increased due to increase demands on agriculture, services, and industries. Carbon Dioxide (CO_2) is a naturally occurring greenhouse gas and accounts for majority of GHG emissions borne from human activities. While there are natural carbon sinks found in forests, soils, and the ocean, which absorb CO_2, there is still a risk that they may not be effective should there GHG emissions levels go out of control.

Ocean Acidification (OA) occurs due to the increased absorption of CO2 which reacts with the saltwater (1) that produces carbonic acid (HCO_3^-) and excess hydrogen ions (H^+) that lower the pH level of our oceans (2). This HCO_3^- further breaks down into carbonate ions (CO_3^{2-}) and H^+ over a period (3). Subsequently, these excess H^+ also undergoes chemical reaction with carbonate ions (CO_3^{2-}) present in seashells, corals, and other similar marine organisms in such a way that it breaks down and weakens the structures of their exoskeletons (4). The chemical reaction that shows the breakdown of the process of ocean acidification [1]:

$$CO_2\,(g) = CO_2\,(l) \quad (1)$$

$$CO_2\,(aq) + H_2O\,(l) = H^+(aq) + HCO_3^-\,(aq) \quad (2)$$

$$HCO_3^-\,(aq) = H^+\,(aq) + CO_3^{2-}\,(aq) \quad (3)$$

$$CaCO_3\,(s) = Ca^{2+}\,(aq) + CO_3^{2-}\,(aq) \quad (4)$$

where: g – gas, l – liquid, aq - aqueous solution, s - solid

In marine carbonate systems, it is important to know the different parameters that identifies capacity of our oceans and seas against ocean acidification. While we know that the chemical changes as shown in equations 1 to 4 describe the process of ocean acidification, directly measuring the concentrations of these elements prove to be difficult. Hence, we analyze these changes using different variables. This includes the parameters of pH, dissolved inorganic carbon, total alkalinity, and partial pressure of CO_2 [2]. These parameters with the inclusion of physical properties of saltwater such as salinity and temperature are key to computing for the concentration of the water-soluble calcium

carbonate in the form of Aragonite, which is a key component of building exoskeleton of crustaceans and corals. This concentration of Aragonite, commonly called Aragonite Saturation rate (Ω_{Ar}) which can be computed with the following formula [3]:

$$\Omega_{Ar} = \frac{[Ca^{2+}]\,[CO_3^{2-}]}{K^*_{sp'}} \quad (5)$$

Where: Ω_{Ar} – Aragonite Saturation Rate;
$[Ca^{2+}]$ – Calcium Ion Concentration;
$[CO_3^{2-}]$ – Carbonate Ion Concentration;
$K^*_{sp'}$ – equilibrium thermodynamic solubility product

Marine protected areas are defined as centers of marine biodiversity which houses rare species of marine flora and fauna. They also serve as areas of marine conservation and different marine ecosystem services, and cultural heritage. Ocean acidification not only lowers the ocean pH level, it also has various effects on different marine organisms which includes changes in behavioral patterns, stunted growth, increase susceptibility against diseases and failure in reproduction [4]. With such effects in a marine ecosystem, countries with vast marine resources are at risk for food insecurity and loss of livelihood. Therefore, it is imperative to monitor the state of our oceans.

The combination of Remote Sensing and Machine Learning has been used for a long time in the field of ocean studies. Analysis of Ω_{Ar} can be attributed to the increase in levels of CO_2 emissions due human activities [5]. A study on the Ω_{Ar} of the Great Barrier Reef using regional coupled circulation-biogeochemical model (eReefs) found that the central reef waters have already experienced a decrease in the aragonite saturation and is expected to expand northward and westward in future projections [6]. Another research was able to estimate carbonate levels in coral reefs using changes in benthic community using remote sensing and in-situ validation and they were able to estimate the decline in aragonite saturation levels in the study area [7]. A remote sensing inversion model of pH combined with deep learning based on global ocean acidification databases have provided a new method to calculate ocean acidification parameters and recognized the need to have more training data to increase the accuracy of machine learning model predictions [8].

The Philippines is a country located in the southeastern part of Asia. It is situated on the "Coral Triangle," marine biodiversity area which includes the countries of Malaysia, Brunei, Indonesia, Timor Leste, Papua New Guinea, and Solomon Islands. This area houses the greatest number of corals, 6 of the 7 marine turtles, mangrove forest in the world, and around 2,000 species of reef fish [9]. The country has around 2.5 million hectares of coral reefs with around 500 species, 2.73 million hectares of seagrass with 19 different species, 5 out of 7 marine turtle species, 28 species of marine mammals, as well as around 168 cartilaginous fishes. The estimated marine monetary value of coastal and marine ecosystem is estimated to be around PHP 15.3 trillion (USD 2.78 billion) [10].

There are very limited studies on the actual state of Philippine seas against ocean acidification. However, based on available studies, pH level is averaged around 8.02. However, there are places that recorded pH levels less than 8.0 and Ω_{Ar} around 3.0. These points with lower than usual carbonate system parameters are located where there are human activities (aquaculture, mariculture, etc.) [11]. Seeing that the threat of ocean acidification especially to marine protected areas will have a huge impact, this study was conducted to establish a time-series data that contains the different ocean acidification parameters which include temperature, salinity, pH which reflects the concentration of H^+ ions present and determines the acidity of the water, total alkalinity (A_T) which is described as the buffering capacity of a water body, partial pressure of CO_2 (pCO_2) which measures the concentration of dissolved CO_2, dissolved inorganic carbon (DIC) which is the measures the total carbonates present, and aragonite saturation rate which is a metric for measuring available water-soluble $CaCO_3$ essential to corals, shellfish and other similar marine species. These parameters can establish benchmarks for studies on mitigation and prevention of ocean acidification in the area. This study will utilize a combination of remote sensing and machine learning to identify these parameters.

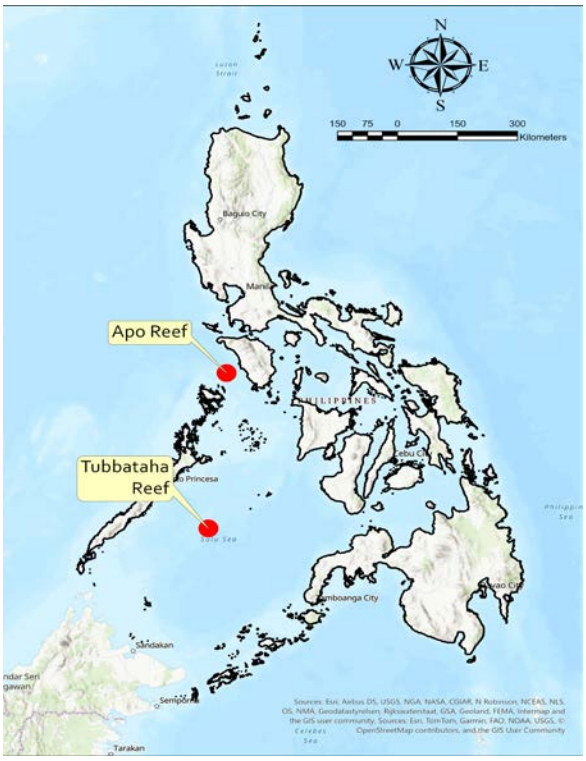


Fig. 1 Map showing locations of study area

II. METHODOLOGY

STUDY SITES

The study focuses on Tubbataha Reef Natural Park (TRNP) and Apo Reef Natural Park (ARNP) which are two of the biggest coral reef system in the Philippines (see Figure 1). ARNP is a triangular shaped coral reef system situated in the middle of Sulu Sea at approximately 150km away from Puerto Princesa City, Palawan in the southwestern part of Philippines. It is estimated to be around 968km^2 and is divided into three areas namely, Jessie Beazley Reef in the northern part, North Atoll located at the right part and South Atoll located at the southern part of the reef system. It has been declared as part of Ramsar Sites and houses around 360 coral species and the Napoleon Wrasse fish [12]. On the other hand, Apo Reef Natural Park is located in the Mindoro Strait between the provinces of Palawan and Occidental Mindoro. The 27.5 km^2 reef system is divided into three islands named Apo Island, Apo Menor Island, and Cavos del Bajo. It is home to 160 flora, 64 birds and 10 mammal species.

DATA COLLECTION AND PROCESSING

This study uses MODIS-AQUA Level 3 Monthly (3km resolution) data coming from NASA Ocean Color Website (oceancolor.gsfc.nasa.gov) covering the period of July 2002 to December 2022 (samples of downloaded satellite imagery in Fig. 2a – 2c). In-situ measurements for temperature, salinity and pH for both study sites were taken from the Water Quality Reports of ARNP from 2021-2022 and Ecosystems Report of TRNP from 2014 – 2022[12,13], respectively. Sample points from the reports are then used as points of reference for the Pixel Extraction Tool of SEADAS 8.40. Around 78 points were collected for the training data based on the available data. The data collected from this process are collected into a csv file (See Table I for list of downloaded products).

TABLE I
DOWNLOADED OCEAN COLOR PRODUCTS AND OA PARAMETERS ASSOCIATED WITH PRODUCT

Products downloaded	Ocean Acidification Parameters
Sea Surface Temperature (SST)	Temperature, pH, Total Alkalinity [14]
Chlorophyll Concentration (Chlorophyll-a)	pH [15]
Remote Sensing Reflectance at 412nm … 678nm (Rrs 412nm – Rrs 678nm)	Salinity [16,17]

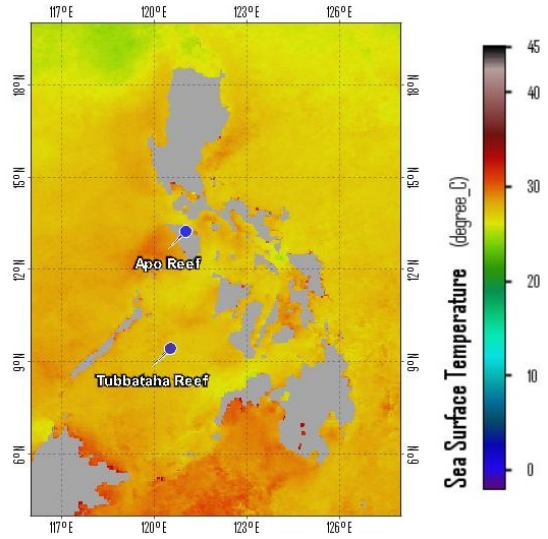


Fig 2a. Sample downloaded SST data

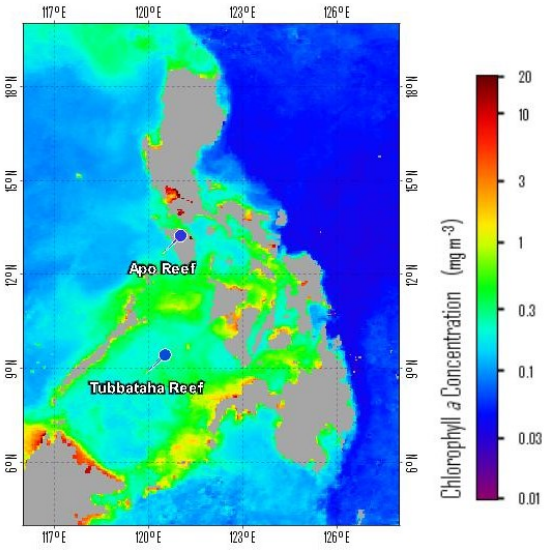


Fig 2a. Sample downloaded Chl-a data

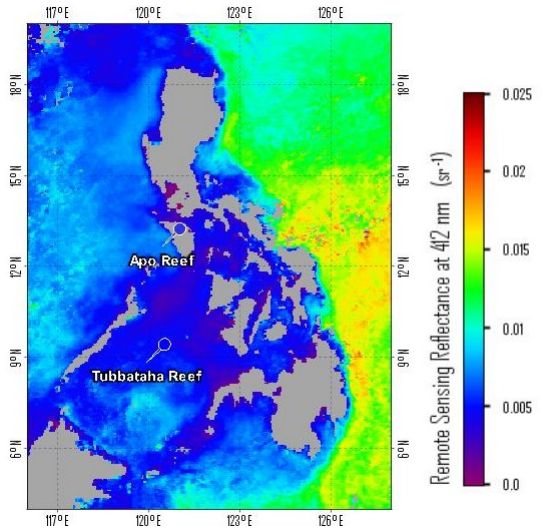


Fig 2c. Sample downloaded R_{RS} 412nm data

DETERMINING OPTIMUM MACHINE LEARNING MODEL FOR PREDICTION OF SALINITY AND pH

Waikato Environment for Knowledge Analysis (WEKA) is a data mining software developed by University of Waikato which features different machine learning (ML) tools that are useful for data mining. It has built-in machine learning models which makes it easy to use especially for researchers with beginner-level knowledge on machine learning and data mining [18]. Four kinds of built-in ML models in WEKA which are Multiple Linear Regression (MLR), Multilayer Perceptron (MLP), Support Vector Machine (SVM), and Random Forest (RF) are used for this research. The collected training data consisting of SSTs, chl-a, and R_{RS} including in-situ temperature, salinity, and pH is used against these four models. The collected data of different RRS from 412nm to 678nm (10 bands) are used to determine salinity while the data of SST and Chl-a are used for estimating pH. The ML Model with the best performance shall be used for the estimation of salinity and pH (see Fig. 3).

The training data are trained in the software using sampling methods whole sample as training data, 10-fold cross-validation, and 80-20 data test split train. For MLR, the data are trained using Akaike criterion for model selection, M5 method, and Greedy method using Akaike information metric. For classification using MLP, the data undergoes 5 nodes and 1 hidden layer. The training data is normalized, standardized or non-standardized/normalized for a PolyKernel SVM. For the RF model, classification parameter uses 100 iterations and maximized forest depth.

We will then validate the different ML models using coefficient of determination (commonly referred as R^2) and root mean square error (RMSE). Its formulae are shown below:

$$R^2 = \frac{n(\Sigma xy)-(\Sigma x)(\Sigma y)}{\sqrt{[n\Sigma x^2-(\Sigma x^2)][n\Sigma y^2-(\Sigma y)^2]}}^2 \quad (6)$$

$$RMSE = \sqrt{\sum_{i=1}^{n}\frac{(\widehat{y_i}-y_i)^2}{n}} \quad (7)$$

Where: r – correlation coefficient
x, y – data points
RMSE – root mean squared error
n – number of observations
y_i – y_n – observed values
$\bar{y}_i$ – $\bar{y}_n$ – predicted values

A_T is determined using the relation between salinity and temperature using empirical formula [14]. The equation used is applicable for tropical area located between 30°S to 30°N and with temperatures more than 20°C and salinity ranging from 31 to 38psu. Predicted data that do not fit the criteria is not included into the time series data.

$$A_T = 2{,}305 + 58.66(SSS-35) + 2.32\,(SSS-35)^2 - 1.41(SST-20) - 0.04(SST-20)^2 \quad (8)$$

Where: SSS – Salinity at sea surface (in psu)
SST – Temperature at sea surface (in °C)
A_T – Total alkalinity (in µmol/kg)

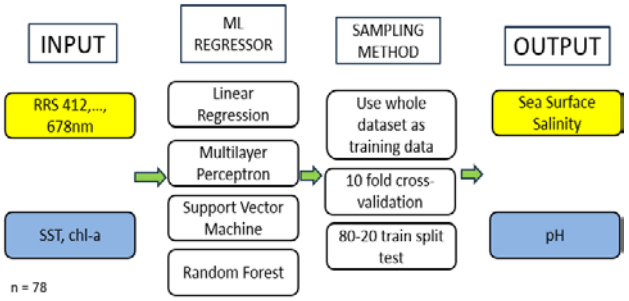


Fig. 3 Process of choosing the best ML model for salinity using RRS data (yellow) and pH using SST and chl-a (blue)

CALCULATING OCEAN ACIDIFICATION PARAMETERS USING CO2CALC MARINE CARBONATE SYSTEM CALCULATOR

Marine carbonate systems can be understood using at least two of these parameters: A_T, DIC, pCO_2, and pH. Meanwhile, the physical properties of saltwater which are salinity, pressure, and temperature help in knowing the CO_2 solubility and the coefficient constants for the carbonate system equations [19]. The CO2Calc version 1.20 is used for determining other parameters for ocean acidification in this study [20].

Using the Salinity, Temperature, pH and A_T as inputs, they are arranged into a csv file for batch processing in the CO2Calc software. Outputs for this software include DIC (characterized by Total CO_2 in the software, in µmol/kg), pCO_2 (in µatm), and Ω_{Ar} (no unit). In addition, in order to compute for the other parameters, the software uses coefficient constants based on other studies readily available in the software (see Table II). You can refer to Table 2 for the table of constants.

TABLE II
CALCULATION CONSTANTS USED IN CO2CALC SOFTWARE [21]

Particulars	Constants
CO_2 Constant	K1, K2 from Mehrbach et al., 1973 refit by Dickson and Millero, 1987
$KHSO_4$	Dickson
Boron Total	Uppstrom, 1974
Air-sea Flux	Wanninkhof, 1992

TABLE III
VALIDATION OF ML PERFORMANCE FOR PH AND SALINITY

ML	Parameters	Validation	pH			Salinity		
			Using whole dataset as training data	10-fold cross-validation	80%-20% split	Using whole dataset as training data	10-fold cross-validation	80%-20% split
MLR	No Attribute Selected	R^2	0.31	0.22	0.58	0.20	0.16	0.40
		RMSE	0.38	0.40	0.38	2.55	2.62	2.34
	M5 Method, Greedy Method	R^2	0.31	0.22	0.59	0.19	0.12	0.15
		RMSE	0.38	0.40	0.37	2.57	2.69	2.71
SVM	Normalized Traing Data	R^2	0.29	0.20	0.61	0.21	0.13	0.54
		RMSE	0.40	0.41	0.41	2.56	2.74	2.06
	Standardized Training Data	R^2	0.28	0.18	0.61	0.35	0.25	0.69
		RMSE	0.40	0.41	0.41	2.34	2.53	1.90
	Not Standardized/ Normalized	R^2	0.30	0.18	0.61	0.08	0.13	0.01
		RMSE	0.40	0.41	0.41	2.93	3.31	2.94
MLP	1 Hidden Layer; 5 Nodes	R^2	0.40	0.28	0.83	0.46	0.14	0.60
		RMSE	0.44	0.40	0.32	2.20	2.95	2.16
RF	Iteration - 100; Max Depth -no limit	R^2	0.94	0.35	0.82	0.94	0.24	0.13
		RMSE	0.13	0.37	0.26	0.89	2.51	2.79

III. RESULTS AND DISCUSSION

Checking into the correlation between the different variables, we see that there is no significant correlation between the different R_{RS} with Salinity [18]. However, we see strong correlations between the different R_{RS} wavelengths. Likewise, we see that there is good correlation between pH and temperature while there is no significant correlation between Chlorophyll-a concentration and pH [15].

In terms of the performance of the different ML Models using the training data, we see that using the whole dataset as training data is not advisable as it makes it hard to validate the predictions made if done using this sampling method (see Table III). As shown in the performance of the Random Forest model wherein it has high R^2 of 0.94 and lower RMSE when the whole training data is used but failed in performance for both 10-fold cross validation and 80-20 training split. This suggests overfitting of the data in RF model [22].

Generally, k-fold cross-validation provides a more robust ML model and addresses overfitting and is suitable for small datasets [23]. However, its performance in this dataset is not as compared to 80:20 train-test split for both pH and Salinity models. In this situation, increasing the number of datasets through more in-situ measurements may help to increase the performance in k-fold cross-validation sampling method. Basing the existing dataset, we have chosen to utilize SVM with standardized training data with R^2 of 0.69 and RMSE of 1.90 to use for salinity, while we used MLP with R^2 of 0.83 and RMSE of 0.32 for pH estimation.

In this research, we have assumed that the SST data from Ocean color website as the remote sensing derived temperature of our study site, the measured SST from the study area is from around 28 – 32 °C with average of 30 °C which is typical measurement for tropical waters. Seasonal variation is also observed with lower temperatures observed

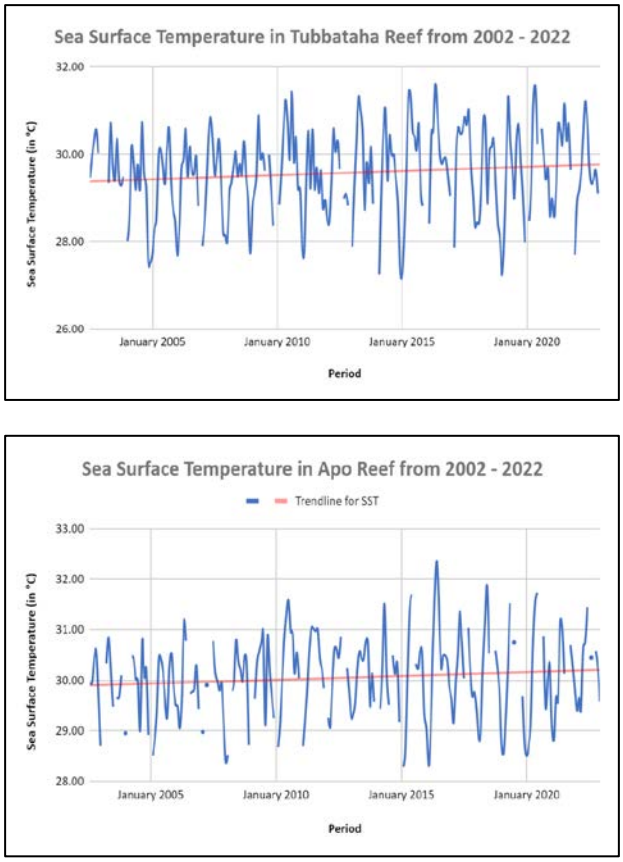
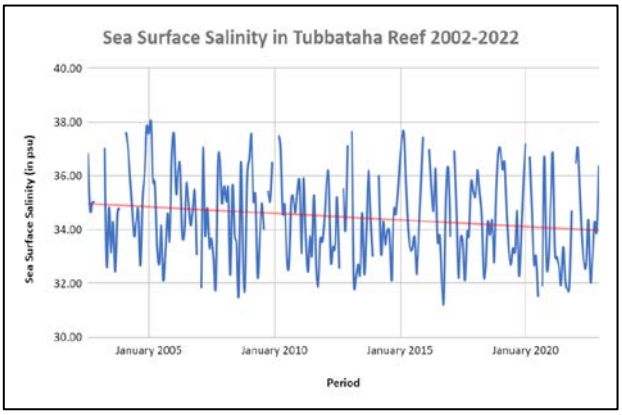


Fig. 4 Time series data of temperature in Tubbataha Reef (above) and Apo Reef (below)

during November to February while higher temperatures are observed in March to May.

The recorded salinity in both study sites is averaged at around 34 psu with ranges of 31-37 psu. Higher salinity levels are observed during November to February while lower levels of salinity are observed during March to May.



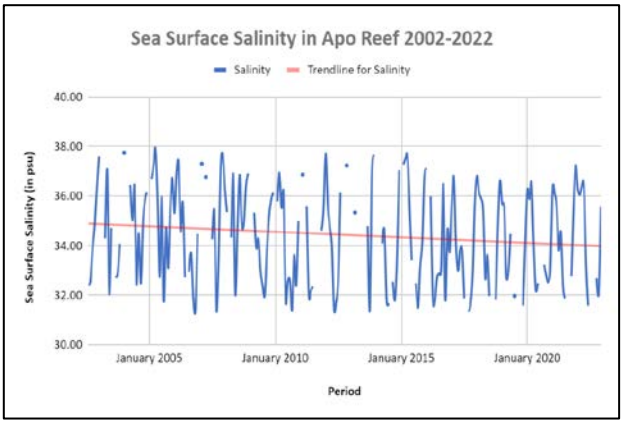
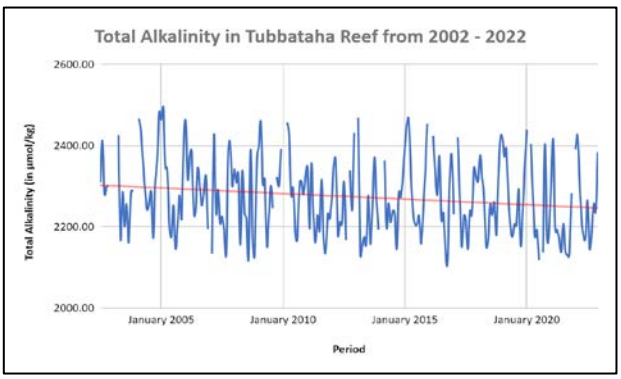


Fig. 5 Time series data of salinity in Tubbataha Reef (above) and Apo Reef (below)

In calculating A_T, we take into consideration the parameters of equation 5 wherein SST should be >20°C and SSS from 31 psu <SSS<38 psu. Predicted data that do not fit the criteria is discarded. Estimated A_T is around 2,270 µmol/kg ranging from 2,100 – 2,500 µmol/kg which is the average A_T in tropical areas. Higher levels of A_T are observed during when there are higher levels of salinity and lower temperature.



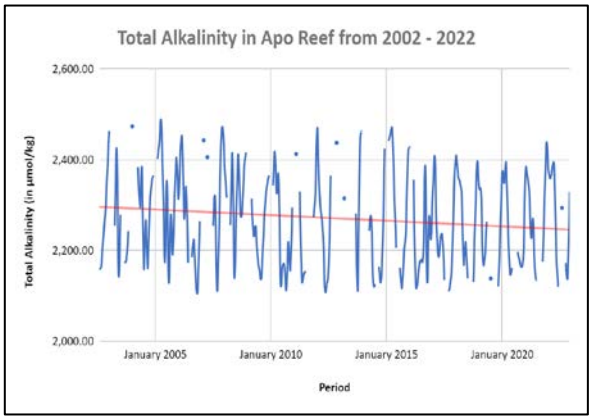
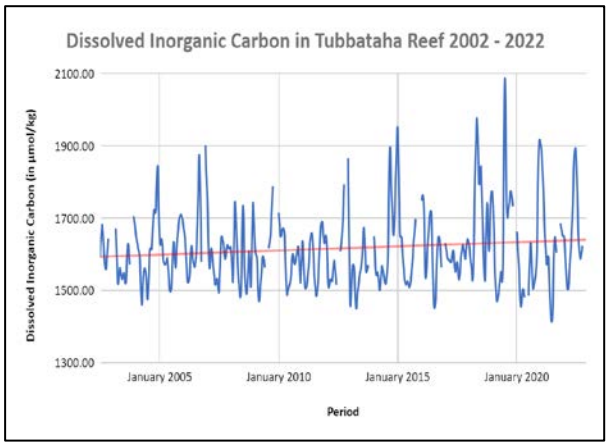


Fig. 6 Time series data of total alkalinity in Tubbataha Reef (lower left) and Apo Reef (above)

Dissolved inorganic carbon measures the total carbonate in the marine system. In our study sites, we observed an average of 1,550 µmol/kg of DIC ranging from 1,200 – 2,200 µmol/kg. An lower value of DIC is observed where there is higher value of observed total alkalinity.



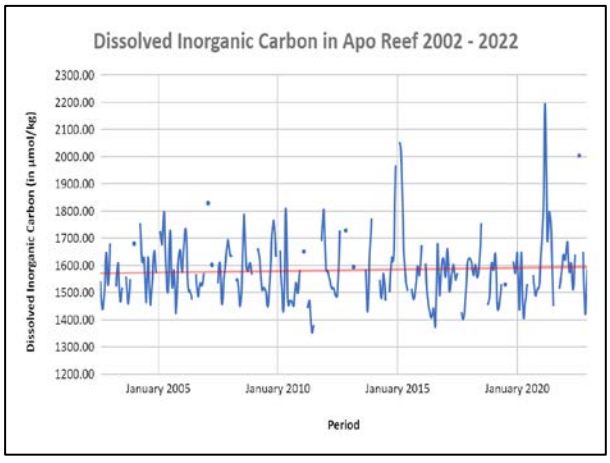
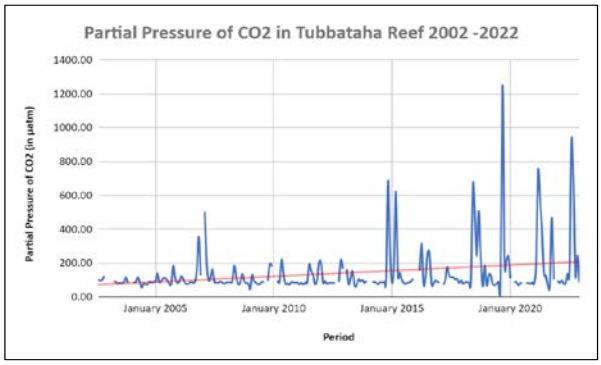


Fig. 7 Time series data of dissolved inorganic carbon in Tubbataha Reef (above) and Apo Reef (below)

Partial Pressure of CO_2 or pCO_2 measures the saturation of CO_2 in a marine system. It follows Henry's law which states that the partial pressure of a dissolved gas is equal to the atmospheric pressure it exerts over the water surface. In our study site, we have recorded around 120 µatm with minimum observed at around 15 µatm to 1,200 µatm. We observed the higher pCO_2 levels during periods where high DIC and low A_T are observed. Looking at Fig. 8, we notice that there are sharp increases in the pCO_2 levels. This may be attributed to the upwellings which commonly occur at the study site [24] however this still merits further investigation.



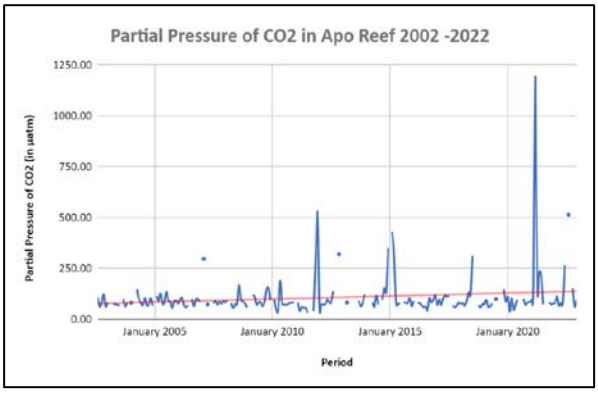
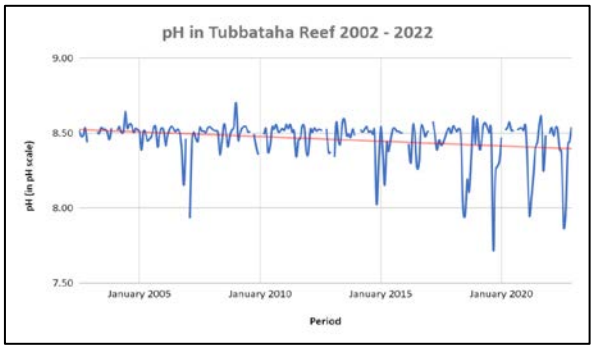


Fig. 8 Time series data of partial pressure of CO2 in Tubbataha Reef (above) and Apo Reef (below)

Looking into the pH level of the study sites, we observed that average pH level of around 8.5 with recorded ranges from 7.6 – 8.8. Average pH level is still greater than the global recorded average observed by UN last 2020 which is 8.05. Lower levels of pH are observed during high concentration of pCO_2, higher values of DIC, and lower levels of A_T.



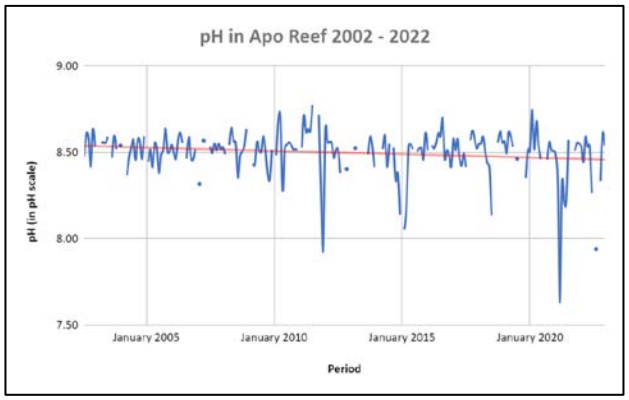
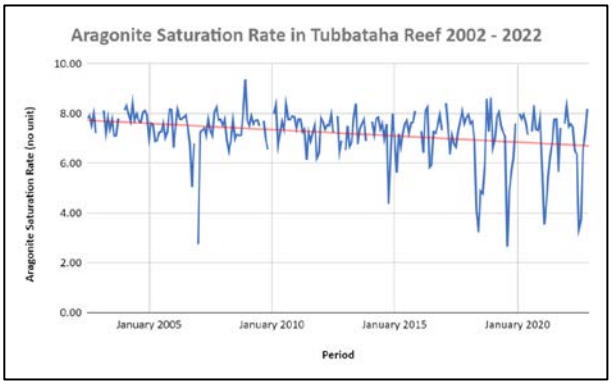


Fig. 9 Time series data of pH in Tubbataha Reef (lower left) and Apo Reef (above)

Aragonite saturation rate is a good metric to identify marine biodiversity present in a marine environment. An Ω_{Ar} >1 means that it is conducive for marine biodiversity especially for crustaceans and coral reefs which rely on aragonite for shell building and skeletal growth. Both study sites observed Ω_{Ar} of 7.5 with lowest recorded of 2 and highest value of 9. Observing the different parameters, we can see that lower Ω_{Ar} is observed when there is lower pH levels and A_T, higher concentrations of DIC and pCO_2.



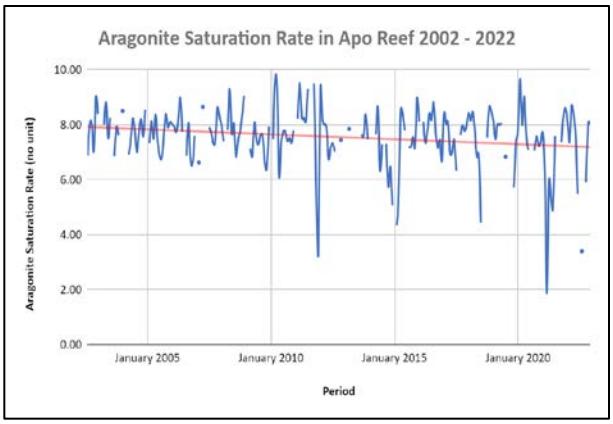


Fig. 10 Time series data of aragonite saturation rate in Tubbataha Reef (above) and Apo Reef (below)

TABLE IV
SUMMARY OF OBSERVED OCEAN ACIDIFICATION PARAMETERS IN TUBBATAHA REEF AND APO REEF (BELOW)

	Temperature (in °C)	Salinity (in psu)	Total Alkalinity (in µmol/kg)	Dissolved Inorganic Carbon (in µmol/kg)	Partial Pressure of CO2 (in µatm)	pH (in pH scale)	Aragonite Saturation Rate
Tubbataha Reef							
Average	30.05	34.44	2,271.42	1,583.13	107.35	8.5	7.55
Standard Deviation	0.79	1.9	106.7	122.31	104.91	0.14	1.08
Minimum	28.3	31.33	2,002	1,207.65	15.93	7.63	1.88
Maximum	32.36	37.92	2,486.40	2,194.06	1193.21	8.78	9.66
Apo Reef							
Average	29.58	34.47	2,274.31	1,617.51	140.82	8.46	7.21
Standard Deviation	0.95	1.64	93.42	107.85	146.02	0.14	1.02
Minimum	27.23	31.41	2,114.31	1,429.22	41.55	7.72	2.65
Maximum	31.6	37.98	2,492.20	2,086.75	1249.98	8.7	9.37

Table IV summarizes the observed values of both study sites. Looking at the over-all data of ocean acidification parameters, we see that while pH and Ω_{Ar} are above the average global levels, there is noticeable gradual decrease in its level over since 2002 to present. This clearly shows that the area is vulnerable to ocean acidification and climate change.

IV. CONCLUSION

Given that the study areas are very big and getting in-situ measurements on a regular basis is very costly and time-consuming, we can see that remote sensing with machine learning is a cost-effective way to ascertain the different parameters of ocean acidification. This partnered with existing in-situ measurements can help create a model that can make accurate estimations of ocean acidification in a localized area. We see in this study that using one type of satellite and sensor can create data on ocean acidification parameters. For future studies, it is recommended to utilize combinations of different satellites and sensors as well as increase the number of in-situ measurements to improve the machine learning model performance. We also suggest to increase the number of training data to increase the accuracy of the machine learning models.

Based on the time series data, we see that pH and Ω_{Ar} levels will gradually decrease and it will affect not only on marine biodiversity but also on food security. We can also suggest for future studies on the effects of external forces (El Niño-Southern Oscillation, anthropogenic activities, upwelling, hydrometeorological hazards, etc.) in how it accelerates ocean acidification.

In a country which is highly vulnerable against climate change and ocean acidification, there is a need to start doing regular monitoring against its effects. The increased awareness in evidence-based research helps to justify the need to address these issues. This may include updating existing laws such as the Expanded National Integrated Protected Areas System Act or enacting new laws that include risk and vulnerability assessments against the effects of climate change. Appropriating funds for installation of on-site measurements which can be monitored by the local environmental authorities and capacity building on the use of remote sensing and machine learning is also highly recommended to ensure prevention and mitigation against ocean acidification and climate change.

ACKNOWLEDGMENT

The authors would like to acknowledge the NASA Ocean Color Website for the readily available data and ease of access. We would also like to thank the Philippine Department of Environment and Natural Resources – Biodiversity Management Bureau for providing the authors with in-situ data for model training and analysis. Lastly, the authors would also like to thank the Thailand International Cooperation Agency for their support to this research.

REFERENCES

[1] Riebesell U, Fabry VJ, Hansson L. Guide to best practices for ocean acidification research and data reporting. 2011.
[2] Millero FJ. Thermodynamics of the carbon dioxide system in the oceans. Geochimica et Cosmochimica Acta. 1995;59(4):661-77.
[3] Feely RA, Byrne RH, Acker JG, Betzer PR, Chen C-TA, Gendron JF, et al. Winter-summer variations of calcite and aragonite saturation in the Northeast Pacific. Mar Chem [Internet]. 1988;25(3):227–41. Available from: http://dx.doi.org/10.1016/0304-4203(88)90052-7
[4] Baag S, Mandal S. Combined effects of ocean warming and acidification on marine fish and shellfish: A molecule to

ecosystem perspective. Science of The Total Environment. 2022;802:149807.
[5] Schoenrock KM, Schram JB, Amsler CD, McClintock JB, Angus RA, Vohra YK. Climate change confers a potential advantage to fleshy Antarctic crustose macroalgae over calcified species. Journal of Experimental Marine Biology and Ecology. 2016 Jan 1;474:58-66.
[6] Mongin M, Baird ME, Tilbrook B, Matear RJ, Lenton A, Herzfeld M, et al. The exposure of the Great Barrier Reef to ocean acidification. Nature Communications. 2016;7(1):10732.
[7] Shaw EC, Hamylton SM, Phinn SR. Incorporating benthic community changes into hydrochemical-based projections of coral reef calcium carbonate production under ocean acidification. Coral Reefs [Internet]. 2016;35(2):739–50. Available from: http://dx.doi.org/10.1007/s00338-016-1407-2
[8] Jiang Z, Song Z, Bai Y, He X, Yu S, Zhang S, et al. Remote Sensing of Global Sea Surface pH Based on Massive Underway Data and Machine Learning. Remote Sensing. 2022;14(10):2366.
[9] Asaad I, Lundquist CJ, Erdmann MV, Costello MJ. Delineating priority areas for marine biodiversity conservation in the Coral Triangle. Biological Conservation. 2018;222:198-211.
[10] Azanza, R. V., Alino, P. M., Cabral, R. B., Juinio-Menez, M. A., Pernia, E. M., Mendoza, R. U., & Siriban, C. S. (2017). Valuing and managing the Philippines' marine resources toward a prosperous ocean-based blue economy. Public Policy, 18, 1-26.
[11] Reyes M, Pavia R, van Hooidonk R. Ocean acidification in the Philippines and the potential role of water pollution management in mitigating an unaddressed threat. Regional environmental change. 2023;23(3):107.
[12] Dygico M, Songco A, White AT, Green SJ. Achieving MPA effectiveness through application of responsive governance incentives in the Tubbataha reefs. Marine Policy. 2013;41:87-94.
[13] Apo Reef Natural Park Protected Area Management Office. Submission on the Annual Water Quality Monitoring Report for Apo Reef Natural Park for CY 2021 - 2022. 2022 Aug.
[14] Alaba, M.R., Gedoria, G., Alarcon, R., and Songco A. 2022. TRNP Ecosystem Research and Monitoring Report 2022. Tubbataha Protected Area Management Board.
[15] Lee K, Tong LT, Millero FJ, Sabine CL, Dickson AG, Goyet C, et al. Global relationships of total alkalinity with salinity and temperature in surface waters of the world's oceans. Geophysical Research Letters. 2006;33(19).
[16] Nakano Y, Watanabe YW. Reconstruction of pH in the Surface Seawater over the North Pacific Basin for All Seasons Using Temperature and Chlorophyll-a. Journal of Oceanography. 2005;61(4):673-80.
[17] Qing S, Zhang J, Cui T, Bao Y. Retrieval of sea surface salinity with MERIS and MODIS data in the Bohai Sea. Remote Sensing of Environment. 2013;136:117-25.
[18] Daqamseh ST, Al-Fugara Ak, Pradhan B, Al-Oraiqat A, Habib M. MODIS Derived Sea Surface Salinity, Temperature, and Chlorophyll-a Data for Potential Fish Zone Mapping: West Red Sea Coastal Areas, Saudi Arabia. Sensors. 2019;19(9):2069.
[19] Eibe Frank, Mark A. Hall, and Ian H. Witten (2016). The WEKA Workbench. Online Appendix for "Data Mining: Practical Machine Learning Tools and Techniques", Morgan Kaufmann, Fourth Edition, 2016.
[20] Land PE, Shutler JD, Findlay HS, Girard-Ardhuin F, Sabia R, Reul N, et al. Salinity from Space Unlocks Satellite-Based Assessment of Ocean Acidification. Environmental Science & Technology. 2015;49(4):1987-94.
[21] Robbins, L.L., Hansen, M.E., Kleypas, J.A., and Meylan, S.C., 2010, CO2calc—A user-friendly seawater carbon calculator for Windows, Max OS X, and iOS (iPhone): U.S. Geological Survey Open-File Report 2010–1280, 17 p
[22] Mitchell T. Machine Learning. New York, NY: McGraw-Hill Professional; 1997.
[23] Raschka S. Model evaluation, model selection, and algorithm selection in machine learning. arXiv preprint arXiv:1811.12808. 2018 Nov 13.
[24] Chen C-TA, Hou W-P, Gamo T, Wang SL. Carbonate-related parameters of subsurface waters in the West Philippine, South China and Sulu Seas. Marine Chemistry. 2006;99(1):151-61.

Poster presentation

Preparation of Amine-Functionalized Rice Husk Ash for Selective Adsorption Polyphenols from Green Tea

Poungmanee Phromchana[1, a)], Nadrada Maneesri [1, b)], Paphada Pathomnatikul[1, c)], Jinda Chuemue[2, d)] and Sakdinun Nuntang[2, e)]

[1]Montfort College, Chiang Mai, 50000, Thailand,
[2]Industrial Chemistry Innovation Program, Faculty of Science, Maejo University, Chiang Mai 50290, Thailand

[a)] poungmaneeph@gmail.com,
[b)] noodee1246@gmail.com
[c)]prim.paphada@gmail.com
[d)]jinda05062544@gmail.com,
[e)]Corresponding author: Sakdinun.nt@gmail.com

Abstract. This study's purpose is to prepare amine-functionalized rice husk ash (RHA-NH_2) as an adsorbent for selective adsorption polyphenols from green tea solution. The RHA–NH_2 were synthesized successfully via post grafting method using (3-Aminopropyl)-triethoxysilane (APTES), [3-(2-Aminoethylamino)propyl]-trimethoxysilane (AEA) and 3-[2-(2-Amino ethylamino)ethylamino] propyl-trimethoxysilane (AEEA) as amine group precursors. Surface and structural characteristics of RHA–NH_2 were examined using X-ray diffraction, N_2 adsorption–desorption analysis and Fourier transform infrared spectroscopy. From the adsorption experiment, the RHA-NH_2 adsorbents exhibited the selective adsorption of polyphenol more than caffeine. In addition, the RHA–NH_2 functionalized with AEEA exhibited effectiveness for polyphenol and caffeine removal. The maximum adsorption capacity of this RHA-NH_2-AEEA for polyphenol was 39.73 mg/g. Therefore, the RHA-NH_2-AEEA was an interesting adsorbent to separate polyphenol and caffeine in green tea solution.

***Index Terms*— Rice husk ash, Amine, Functionalize, Polyphenol, Adsorption**

I. INTRODUCTION

Green tea is the most famous drink in the world. The main bioactive components in green tea are catechins (polyphenols) and alkaloids (Fig. 1). Catechins, which constitute up to 30% of dry green tea leaves, possess various biological, physiological, and pharmaceutical effects, such as antioxidative and anticancer. In addition, green tea leaves contain 2–4.5 % of caffeine (Fig. 1), which stimulates respiration and the central nervous system and has been clinically used as a central stimulant and an adjunct in numerous analgesic preparations. According to regulations issued by the US Food and Drug Administration, the concentration of caffeine as an ingredient in food and beverages should be limited to 200 parts per million (ppm) (0.02%). Therefore, an effective, industry operational preparation method should be developed to obtain enriched catechins extract with as little caffeine content as possible for use in pharmaceuticals, functional foods, and beverages. Separation of tea polyphenols and caffeine are usually carried out by partition with water/chloroform. Organic solvent extraction method has the shortcomings of high cost, environmental pollution, and high toxicity of some of these organic solvents [1].

Recently, there has been an interest in using the adsorbents to separate tea polyphenols and caffeine from aqueous solution. Especially, mesoporous silica materials (MCM-41, SBA-15, and HMS) have been reported to be used as adsorbent for adsorption polyphenols in the solution [2]. Since, these materials possess high specific surface area (> 500 m^2 g^{-1}) and high mesoporosity. In addition, they possess high amount of silanol groups (Si–OH) on their surface. Therefore, the molecule of polyphenols could interact with silanol groups on their surfaces. However, those adsorbents are expensive, which resulted in high operating costs.

The low-cost adsorbent is interesting for use in the adsorption process and has a higher usage trend. Rice husk (RH) is an agricultural waste that produces about one-fifth of the annual rice production of the world, or around 550 million metric tons [3]. It can be applied for several uses such as energy production and as an adsorbent. RH has the potential to be accounted as an adsorbent because its main components are carbon and silica. In addition, rice husk ash (RHA) produced by heating rice husk at 700 °C possesses a higher amount of silica content (~84.3%) and has been used as a potential adsorbent in various fields [4].

To enhance the adsorption capacity and selectivity of silica adsorbents for substances, surface modification is useful by utilizing interactions between adsorbents and adsorbates. Therefore, the modified silica surface has been regarded as an effective adsorbent due to its high surface area which allows the binding of many surface groups. The amine-functionalized silica (SiO_2-NH_2) has received

substantial attention because the amine group has a positive impact on the performance of the adsorption systems of heavy metals, dyes, and other organic compounds [5]. Therefore, rice husk ash with high amount of silica content is interesting to modify its surface with amine groups to enhance its adsorption performance to applied as an effective low-cost adsorbent.

In this work, we prepared amine-functionalized rice husk ash (RHA-NH_2) as adsorbents. Effects of types of organic amines with different -NH_2 group functionalized on RHA surface are investigated. The synthesized adsorbents were used as adsorbent for selective adsorption studies of tea polyphenols and caffeine from aqueous solution.

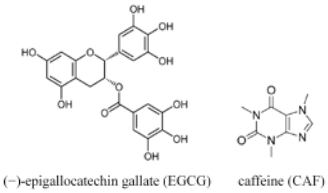


Fig. 1. Chemical structures of (−)-epigallocatechin gallate (EGCG) and caffeine (CAF).

II. METHODOLOGY

1. Preparation of amine-functionalized rice husk ash (RHA-NH_2)

The preparation of amine-functionalized rice husk ash was applied synthesis method reported by Imyim et al. [3]. First, 40 g of rice husk was submerged for 4 h in 1M HCl. The solid was rinsed 5 times with DI water and dried in a 110 ºC oven for 2 days. The material was generated into rice husk ash by heat-treating in a muffle furnace for 6 h at 700 ºC using ceramic crucibles. The rice hush treated with acid was designated as A-RHA.

A-RHA is categorized into 3 types, each amounting to 5 g. They were refluxed in toluene (100 mL) at 70 ºC for 2 h, and then slowly added amino silanes by a dropper into each portion such as 4 g of (3-Aminopropyl)-triethoxysilane (APTES), 4.03 g of [3-(2-Aminoethylamino)propyl]tri methoxysilane (AEA) and 4.80 g of 3-[2-(2-Amino ethyl amino)ethylamino]propyl-trimethoxysilane (AEEA). The mixtures were continually refluxed for 24 hours. The solids were then filtered, washed three times with ethanol and dichloromethane, and left to dry at room temperature. The obtained samples were designed as RHA-NH_2-APTES RHA-NH_2-AEA and RHA-NH_2-AEEA, respectively. The synthetic route for the modification of RHA with amino silanes is shown in Fig 2.

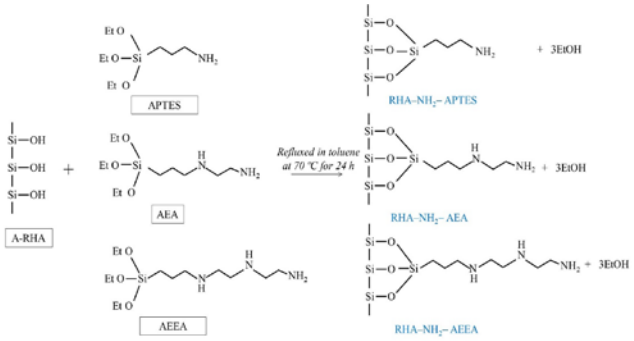


Fig. 2. Synthetic route for the modification of RHA with amino silanes.

2. Physical Measurements of adsorbents

The physical measurements of the synthesized adsorbents were analyzed using multiple techniques. Standard procedures were used to characterize the prepared physical and chemical properties of the adsorbents. X-ray powder diffraction (XRD) patterns were obtained on a PAN analytical Empyrean III X-ray diffractometer employing CuKα radiation and an X-ray power of 40 kV and 40 mA. N_2 adsorption-desorption measurements were carried out at 196 °C using a BEL Japan BELSORP-mini II instrument to determine the material's textural characteristics. All samples were pretreated at 150 °C for 2 h and then measured for exact weight prior to the adsorption. The presence of functional groups was determined using Fourier-transform infrared (FTIR) spectra (Perkin Elmer spectrum two: 900–4000 cm^{-1} in transmittance mode by KBr pellet technique). The number of organic contents in the silica structure was evaluated using an Elementar Vario Micro Cube CHNS elemental analyzer.

3. Preparation of green tea extract

Green tea (25 g) was refluxed with 200 mL of ethanol-water (60:40, v/v) solution at 90 °C for 1 h in triplicate. All the extracts were filtrated and evaporated to a fluid extract by removing the solvent under reduced pressure in a rotary evaporator at 90 °C. The obtained extracts were placed in a vacuum drying oven and dried at 50 °C until the mass was constant. The adsorption capacity of green tea polyphenol and caffeine on adsorbent was studied according to the following operation mode: 100 mL of extract solution (C_i = 2.0 mg/mL) [6].

4. Adsorption study

The adsorption was measured by batch method. The adsorbent (0.05 g) was added into an Erlenmeyer flask containing 25 ml of aqueous solution of green tea with certain initial concentration. The flask was sealed and shaken for 24 h at 25 °C. The equilibrium concentration of tea polyphenols or caffeine of the supernatant was determined by UV spectrophotometry at 728 nm of green tea polyphenols and at 272 nm of caffeine. The amount of green tea polyphenols or caffeine adsorbed per unit adsorbent (q_e, mg adsorbate per g adsorbent) was calculated according to a mass balance on the adsorbate concentration using Eq. (1):

$$q_e = \frac{(C_i - C_e)V}{m} \quad (1)$$

where C_i is the initial adsorbate concentration (mg L^{-1}), C_e is the equilibrium adsorbate concentration in solution (mg L^{-1}), V is the volume of the solution (L), m is the weight of adsorbent in g. All the experiments were performed duplicate and the average value from results was taken.

III. RESULTS AND DISCUSSION

The RHA consists mainly of silica (>90%) and some metallic impurities such as Fe^{3+}, Ca^{2+}, Na^{+}, K^{+}, and Mg^{2+} which influence the purity and color of silica [7]. The minor metallic impurities could be eliminated by HCl leaching treatments before the combustion process and designated as A-RHA [8]. After that, the A-RHA was functionalized with various aminosilane to modify its surface and designated as RHA-NH_2. The obtained materials were characterized by the XRD technique to investigate the structural properties shown in Fig. 3. XRD patterns of A-RHA and RHA-NH_2 materials were similar. The broad diffused peak between 15°–35° 2θ diffraction angle indicated that amorphous silica was formed. However, the heating A-RHA at 700 °C exhibited some structure change of amorphous silicon dioxide to a crystalline structure (observed at 2θ around 22° and 26°), such as quartz, cristobalite, or tridymite [3].

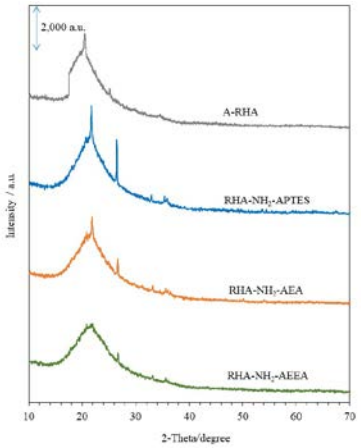


Fig. 3. XRD patterns of A-RHA and RHA-NH_2 adsorbents.

FTIR spectra of A-RHA and RHA-NH_2 adsorbents in the wavenumber range of 900–4000 cm^{-1} are shown in Fig. 4. The FTIR spectrum of these materials exhibited a broad band between 2800 and 3700 cm^{-1} which indicated the presence of both OH groups and the Si-OH group. In addition, the band at 1100 cm^{-1} is attributed to the stretching of Si–O–Si. However, the small peaks of RHA-NH_2 materials found at 2940 and 2866 cm^{-1} were assigned to the C–H stretching of methylene groups [4-5]. In addition, the FTIR spectrum of RHA-NH_2 observed in the range of 1465–1642 cm^{-1} is attributed to the bending vibration of the N–H group [6]. The appearance of these bands confirmed the successful functionalization of amino silanes onto the RHA surface.

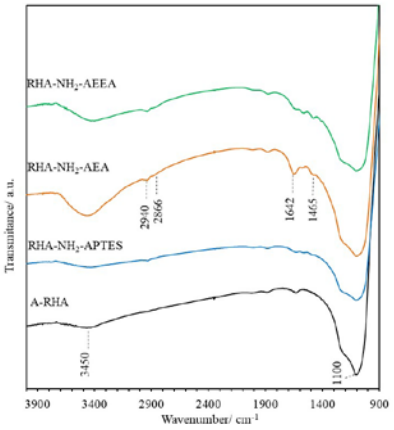


Fig. 4. FTIR spectra of A-RHA and RHA-NH_2 adsorbents.

Textural properties of synthesized materials were determined by N_2 adsorption–desorption analysis as shown in Table I. Based on its textural properties, A-RHA is suitable for use as an adsorbent due to its high specific surface area, pore diameter, and pore volume. After the surface of A-RHA with various amino silanes was modified, the surface area, pore diameter, and pore volume were reduced in comparison to pristine A-RHA. Additionally, the RHA-NH_2 materials exhibited an increase in CHN contents (Table II). These results suggested that the aminosilanes had been incorporated into the pore structure of A-RHA.

TABLE I
TEXTURAL PROPERTIES OF THE ADSORBENTS

Sample[a]	S_{BET}[b] (m^2 g^{-1})	D_p[c] (nm)	Vt[d] (cm^3 g^{-1})
A-RHA	136	3.71	0.34
RHA-NH_2-APTES	44	3.28	0.20
RHA-NH_2-AEA	45	3.28	0.19
RHA-NH_2-AEEA	45	3.28	0.19

[a] Dried sample
[b] BET surface area
[c] Pore diameter calculated using BJH equation
[d] Total pore volume

TABLE II
ELEMENTAL ANALYSIS OF THE ADSORBENTS

Sample[a]	CHN contents[b]		
	% C	% H	% N
A-RHA	0.22	0.22	0.08
RHA-NH_2-APTES	1.63	0.42	0.55
RHA-NH_2-AEA	3.34	0.76	1.28
RHA-NH_2-AEEA	4.36	1.03	1.86

[a] Dried sample
[b] Determined by CHNS analyzer

Figure 5 shows the polyphenol (EGCG) and caffeine (CAF) adsorption capacity of synthesis adsorbents in the green tea solutions. Their EGCG adsorption capacity was higher than their CAF adsorption capacity, indicating that these adsorbents favored the adsorption of EGCG rather than CAF. Since, the EGCG has several hydroxyl groups as shown in Fig.1 that they are good interact with the functional groups such as silanol group (-SiOH) and/or amine group (-NH_2) on RHA surface. In addition, the RHA-NH_2 adsorbents exhibited a greater adsorption capacity than pristine A-RHA, suggesting that EGCG and CAF were strongly adsorbed onto the RHA-NH_2 surface, resulting in an increase in adsorption capacity. Moreover, RHA-NH_2-AEEA with a high amount of amine group exhibited the highest adsorption capacity of EGCG and CAF at 39.73 and 1.89 mg/g, respectively. Since the amount of amine group could enhance the H-bond interaction between EGCG and CAF and surface of adsorbent, resulted in an increasing of EGCG and CAF removal. Therefore, RHA-NH_2-AEEA was a suitable adsorbent for use in further studies.

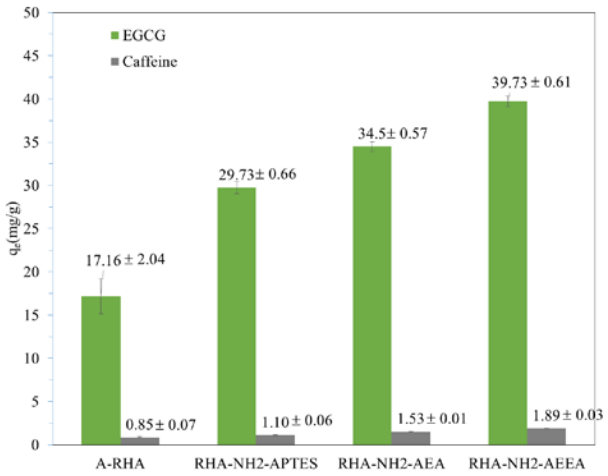


Fig. 5. EGCG and CAF adsorption capacity of A-RHA and RHA-NH_2 adsorbents from green tea solutions (adsorbent dose: 0.05 g, aqueous solution: 25 mL, agitation speed: 120 rpm, time: 24 h, temperature: room temperature).

IV. Conclusion

Rice husk ash was pre-treated by HCl and thermally treated at 700 °C and chemically modified it surface with 3 types of aminosilanes. It was found to be a suitable adsorbent for adsorption of EGCG and CAF from green tea solution compared to pristine rice husk ash. Characterization analysis revealed that RHA-NH_2 possessed an amorphous silica structure with an amine group incorporated into their surface. Moreover, they possessed a high surface area and pore volume, making them an efficient adsorbent. According to the adsorption results, the RHA-NH_2 adsorbents exhibited the selective adsorption of EGCG more than CAF. In addition, the RHA–NH_2 functionalized with AEEA exhibited effectiveness for EGCG and CAF removal. The maximum adsorption capacity of this RHA-NH_2-AEEA for EGCG and CAF were 39.73 and 1.89 mg/g, respectively. As a result, the RHA-NH_2-AEEA was an interesting adsorbent to separate EGCG and CAF in green tea solution.

Acknowledgment

The authors would like to acknowledge all staff of Industrial Chemistry Innovation Program, Faculty of Science, Maejo University for supporting the facilities to this study.

References

[1] Zhao, R., Y. Yan, L. Mingxian and Y. Husheng (2008). Selecive Adsorption of Tea Polyphenol from Aqueous Solution of the Mixture with Caffeine on Macroporous Crosslinked Poly(N-vinyl-2-pyrrolidinone), Reactive & Functional Polymers, Vol. 68, pp. 768-774.

[2] Cotea, V.V., C.E. Luchiana, N. Bilba and M. Niculana (2012). Mesoporous Silica SBA-15, a New Adsorbent for Bioactive Polyphenols from Red Wine, Analytica Chimica Acta, Vol. 732, pp. 180-185.

[3] Imyim, A., & Prapalimrungsi, E. (2010). Humic acids removal from water by aminopropyl functionalized rice husk ash. Journal of Hazardous Materials, 184(1), 775-781.

[4] Feng, Q., Lin, Q., Gong, F., Sugita, S., & Shoya, M. (2004). Adsorption of lead and mercury by rice husk ash. Journal of Colloid and Interface Science, 278(1), 1-8.

[5] Abu Rumman, G., Al-Musawi, T. J., Sillanpaa, M., & Balarak, D. (2021). Adsorption performance of an amine-functionalized MCM–41 mesoporous silica nanoparticle system for ciprofloxacin removal. Environmental Nanotechnology, Monitoring & Management, 16, 100536.

[6] Ma, G., J. Zhang, L. Chen, T. Liu, L. Yu, X. Liu and C. Lu (2014). Amino-functionalized Ordered Mesoporous Silica SBA-15, a Rapid and Efficient Adsorbent for the Adsorption of (-)-epigallocatechin Gallate from Green Tea Extract, RSC Advances, Vol. 4, pp. 41341-41347.

[7] Xu, W., Wei, J., Chen, J., Zhang, B., Xu, P., Ren, J., & Yu, Q. (2018). Comparative study of water-leaching and acid-leaching pretreatment on the thermal stability and reactivity of biomass silica for viability as a pozzolanic additive in cement. Materials, 11(9), 1697.

[8] Rafiee, E., Shahebrahimi, S., Feyzi, M., & Shaterzadeh, M. (2012). Optimization of synthesis and characterization of nanosilica produced from rice husk (a common waste material). International nano letters, 2, 1-8.

Development of Probiotic Legume Beverages with their Potential Antioxidant and Antidiabetic Properties

Suree Nanasombat

Department of Biology, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand

Corresponding author: suree.na@kmitl.ac.th

Abstract. The present study aimed to develop probiotic germinated and non-germinated legume beverages. Phytochemical properties of five germinated and non-germinated legume ethanolic extracts including mung bean (*Vigna radiata*), black gram (*Phaseolus mungo*), kidney bean (*Phaseolus vulgaris*), soybean (*Glycine max*) and peanut (*Arachis hypogaea*) were studied for their antidiabetic activity (α-amylase and α-glucosidase inhibition activity), antioxidant activity, total phenolic and indigestible polysaccharide content. The extract of non-germinated kidney bean had the strongest antioxidant activity with high reducing capacity of 0.396 mmol Fe(II)/g extract by ferric reducing antioxidant power (FRAP) assay and highest total phenolic content (477.17 mg GAE/g extract). The extract with higher antidiabetic activity was the extract of germinated soybean (23.3% α-amylase inhibitory activity and 10.2% α-glucosidase inhibitory activity) as compared to other extracts. The extracts with high amount of indigestible polysaccharide were non-germinated peanut (614.7 mg/g extract), germinated mung bean, black gram and soybean (443.48-580.19 mg/g extract). Then, the germinated and non-germinated kidney bean and soybean were selected to produce probiotic beverages. Survival of *Lactobacillus acidophilus* in these legume beverages during storage at 4 ℃ for 14 days was studied. Viable probiotic counts in germinated soybean beverage was higher than other beverages. Non-germinated kidney bean beverage had reducing capacity of 0.32 mmol Fe(II)/ 100 ml beverage and the highest phenolic content of 1,838.2 mg GAE/100 ml beverage. Of all, non-germinated kidney bean had the highest α-amylase inhibitory activity. These results suggested that germinated soybean and non-germinated kidney bean probiotic beverages were the potential non-dairy based beverages.

Index Terms— α-amylase, α-glucosidase, kidney bean, soybean, mung bean, black gram

I. INTRODUCTION

Legumes such as black gram, peanut, soybean, broad bean, chickpea, lentil and other beans are normally used as ingredients in foods of human in several areas in the world. These beans are rich in protein (16-50%), dietary fibre (10-23%) and are source of slow release carbohydrate, mono and polyunsaturated fatty acid, vitamin and minerals [1]. The legumes are usually consumed after processing to enhance taste and bioavailability by destroying their trypsin inhibitor, growth inhibitor and hemagglutinins. Germination is a key procedure to make the inaccessible proteins available [2]. Germinated legumes have been reported to rich in phenolic compounds such as phenolic acids, flavonoids, lignin, anthocyanin pigments (delphinidin, cyaniding, pelargonidin, malvidin and petunidin), vitamin C, riboflavin, niacin [3]. Legume consumption has been reported to decrease risk of suffering from obesity, type 2 diabetes and cancer in human as their constituents are antioxidant compounds which are good source of bioactive compounds [1, 2].

Reactive oxygen species (ROS) including superoxide anion hydrogen peroxide, peroxyl radical, peroxynitrite and singlet oxygen are formed in human body due to cell metabolic activity. The increased concentration of ROS can induce damage to cellular DNA, proteins, lipids and carbohydrates which leads to development of non-communicable diseases including cancers, cardiovascular diseases, Alzheimer's disease, diabetes and others [4]. Diabetes mellitus is a group of metabolic disorder characterized by elevated level of blood glucose which occurs by malfunction of insulin or decreasing of insulin reaction causing failure of some organs such as heart, nerves, kidneys and eyes. There are two types of diabetes including type 1 diabetes occurring from damage of pancreatic beta-cell by autoimmune destruction, leading to insulin deficiency. Patients need insulin injection. Type 2 diabetes is more common form of diabetes as compared to type 1 diabetes. Resistance to the action of insulin together with deficiency in insulin secretion leads to pathogenesis of the type 2 diabetes [5]. Several synthetic medicines include sulfonylureas used for activation of insulin secretion from β cell of pancreas, biguanide used for decreasing of insulin resistance and acarbose used for inhibition of α-amylase. α-Amylase and α-glucosidase are the principal enzyme involved in carbohydrate digestion [6] Alternative way to decrease blood glucose is to prevent carbohydrate absorption after dining. Polysaccharides are a complex nutrient which will be hydrolysed by α-amylase to oligosaccharides, then further hydrolysed by α-glucosidase to glucose before absorption in intestinal epithelium. Therefore, inhibition of α-amylase and α-glucosidase can affect decreasing of blood glucose level [7]. Thus, it is interesting to investigate some

legumes with antioxidant and antidiabetic activities to formulate novel health products.

II. METHODOLOGY

Extraction of germinated and non-germinated legumes

Five legumes including mung bean (*V. radiata*), black gram (*P. mungo*), kidney bean (*P. vulgaris*), soybean (*G. max*) and peanut (*A. hypogaea*) used in this study were purchased from supermarket in Bangkok, Thailand. These legumes were germinated using the method as described by Pradeep and Sreerama [8] with some modification. Briefly, all legumes were washed with tap water and soaked in water in the ratio of 1:3 w/v for 16 h. Then, it was placed on a wet cheese cloth allowing to germinate at 30 ± 2°C for 48 h, and dried in hot air oven at 60°C for two days before grinding into powder. To prepare ethanolic extract of germinated and non-germinated legumes, 10 g dried powder of each legume was soaked with 100 ml of 80% ethanol, and then shaked at 200 rpm for 48 h. The extract was evaporated and dried in vacuum oven at 40°C.

Phytochemical analysis of legume extracts

Antidiabetic activity and antioxidant activity assay

α-Amylase inhibition assay and α-glucosidase inhibition assay of germinated and non-germinated legume extracts were performed according to the method as described by Sancheti et al. [9] and Kim et al. [10], respectively. Percentage of these assays were calculated. Antioxidant activity of all extracted samples was performed by ferric reducing antioxidant power (FRAP) assay using the method as described by Lado et al. [11]. Reducing capacity of all samples was calculated as mmol Fe(II)/g extract.

Determination of total phenolic content

Total phenolic content (TPC) of each legume extract was determined according to the method as described by Zhou et al. [12]. The sample (0.1 ml of 1 mg/ml crude extract in 30% ethanol) was transferred to a test tube, then 3 ml ultra-pure water was added. Folin-Ciocalteu's phenol reagent (UN3264, VWR chemical, European Commission) (250 µl) was added, shaken thoroughly, and allowed to stand for 2 min at 25°C in the dark. Then, 1 ml of 15% Na_2CO_3 was added, mixed thoroughly and left for 1 min before adjusting the volume to 10 ml with ultrapure water and mixed thoroughly. After incubation for 90 min at 25°C in the dark, the absorbance was measured at 760 nm using UV-visible spectrophotometer (UV1601, Shimadzu Scientific Instruments (Oceania) Pty. Ltd., Australia). Standard curve of gallic acid (Fluka, Sigma-Aldrich, Spain) at 10-1,000 µg/ml concentration was prepared and used for calculation of total phenolics in the sample. The results were expressed as mg gallic acid equivalents (GAE)/g extract.

Determination of indigestible polysaccharide content

Indigestible polysaccharide content in each legume extract was analysed using the method as described by Wichienchot et al. [13]. Total sugar content was analysed by phenol-sulfuric method using the protocol as described by Dubois et al. [14]. Indigestible polysaccharide content (mg/g extract) was calculated as by subtraction of total sugar content after acid and enzyme digestion out of reducing sugar content before digestion.

Development of probiotic legume beverages

Preparation of starter culture

Lactobacillus acidophilus TISTR 1034 was grown in 10 ml MRS broth (deMan Rogosa Sharpe medium, Difco Laboratories, USA) and incubated at 37°C for 24 h in microaerophilic condition. The cells were harvested by centrifugation at 3,000 rpm for 15 min, washed twice and resuspended in 10 ml 0.85% NaCl. The cell turbidity was adjusted using McFarland No. 8 to get the final cell concentration of 10^8 CFU/ml.

Preparation of probiotic legume beverages

Two suitable types of legumes, with strong antioxidant activity and high total phenolic content were selected to develop probiotic germinated and non-germinated legume beverages. To prepare non-germinated legumes, 250 g of each legume was washed and soaked with water in the ratio of 1:3 w/v for 18 h. For preparation of germinated legume, each legume was prepared as above mentioned. After soaking, all legume seeds were left to germinate on wet cheese cloth at 30°C for 48 h. Then, these germinated and non-germinated legumes were grinded and added with 500 ml water. After filtering, the liquid volume was adjusted to 2,000 ml with water and heated to 80-85°C for 20 min before adding with 3.75 g table salt and filling in a sterile bottle. After cooling down, *L. acidophilus* was added in the ratio of 1:30 v/v. Each probiotic legume beverage was stored at 4°C for 14 days. At 0, 7 and 14-day storage, the pH value of each beverage was measured using Benchtop pH meter (Starter 3000, Ohaus, USA), and total lactic acid bacterial count was analysed by spiral plate technique, plating out onto MRS agar using spiral plater, model Autoplate 4000 (Spiral Biotech, USA). Moreover, antioxidant activity by FRAP method, and α-amylase and α-glucosidase inhibitory activity of 14-day stored beverage were also analysed using the same method as mentioned above. The results were recorded as mmol Fe(II)/100 ml beverage for antioxidant activity by FRAP method and mg gallic acid equivalents (GAE)/ 100 ml beverage for total phenolics. Statistical analysis of pH value and *L. acidophilus* survival data were performed by analysis of variance (ANOVA) and Duncan's multiple range test to compared between treatment means using 20.0 version of SPSS statistical package.

III. RESULTS AND DISCUSSION

Phytochemical properties of legume extracts

For α-amylase inhibitory activity of legume extracts, the results showed that all germinated and non-germinated legume extracts had α-amylase inhibitory activity in the range of 13.4-23.3%. The highest α-amylase inhibitory activity was found in germinated soybean extract (23.3%). Non-germinated mung bean, black gram and peanut extracts

had higher α-amylase inhibitory activity as compared to its germinated legume extract. However, germinated soybean and kidney bean extracts had higher α-amylase inhibitory activity, comparing with its non-germinated one. Considering to α-glucosidase inhibitory activity, this activity of both germinated and non-germinated forms of extract was not strong (4.6-10.5%). Germinated mung bean, soybean and peanut extracts had higher α-glucosidase inhibitory activity, compared to its non-germinated one. However, non-germinated black gram and kidney bean extracts possessed higher α-glucosidase inhibitory activity than its germinated form. These two enzymatic inhibitory activity was found relatively strong in acarbose, the positive control (54.8-54.9%) (Table 1).

Non-germinated black gram and kidney bean extracts had strong antioxidant power of 0.196 and 0.396 mmol Fe (II)/g extract by FRAP method, respectively which were higher than the antioxidant power of BHT (0.12 mmol Fe (II)/g extract). However, the germinated extract of soybean possessed higher antioxidant power as compared to its non-germinated extract (Table 1).

All legume extracts had high total phenolics. The non-germinated black gram, kidney bean and peanut extracts had higher total phenolics as compared to their germinated one. The extract with highest total phenolics was non-germinated kidney bean (477.17 mg GAE/g extract), followed by non-germinated black gram (285.41 mg GAE/g extract) (Table 1).

TABLE 1
ANTIDIABETIC AND ANTIOXIDANT ACTIVITIES, AND TOTAL PHENOLIC CONTENT OF LEGUME EXTRACTS

Legume extracts		Diabetes inhibition		FRAP assay (mmol Fe (II)/ 100 ml beverage) ± SD[z]	TPC (mg GAE/100 ml) ± SD
		α-amy. Inh. (%)	α-glu. Inh. (%)		
M. bean	Ger	14.1±0.2	9.2±0.01	0.006±0.01	204.17±3.61
	Non	16.1±0.3	6.6±0.05	0.012±0.05	189.58±3.61
B. gram	Ger	15.2±0.7	9.2±0.04	0.042±0.04	195.83±6.25
	Non	16.7±1.4	10.5±0.7	0.196±0.07	285.42±3.61
K. bean	Ger	17.8±0.2	8.7±0.03	0.044±0.00	197.92±3.61
	Non	13.4±0.1	9.3±0.02	0.396±0.02	477.17±9.68
Soy-bean	Ger	23.3±0.4	10.2±0.5	0.089±0.05	247.92±3.6
	Non	19.6±0.3	8.5±0.02	0.037±0.00	189.58±3.61
Pea-nut	Ger	15.3±0.0	7.2±0.02	0.006±0.02	141.67±3.61
	Non	17.2±0.3	4.6±0.03	0.005±0.00	193.75±0.00
BHT		-	-	0.12±0.03	-
Acarbose		54.8±0.2	54.9±0.3	-	-

[z]Data are mean of three replications.
M. bean, Mung bean; B. gram, Black gram; K. bean, Kidney bean; Ger, germinated; Non, non-germinated; α-amy. Inh., α-amylase inhibitory activity; α-glu. Inh., α-glucosidase inhibitory activity

In the current study, germinated soybean extract possessed quite strong α-amylase inhibitory activity. The results are in agreement with those reported by Tan et al. [15] reported that fraction V of soybean extract possessed strong α-amylase inhibitory activity with IC_{50} of 0.25 mg/ml. These was probably due to high phenolic content in soybean. Ademiluyi and Obeh [16] reported that soybean extract had high total phenolics. They revealed that bound phenolic extract of soybean (IC_{50} 320 µg/ml) had stronger α-amylase inhibitory activity as compared to free phenolic extract (IC_{50} 526.32 µg/ml).

Non-germinated peanut, germinated mung bean and germinated black gram extracts were found to contain high indigestible polysaccharide (511.66-614.71 mg/g extract), but the legumes with relatively high indigestible polysaccharide were non-germinated kidney bean, germinated peanut and germinated soybean (443.48-459.97 mg/g extract. Mostly, polysaccharides in germinated legumes had higher ability to tolerate digestion by acid and enzyme, compared to non-germinated legume, except for kidney bean and peanut. The highest amount of reducing sugar and total sugar was found in germinated black gram extract (Table 2).

Indigestible oligosaccharides containing 3-10 degree of polymerization is normally found in plant food. When these oligosaccharides were consumed, some indigestible polysaccharides become nutrients of probiotic bacteria such as *Bifidobacterium* and several species of *Lactobacillus*. Consumption of galactooligosaccharides, Fructo-oligosaccharides and inulin can help to increase the number of probiotic bacteria in colon and decrease the number of pathogenic bacteria [17].

TABLE 2
INDIGESTIBLE POLYSACCHARIDE CONTENT OF LEGUME EXTRACTS

Type of legumes		Reducing sugar (mg/g extract) ± SD[z]	Total sugar (mg/g extract) ± SD[z]	Indigestible polysaccharide (mg/g extract) ± SD[z]
M. bean	Ger	76.42 ± 4.53	656.62 ± 5.66	580.19 ± 5.71
	Non	61.96 ± 2.14	115.25 ± 0.47	53.29 ± 2.15
B. Gram	Ger	150.23 ± 150.23	661.90 ± 5.69	511.66 ± 4.50
	Non	62.38 ± 2.24	247.22 ± 1.41	184.84 ± 1.08
K. Bean	Ger	57.38 ±1.57	442.82 ± 0.92	385.44 ± 2.47
	Non	78.74 ± 1.53	538.71 ± 1.06	459.97 ± 1.63
Soy-Bean	Ger	46.84 ± 2.62	490.32 ± 1.55	443.48 ± 1.61
	Non	42.73 ± 3.80	238.27 ± 1.73	195.54 ± 4.70
Pea-Nut	Ger	42.73 ± 3.80	486.81 ± 1.49	444.07 ± 2.34
	Non	42.20 ± 4.96	656.90 ± 3.63	614.71 ± 2.63

[z]Data are mean of three replications.
M. bean, Mung bean; B. gram, Black gram; K. bean, Kidney bean; Ger, germinated; Non, non-germinated

Survival of probiotic bacteria in legume beverages during cold storage

Survival study of *Lactobacillus acidophilus* TISTR 1034 in four-type legume beverages during storage at 4°C. After 14-day storage, all beverages had less than 50% survival of

probiotic bacteria (Table 3). Survival of probiotic bacteria in germinated kidney bean beverage was the lowest (13.87%).

Before storage, pH value of all beverages was in the range of 6.54-6.60. After 7-day storage, the pH value of germinated and non-germinated kidney bean and soybean beverages slightly decreased to 5.99-6.32, especially in non-germinated kidney bean which slightly pH change. After 14-day storage, both non-germinated kidney bean and soybean had slightly pH change to 5.94-6.40 (Table 4).

Although the survival of *L. acidophilus* decreased after storage, it did not too low. To gain health benefit, probiotic products should have viable count of probiotic bacteria of at least 1 0 ^{6} -1 0 ^{7}CFU/ml [18]. The reason that viable *L. acidophilus* TISTR 1034 decreased as storage time increase may associate with oxygen. Oxygen affects toxicity to the cells of probiotic bacteria. Probiotic bacteria faced low oxygen concentration, resulting in loss of viability [19].

TABLE 3
SURVIVAL OF *Lactobacillus acidophilus* IN LEGUME BEVERAGES DURING CHILLED STORAGE

Legume beverages	Survival (%)		
	Day 0	Day 7	Day 14
Germinated kidney bean	100±0.00^{a}	32.21±26.32^{ab}	13.87±9.45^{b}
Non-germinated kidney bean	100±0.00^{a}	23.28±3..82^{b}	21.20±4.89^{b}
Germinated soybean	100±0.00^{a}	25.66±7.06^{b}	23.28±5.45^{b}
Non-germinated soybean	100±0.00^{a}	21.84±4.65^{b}	15.82±7.46^{b}

$^{a, b}$ Different letter in different row of the same column indicated significant difference (P<0.05).

TABLE 4
CHANGE OF PH IN LEGUME BEVERAGES DURING CHILLED STORAGE

Legume beverages	pH value		
	Day 0	Day 7	Day 14
Germinated kidney bean	6.57±0.05^{ab}	5.99±0.67^{a}	6.02±0.38^{ab}
Non-germinated kidney bean	6.60±0.17^{a}	6.32±0.30^{a}	6.40±0.12^{a}
Germinated soybean	6.54±0.30^{ab}	6.28±0.10^{a}	5.87±0.14^{b}
Non-germinated soybean	6.55±0.39^{ab}	6.14±0.28^{a}	5.94±0.17^{ab}

*Different letter in different row of the same column indicated significant difference (P<0.05).

Phytochemical properties of legume beverages

Study of antioxidant activity in legume beverages after 14-day storage at 4°C revealed that non-germinated kidney bean beverage had stronger antioxidant power (0.32 mmol Fe (II)/ 100 ml beverage) as compared to BHT. This beverage had the highest total phenolics (1,838.2 mg GAE/100 ml). Germinated kidney bean had highest α-glucosidase inhibitory activity, while non-germinated kidney bean possessed strongest α-amylase inhibitory activity (Table 5).

IV. Conclusion

The potential legume extract with strongest antioxidant activity was kidney bean, while the extract with high antidiabetic activity was the extract of germinated soybean. However, non-germinated peanut contained high amount of indigestible polysaccharide. Viable probiotic counts in germinated soybean beverage was higher than other beverages. Kidney bean beverage had the strongest reducing capacity and the highest phenolic content, while non-germinated kidney bean had the highest α-amylase inhibitory activity. These findings are important for further development of the functional probiotic legume-based beverage formulations and could be exploited to design these beverages with consumer acceptability.

TABLE 5
PHYTOCHEMICAL PROPERTIES OF GERMINATED AND NON-GERMINATED LEGUME BEVERAGES AFTER 14-DAY CHILLED STORAGE

Legume beverages		Diabetes inhibition		FRAP assay (mmol Fe (II)/ 100 ml beverage) ± SD ^{z}	TPC (mg GAE/100 ml) ± SD
		α-Amylase Inhibitory Activity (%)	α-Glucosidase inhibitory activity (%)		
K. Bean	Ger	5.76±0.17	17.56±0.21	0.30±0.01	1,690.0±5.56
	Non	11.34±0.26	15.74±0.24	0.32±0.02	1,838.2±3.21
Soy-Bean	Ger	3.18±0.19	12.10±0.15	0.31±0.03	1,556.7±5.56
	Non	3.21±0.17	12.13±0.22	0.26±0.02	1,456.7±5.56
BHT		-	-	0.08±0.03	-
Acarbose		72.00±0.20	57.60±0.10	-	-

^{z}Data are mean of three replications.
K. bean, Kidney bean; Ger, germinated; Non, non-germinated

References

[1] A. Dutta, A. Trivedi, C. P. Nath, D. S. Gupta, K. K. Hazra, "A comprehensive review of grain legumes as climate-smart crops: Challenges and prospects". Environmental Challenges, 7: 100479, 2022.
[2] J. de Paiva Gonçalve, K. Gasparini, E. A. de Toledo Picoli, M. D.-B. L. Costa, W. L. Araujo, A. Zsögön, D. M. Ribeiro, "Metabolic control of seed germination in legumes". Journal of Plant Physiology, 295: 154206, 2024.
[3] G. Mazza, E. Miniati, "Anthocyanins in fruits, Vegetables, and Grains". CRC Press: Boca Raton, FL, pp. 2-10., 1993.
[4] A. K. Keshari, A. K. Verma, T. Kumar, R. Srivastava, "Oxidative stress: A review". The International Journal of Science & Technoledge, 3(7): 155-162, 2015.
[5] A. Ullah, A. Khan, I. Khan, "Diabetes mellitus and oxidative stress-A concise review". Saudi Pharmaceutical Journal, 24: 547-553, 2016.
[6] M. A. Huneif, D. B. Alshehri, K. S. Alshaibari, M. Z. Dammaj, M. H. Mahnashi, S. U. Majid, M. A. Javed, S. Ahmad, U. Rashid, A. Sadiq, "Design, synthesis and

bioevaluation of new vanillin hybrid as multitarget inhibitor of α-glucosidase, α-amylase, PTP-1B and DPP4 for the treatment of type-II diabetes". Biomedicine & Pharmacotherapy, 150: 113038, 2022.
[7] A. O. Ademiluyi, G. Oboh, "Soybean phenolic-rich extracts inhibit key enzymes linked to type 2 diabetes (α-amylase and α-glucosidase) and hypertension (angiotensin I converting enzyme) in vitro." Experimental and Toxicologic Pathology, 65: 305-309, 2013.
[8] P.M. Pradeep, Y.N. Sreerama, "Impact of processing on the phenolic profiles of small millets: Evaluation of their antioxidant and enzyme inhibitory properties associated with hyperglycemia." Food Chemistry, 169: 455-463, 2015
[9]. S. Sancheti, S. Sancheti, S.-Y. Seo. "Antidiabetic and antiacetylcholinesterase effects of ethyl acetate fraction of *Chaenomeles sinensis* (Thouin) Koehne fruits in streptozotocin-induced diabetic rats." Experimental and Toxicologic Pathology, 65: 55-60, 2013.
[10] G.N. Kim, J.G. Shin, H.D. Jang, "Antioxidant and antidiabetic activity of Dangyuja (*Citrus grandis* Osbeck) extract treated with *Aspergillus saitoi*." Food Chemistry, 117: 35-41, 2009.
[11] B. Lado, M. Then, I. Varga, E. Szóke, K. Szentmihalyi. "Antioxidant property of volatile oils determined by the ferric reducing ability." A Journal of Biosciences, 59: 354-358, 2004.
[12] Z. Zhou, X. Chen, M. Zhang, C. Blanchard, "Phenolics, flavonoids, proanthocyanidin and antioxidant activity of brown rice with different pericarp colors following storage." Journal of Stored Products Research, 59: 120-125, 2014.
[13] S. Wichienchot, P. Thammarutwasik, A. Jongjareonrak, W. Chansuwan, P. Hmadhlu, T. Hongpattarakere, A. Itharat, B. Ooraikul, "Extraction and analysis of prebiotics from selected plants from southern Thailand." Songklanakarin Journal Science Technology, 33(5): 517-523, 2011.
[14] M. Dubois, K. A. Gilles, J. K. Hamilton, P. A. Rebers, F. Smith, "Colorimetric method for determination of sugars and related substances." Analytical Chemistry, 28(3): 350-356, 1956.
[15] Y. Tan, S. K. C. Chang, Y. Zhang, "Comparison of α-amylase, α-glucosidase and lipase inhibitory activity of the phenolic substances in two black legumes of different genera". Food Chemistry, 214: 259-268, 2017.
[16] A. O. Ademiluyi, G. Oboh, "Soybean phenolic-rich extracts inhibit key enzymes linked to type 2 diabetes (α-amylase and α-glucosidase) and hypertension (angiotensin I converting enzyme) in vitro." Experimental and Toxicologic Pathology, 65: 305-309, 2013.
[17] D. Mudgil, S. Barak, "Composition, properties and health benefits of indigestible carbohydrate polymers as dietary fiber: A review." International Journal of Biological Macromolecules, 61: 1-6, 2013.
[18] P. M. Manoj, J. R. Mohan, B. Y. Khasherao, R. Shams, K. K. Dash. "Fruit based probiotic functional beverages: A review". Journal of Agriculture and Food Research, 14: 100729, 2023.
[19] M.K. Tripathi, S. K. Giri, "Probiotic functional foods: Survival of probiotic during processing and storage." Journal of Functional Foods, 225-241, 2014.

Application of Recombinant Human scFv Antibody for Determination the Nodule Occupancy of Rhizobia from Dual Culture Inoculant

Chayanan Singtothong[1, a)], Kuntalee Rangnoi[2], Pakpoom Boonchuen[1], Kamonluck Teamtisong[3], Jenjira Wongdee[2], Teerana Greetatorn[2], Pongpan Songwattana[2], Pongdet Piromyou[2], Monchai Manassila[4], Nantakorn Boonkerd[1], Neung Teaumroong[1], Montarop Yamabhai[1, *] and Panlada Tittabutr[1, *, a)]

[1]School of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand
[2]Institute of Research and Development, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand
[3]The Center for Scientific and Technological Equipment, Suranaree University of Technology, Nakhon Ratchasima 30000, Thailand
[4]Soil Microbiology Research group, Soil Science Research Group, Department of Agriculture, Bangkok 10900, Thailand.

[a)] Corresponding author: coseter123@gmail.com
[b)] panlada@sut.ac.th

Abstract. *Rhizobium*, a key nitrogen-fixing bacterium, significantly influences legume production as a biofertilizer. Determining nodule occupancy of applied rhizobial inoculant is crucial for enhancing nodulation and nitrogen fixation efficiency in field applications. This study introduces the use of recombinant human single chain fragment variable (scFv) antibodies as a technique for the precise detection of rhizobial cells in inoculants and continuous monitoring of nodule occupancy in dual-culture systems. The study utilized a dual inoculant between *Bradyrhizobium* sp. SUTN9-2 and *Bradyrhizobium* sp. DASA03028 as a model. While the specific scFv antibody E10 targeting SUTN9-2 was previously characterized, a new scFv antibody, F11, was selected for DASA03028 using biopanning from a naïve human scFv antibody phage display library. These recombinant scFv antibodies were successfully employed with the enzyme-linked immunosorbent assay (ELISA) technique to detect each strain of *Bradyrhizobium* in liquid inoculant at various timepoints (5, 7, 15, and 30 days). Once each strain is applied to the mung bean seed under a dual-culture system, these recombinant scFv antibodies could be an efficient tool to determine the nodule occupancy. The result showed that both strains had the similar ability to compete with each other for nodulation. This study demonstrates the feasibility and potential benefits of using recombinant human scFv antibodies for *Bradyrhizobium* detection and nodule occupancy determination. This technique would help optimize the mixed-culture inoculant formulations to increase biofertilizer efficiency in sustainable agriculture further.

Keywords: ***Recombinant scFv, phage display, Bradyrhizobium, legume inoculant, Nodule occupancy***

I. INTRODUCTION

Rhizobium, a nitrogen-fixing bacterium, plays a pivotal role in legume plant production as a biofertilizer. To achieve high efficiency crop production and reduce cost of chemical fertilizer, the application of nitrogen biofertilizer *Bradyrhizobium* is necessary. *Bradyrhizobium* sp. is one of soil bacteria that can convert atmospheric nitrogen into ammonia and provide nitrogen to plants for growth requirements (Bogino et al., 2006). To enhance nodulation and nitrogen fixation success in the field application, determining of nodule occupancy provides valuable information for optimizing inoculant selection to improve crop productivity and mitigate risks of unsuccessful rhizobial inoculation. Therefore, developing a robust technique to determine nodule occupancies is important. These may ensure the successful nodulation and symbiosis with host legume in the field application. The serological technique, especially the Enzyme-linked immunosorbent assay (ELISA) technique, has been used as the basic tools for evaluation and monitoring the rhizobia. ELISA is one of the popular enzyme immunoassays used in detecting antigens. Advantages of this technique are the specificity, rapidity, and no microscopic equipment needed. (Vu et al., 2017). However, the traditional way of producing antibodies must use the application of immune system in rabbits to produce polyclonal antibodies (Wilson et al., 1988) for each *rhizobium*. The main problem with using polyclonal antibodies to detect and monitor specific *rhizobium* is that they often cross-react with other rhizobial strains within the same species (Vu et al., 2017). Therefore, the application of phage display technology was used for the generation of monoclonal antibody for specific detection of *Bradyrhizobium*. In this research, the *Bradyrhizobium* sp.

strain DASA03028 provided from DOA is selected to develop the specific antibody using phage display technology and further developed as a mixed culture with *Bradyrhizobium* sp. strain SUTN9-2 as inoculant production. *Bradyrhizobium* sp. SUTN9-2 has been used as selected strain for inoculants, which is produced under the Laboratory of Inoculant Production research group, Suranaree University of Technology (SUT) (Noisangiam et al., 2012). This strain produces plant hormone, indole acetic acid (IAA) (Greetatorn et al., 2019), and contains 1-aminocyclopropane-1-carboxylic acid (ACC) deaminase activity that can decrease plant stress under water deficit conditions (Sarapat et al., 2020). The specific antibody for the strain SUTN9-2 using phage display technology has been produced and characterized (Khaing et al., 2021). The strain SUTN9-2, in combination with the *Bradyrhizobium* sp. DASA03028, a dual culture inoculant for mung bean, was used as a model in this study. This research aimed to develop a robust technique for detecting cells in liquid inoculant and further monitoring nodule occupancy of each strain after application to the legume. Therefore, specific and precise technique is required in order to improve crop productivity and ensure nodulation efficiency with legumes.

II. Methodology

Antigen preparation

Bradyrhizobium sp. DASA03028 obtained from Department of Agriculture (DOA) was used for antigen preparation. To prepare the antigen, *Bradyrhizobium* was cultured in Yeast extract Mannitol Broth medium (YMB) in the incubator shaker at 28°c for 7 days. Then, the cells were washed 3 times with sterilized 0.85% (w/v) saline buffer and the cell concentration was adjusted to 10^9 cells/ml and stored at -20°c until use. Then, the total protein of the antigen was measured by Bradford technique (Bradford, 1976).

Affinity selection of phage displayed scFv antibodies

The biopanning method selected phage clones that showed specific binding to *Bradyrhizobium* sp. DASA03028. Each round of biopanning antigen targets was immobilized on an immunotube containing 20 µg of total protein in 400 µl of 100 mM $NaHCO_3$ at pH 8.5, then stored at 4˚C overnight. Then the biopanning method was performed as previously described by Khaing et al., (2021) using Phage antibody library (Yamo-I Library) of 10^{12} PFU (Pansri et al., 2009). To continue with the next round of biopanning, 1 ml of 2xYT media was added onto infected *E. coli* TG1 agar plates, then the cells were scraped off from the surface of the agar plate. The scraped cells were kept in 15% (v/v) glycerol stock at -70˚C, and 10 µl of scraped bacteria was added into 10 ml of 2xYT supplemented with 100 µg/ml ampicillin and 1% (w/v) glucose before incubation at 37°C with shaking for 2 hours. The cells were then infected with helper phage, and the eluted phage was amplified as previously described (Khaing et al., 2021). The selected phage scFv clone after 4 rounds of bio-panning was verified by ELISA.

Production of soluble scFv antibody

For cloning of scFv antibody, the plasmid of selected phagemid in pMod1 vector (clone F11 specific to *Bradyrhizobium* sp. DASA03028 (PP346337)) was extracted using the FavorPrep™ Plasmid Extraction Kit. The gene of selected scFv antibody was sequenced and inserted into multiple cloning sites of pET21d (+) vector. The scFv genes were cloned into pET-21d (+) expression vectors and designated as previously described (Khaing et al., 2021). Finally, the recombinant plasmid F11/pET21d+ was obtained. To produce soluble scFv antibody from F11, the F11/pET21d+ plasmid was transformed into SHuffle® T7 competent *E. coli* (NEB, USA). Then, the scFv genes were expressed as in previous studies (Khaing et al., 2021; Lobstein et al., 2012). To collect and purify the recombinant antibody, the induced *E. coli* culture was lysed and disrupted by intermittent sonication. The soluble protein was purified using 1 ml of His-Trap column (GE Health care, USA) as described in previous study (Khaing et al., 2021). In cases of E10, a recombinant antibody specific for strain SUTN9-2 was produced, as described by Khaing et al. (2021).

Detection of *Bradyrhizobuim* in the liquid inoculant

To investigate the binding property of scFv antibody in relation to the liquid inoculant at various timepoints of culture, the ELISA plate was immobilized with cell liquid culture at the age of 5, 7, 15, and 30 days after growing in Yeast Extract Broth (YMB) medium, and then the cell was dissolved with 100 mM $NaHCO_3$ and adjusted to 10^9 CFU/ml. Afterward, the ELISA was performed as described previously by Khaing et al. (2021) with the concentration of scFv antibody at 20 µg /ml.

Determination of nodule occupancy by scFv antibody ELISA

This experiment used mung bean (Vigna radiata) Chainat 72 (CN72) for cultivation. The legume seeds were surface sterilized and allowed to germinate. The germinated seeds were planted in Leonard's jars and inoculated with bradyrhizobia at a concentration of 10^9 cells/seed for each treatment. For dual culture, the amounts of *Bradyrhizobium* sp. DASA03028 and SUTN9-2 cells were adjusted to a 1:1 ratio. All plants were grown in N-free medium (Somasegaran & Hoben, 2012) and grown in lightroom for 1 month, which provided a flux density of light of about 450/µEs^{-1} m^{-2} with a 12–12 h light–dark cycle. After one month of growth, nodules from each experiment were collected and ground in 100 mM sodium carbonate buffer pH 8.5. The bacteroid suspension was centrifuged at 6,000 rpm for 1 minute to separate plant tissue. Subsequently, 150 µl of bacteroid suspension was immobilized in each well of microplate. ELISA was performed using specific scFv for each *Bradyrhizobium* strain, following previously described methods.

III. Results and Discussion

Selection phage displayed scFv antibodies

The biopanning procedure was performed to select the affinity of *Bradyrhizobium* sp. DASA03028 using naïve

human library (Yamo library) (Pansri et al., 2009). **Table I** provides an enrichment of F11 scFv antibody during the biopanning process. After the 4th round of biopanning, the colony obtained 3.74×10^6 clones against DASA03028; 16 clones were selected, and their affinity was verified through phage ELISA. Among these clones, one positive clone (F11). Naïve human library, the YAMO-I library was used in this study. This library was constructed using samples from 140 non-immunized (naïve) donors (Pansri et al., 2009) for screening of specific antibody *Bradyrhizobium* sp. DASA03028 from DOA, resulting positive in the selection of a specific scFv antibody clone (F11) after 4 rounds of bio-panning. This result further confirms the potential to isolate antibodies specific to bacteria in the genus *Bradyrhizobium* by screening from a naïve human library. This approach has previously successfully identified specific antibodies for *Bradyrhizobium* sp. strains DOA9 and SUTN9-2. (Vu, 2017; Khaing, 2020). The success in using antibodies to identify *Bradyrhizobium* sp. may also indicate the presence of specific epitope proteins on the cell surface of each bradyrhizobial strain (Bolaños et al., 2004). It is possible that the surface proteins of certain bradyrhizobia may resemble those of human pathogens, leading to the successful isolation of scFv antibodies from the naïve human library. Then, large quantity of soluble scFv antibody was produced in *E. coli* SHuffle T7 (C3029) as a host organism for promoting disulfide bond formation in the cytosol (Ren et al., 2016) and engineered to produce correctly disulfide bonded active proteins to high yields within its cytoplasm (Lobstein et al., 2012).

TABLE I Selective enrichment of F11 scFv antibodies against to *Bradyrhizobium* sp DASA03028 during the biopanning process

Rounds	Colonies obtained (CFU/ml)	Colonies pickup for ELISA (Colony)	Positive clones at phage ELISA
1st	$1.46\text{x}10^3$	-	-
2nd	$2.47\text{x}10^4$	-	-
3rd	$1.14\text{x}10^6$	-	-
4th	$3.74\text{x}10^6$	16	1(F11)

Detection of bradyrhizobia in the liquid inoculant

To verify the binding of antibodies against strain SUTN9-2 and DASA03028, the cells, after grown in liquid culture media at various time points of 5, 7, 15, and 30 days after inoculation (dai) determined using ELISA. For DASA03028 with the recombinant F11 scFv antibody the highest signal was observed at 30 dai **(Fig. 1A)**. In contrast, the recombinant E10 scFv antibody, specific to strain SUTN9-2, the cells from liquid culture displayed the highest signal at 15 dai **(Fig. 1B)**. Notably, the signal consistently remained over two-fold higher than that of the non-target antigen (1% BSA), even when the cell age was younger or older than 15 dai. These results confirm that the possibility of using recombinant Human scFv antibody for *Bradyrhizobium* detection is the same as previously successful for the detection of *Bradyrhizobium* in mung bean (*Vigna radiata*), *Aeschynomene americana* (Khaing et al. 2021), and siratro (*Macroptilium atropurpureum*) (Vu et al. 2017). Here, we observe that the bacterial cell age may influence on the binding of antibodies to antigen targets. This result could be due to differences in the nature of the target antigen on the bacterial cell surface. Generally, the surface of bradyrhizobia is known to comprise various components, including extracellular polysaccharides (EPS), capsular polysaccharides (CPS), flagellar proteins, and lipopolysaccharides (LPS) (Carlson, 1984), while some components are heat-labile (Somasegaran & Hoben, 2012). It is possible that strain-specific cell surface molecules may alter their structure or chemical composition as the cells growth. Therefore, it is interesting for further experiments to investigate the epitope on the bradyrhizobial cell membrane that interacts with this antibody, aiming to understand the factors influencing the binding of this scFv.

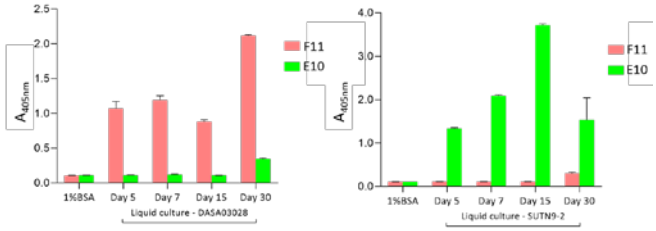


Fig. 1. The binding of soluble scFv F11 antibody (pink bar) and E10 (green bar) which specific for *Bradyrhizobium* sp. DASA03028 and SUTN9-2, respectively in the liquid culture of (A) DASA03028 and (B) SUTN9-2 at 5, 7, 15 and 30 days after inoculation in YMB medium. BSA was used as non-target antigen. Values are the mean of triplicate wells. Error bars show the standard deviation for each set of data.

Determination of nodule occupancy by scFv antibody using ELISA technique

In this study, the experiment was performed on a mung bean to evaluate the application of scFv antibodies for nodule occupancy determination using a dual culture inoculant system. The results are shown in **Table II.** The observed nodule occupancies suggest competition between these strains for symbiosis with the mung bean host. This competition may be influenced by various factors, including strain-specific interactions with the host plant, differences in nodulation efficiency, or variations in nitrogen fixation capabilities (Clúa et al., 2018). Understanding these competitive dynamics can inform the development of mixed or co-inoculant formulations that take advantage of the strengths of both strains to improve overall symbiotic performance. This result indicates the possible application of scFv antibodies for further detection of nodulation competitiveness in the field practice to ensure the nodulation efficiency over the indigenous rhizobial strains and reduce the risk of nodulation failure or inefficient symbiosis.

TABLE II The analysis of mung bean nodule occupancy of DASA03028 and SUTN9-2

Plant inoculates with	Nodule occupancy %		
	DASA03028	SUTN9-2	Co-occupied
DASA03028	100±0	0±0	0±0
SUTN9-2	0±0	100±0	0±0
Mixed culture (1:1)	28.2±8.7	22.3±18.3	49.5±19.8

IV. Conclusion

Our results indicate the potential of scFv antibodies for the specific detection of bradyrhizobia and determining nodule occupancy in mung beans using a dual-culture inoculant system. The observed competitive dynamics between the *Bradyrhizobium* strains suggest that strain-specific interactions with the host plant and variations in nodulation and nitrogen fixation capabilities may influence symbiotic performance. This understanding can guide the development of mixed or dual inoculant formulations to optimize nitrogen availability while minimizing nodulation failures in field applications.

Acknowledgment

This work was supported and funded by the One Research One Graduate (OROG) of The Thailand Research Fund, Agricultural Research Development Agency (Public Organization) (ARDA) (HRD6405070 / M6200121-2/2564), and the School of Biotechnology, Institute of Agricultural Technology, Suranaree University of Technology.

References

[1] Bogino, P., Banchio, E., Rinaudi, L., Cerioni, G., Bonfiglio, C., & Giordano, W. (2006). Peanut (*Arachis hypogaea*) response to inoculation with *Bradyrhizobium* sp. In soils of Argentina. Annals of Applied Biology, 148(3), 207–212. https://doi.org/10.1111/j.1744-7348.2006.00055.xL.

[2] Bolaños, L., Redondo-Nieto, M., Rivilla, R., Brewin, N. J., & Bonilla, I. (2004). Cell surface interactions of *Rhizobium* bacteroids and other bacterial strains with symbiosomal and peribacteroid membrane components from pea nodules. Molecular Plant-Microbe Interactions, 17(2), 216–223. https://doi.org/10.1094/MPMI.2004.17.2.216

[3] Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein – dye binding. Analytical Biochemistry,72(1–2),248–254. https://doi.org/10.1016/0003-2697(76)90527-3

[4] Carlson, R. W. (1984). Heterogeneity of *Rhizobium* lipopolysaccharides. Journal of Bacteriology, 158(3), 1012–1017. https://doi.org/10.1128/jb.158.3.1012-1017.1984

[5] Clúa, J., Roda, C., Zanetti, M. E., & Blanco, F. A. (2018). Compatibility between legumes and rhizobia for the establishment of a successful nitrogen-fixing symbiosis. Genes, 9(3), 125.

[6] Greetatorn, T., Hashimoto, S., Sarapat, S., Tittabutr, P., Boonkerd, N., Uchiumi, T., & Teaumroong, N. (2019). Empowering rice seedling growth by endophytic *Bradyrhizobium* sp. SUTN9-2. Letters in Applied Microbiology,68(3),258–266. https://doi.org/10.1111/lam.13114

[7] Khaing, K., Rangnoi, K., Michlits, H., Boonkerd, N., Teaumroong, N., Tittabutr, P., & Yamabhai, M. (2021). Application of recombinant human scFv antibody as a powerful tool to monitor nitrogen fixing biofertilizer in rice and legume. Microbiology Spectrum, 9(3), e02094-21. https://doi.org/10.1128/Spectrum.02094-21

[8] Lobstein, J., Emrich, C. A., Jeans, C., Faulkner, M., Riggs, P., & Berkmen, M. (2012). SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm. Microbial Cell Factories, 11(1), 753. https://doi.org/10.1186/1475-2859-11-56.

[9] Noisangiam, R., Teamtisong, K., Tittabutr, P., Boonkerd, N., Toshiki, U., Minamisawa, K., & Teaumroong, N. (2012). Genetic diversity, symbiotic evolution, and proposed infection process of *Bradyrhizobium* strains isolated from root modules of *Aeschynomene americana* L. in thailand. Applied and Environmental Microbiology, 78(17), 6236–6250. https://doi.org/10.1128/AEM.00897-12

[10] Pansri, P., Jaruseranee, N., Rangnoi, K., Kristensen, P., & Yamabhai, M. (2009). A compact phage display human scFv library for selection of antibodies to a wide variety of antigens. BMC Biotechnology, 9(1), 6.

[12] Ren, G., Ke, N., & Berkmen, M. (2016). Use of the SHuffle strains in production of proteins. Current Protocols in Protein Science, 85(1), 5.26. 1-5.26. 21. https://doi.org/10.1002/cpps.11

[13] Sarapat, S., Songwattana, P., Longtonglang, A., Umnajkitikorn, K., Girdthai, T., Tittabutr, P., Boonkerd, N., & Teaumroong, N. (2020). Effects of increased 1-aminocyclopropane-1-carboxylate (ACC) deaminase activity in *Bradyrhizobium* sp. SUTN9-2 on mung bean symbiosis under water deficit conditions. Microbes and Environments, 35(3), ME20024.

[14] Somasegaran, P., & Hoben, H. J. (2012). Handbook for rhizobia: Methods in legume-*Rhizobium* technology. Springer Science & Business Media.

[15] Vu, N. X., Pruksametanan, N., Srila, W., Yuttavanichakul, W., Teamtisong, K., Teaumroong, N., Boonkerd, N., Tittabutr, P., & Yamabhai, M. (2017). Generation of a rabbit single-chain fragment variable (scFv) antibody for specific detection of *Bradyrhizobium* sp. DOA9 in both free-living and bacteroid forms. Plos One, 12(6), e0179983. https://doi.org/10.1371/journal.pone.0179983

[16] Wilson, P. O., Barber, P. C., Hamid, Q. A., Power, B. F., Dhillon, A. P., Rode, J., Day, I. N., Thompson, R. J., & Polak, J. M. (1988). The immunolocalization of protein gene product 9.5 using rabbit polyclonal and mouse monoclonal antibodies. British Journal of Experimental Pathology, 69(1), 91–104. PubMed.

Antibacterial Activities of Kombucha Derived from Oolong Tea, Butterfly Pea, Roselle, and Peppermint

Engkarat Kingkaew[1], Jutamas Pracumtong[1] and Duangjai Ochaikul[1,a)]

[1]Department of Biology, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand 10520

[a)] Corresponding author: daungjai.oc@kmitl.ac.th

Abstract. Kombucha is a traditional beverage obtained through the fermentation of tea with sugar, utilizing a symbiosis of acetic acid bacteria and yeasts. The present study aimed to develop new herbal kombucha beverages using ingredients such as roselle, butterfly pea, peppermint, and oolong tea. The antibacterial activity of these kombuchas was investigated against several pathogenic microorganisms. The results showed that the pH steadily decreased from 4.01 to 2.25 during the fermentation process, while the total acidity of Kombucha increased over the 15-day fermentation period. Most of the tested microorganisms (including *Bacillus cereus*, *Staphylococcus aureus*, *Escherichia coli*, *Salmonella* Typhimurium, and *Pseudomonas aeruginosa*) demonstrated sensitivity to the antibacterial activity of these kombuchas. Interestingly, the herbal kombuchas exhibited effective antibacterial activity against *Pseudomonas aeruginosa* even after thermal denaturation at 85°C for 15 minutes.

***Index Terms*—Antibacterial activity, Fermentation, Herbal beverage, Kombucha**

I. INTRODUCTION

Kombucha is a slightly sweet and acidic beverage typically crafted through the fermentation of sweetened black tea with the aid of a symbiotic culture of bacteria and yeast (SCOBY). This consortium predominantly comprises acetic acid bacteria and various yeast species [1]. The symbiotic relationship within the Kombucha colony/mat involves a blend of yeasts and bacteria, whose composition is notably diverse, as elucidated by Jayabalan, Malbaša [1]. The prevalent bacteria in this culture hail from the *Acetobacter* and *Gluconobacter* genera, with *Acetobacter xylinum* being the primary bacteria. In addition to acetic acid bacteria, a wide array of yeast species belonging to *Saccharomyces*, *Saccharomycodes*, *Schizosaccharomyces*, *Zygosaccharomyces*, *Brettanomyces*, *Candida*, *Kloeckera*, *Pichia*, *Mycotorula*, and *Mycoderma* are commonly found in Kombucha. The symbiotic relationship revolves around yeasts converting sucrose into glucose and fructose via invertase, leading to ethanol production through metabolic pathways. Simultaneously, acetic acid bacteria convert glucose into gluconic acid and ethanol into acetic acid [1]. The fermentation process yields two distinct components: a floating cellulosic pellicle layer and sour liquid broth, also known as fermented broth [2].

Kombucha has shown promising prophylactic potential in various areas, including weight loss, metabolic diseases, arthritis, indigestion, cancer, and acquired immunodeficiency syndrome (AIDS) [3-5]. The regular consumption of Kombucha has been shown to significantly inhibit weight gain and extend lifespan [6]. These prophylactic and antioxidant effects are attributed to numerous active components generated during fermentation, such as organic acids (predominantly acetic, gluconic, and glucuronic acid), sugars (sucrose, glucose, and fructose), water-soluble vitamins (B1, B6, B12, and C), amino acids, biogenic amines, purines, pigments, lipids, proteins, hydrolytic enzymes, ethanol, carbon dioxide, polyphenols, minerals, anions such as D-saccharic acid-1,4-lactone, and metabolic byproducts of yeasts and bacteria [7].

Kombucha has demonstrated antimicrobial activity against various strains, including several strains of *Candida*, *Haemophilus influenzae*, and *Escherichia coli* [8], as well as *Staphylococcus epidermidis*, *S. aureus*, *Micrococcus luteus*, *Salmonella* Typhimurium, *E. coli*, *Listeria monocytogenes*, and *Pseudomonas aeruginosa* [9]. Kombucha contains numerous metabolites with antimicrobial properties, notably organic acids such as acetic acid and catechins, in addition to antibiotic substances that are effective against both Gram-positive and Gram-negative bacteria [10].

The rising demand for nondairy nutraceuticals, driven by an increase in lactose intolerance cases and the growing trend toward veganism, underscores the importance of developing new, safe, nondairy probiotic products with prophylactic benefits [11, 12].

Kombucha and fermented herb drinks present viable alternatives to traditional dairy products for individuals with lactose intolerance [13]. This study aimed to demonstrate the *in vitro* antimicrobial activity of the herbal Kombucha.

II. METHODOLOGY

1. Preparation of herbal plants

Butterfly pea flowers (*Clitoria ternatea*), roselle (*Hibiscus sabdariffa* L.), and peppermint (*Melissa officinalis*) were dried at 50°C in a hot air oven for 1–2 days. The dried herbal plants were then collected and stored in a desiccator. Additionally, dried oolong tea (*Camellia sinensis*) was purchased from a local supermarket in Bangkok, Thailand.

2. Preparation of herbal kombucha

Ten grams of each herbal plant was added to one liter of water and boiled for 5 minutes. The solid parts of the plants were then discarded, and the liquid was supplemented with 1% (w/v) sucrose. Subsequently, 300 ml of the mixture was poured into glass jars and inoculated with 10% (v/v) of the previously fermented Kombucha liquid starter and 3% (w/v) of the previously fermented Kombucha cellulose starter. The inoculated mixtures in the jars were incubated at room temperature for 15 days. Samples were collected at 0, 5, 10, and 15 days and analyzed.

3. pH determination

The pH of each sample was measured in triplicate using a pH meter (Mettler Toledo S220, United States).

4. Total acid content assay

The total acid content of each sample was determined using the AOAC method (2000). Each sample was titrated with 0.1 N NaOH to an endpoint, and the volume in millilitres of NaOH was recorded. This volume was then used to calculate the total acid content.

5. Antibacterial activity

Antibacterial activity was assessed using the agar well diffusion method as described by Sreeramulu, Zhu [5]. Pathogenic microorganisms, such as *Bacillus cereus* (*Bac.*), *Staphylococcus aureus* (*Sta.*), *Escherichia coli* (*Esc.*), *Salmonella* Typhimurium (*Sal.*) and *Pseudomonas aeruginosa* (*Pse.*), and the cells were cultured on Mueller Hinton agar (MHA) for 24 hours. Cell suspensions were prepared using 0.85% NaCl and adjusted to a turbidity of 0.5 McFarland (equivalent to 1.5 x 10^8 CFU/ml), and 6 mm diameter wells were created using a sterile metallic tube. Sterile supernatant was obtained by filtering the centrifuged infusions through a sterile 0.22 µm microfilter. Sterile-fermented Kombucha samples (30 µl) were then transferred to the wells created on pre-inoculated plates containing pathogenic microorganisms. The plates were incubated at 37°C for 24 hours. Unfermented and acidified infusions were used as controls. The heat-denatured samples were prepared by heating at 85°C for 15 minutes. The samples were aseptically filtered, and their antibacterial activity was tested.

6. Statistical analysis

All results are presented as the mean ± standard deviation (SD). Statistical comparisons were conducted using one-way analysis of variance (ANOVA) followed by Duncan's new multiple range test (DMRT) at a significance level of $p < 0.05$. All the statistical analyses were performed using the SPSS (version 22.0) program for Windows.

III. RESULTS AND DISCUSSION

pH

The pH of the herbal kombucha decreased during fermentation (Fig. 1). The initial pH values of oolong tea kombucha, butterfly pea kombucha, rosella kombucha, and peppermint kombucha were recorded as 3.63±0.02, 3.95±0.02, 2.80±0.02, and 4.01±0.09, respectively. Following the fermentation process, the final pH values of oolong tea kombucha, butterfly pea kombucha, rosella kombucha, and peppermint kombucha were recorded as 3.18±0.04, 2.90±0.04, 2.25±0.01, and 2.96±0.03, respectively. It exhibited a gradual decline from 4.01 to 2.25. This decrease can be attributed to the increasing levels of organic acids produced by bacteria and yeasts in the starter culture. Notably, Roselle kombucha had the lowest pH (2.25 ± 0.01).

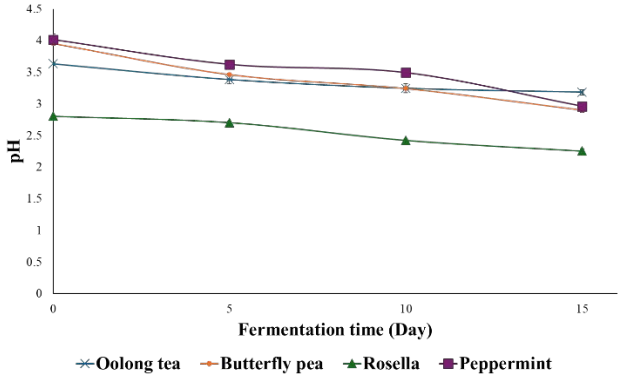


Fig. 1. The pH dynamics of each herbal kombuchas.

Total acid content

The total acid content exhibited a gradual increase over the fermentation period (Fig. 2). Remarkably, roselle kombucha showed the highest acid production, reaching 1.68±0.99 at the 15-day mark. This increase in acidity can be attributed to the metabolic activity of yeast and bacteria within the herbal kombucha. Specifically, these microorganisms metabolize glucose into organic acids, such as acetic acid and gluconic acid [14]. Consequently, as the pH decreased during fermentation, the total acid content increased, reflecting the dynamic biochemical transformations occurring within the Kombucha matrix. The total acid content of the herbal kombucha is shown in Fig. 2.

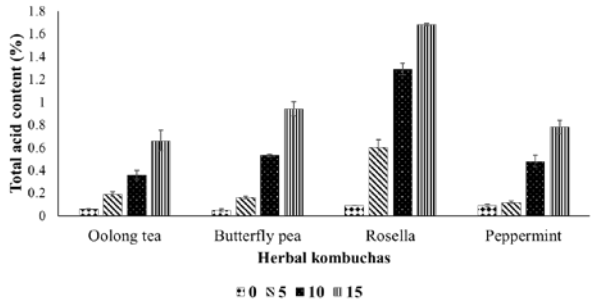


Fig. 2. The total acid content (%) of each herbal kombucha.

Antibacterial activity of fermented herbal kombucha

The antibacterial efficacy of the herbal kombucha beverages against the tested pathogenic microorganisms is detailed in Table 1. The antibacterial activity of the prepared kombucha fermented beverages (KFBs), heat-denatured kombucha beverages (HKBs) and neutralized kombucha beverages (NKBs) was demonstrated. The antibacterial activity can be attributed to the acidity of organic acids, including acetic, citric, and gluconic acids, present in kombucha beverages [15]. These organic acids, generated during fermentation through the conversion of sucrose by yeasts and bacteria within the Kombucha consortium, decrease the pH to final levels typically ranging from 3.18 to 2.25, as observed in this study.

Table 1 Antimicrobial activity of herbal kombucha.

Herbal plants	Treatment(s)	Inhibition Zone Diameter (mm)				
		Gram-positive bacteria		Gram-negative bacteria		
		B.	*Sta.*	*Esc*	*Sal.*	*Pse.*
Oolong tea	Unfermented infusion	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Fermented infusion	16.32±0.22	20.28±1.02	15.29±0.23	11.02±0.65	19.24±1.20
	Acidified infusion	17.98±0.98	22.25±1.15	13.30±0.85	14.21±3.57	21.54±2.80
	Neutralized	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Heat-denatured	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	12.32±0.00
	Antibiotic	12.05±0.60	27.85±0.95	15.34±1.11	17.26±0.45	16.36±0.45
Butterfly Pea flower	Unfermented infusion	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Fermented infusion	10.25±0.15	11.53±0.66	11.02±0.66	10.06±0.50	12.33±0.62
	Acidified infusion	11.30±1.68	13.32±0.52	12.64±0.18	11.32±1.25	13.78±0.92
	Neutralized	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Heat-denatured	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	12.08±0.82
	Antibiotic	12.65±0.47	29.52±0.39	17.30±0.71	17.35±0.16	16.70±0.67
Roselle	Unfermented infusion	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Fermented infusion	13.90±0.54	13.32±0.49	13.85±0.59	12.65±0.32	13.69±0.41
	Acidified infusion	22.85±3.98	21.53±4.12	20.73±2.28	21.69±5.02	17.86±0.44
	Neutralized	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Heat-denatured	11.61±1.34	12.30±0.03	10.80±0.07	10.59±0.13	12.25±1.25
	Antibiotic	12.99±0.52	26.86±1.22	16.41±0.62	17.11±0.74	12.28±0.48
Peppermint	Unfermented infusion	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Fermented infusion	11.40±0.75	15.36±1.13	13.71±0.65	11.25±1.24	13.42±1.09
	Acidified infusion	21.09±1.56	22.89±0.62	19.26±1.67	19.52±2.01	19.36±1.31
	Neutralized	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00
	Heat-denatured	6.00±0.00	6.00±0.00	6.00±0.00	6.00±0.00	13.30±0.08
	Antibiotic	13.14±0.50	28.45±0.64	17.44±0.68	17.64±0.61	17.21±0.48

B. = Bacillus cereus; *Sta. = Staphylococcus aureus*; *Esc. = Escherichia coli*; *Sal. = Salmonella* Typhimurium; and *Pse. = Pseudomonas aeruginosa*

This is strongly corroborated by the previous research indicated that an increased concentration of organic acids produced in fermented Kombucha broth leads to a decrease in pH from 5 to 2.5 [15, 16]. Furthermore, Greenwalt et al. [17] demonstrated that the antimicrobial efficacy of Kombucha against pathogenic microorganisms is primarily attributed to acetic acid. Organic acid molecules can induce cytoplasmic acidification and disrupt bacterial cells [18]. The antibacterial activity exerted by compounds in kombucha may also be explained by the osmotic pressure of solutes present in the hypertonic medium compared to the external aqueous medium, facilitating the diffusion of bioactive materials across cell membranes through selective permeability. The lipophilic nature of certain solutes enables them to adhere to bacterial cell membranes, leading to cell death [19, 20].

Some inhibitory biological compounds are sensitive to heat, as evidenced using the HKB preparation as an alternative to KFB to characterize the nature of these compounds produced by Kombucha. According to the findings of this study, the inhibitory activity of kombucha infusions is not primarily attributed to heat-sensitive molecules; hence, the antibacterial effectiveness of the HKB preparation may be slightly diminished. This is strongly supported by the studies of Sreeramulu et al. [5] and Battikh et al. [9], which explain the relatively weak antimicrobial activity exhibited by the NKB preparation.

This observation suggested that the active compounds of the selected herbs, along with the fermentation of Kombucha, result in the production of numerous fermentation end products, such as alkaloids [21, 22], esters, fatty acids [23, 24], 5-hydroxymethyl-2-furaldehyde [25-27], unsaturated lactones, hydroxylactones [28, 29], and heterocyclic compounds. These compounds have been reported to exhibit antibacterial activity. Additionally, the variable antibacterial activity mainly depends on the microorganisms present in the symbiotic culture used for kombucha fermentation, as well as fermentation time and temperature [1, 5].

IV. Conclusion

Our study revealed an intriguing finding: the robust antibacterial properties exhibited by herbal kombucha, particularly the variant infused with roselle. It appears that during the intricate process of kombucha fermentation, a synergistic interplay between bacteria and yeasts gives rise to potent antimicrobial metabolites. This finding not only piques scholarly interest but also warrants further investigation. Delving into the intricate mechanisms underlying these microbial interactions holds promise for uncovering a wealth of novel antimicrobial compounds, with implications for future therapeutic endeavors. Thus, our exploration of kombucha and its microbial intricacies underscores the importance of ongoing research to unravel its profound therapeutic potential.

Acknowledgment

This research was supported by the School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand.

References

[1.] R. Jayabalan, et al, "A review on kombucha tea—microbiology, composition, fermentation, beneficial effects, toxicity, and tea fungus", Comprehensive Reviews in Food Science and Food Safety, 13(4): 538-550, 2014.

[2.] S.A. Villarreal-Soto, et al, "Understanding kombucha tea fermentation" A Review. Journal of Food Science, 83(3): 580-588, 2018.
[3.]A.R. Al-Mohammadi, et al, "Chemical constitution and antimicrobial activity of kombucha fermented beverage," Molecules, 26(16), 2021.
[4.]C.J. Greenwalt, K.H. Steinkraus, and R.A. Ledford, " Kombucha, the fermented tea: microbiology, composition, and claimed health effects," Journal of Food Protection, 63(7): 976-981, 2000.
[5.]G. Sreeramulu, Y. Zhu, and W. Knol, "Kombucha fermentation and its antimicrobial activity," Journal of Agricultural and Food Chemistry, 48(6): 2589-2594, 2000.
[6.] A.M. Hartmann, et al, "Effects of chronic kombucha ingestion on open-field behaviors, longevity, appetitive behaviors, and organs in c57-bl/6 mice: a pilot study," Nutrition, 16(9): 755-761, 2000.
[7.] J. Martínez. Leal, et al, "A review on health benefits of kombucha nutritional compounds and metabolites," CyTA-Journal of Food, 16(1): 390-399, 2018.
[8.] E. Ivanišová, et al, "The evaluation of chemical, antioxidant, antimicrobial and sensory properties of kombucha tea beverage," Journal of Food Science and Technology, 57(5): 1840-1846, 2020.
[9.] H. Battikh, A. Bakhrouf, and E. Ammar, "Antimicrobial effect of kombucha analogues," LWT-Food Science and Technology, 47(1): 71-77, 2012.
[10.] M.I. Watawana, et al, "Health, wellness, and safety aspects of the consumption of kombucha," Journal of Chemistry, : 591869, 2015.
[11.] C. Enan, et al, "Antimicrobial activity of *Enterococcus faecium* NM2 isolated from urine: Purification, characterization and bactericidal action of enterocin NM2," Asian Journal of Applied Sciences, 7(7): 621-634, 2014.
[12.] G. Enan, M.E.F. Abdel-Haliem, and E. Tartour, "Evaluation of the antimicrobial activity, starter capability and technological properties of some probiotic bacteria isolated from Egyptian pickles," Life Sci. J, 11(11): 976-985, 2014.
[13.] F.C. Prado, et al, "Trends in non-dairy probiotic beverages," Food Research International, 41(2): 111-123, 2008.
[14.] T.-Y. Sun, J.-S. Li, and C. Chen, "Effects of blending wheatgrass juice on enhancing phenolic compounds and antioxidant activities of traditional kombucha beverage," Journal of food and drug analysis, 23(4): 709-718, 2015.
[15.] S.A. Villarreal-Soto, et al, "Understanding kombucha tea fermentation: a review," Journal of food science, 83(3): 580-588, 2018.
[16.] E. Lončar, et al, "Influence of working conditions upon kombucha conducted fermentation of black tea," Food and Bioproducts Processing, 84(3): 186-192, 2006.
[17.] C.J. Greenwalt, R.A. Ledford, and K.H. Steinkraus, "Determination and characterization of the antimicrobial activity of the fermented tea kombucha," LWT-Food Science and Technology, 31(3): 291-296, 1998.
[18.] T. Kaewkod, S. Bovonsombut, and Y. Tragoolpua, "Efficacy of kombucha obtained from green, oolong, and black teas on inhibition of pathogenic bacteria, antioxidation, and toxicity on colorectal cancer cell line," Microorganisms, 7(12): 700, 2019.
[19.] A. Osman, et al, "Catfish glycoprotein, a highly powerful safe preservative of minced beef stored at 4 C for 15 days," Foods, 9(8): 1115, 2020.
[20.] A. Osman, et al, "Dietary supplementation of soybean glycinin can alter the growth, carcass traits, blood biochemical indices, and meat quality of broilers," Poultry science, 99(2): 820-828, 2020.
[21.] J.S. Teh, "Toxicity of short-chain fatty acids and alcohols towards *Cladosporium resinae*," Applied microbiology, 28(5): 840-844, 1974.
[22.] A. Mithöfer, and W. Boland, "Plant defense against herbivores: chemical aspects," Annual review of plant biology, 63: 431-450, 2012.
[23.] A.-R. Al-Mohammadi, et al, "Chemical constitution and antimicrobial activity of kefir fermented beverage," Molecules, 26(9): 2635, 2021.
[24.] L. Zhao, et al, "In vitro antibacterial activities and mechanism of sugar fatty acid esters against five food-related bacteria," Food chemistry, 187: 370-377, 2015.
[25.] J.M. Beale, J. Block, and R. Hill, "Organic medicinal and pharmaceutical chemistry," Philadelphia: Lippincott Williams & Wilkins, 2010.
[26.] J.L. Holtzman, et al, "The kinetics of the aerobic reduction of nitrofurantoin by NADPH-cytochrome P-450 (c) reductase," Molecular Pharmacology, 20(3): 669-673, 1981.
[27.] D. Trombetta, et al, "Study on the mechanisms of the antibacterial action of some plant α, β-unsaturated aldehydes," Letters in applied microbiology, 35(4): 285-290, 2002.
[28.] K. Wińska, et al, "Antimicrobial activity of new bicyclic lactones with three or four methyl groups obtained both synthetically and biosynthetically," Journal of Saudi Chemical Society, 22(3): 363-371, 2018.
[29.] D. Mares, "Antimicrobial activity of protoanemonin, a lactone from ranunculaceous plants," Mycopathologia, 98(3): 133-140, 1987.

Empirical Correlations Between Standard Penetration Test N Value and Shear Wave Velocity from Multichannel Analysis of Surface Wave for Earthfill Dam in Northern Thailand

Orawan Lunanta[1, 2, a)] and Siriporn Chaisri[3, b)]

[1] Department of Geological Sciences, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand.
[2] Geology Division, Bureau of Engineering Topographical and Geotechnical Survey, Royal Irrigation Department, Ministry of Agriculture and Cooperatives, Bangkok, 10300, Thailand.
[3] Department of Physics and Materials Sciences, Faculty of Science, Chiang Mai University, Chiang Mai, 50200, Thailand.

[a)] Corresponding author: orawan_lunanta@cmu.ac.th
[b)] Corresponding author: siriporn.chaisri@cmu.ac.th

Abstract: The soil stiffness of the subsurface can be obtained through soil mechanical survey methods such as drilling and in-situ testing like the Standard Penetration Test (SPT) N values. The Royal Irrigation Department (RID) engineers often require N values for the earthfill dam inspections and their geotechnical maintenance. To minimize the risk of damage, the SPT will only be performed on the dam crests where essential for maintenance or reconstruction. This necessitates estimating N values of the subsurface from available non-invasive geophysical techniques. Among various geophysical methods, the multichannel analysis of surface waves (MASW) method is widely used involving the measurement of in-situ shear wave velocity (V_s) of subsurface which also provides valuable information about soil stiffness for geotechnical investigations. This study establishes empirical correlations between V_s values obtained from the MASW method and uncorrected SPT N values for compacted soil in both homogeneous and zoned type of earthfill dams in Northern Thailand, using ordinary linear regression (OLR) and orthogonal regression (OR) methods. The new equations indicate an acceptable ability to estimate N values based on V_s values. However, caution is advised when applying empirical correlations, especially considering the specific soil conditions and deeper zones within the dam. These correlations provide valuable insights for predicting N values and guiding engineers in assessing the structural stability and potential damage of earthfill dams in Northern Thailand.

***Index Terms*— Earthfill dam, Multichannel Analysis of Surface Wave (MASW), N value, Ordinary linear regression (OLR), Orthogonal regression (OR), Shear wave velocity (V_s), Standard Penetration test (SPT)**

I. INTRODUCTION

In the past, Royal Irrigation Department (RID) assessed the stability of earthfill dams through visual inspection. Conventionally, RID safety engineers investigate dams by inspecting for fractures, seepages, and settlements. These characteristics are apparent in the exteriors of long-utilized reservoirs. The structural stability of a dam is influenced by the dam body, particularly the impervious clay core and the contact between the embankment and the foundation. These factors are invisible and cannot be assessed by a traditional visual inspection. Since 2014, RID has utilized three geophysical techniques: electrical resistivity, multichannel analysis of surface waves (MASW), and ground penetrating radar (GPR) investigations to enhance their comprehension of the dam's construction.

Seismic methods are widely utilized in geotechnical engineering work to evaluate the stability of earthfill dams, characterize subsurface conditions, and determine material properties [1-3]. The body wave velocities of geomaterials, shear wave velocity (V_s), and compressional wave velocity (V_p), are directly related to the elastic moduli of the medium which the waves propagate through [4] which V_s is primarily a key indicator of the stress-strain behaviour of soil due to its relation to the small-strain shear modulus [5] as,

$$G = \rho V_s^2 \quad (1)$$

where G is the shear modulus, and ρ is the density of the soil [3].

Shear wave velocity (V_s) can be measured or estimated through laboratory or field tests. In situ field measurements include invasive techniques, such as down-hole and cross-hole seismic surveys. These methods require drilling boreholes to an appropriate depth. On the other hand, non-invasive geophysical techniques like spectral analysis of surface waves (SASW) and MASW can be used to estimate V_s at a specific location. These methods do not require drilling equipment and only rely on acoustic sensors to determine the V_s of the subsurface. Due to these reasons, RID engineers prefer employing a non-destructive geophysical method, i.e., MASW. The MASW is a seismic surface wave technique developed specifically for near-surface applications at shallow depths of a few tens of meters [6]. This method has several advantages: fast field data acquisition, cost-effectiveness, simplified data processing, and easy elimination of noise sources. MASW does not require any heavy machinery and does not create lasting

marks on the surface of the test site. Moreover, it is environmentally friendly.

The standard penetration test (SPT) is commonly used in geotechnical engineering surveys to evaluate various soil design parameters. The SPT provides a measure of soil penetration resistance by counting the number of hammer blows, known as the 'N value' (SPT blow counts, or standard penetration resistance). Determining the N value involves driving a standard split spoon sampler into the ground with a drop hammer weighing 63.5 kg falling free from a height of 76 cm. The sampler inserts 15 cm into the soil at the bottom of a borehole, and the number of blows (N) needed to drive it a further 30 cm is measured [7]. The N value is utilized to approximate the homogeneity, consistency, compressibility, and strength of soils. It has been reported to show correlations of soil engineering properties, especially soil stiffness in terms of shear modulus (G). Although this approach allows the study of intrinsic properties, they inevitably cause structural damage to the inner structure of the dam.

In most practical cases, RID engineers require the N value for structural inspection and geotechnical design purposes. Due to the fact that the SPT is a destructive method and cannot be carried out at every test location, we need a method to estimate the N value from available seismic data. According to (1), the shear modulus is directly related to the V_s, and since the shear modulus correlates with the SPT N value, we can estimate the N value using V_s. Several empirical relations were derived for V_s and N value correlation in the general form, described as, [5, 8, 9, 10]

$$V_s = aN^b \qquad (2)$$

where a and b are fitted coefficients depending on locations and types of soils.

Numerous studies demonstrate the correlation between V_s and soil geotechnical parameters, such as standard penetration resistance (N value), in specific areas. To evaluate the dynamic properties of soil, there are the report studies that established an empirical correlation among the V_s and N values as in [11-16]. A large dataset from 250 sites throughout Japan was analyzed. They developed V_s correlation equations based on N value, soil type, and geological age [14]. To determine the statistical relationships between SPT and V_s for Tehran city's clayey, silty, and fine-grained soil, as discussed in [17]. The proposed equations using uncorrected SPT N values offer a slightly better fit than those using energy-corrected values. They suggest that employing equations based on unadjusted blow-counts (SPT N) for all soil types is suitable for indirectly estimating V_s as discussed in [18]. In Turkey, the study conducted a graphical comparison and validation using the consistency ratio to propose a statistical correlation between V_s and N values for all soil types, silty and clayey. The finding showed that the newly established relationships demonstrated good prediction performance [19]. As detailed in [20], the study aimed to investigate the correlation between V_s, SPT N, and soil characteristics as a function of depth. To study seismic site characterization in the northeastern region of India, researchers employed the MASW technique at 27 sites. They selected 27 SPT bore log profiles for statistical analysis near these sites. Site-specific data were utilized to establish correlations between V_s and N values (both corrected and uncorrected). These correlations were subsequently predicted V_s based on the available SPT N values [21]. The new regression equation demonstrates good prediction performance in establishing correlations between N value and V_s for different categories of soils in Western India. V_s were measured using MASW. These correlations are based on corrected and uncorrected N values and were developed using simple regression analysis, as described in [22]. The study investigated the relationship between the V_s and SPT N value for Dhaka City in Bangladesh. Empirical correlations were found to predict V_s from the SPT N data. A total of 152 V_s and SPT N data pairs were used to derive empirical correlations for all soils, sandy soils, and clayey soils. The proposed correlations between the V_s and uncorrected SPT N values were expressed using nonlinear regression equations, as detailed in [23]. An empirical correlation between V_s and SPT N values for various soil types in the Philippines was developed. Nonlinear regression was employed to formulate models applicable to Philippine soils. These equations were then compared with existing research to validate their effectiveness [24].

Many researchers have proposed correlations mostly between N value and V_s for different soils (all soil, sand, gravel, silt, and clay). There have been attempts to incorporate additional dependent variables in describing the relationships. Table I gives an example of the existing empirical relationships.

Though many empirical relationships between V_s and N values exist, these relationships are specific to particular areas and cannot be applied to all regions. Nevertheless, in Thailand, RID engineers have attempted to compare the measured N value from SPT, with the N value estimated from V_s using a default formula available in the software. Fig. 1. shows the discrepancies in the N values obtained from two different approaches. Due to the significant uncertainties arising from the default parameters, a correlation between V_s and N values for dam investigation in Thailand must be examined to optimize the good conversion between them for the dam structural stability inspection.

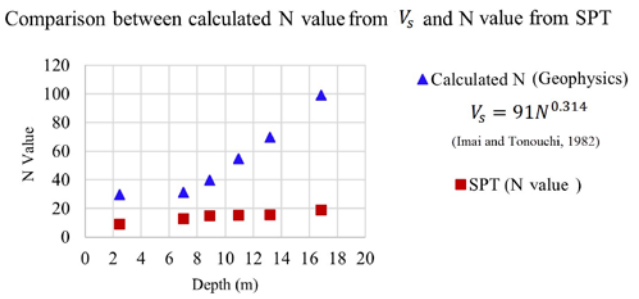


Fig. 1. Comparison between calculated N values from V_s and N values from SPT, data from some earthfill dam in Northern Thailand.

The main purpose of this article is to estimate empirical regression equations between the V_s and SPT N values based on the geotechnical and geophysical studies carried out on earthfill dams in Northern Thailand. These equations were

applied to evaluate N values from V_s because RID engineers attempt to avoid SPT N values acquisition on the dam crest.

Table I
EMPIRICAL EQUATION EXAMPLES BETWEEN V_s AND N VALUE IN FORM OF $V_s = aN^b$.

No.	Authors	Country	All soil	Sand	Clay
1	Ohba and Tourima 1970)	Japan	$85.34N^{0.31}$		
2	Ohsaki and Iwasaki (1973)	Japan	$81.38N^{0.39}$	$59.4N^{0.47}$	
3	Imai (1977)	Japan	$91N^{0.337}$	$80.6N^{0.331}$	$80.2N^{0.292}$
4	Seed and Idriss (1981)	USA	$61N^{0.5}$		
5	Imai and Tonouchi (1982)	Japan	$97N^{0.314}$		
6	Sykora and Stokoe (1983)	USA		$100.5N^{0.49}$	
7	Lee (1990)	Taiwan		$57.4N^{0.49}$	$114.43N^{0.3}$
9	Pitilakis et al. (1992)	Greece		$162N^{0.17}$	
10	Athanasopoulos (1995)	Greece	$107.6N^{0.36}$		$76.55N^{0.445}$
11	Iyisan (1996)	Turkey	$51.5N^{0.516}$		
13	Jafari et al. (2002)	Iran			$27N^{0.73}$
14	Hasancebi and Ulusay (2007)	Turkey	$90N^{0.309}$	$90.82N^{0.319}$	$97.89N^{0.26}$
15	Dikmen (2009)	Turkey	$58N^{0.39}$	$73N^{0.33}$	$44N^{0.48}$
16	Maheswari et al. (2010)	India	$95.64N^{0.301}$	$100.53N^{0.265}$	$89.31N^{0.35}$
17	Hafezi Moghadas et al. (2010)	Iran	$99N^{0.53}$	$80N^{0.58}$	$45N^{0.72}$
18	Chatterjee and Choudhury (2013)	India	$78.21N^{0.38}$	$54.82N^{0.53}$	$77.11N^{0.39}$
19	Sil and Haloi (2017)	Any region	$75.48N^{0.38}$	$79.22N^{0.37}$	$99.71N^{0.34}$
20	Gautam (2017)	Nepal	$115.8N^{0.251}$	$78.7N^{0.352}$	
21	Rahman et al. (2018)	Bangladesh	$97.306N^{0.33}$	$82.01N^{0.381}$	$100.58N^{0.3}$
22	Daag et al. (2022)	Manila, Philippines	$56.8N^{0.486}$	$45.07N^{0.553}$	$70.26N^{0.42}$

II. STUDY AREA

Our work focuses on the impervious core of earthfill dams in Northern Thailand. The data have been acquired from eight earthfill dams with their locations indicated in Fig.2. Each dam was selected for this study based on data availability on both seismic surface wave data and SPT N value data on the dam crest was constructed by the RID. According to RID [25], most water-storage dams in Thailand are built from compacted soil, known as earthfill dams. A rolled earthfill is a compacted layer of man-made soil, specifically designed to fulfil engineering requirements. Such specifications are determined based on the structural characteristics and materials used for construction. The studied dams were divided into homogeneous dams and zoned type dams, as shown in Fig. 3. A homogeneous dam is built from a single type of solid clay, such as clayey gravel (GC), clayey sand (SC), and lean clay (CL) soils. In contrast, zoned type dams are composed of different types of soil. The core zone consists of impervious soil types such as GC, SC, and CL soils. Moreover, in the random zone or shell zone next to the dam's core, semi-impervious soils, such as soil types, are mixed with gravel or sand as construction materials. Based on the classified dam components proposed by [26], the data from 37 dam projects in Thailand were used for statistical analysis. Those dams were classified by dam component according to [27-28] standards. They are (1) homogeneous zone, (2) core zone, and (3) random zone. Table II. is a guideline for the general classification of dam material.

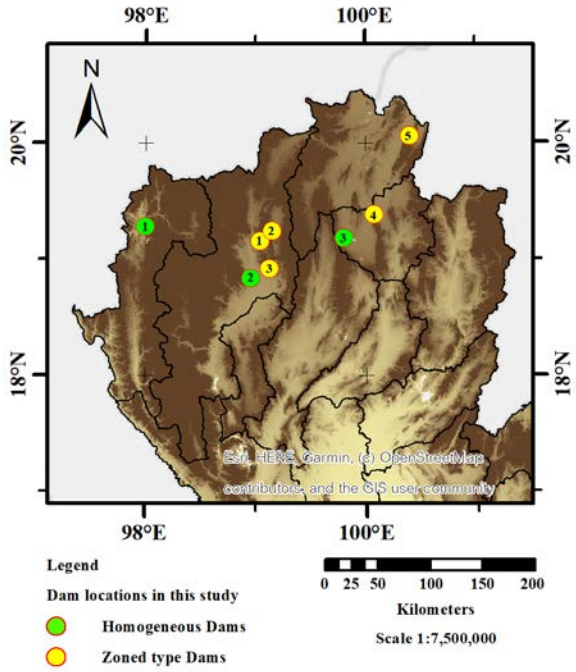


Fig. 2. Location map of the 8 earthfill dams used in this study.

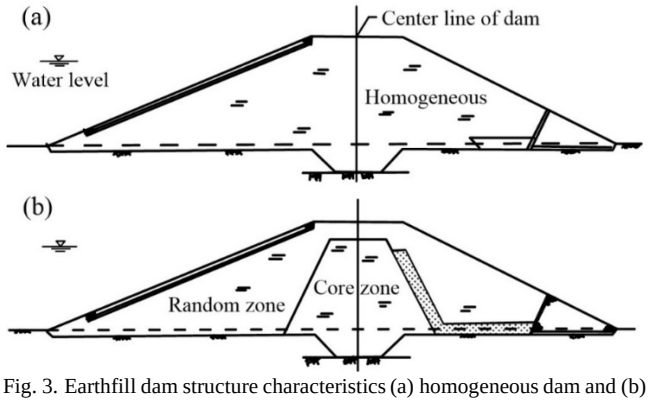


Fig. 3. Earthfill dam structure characteristics (a) homogeneous dam and (b) zoned type dam (modified from [25]).

The standard rule for earthfill dam construction is that the degree of compaction at the core must be equal to or above 98% standard proctor compaction test (SCPT). For the outer shell, the compaction level must be 95% (SPCT) minimum [25]. Due to the variation in soil type, composition, and

unique layer structure, the correlation between V_s and N differs from other study areas.

Table II
CLASSIFICATION OF DAM TYPE BASED ON CLASSIFICATION OF DAM MATERIAL [29].

Dam component	Soil classification (USCS)
Homogeneous	GC, GM, GC-GM, SC, SM, SC-SM, SM-SC, CL, CL-ML, ML
Core zone	GC, SC, CL
Random zone	GM, GC-GM, SM, SC-SM, SM-SC, CL-ML, ML

Note: GC = clayey gravel, SC = clayey sand, CL = lean clay, GM = silty gravel, GM-GC = clayey silty gravel, SM = silty sand, SC-SM = silty clayey sand, CL-ML = silty clayey, and ML = silt

III. METHODOLOGY

A. Data collection

The reliable SPT data used in this study were collected from soil and geology investigation reports of earthfill dams, conducted by the Soil Science and Geological branch, Engineering Division of Regional Irrigation Offices 1 and 2. The SPT investigations were carried out according to the standardized approach provided in ASTM-D1586 [30], while the soil description and identification were performed using a Unified Soil Classification System in ASTM-D2487 [31], employing a visual manual procedure. The objective of the SPT investigation was to explore suitable methods for improving the earthfill dam. The geophysical surveys were conducted using active MASW techniques by the Geophysics branch of the Geology Division at the Bureau of Engineering Topographical and Geotechnical Survey. V_s was estimated at shallow depths, with appropriate target frequency, geophone spacing, and survey line length, enabling investigation to a depth of 30 meters. This study utilized SPT N data from 17 boreholes and 18 MASW sites lying close positioned along the crest of the dams.

B. Data processing

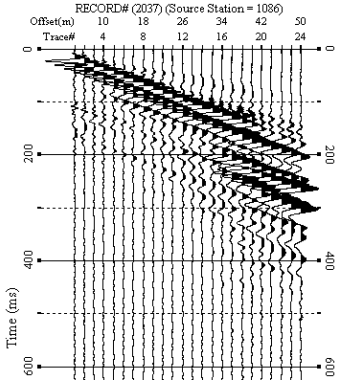
The Shear wave velocity models for each location were generated using the ParkSEIS software. This software was utilized to process and analyze the MASW data. By using a 2D wavefield transformation, the MASW technique converts raw field data from a time-offset (t-x) domain into a frequency-phase velocity (f-v) domain in which dispersion patterns are detectable through the wavefield maxima. These dispersion properties of surface waves can be extracted as a dispersion curve, which is subsequently inverted to derive a one-dimensional (1D) V_s model [32]. The final V_s model was obtained by iteratively updating it until the theoretical dispersion curve aligned with the observed dispersion curve. An overview of the main steps of the MASW method is provided in Fig. 4.

C. Correlation between V_s and SPT N values

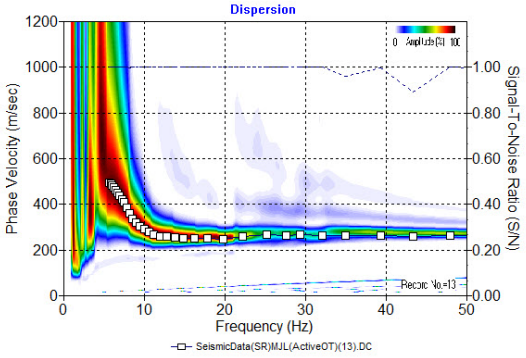
The shear wave velocity and uncorrected SPT N values from eight locations were analyzed to establish correlation data from up to a depth of 30 m. Data from three locations for homogeneous dams were considered, comprising 86 V_s and N Value datasets, with 69 datasets being fine-grained soil and 17 being coarse-grained soil. Conversely, data from five locations were analyzed for zoned type dams, totaling 95 V_s and N Value datasets, with 75 fine-grained and 20 coarse-grained soils. Fine-grained soils included CL, ML, CH, and CL-ML, while coarse-grained soils comprised SC, SM, SP, GM, SP-SC, SC-GC, and GW-GM.

Data preconditioning involves removing data points acquired from a depth shallower than one meter, which consists of asphalt and a sub-base layer of compacted aggregate material. It also entails filtering out V_s values and SPT N values that exhibit excessive distribution at the same depth.

(a) Multichannel record



(b) Dispersion analysis



(c) Inversion: 1D V_s model

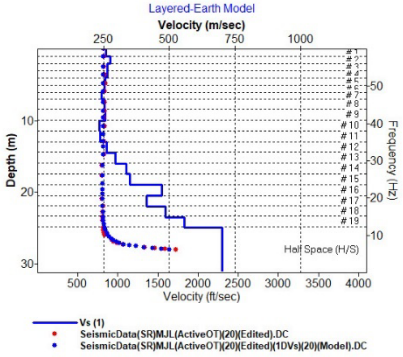


Fig. 4. Main steps of the MASW method. The seismic record in time-space domain (a) is transformed into phase velocity spectrum (b). The dispersion curve will be extracted and then inverted for (c) a final 1D V_s model.

A power regression analysis derives coefficients from the ordinary linear regression (OLR) and orthogonal regression

(OR) for the earthfill dams. Furthermore, it involves consideration of the scale error percentage (Er), where V_s values obtained from SPT N values that deviate from the processed V_s by more than 15% are filtered out. The scale error percentage is defined as [19],

$$Er = 100 \times \frac{(Vs_{cal} - Vs_{proc})}{Vs_{cal}} \quad (3)$$

where Vs_{proc} and Vs_{cal} are the processed and calculated V_s using the proposed regression equations, respectively.

IV. RESULT AND DISCUSSION

The resulting equation and R^2 value derived for the dams are shown in Fig.5 and Fig.6, which illustrate the correlation between V_s and SPT N values for homogeneous and zoned type dams, respectively.

For homogeneous dams using OLR

$$V_s = 120.77N^{0.338} \ (R^2 = 0.758) \quad (4)$$

For homogeneous dams using OR

$$V_s = 116.34N^{0.351} \ (R^2 = 0.810) \quad (5)$$

For zoned type dams using OLR

$$V_s = 93.70N^{0.365} \ (R^2 = 0.803) \quad (6)$$

For zoned type dams using OR

$$V_s = 91.25N^{0.375} \ (R^2 = 0.849) \quad (7)$$

According to the analysis, the difference in equations between homogeneous dams and zoned type dams is likely due to the specific soil conditions of each dam type. The functions obtained from both OLR and OR give comparable coefficients, resulting in similar N values. Therefore, both methods could be employed to determine a suitable function. Moreover, considering the equation validated based on the normal consistency ratio (C_d) of the difference between processed and calculated V_s values to measured SPT N values. It can be calculated as

$$C_d = \frac{(Vs_{proc} - Vs_{cal})}{N} \quad (8)$$

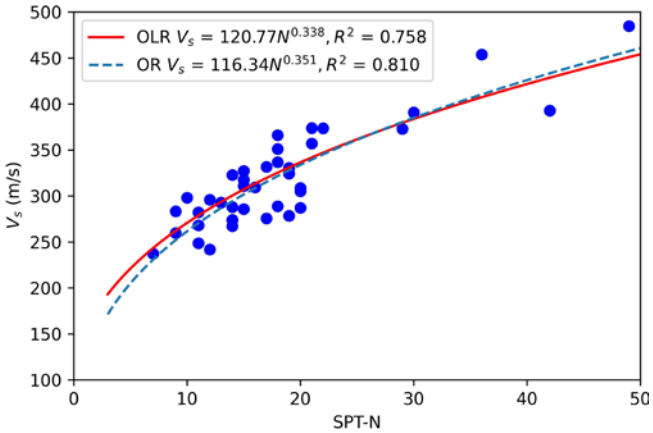


Fig.5. Correlation between V_s and SPT N values for homogeneous dams. The legend indicates a solid red line for OLR and a dashed blue line for OR

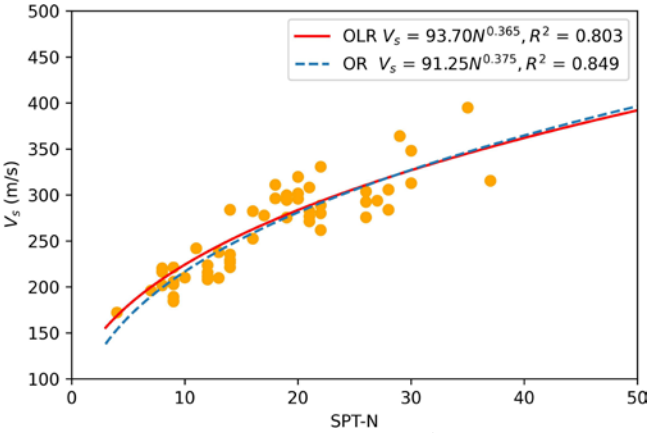


Fig.6. Correlation between V_s and SPT N values for the zoned type dams. The legend indicates a solid red line for OLR and a dashed blue line for OR

Fig. 7. presents a comparison of C_d and N values for all dams and two regression methods. Most of the scatter points are close to zero, indicating that the proposed equations have good performance in predicting V_s, except for small SPT N values of less than 15. Hence, it is crucial to consider the specified small SPT N values (SPT N < 15) before applying them to derive the equations. Additionally, In the shallow zone, the N values determined by SPT are more accurate than those in the deeper zone. Errors that could arise from SPT N value are low in the upper soil layers and gradually rise with increasing depth. The rise in N value may be related to overburden stress. Therefore, it could be either depth-independent or depth-dependent variables. Therefore, for a larger dataset, additional considerations regarding the measurement depth should be taken into account to improve the correlation.

V. CONCLUSION

This study derived a new correlation equation between V_s and SPT N values by categorizing the types of dams in Northern Thailand. Both OLR and OR methods were employed to establish the correlation. The N value obtained from V_s by the empirical correlations can provide useful engineering parameters of soils. Since all earthfill dams in Thailand are constructed under the common RID standard, the results from this study can be applied to the other Northern dams in Thailand. The correlations between V_s and the N value derived in this work can guide engineers in inspecting structural stability, investigating potential damage, and designing new earthfill dams in the future.

ACKNOWLEDGMENT

The authors would like to thank the Geophysics branch of the Geology Division in the Bureau of Engineering Topographical and Geotechnical Survey of the Royal Irrigation Department for allowing the use of data and software and the Soil Science and Geological branch, Engineering Division of Regional Irrigation Offices 1 and 2 for providing data for this research work. Moreover, the

Department of Geological Sciences, Faculty of Science, Chaing Mai University, for all research facilities.

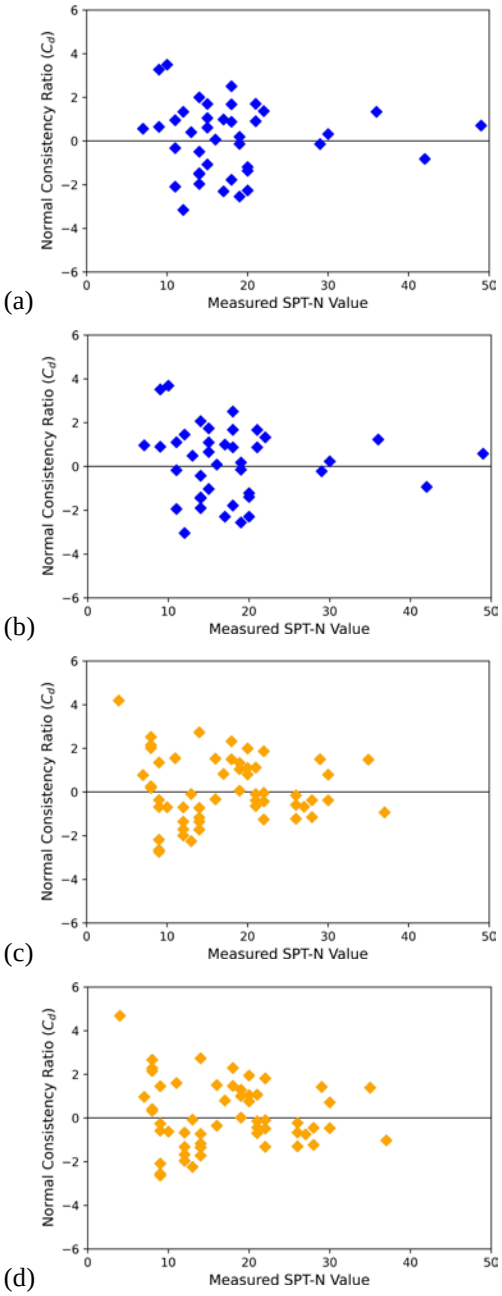


Fig.7. Normal consistency ratio for (a) homogeneous dams (OLR) (b) homogeneous dams (OR) (c) zoned type dams (OLR), and zoned type dams (OR)

REFERENCES

[1] A. M. E. Mohamed, A.S.A. Abu El Ata, F. Abdel Azim, and M.A. Taha, "Site-specific shear wave velocity investigation for geotechnical engineering applications using seismic refraction and 2D multi-channel analysis of surface wave," NRIAG Journal of Astronomy and Geophysics, vol. 2, no. 1, pp. 88-101, 2013.

[2] E. Cardarelli, M. Cercato, and G. De Donno, "Characterization of an earth-filled dam through the combined use of electrical resistivity tomography, P- and SH-wave seismic tomography and surface wave data," Journal of Applied Geophysics, vol. 106, pp. 87-95, 2014.

[3] D. Park and T. Kishida, "Shear wave velocity profiles of fill dams," Soil Dynamics and Earthquake Engineering, vol. 104, pp.250-258, 2018.

[4] M. E. Everett, "Near-Surface Applied Geophysics," Cambridge University Press, 2013.

[5] B. R. Wair, J. T. Dejong, and T. Shantz, "Guidelines for estimation of shear wave velocity Profiles," PEER report, University of California, 2012.

[6] C. B. Park, R. D. Miller, and J. Xia, "Multichannel analysis of surface waves," Geophysics, vol.64, 800-808, 1999.

[7] M. Budhu, "Soil mechanics and foundations" John wiley & Sons, Inc.,2010.

[8] B.Kirar, B. K. Maheshwari, and P. Muley, "Correlation between shear wave velocity (V_s) and SPT resistance (N) for Roorkee Region," International Journal of Geosynthetics and Ground Engineering, vol. 2, no. 1, 2016.

[9] A. Sil and J. Haloi, "Empirical correlations with standard penetration test (SPT)-N for estimating shear wave velocity applicable to any region," International Journal of Geosynthetics and Ground Engineering, vol.3, 2017.

[10] S. Bandyopadhyay, A. Sengupta, and G. R. Reddy, "Development of correlation between SPT-Nvalue and shear wave velocity and estimation of non-linear seismic site effects for soft deposits in Kolkata City," Geomechanics and Geoengineering, vol.16, no.1, pp. 1-19, 2021.

[11] S. Ohba, and I. Toriumi, "Dynamic response characteristics of Osaka plain," Annual Meeting AIJ (in Japanese), 1970.

[12] T. Imai, "P- and S-wave velocities of the ground in Japan," 9th International Conference on Soil Mechanics and Foundation Engineering, vol.2, pp. 257-260, 1977.

[13] H. Seed, and I. Idriss, "Evaluation of liquefaction potential of sand deposits base on observations and performance in previous earthquakes," ASCE National convention, pp. 481-544, 1981.

[14] T. Imai, and K. Tonouchi, "Correlation of N-value with S-wave velocity and shear modulus," 2nd European Symposium on Penetration Testing, pp. 57-72, 1982.

[15] D.E. Sykora, and K. H. Stokoe, "Correlation of in-situ measurement in sands of shear wave velocity," Soil dynamic earthquake engineering, vol. 20, no. 1, pp. 125-136, 1983.

[16] S.H.H. Lee, "Regression models of shear wave velocities in Taipei basin," Journal of the Chinese Institute of Engineers, vol. 13, no. 5, 519-532, 1990.

[17] M. K. Jafari, A. Shafiee, and A. Razmkhah, "Dynamic properties of fine-grained soils in south of Tehran," Journal of Seismology and Earthquake Engineering, vol. 4, no.1, pp. 25-35, 2002.

[18] N. Hasancebi, and R. Ulusay, "Empirical correlations between shear wave velocity and penetration resistance for ground shaking assessments," Bulletin of Engineering Geology and the Environment, vol. 66, pp. 203-213, 2007.

[19] U. Dikmen, “Statistical correlations of shear wave velocity and penetration resistance for soils,” Journal of Geophysics and Engineering, vol.6, pp. 61-72, 2009.
[20] M. K. Akin, S. L. Kramer, and T. Topal, “Empirical correlations of shear wave velocity (V_s) and penetration resistance (SPT-N) for different soils in an earthquake-prone area (Erbaa-Turkey),” Engineering Geology, vol. 119, no. 1, pp. 1-17, 2011.
[21] A. Sil, and T.G. Sitharam, “Dynamic site characterization and correlation of Shear wave velocity with standard penetration test ‘N’ values for the city of Agartala, Tripura State, India” Pure and Applied Geophysics, vol.171, no. 8, pp. 1859-1876, 2014
[22] S. Thokchom, B. K. Rastogi, N. N. Dogra, V. Pancholi, B. Sairam, F. Bhattacharya, and V. Patel, “Empirical correlation of SPT blow counts versus shear wave velocity for different types of soils in Dholera, Western India,” Natural Hazards, vol. 86, no. 3, pp. 1291-1306, 2017.
[23] M.Z, Rahman, A. S. M. M. Kamal, and S. Siddiqua, “Near-surface shear wave velocity estimation and V_s30 mapping for Dhaka City, Bangladesh,” Natural Hazards, vol. 92, no. 3., pp, 1687-1715, 2018.
[24] A. S. Daag, O. P. C. Halasan, A. A. T. Magnaye, R. N. Grutas, and R. U., “Solidum, Empirical correlation between standard penetration resistance (SPT-N) and shear wave velocity V_S for soils in Metro Manila, Philippines,” Applied Sciences, vol. 12, no. 16, 2022.
[25] Royal Irrigation Department, “Standard drawings for water storage dams and assembly buildings (แบบมาตรฐานเขื่อนกักเก็บน้ำและอาคารประกอบ),” Internal report of Royal Irrigation Department, Thailand, 2002.
[26] B. Kumma, “The Influence of dam components and their properties of small and medium earth dams in Thailand on their stability during earthquake,” M.A. thesis, Kasetsart University, 2010.
[27] USBR, “Design of small dams,” U.S. Government Printing Office, 1987.
[28] USBR, “Earth manual Part 1,” U.S. Government Printing Office, 1998.
[29] S. Solalump, “Dam safety engineering,” Kasetsart University, 2011.
[30] ASTM-D1586; “Standard Test Method for Standard Penetration Test (SPT) and Split-Barrel Sampling of Soils,” ASTM International: West Conshohocken, PA, USA, 2019.
[31] ASTM-D2487; Standard Practice for Classification of Soils for Engineering Purpose (Unified Soil Classification System). ASTM International: West Conshohocken, PA, USA, 2020.
[32] C. Park, “MASW for geotechnical site investigation,” The Leading Edge, vol. 32, no. 6., 2013.

Assessment of Antioxidant Activity, Total Flavonoid Content, and Total Phenolic Compounds from Sequential Exhaustive Extraction of *Wolffia globosa* Cultivated in a Controlled Greenhouse Environment.

Pakjira Patchim[1, b)], Chutwadee Krisanapun[1], Sukanya Mingyai[2], Phuwanai Palakachen[3], Arnatchai Maiuthed[4,5] and Pattarawit Rukthong [1, 6, a)]

[1] Faculty of Pharmacy, Srinakharinwirot University, Nakhonnayok 26120, Thailand.
[2] Faculty of Agricultural Product Innovation and Technology, Srinakharinwirot University, Nakhonnayok 26120, Thailand.
[3] Faculty of Engineering, Srinakharinwirot University, Nakhonnayok 26120, Thailand.
[4] Department of Pharmacology, Faculty of Pharmacy, Mahidol University, Bangkok 10400, Thailand.
[5] Centre of Biopharmaceutical Science for Healthy Ageing, Faculty of Pharmacy, Mahidol University, Bangkok 10400, Thailand.
[6] Center for Excellence in Plant and Herbal Innovation Research, Strategic Wisdom and Research Institute, Srinakharinwirot University, Nakornnayok 26120, Thailand.

[a)] Corresponding author: pattarawit@g.swu.ac.th
[b)] pakjira0804@gmail.com

Abstract. *Wolffia globosa*, a small aquatic plant within the Lemnaceae family, is renowned for its rich nutrient profile, abundance of antioxidants, and polyphenols. This study aimed to assess the antioxidant capacity, Total Phenolic Content and Flavonoid Content of *W. globosa* cultivated in a controlled environment, utilizing a serial exhaustive extraction method known for its effectiveness in extracting metabolites from various plant materials. Through successive extractions with solvents of increasing polarity, ranging from nonpolar (hexane) to polar (water), a broad spectrum of phytochemicals was extracted, significantly impacting antioxidant potential. The results unveiled varying levels of total phenolics and total flavonoids in the extracts and fractions, ranging from 32.91±1.65 to 52.77±0.43 GAE mg/g and 46.59±0.39 to 175.95±1.12 QE mg/g, respectively. These phenolic and flavonoid compounds play pivotal roles in antioxidant activity by neutralizing free radicals through hydrogen atom donation. Antioxidant activities were assessed using DPPH, ABTS, and FRAP assays. Ethanol exhibited the highest radical scavenging activity in the DPPH assay, measuring at 435.84±4.34 TE µmol/g. Meanwhile, hexane demonstrated the highest antiradical power in the ABTS assay at 5.49±0.006 TE µmol/g. Additionally, ethyl acetate extracts demonstrated the highest ferric reducing antioxidant power in the FRAP assay with 14.48±0.04 TE µmol/g. This study underscores the impact of solvent polarity on the antioxidant properties of *W. globosa* extracts, highlighting the importance of solvent selection in extraction methodologies. Further exploration is needed to identify the active constituents within these extracts, with the aim of harnessing *W. globosa* extract for potential applications in functional foods and nutritional cosmetics.

***Index Terms*—Antioxidant, Biological activity, Greenhouse, Serial exhaustive extraction, Wolffia globose**

I. INTRODUCTION

Wolffia globosa, commonly known as Duckweed or Watermeal, is a member of the Lemnaceae family. These tiny, floating green plants, part of the duckweed group, are esteemed for their high nutritional value and abundant antioxidant content[(1)] . *W. globosa* typically contains protein and fiber, alongside a significant amount of beta-carotene, and is also rich in essential amino acids[(2) (3)]. A recent study by Yaskolka Meir et al. (2021) delved into the composition of *Wolffia globosa* (Mankai), uncovering a diverse range of polyphenols, particularly flavonoids and phenolic acids, known for their potent antioxidant properties. These compounds play a pivotal role in mitigating cellular damage caused by free radicals, thus reducing the risk of infections and various diseases. Duckweed is a staple food in many Asian countries and is also utilized in the pharmaceutical industry regions of Korea, Japan, and China, employing it to address inflammation, urticaria, and various skin conditions in traditional medicine[(4, 5)]. However, despite its nutritional benefits, duckweed cultivation faces challenges in managing the quality and quantity of essential substances[(6)]. Cultivating Wolffia within enclosed systems, such as greenhouses, offers a solution by providing controlled environmental conditions, enabling better regulation of the quality and quantity of vital substances in raw materials. Factors such as storage processes and appropriate extraction techniques significantly impact the integrity of these substances.

This study investigates the extraction of essential compounds from *W. globosa* cultivated in a controlled

greenhouse environment using the Sequential Exhaustive Extraction method, a well-established technique for extracting metabolites from various plant materials. This progressive extraction process involves the use of solvents with increasing polarity, ranging from non-polar solvents like hexane to more polar ones like methanol, to effectively extract active compounds. Additionally, the study assesses the antioxidant activity, Total Phenolic Content, and Flavonoid Content of the extracted compounds. The findings may contribute to promoting duckweed cultivation in greenhouse systems and developing methods for extracting antioxidant-rich substances from *W. globosa* for potential applications in functional foods and nutritional cosmetics, thereby enhancing the value of *W. globosa*.

II. METHODOLOGY

Dried *Wolffia globosa* specimens cultivated in a Controlled Environment Greenhouse located in Khao Phra, Nakhon Nayok, Thailand, were utilized for extraction employing the Serial Exhaustive Extraction Method as adapted by Ngouana et al. (2021)[7]. Initially, 1 kg of dried *W.globosa* underwent fermentation in 80% ethanol, kept in darkness for 96 hours. Subsequently, the extracts were filtered and concentrated using a vacuum rotary evaporator. The resulting crude extract underwent further extraction with hexane in an ultrasonic cleaner. Residues remaining after each extraction step were then subjected to sequential extraction using solvents of increasing polarity, including dichloromethane, ethyl acetate, and water, respectively (Fig. 1), followed by evaporation using a vacuum rotary evaporator. The obtained extracts were stored at -20°C. Solvent extraction is contingent on the polarity of the solvent and the crucial substances present in plants. Significant compounds can dissolve in solvents with comparable polarity. In this study, the solvents utilized for the extraction of active components are detailed, as depicted in Table I[8].

TABLE I
Selection of Solvent for Active Constituent Extraction for Serial Exhaustive Extraction

Hexane	Dichloromethane	Ethyl acetate	Water
Non-polar	Semi-polar		Polar
Alkaloids	Terpenoids		Anthocyanins
Terpenoids	Flavonols		Starches
Coumarins	Alkaloids		Tannins
Fatty acids	Tannins		Saponins
Flavonoids	Polyphenols		Terpenoids
			Polypeptides
			Lectins

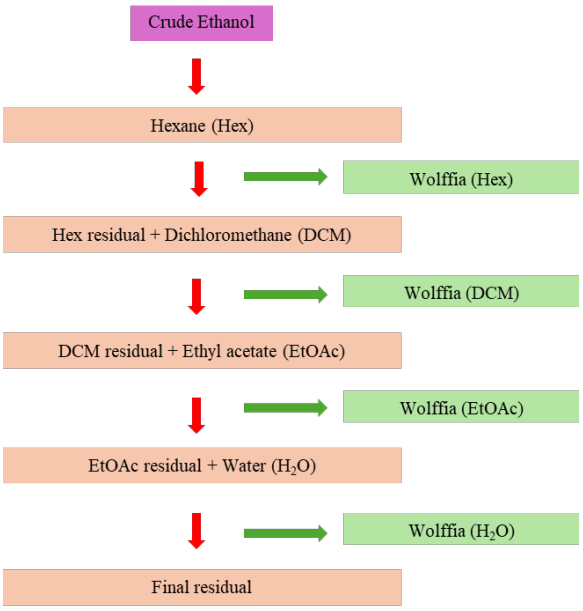


Fig. 1. Showing the serial exhaustive extraction method with solvents of increasing polarity from nonpolar (hexane) to more polar solvent (water) to extract of *Wolffia globosa*

The DPPH method was conducted following the procedure outlined by Panyatip et al. (2022)[9], with slight adjustments. In summary, 100 µL of extract *W.globosa* were mixed with 100 µL of 200 µM DPPH solution and incubated at room temperature for 30 minutes in darkness. Following incubation, the absorbance was measured at 517 nm using a microplate reader. The results were expressed as radical scavenging activity, defined with antioxidant capacity equivalent of Trolox.

The ABTS method was modifications according to the method of Daniel C Moreira, (2019)[10]. The solution $ABTS^{\cdot+}$ prepared by aqueous solution 7 mM ABTS with aqueous solution 245 mM APS. mixing and Incubate solution for 12–16 hr. at room temperature in the dark. The $ABTS^{\cdot+}$ solution was 100 times diluted before use. samples 10 µL was mixed with 190 µL $ABTS^{\cdot+}$ solution incubate it for 5 minutes in the dark. The absorbance was determined at 729 nm by a microplate reader. The final reports quantified the radical scavenging activity as Trolox Equivalent Antioxidant Capacity.

The FRAP method adopted according to the method of Sachett et al.(2023)[11]. The TPTZ reagent prepared by Acetate buffer (300 mM), TPTZ (2, 4, 6-tripyridyl-Striazine, 10 mM) in HCl (40 mM) and Iron (III) chloride hexahydrate (20 mM) were mixed in a ratio of 10:1:1 prepared freshly and warmed at 37°C before use. The samples for 5ul were added with 150 µl FRAP reagent then incubated at 37 °C for 15 minutes and measured the absorbance at 593 nm by a microplate reader. The final reports were expressed as the concentration of antioxidants having a ferric reducing ability equivalent to Trolox.

The Total Phenolic Content (TPC) in the extracts was assessed utilizing the Folin-Ciocalteu method, as outlined by Singleton and Rossi (1965)[12], with slight modifications. The extract or a gallic acid standard (25 µL) was mixed with 25 µL of Folin–Ciocalteu reagent (diluted 1:1 with distilled water). After a 5-minute incubation period, 125 µL of 2% sodium carbonate solution was added to the mixture and incubated at room temperature for 30 minutes in darkness. The absorbance of the resulting mixture was then measured at 725 nm using a microplate reader. The results were expressed as the concentration of milligram gallic acid equivalent per gram dry extract.

The Total Flavonoid Content (TFC) was determined following the method outlined by Horszwald and Andlauer (2011)[13], with minor adjustments. In summary, aliquots of the 50 µL extract or quercetin were mixed with 10 µL of 10% aluminium chloride and 10 µL of 1 M potassium acetate. Subsequently, 150 µL of solvent was added, and incubated at room temperature for 30 minutes in darkness. The absorbance of the resulting mixture was then measured at 415 nm using a microplate reader. The results were expressed as the concentration of milligram quercetin equivalent per gram of dry extract.

III. Results and Discussion

In the initial extraction process, employing 80% ethanol as the solvent yielded 21.467%. Subsequent Serial Exhaustive Extractions transitioned from non-polar to polar solvents, resulting in four extracts with yields ranging from 1.714% to 51.133% (refer to Table II). Notably, hexane (51.13%) and water (38.28%) extractions exhibited higher yields compared to dichloromethane (1.71%) and ethyl acetate (2.71%)

TABLE II
% Yield of extracts from *Wolffia globosa*.

Extract Fraction	Yield (%)
Crude Ethanol	21.47
Hexane	51.13
Dichloromethane	1.71
Ethyl acetate	2.71
Water	38.28

The findings of the DPPH assay conducted on the extracts are summarized in Table III. The analysis revealed that Crude Ethanol displayed radical scavenging activity equivalent to Trolox 435.84 ± 4.34 µmol/g, as did all Serial Exhaustive Extraction extracts. All fractions exhibited a radical scavenging effect equivalent to Trolox. Overall values ranged from 40.46 ± 0.16 to 378.21 ± 1.38 µmol/g, with the water extract showing the highest scavenging effect at 378.21 ± 1.38 TE µmol/g, indicating notable antioxidant activity.

The results of the ABTS assay, presented in Table III, demonstrated that Crude Ethanol exhibited radical scavenging capacity equivalent to Trolox 2.52 ± 0.002 µmol/g. Notably, all fraction showed radical scavenging effects as well. Overall values ranged from 0.43 ± 0.001 to 5.49 ± 0.006 TE µmol/g. Hexane and water extracts demonstrated the highest scavenging activities at 5.49 ± 0.006 TE µmol/g and 4.46 ± 0.031 TE µmol/g, respectively.

The evaluation of radical scavenging activity in extracts commonly employs the DPPH and ABTS methods. DPPH, a stable free radical soluble in organic solvents but not in water, assesses antioxidant capability by donating hydrogen or electrons to free radicals. Structurally akin to inert radicals found in the body, DPPH contains a benzene ring and a nitrile group that hinder electron pairing. DPPH often exhibits slower reactions with antioxidants possessing bulky structures. In contrast, ABTS, with higher water solubility, necessitates interactions with substances possessing potent oxidative abilities to generate free radicals. Like DPPH, ABTS evaluates antioxidant potency by supplying hydrogen or electrons to free radicals but has a wider pH range for reaction and typically reacts more rapidly than DPPH.[14, 15]

TABLE III
ANTIOXIDANT ACTIVITY OF *Wolffia globose*

Extract	DPPH	ABTS	FRAP
	TE (µmol/g)	TE (µmol/g)	TE (µmol/g)
Wolffia Crude Ethanol *	435.84 ±4.34	2.52 ±0.002	8.49 ±0.05
Wolffia (Hex)*	40.46 ±0.16	5.49 ±0.006	3.37 ±0.02
Wolffia (DCM)*	88.58 ±0.32	0.43 ±0.001	9.19 ±0.04
Wolffia (EtOAc)*	118.31 ±1.63	1.31 ±0.005	14.48 ±0.04
Wolffia (H_2O)*	378.21 ±1.38	4.46 ±0.031	10.26 ±0.10
W. arrhiza[16]*	62.89 ±0.58	92.62 ±0.62	85.52 ±1.94
Apple[17]	22.14 ±1.2	32.85 ±1.5	26.98 ±0.9
Common bean[18]	13.17 ± 1.27	39.77 ± 3.78	0.44 ± 0.02

*Values are means of three samples ± SD
Wolffia (Hex): hexane extract of *W. globosa*; Wolffia (DCM): Dichloromethane of W. globose; Wolffia (EtOAc): Ethyl acetate of *W. globosa*; Wolffia (H2O): Water of *W. globosa*

The Ferric Reducing Antioxidant Power (FRAP) assay evaluates antioxidants' ability to reduce iron. The reaction involves the conversion of ferric iron (Fe^{3+}) complexes to ferrous (Fe^{2+})[14, 15], with higher absorbance indicating increased Fe^{2+} formation and thus, higher Fe^{3+} ion reduction. Quantitative FRAP analysis (Table III) indicated that Crude Ethanol exhibited Fe^{3+} reducing activity equivalent to Trolox at 16.23 ± 0.10 µmol/g, along with all Serial Exhaustive Extraction fractions displaying Fe^{3+} removal effects. Overall values ranged from 3.37 ± 0.02 to 14.48 ± 0.04 TE µmol/g. Ethyl acetate and water extracts showed notable activity,

with the highest Fe^{3+} removal at 14.48 ± 0.04 TE µmol/g and 10.26 ± 0.10 TE µmol/g, respectively.

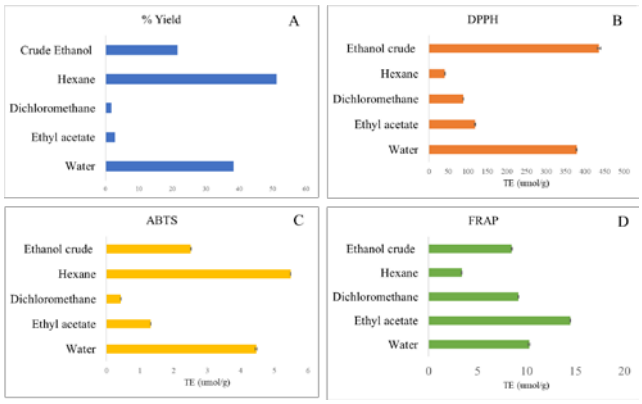


Fig. 2. Solvent effects on the yield and efficiency of antioxidants of *W. globosa* by (A) %Yield; (B) DPPH; (C) ABTS; (D) FRAP. Results expressed as Values are means of three ± SD.

TABLE IIII
Total phenolics content and Total flavonoids content of *Wolffia globosa.*

Extract	Total phenolic	Total flavonoid
	GAE (mg/g)	QE (mg/g)
Wolffia Crude Ethanol *	47.11 ±0.24	175.95 ±1.12
Wolffia (Hex)*	45.67 ±1.27	NA
Wolffia (DCM)*	45.16 ±0.73	77.06 ±0.05
Wolffia (EtOAc)*	52.77 ±0.43	49.53 ±0.37
Wolffia (H_2O)*	32.91 ±1.65	46.59 ±0.39
Spirulina**[19]	15.8 ± 0.9	-

*Values are means of three samples ± SD
**Values are mean of five samples ± SD
Wolffia (Hex): hexane extract of *W. globosa*; Wolffia (DCM): Dichloromethane of *W. globosa*; Wolffia (EtOAc): Ethyl acetate of *W. globosa* ; Wolffia (H_2O): Water of *W. globosa*

The total phenolic content test results of all Fraction extracts ranged from 52.77 ± 0.43 to 32.91 ± 1.65 GAE mg/g, with the highest values observed in Ethyl acetate (52.77 ± 0.43 GAE mg/g). Similarly, the total flavonoid content varied between 77.06 ± 0.05 and 46.59 ± 0.39 QE mg/g, with the highest value found in Dichloromethane (77.06 ± 0.05 GAE mg/g). Notably, flavonoid content was not detected in the Hexane extract, as detailed in Table IIII. Phenolic compounds, such as phenolic acid and flavonoids, are renowned for their antioxidant properties. These compounds act as potent free radical scavengers, capable of trapping metal ions within their molecular structure. They function both as electron donors and hydrogen donors, effectively neutralizing active oxygen species. Consequently, phenolic compounds play a pivotal role as antioxidants in plants.[20, 21]

Examination of the antioxidant activity and phenolic compound levels revealed that utilizing a highly polar solvent for the extraction of *W. globosa* from controlled environment resulted in heightened antioxidant efficacy, This discovery aligns with previous research conducted by Supannee et al[22]. Antioxidant activity was notably detected in *W. globosa* extracts obtained from polar solvents such as ethyl acetate and methanol, which corresponds with the observed Total Phenolic content and Flavonoid content in extracts from *W. globosa* extracted using polar solvents. Conversely, in solvent extracts with low polarity, although elevated levels of Total Phenolic and Flavonoid were detected, the antioxidant activity was comparatively lower. This disparity may be attributed to the types of metabolites present and the synergistic or antagonistic interactions among these metabolites within the extracts. Medicinal plants typically harbor a multitude of compounds that work in concert to exhibit specific functions[23, 24], as elucidated by Anat Yaskolka Meir et al. [25], who identified 198 distinct phenolic metabolites in *W. globosa*. Further exploration and development are essential for the comprehensive characterization of pivotal active metabolites from *W. globosa.*

The Serial Exhaustive Extraction method has garnered attention for its efficacy in enriching active metabolite extraction from diverse plant materials. This method involves a continuous extraction process with solvents of increasing polarity, ranging from hexane to water, thereby facilitating the extraction of a broad spectrum of phytochemicals with varying polarities[26, 27]. Given the diverse chemical composition of phytochemicals, no single solvent is capable of extracting them all[28], necessitating the use of solvents with differing polarities. Overall, the results underscore the significant impact of Serial Exhaustive Extraction on phenolic and flavonoid compound levels. When compared to spirulina, an algae renowned for its high antioxidant properties and widespread use in dietary supplements, *W. globosa* boasts a higher total phenolic content, as well as the antioxidant potential of *W. globosa* extracts, as evidenced by DPPH, ABTS, and FRAP tests. While the radical scavenging activity with ABTS and the Fe^{3+} reducing ability of Serial Exhaustive Extraction from *W. globosa* were lower compared to *Wolffia arrhiza*, the radical scavenging activity with DPPH was more effective than in *W. arrhiza* and other plants such as beans and apples. Hence, this method holds promise in validating preliminary extraction techniques for *W. globosa*. Further refinement is imperative for quality control and precise identification of *W. globosa* extracts.

IV. Conclusion

This study investigated the antioxidant potential of *Wolffia globosa* extracts cultivated in a greenhouse environment using the Serial Exhaustive Extraction method. Through this exploration, notable variations in extraction yield and antioxidant activity, particularly related to phenolic

and flavonoid content, were observed across the extracts. All four extract species exhibited robust free radical scavenging and ferric reducing antioxidant power, with particularly noteworthy results observed in extracts derived from polar solvents. This initial inquiry marks a significant stride in addressing the existing gap in knowledge concerning the extraction and isolation techniques for active compounds from *W. globosa* via Serial Exhaustive Extraction. It underscores the critical importance of selecting appropriate solvents and extraction methods to fully harness the antioxidant capabilities of *W. globosa* and emphasizes the need for further investigation to pinpoint and characterize the phytochemical substances involved. In the subsequent phase, research will focus on identifying potent active ingredients within *W. globosa*. Investigating its mechanism of action and delving into additional biological functions, including its potential anti-tyrosinase and anti-inflammatory properties, is crucial for utilizing *W. globosa* in the development of health-centric food products and nutritional cosmetics. Furthermore, these findings will guide strategies for commercially cultivating *W. globosa* under controlled conditions, optimizing the extraction process of its active compounds, thereby enhancing the agricultural value of *W. globosa*.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to the Center for Excellence in Plant and Herbal Innovation Research (CEPHIR), Strategic Wisdom and Research Institute of Srinakharinwirot University, Reinventing University Projects 2023 in Area-Based and Community-Engaged Research under Thailand's Ministry of Higher Education, Science, Research and Innovation (MHESI), Baan Suan Noo Dee, and the Plant and Herbal Innovation Community Enterprise for their invaluable support, information, and materials that were instrumental to the success of this research. Their generous contributions and collaboration have been crucial in enabling the completion of this study.

REFERENCES

[1] sukhoom rowchai SS. Study on factors effecting growth of wolffia (Wolffia arrhiza (L.) Wimm.). Thai fisheries gazette [Internet]. 2007; 60:[405-13 pp.].
[2] Appenroth K-J, Sree KS, Böhm V, Hammann S, Vetter W, Leiterer M, Jahreis G. Nutritional value of duckweeds (Lemnaceae) as human food. Food Chemistry. 2017;217:266-73.
[3] Kansinee Panwanitdumrong SR. Experimental study on culture of wolffia (Wolffia arrhiza (L.) Wimm.) and it's use for gold fish color quality improvement. Proceedings of 47th Kasetsart University Annual Conference: Fisheries [Internet]. 2009; 47.
[4] Hui T, Zhang Y, Jia R, Hu Y, Wang W, Wang Y, et al. Metabolomic analysis reveals responses of Spirodela polyrhiza L. to salt stress. Journal of Plant Interactions. 2023;18(1):2210163.
[5] Pagliuso D, Palacios Jara CE, Grandis A, Lam E, Pena Ferreira MJ, Buckeridge MS. Flavonoids from duckweeds: potential applications in the human diet. RSC Advances. 2020;10(73):44981-8.
[6] Ruekaewma N, Piyatiratitivorakul S, Powtongsook S. Culture system for Wolffia globosa L. (Lemnaceae) for hygiene human food. Songklanakarin Journal of Science and Technology. 2015;37:575-80.
[7] Ngouana V, Zeuko'o Menkem E, Youmbi DY, Yimgang LV, Toghueo RMK, Boyom FF. Serial Exhaustive Extraction Revealed Antimicrobial and Antioxidant Properties of Platycerium stemaria (Beauv) Desv. BioMed Research International. 2021;2021:1584141.
[8] Kasiramar G, K G. Significant role of soxhlet extraction process in phytochemical research. 2019;7:43-7.
[9] Panyatip P, Padumanonda T, Yongram C, Kasikorn T, Sungthong B, Puthongking P. Impact of Tea Processing on Tryptophan, Melatonin, Phenolic and Flavonoid Contents in Mulberry (Morus alba L.) Leaves: Quantitative Analysis by LC-MS/MS. Molecules [Internet]. 2022; 27(15).
[10] Daniel C Moreira. ABTS decolorization assay – in vitro antioxidant capacity 2019 [Available from: https://www.protocols.io/view/abts-decolorization-assay-in-vitro-antioxidant-cap-14egnxk86l5d/v1.
[11] Piato A. Antioxidant activity by FRAP assay: in vitro protocol 2023 [Available from: https://www.protocols.io/view/antioxidant-activity-by-frap-assay-in-vitro-protoc-j8nlk4n35g5r/v1?step=1.2.
[12] Singleton VL, Joseph A. Rossi, Jr. Colorimetry of Total Phenolics with Phosphomolybdic-Phosphotungstic Acid Reagents. American Journal of Enology and Viticulture. 1965;16(3):144.
[13] Horszwald A, Andlauer W. Characterisation of bioactive compounds in berry juices by traditional photometric and modern microplate methods. Journal of Berry Research. 2011;1:189-99.
[14] Alam MN, Bristi NJ, Rafiquzzaman M. Review on in vivo and in vitro methods evaluation of antioxidant activity. Saudi Pharmaceutical Journal. 2013;21(2):143-52.
[15] Mendonça JDS, Guimarães RCA, Zorgetto-Pinheiro VA, Fernandes CDP, Marcelino G, Bogo D, et al. Natural Antioxidant Evaluation: A Review of Detection Methods. Molecules. 2022;27(11).
[16] Hu Z, Fang Y, Yi Z, Tian X, Li J, Jin Y, et al. Determining the nutritional value and antioxidant capacity of duckweed (Wolffia arrhiza) under artificial conditions. LWT. 2022;153:112477.
[17] Kotha RR, Tareq FS, Yildiz E, Luthria DL. Oxidative Stress and Antioxidants—A Critical Review on In Vitro Antioxidant Assays. Antioxidants. 2022;11(12):2388.
[18] Lopes C, Akel Ferruccio C, de Albuquerque Sales AC, Tavares GM, de Castro RJS. Effects of processing technologies on the antioxidant properties of common bean (Phaseolus vulgaris L.) and lentil (Lens culinaris) proteins and their hydrolysates. Food Research International. 2023;172:113190.

[19] Shukla V, Vashistha M, Singh SN. Evaluation of antioxidant profile and activity of amalaki (Emblica officinalis), spirulina and wheat grass. Indian J Clin Biochem. 2009;24(1):70-5.
[20] Rice-Evans CA, Miller NJ. Antioxidant activities of flavonoids as bioactive components of food. Biochem Soc Trans. 1996;24(3):790-5.
[21] Chiang Mai University. Free radical [Available from: https://archive.lib.cmu.ac.th/full/T/2551/biol0451tp_ch2.pdf.
[22] Tipnee S, Jutiviboonsuk A, Wongtrakul P. The Bioactivity Study of Active Compounds in Wolffia globosa Extract for an Alternative Source of Bioactive Substances. Cosmetics. 2017;4(4):53.
[23] Liu X, Vrieling K, Klinkhamer PGL. Interactions between Plant Metabolites Affect Herbivores: A Study with Pyrrolizidine Alkaloids and Chlorogenic Acid. Frontiers in Plant Science. 2017;8.
[24] Boik J, Kirakosyan A, Kaufman PB, Seymour EM, Spelman K. Interactions of Bioactive Plant Metabolites: Synergism, Antagonism, and Additivity. In: Kirakosyan A, Kaufman PB, editors. Recent Advances in Plant Biotechnology. Boston, MA: Springer US; 2009. p. 213-30.
[25] Yaskolka Meir A, Tuohy K, Bergen M, Krajmalnik-Brown R, Heinig U, Zelicha H, et al. The Metabolomic-Gut-Clinical Axis of Mankai Plant-Derived Dietary Polyphenols. Nutrients. 2021;13:1866.
[26] Mokgotho MP, Gololo SS, Masoko P, Mdee LK, Mbazima V, Shai LJ, et al. Isolation and Chemical Structural Characterisation of a Compound with Antioxidant Activity from the Roots of Senna italica. Evidence-Based Complementary and Alternative Medicine. 2013;2013:519174.
[27] Nawaz H, Shad M, Rehman N, Andaleeb H, Ullah N. Effect of solvent polarity on extraction yield and antioxidant properties of phytochemicals from bean (Phaseolus vulgaris) seeds. Brazilian Journal of Pharmaceutical Sciences. 2020;56.
[28] Wakeel A, Jan SA, Ullah I, Shinwari ZK, Xu M. Solvent polarity mediates phytochemical yield and antioxidant capacity of Isatis tinctoria. PeerJ. 2019;7:e7857.

Microwave Pretreatment Of Sugarcane Bagasse For Bioethanol Production By *Saccharomyces cerevisiae* TISTR 5088 and *Pichia stipitis* TISTR 5806

Engkarat Kingkaew[1, a)], Onwanya Chaimanat[1], Siripak Wongchachom[1] and Duangjai Ochaikul[1]

[1]Department of Biology, School of Science, King Mongkut's Institute of Technology Ladkrabang, Bangkok, Thailand 10520

[a] Corresponding author: engkarat.ki@kmitl.ac.th

Abstract. Currently, the predominant fuel sources utilized in daily life are often derived from pollutants, prompting the development of alternative energy sources, particularly biofuels sourced from agricultural residues. Sugarcane, in particular, is a viable option for conversion into ethanol for fermentation processes. This study employed microwave-assisted pretreatment and hydrolysis with a sodium hydroxide solution. The liquid portion of the hydrolysate obtained from microwave-assisted pretreatment with a sodium hydroxide solution at a concentration of 8% (w/v) and a power of 800 Watts for 2 minutes yielded a maximum sugar concentration of 18.69 ± 0.05 g/L. Subsequently, this hydrolysate liquid fraction underwent enzymatic saccharification for ethanol production using *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806. *S. cerevisiae* TISTR 5088 yielded a maximum ethanol concentration of 3.30 ± 0.08 g/L at 150 rpm and 30°C for 48 hours, whereas *P. stipitis* TISTR 5806 yielded 2.68 ± 0.06 g/L at 150 rpm and 30°C for 48 hours. When these strains were cocultured at a ratio of 2:1 (*S. cerevisiae* TISTR 5088: *P. stipitis* TISTR 5806), the ethanol concentration reached 5.34 ± 0.01 g/L at 48 hours. Thus, coculturing in a 2:1 ratio during fermentation significantly increased ethanol production compared to fermentations with individual strains, demonstrating statistically significant improvements in ethanol yield.

***Index Terms*—Bagasse, Bioethanol, *Saccharomyces cerevisiae*, *Pichia stipitis*, Fermentation**

I. INTRODUCTION

Biobased energy production garners significant attention as an environmentally friendly, renewable, and nonpolluting approach with wide-ranging social, political, and economic implications [1]. One solution in the mitigation portfolio involves utilizing biomass as an alternative fuel or feedstock, displacing the use of fossil fuels and fossil-based products. If sustainably produced, biomass results in an overall emissions reduction [2]. Bioethanol, as a viable alternative to liquid fuels, can either replace petroleum or be blended with it to enhance the fuel octane value without the need for pollutant additives, thus reducing petroleum dependency [3]. However, the development of bioethanol production from biomass encounters challenges, primarily due to the resistance of the biomass to hydrolysis. Various pretreatment methods have been devised to improve the rate and yield of enzymatic saccharification of the recalcitrant structures of lignocellulosic materials [4]. These methods include mechanical, pyrolysis, steam explosion, ammonia fiber explosion (AFEX), carbon dioxide explosion, ozonolysis, dilute-acid hydrolysis, alkaline hydrolysis, the organo-solvent process, biological, pulsed electrical field, microwave, ultrasound, ionic liquids, and others [5-7]. The choice of pretreatment significantly affects subsequent stages of bioethanol production, such as enzymatic hydrolysis and fermentation. Alterations in cellulose and hemicellulose availability during enzymatic hydrolysis and the formation of inhibitors during pretreatment can impact these stages [5]. Therefore, understanding the optimal pretreatment method and its mechanisms is crucial for establishing a successful and cost-effective bioethanol production process. For instance, microwave heating offers homogeneous radiation absorption and rapid heating, which can enhance pretreatment efficiency [4]. Singh, Tuteja [8] demonstrated the influence of microwave heating on the alkali pretreatment of rice straw and hulls, highlighting factors such as alkali concentration, irradiation time, and substrate concentration [8]. Similarly, microwave intensity, irradiation time, and substrate concentration were found to significantly affect the enzymatic saccharification of rice straw [9]. Various studies have investigated the effects of dilute acid or microwave pretreatment on sugarcane bagasse structure by employing techniques such as X-ray diffraction (XRD), scanning electron microscopy (SEM), Fourier transform infrared (FTIR), and thermogravimetric analysis (TGA) to understand its physical and chemical characteristics [10, 11]. While microwave pretreatment shows promise, the mechanisms underlying sugar release and inhibitor formation during pretreatment and enzymatic hydrolysis remain unclear, necessitating further research [12]. On the other hand, dilute-acid pretreatment can effectively remove and hydrolyze hemicellulose sugars, particularly xylose, from biomass, facilitating complete cellulose hydrolysis [6].

Pichia stipitis is a significant yeast capable of consuming pentose under oxygen-limited conditions for ethanol production, unlike common ethanol-producing

microorganisms such as *Saccharomyces cerevisiae* or *Zymomonas mobilis*, which ferment hexoses anaerobically. This distinction is attributed to the metabolic pathway of pentose consumption via the pentose phosphate pathway [13]. One major challenge in bioethanol production from dilute-acid pretreated hydrolysate is the formation of inhibitors that impede microbial growth or ethanol production. Detoxification methods are typically employed to mitigate these inhibitors, albeit at operational and investment costs [14]. Hence, alternative pretreatment methods, especially dilute-alkali pretreatment, which results in the production of no inhibitors or the use of inhibitor-tolerant microorganisms, are preferred.

In this context, sugarcane bagasse pretreatment was optimized using microwave-assisted dilute-alkali methods followed by enzymatic hydrolysis. The reducing sugar content was measured during pretreatment. Subsequently, the ratio of *S. cerevisiae* to *P. stipitis* was investigated. Ethanol production from hemicellulose hydrolysate under optimal conditions was then investigated using single culture and mixed cultues to assess the impact of pretreatment on ethanol yield.

II. Methodology

1. Sample Preparation

Bagasse was obtained from Chonburi Province, Thailand. Bagasse was cut into small pieces approximately 1-2 cm (in size). These pieces were then oven-dried at 70°C for 1-2 days. Subsequently, they were ground into a fine powder using a grinder and sifted through a sieve to achieve a particle size of 850 µm. The powder was then stored in plastic bags at 4°C.

2. Study of microwave power combined with sodium hydroxide (NaOH) concentration for bagasse pretreatment

Ten grams of bagasse obtained from dry grinding was mixed with 100 ml of 2% (w/v) NaOH solution. The mixture was heated using microwave wattages of 100, 200, 400, and 800 watts for 2 minutes. Subsequently, the liquid and solid fractions were separated using a double-layered fine white cloth. The solid fraction was washed with clean water until the final wash had a pH of 7.00 ± 0.10. The washed solid fraction was then oven-dried at 70°C for 48 hours. It was then enzymatically hydrolyzed using ACCELLERASE 1500 enzyme at a concentration of 1 ml/$g_{bagasse}$. The hydrolysate was incubated at 50°C for 48 hours, following the method described by [15]. After completion, the hydrolysate was centrifuged at 4,500 rpm for 20 minutes. The resulting liquid fraction was analyzed for reducing sugar content using the DNS method [16] to select the microwave wattage yielding the highest reducing sugar content for subsequent experiments.

3. Study of the optimal duration for microwave-assisted NaOH pretreatment

Ten grams of bagasse obtained from dry grinding was mixed with 100 ml of 2% (w/v) NaOH solution. The mixture was then heated using microwave power obtained from the investigation in section 2. The time variation in microwave usage was measured for 1, 2, 3, 4, and 5 minutes. The liquid and solid fractions were separated using a double-layered fine white cloth. The solid fraction was washed with clean water until the final wash had a pH of 7.00 ± 0.10. It was then oven-dried at 70°C for 48 hours. Subsequently, enzymatic hydrolysis was performed using the ACCELLERASE 1500 enzyme. The mixture was then centrifuged at 4,500 rpm for 20 minutes. The reducing sugar content of the liquid fraction was analyzed using the DNS method. A suitable duration for yielding a high reducing sugar content was selected for further experiments.

4. Study on the optimal concentration of NaOH combined with microwave pretreatment

Ten grams of dried ground bagasse was mixed with a sodium hydroxide solution at concentrations of 1, 2, 4, 6, 8, and 10% (w/v) in a volume of 100 ml. The mixture was heated using microwave power obtained from the investigation in section 2, and the appropriate duration was determined from the study in section 3. The liquid and solid fractions were separated using a double-layered fine white cloth. The solid fraction was washed with clean water until the final wash had a neutral pH of 7.00 ± 0.10. Subsequently, the sample was oven-dried at 70°C for 48 hours. The dried samples were weighed and then subjected to enzymatic hydrolysis using the ACCELLERASE 1500 enzyme. The mixture was then centrifuged at 4,500 rpm for 20 minutes. The resulting liquid fraction was analyzed for reducing sugar content using the DNS method. A suitable concentration of sodium hydroxide solution yielding a high reducing sugar content was selected for further experiments.

5. Preparation of Inoculum for Ethanol Fermentation

Cultures of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 were prepared for ethanol fermentation. For *S. cerevisiae* TISTR 5088, the culture was grown on YPD media (composed of 10 g/L yeast extract, 20 g/L peptone, 20 g/L glucose, and 25 g/L agar, pH 5.0). Similarly, for *P. stipitis* TISTR 5806, the culture was grown on MGYP media (composed of 5 g/L malt extract, 5 g/L yeast extract, 20 g/L peptone, 5 g/L glucose, 30 g/L xylose, and 25 g/L agar, pH 5.0).

For both cultures, the inoculum was prepared by inoculating the respective cultures into their respective broths. The cultures were aerobically incubated on a shaker at 150 rpm at 30°C for 24 hours. The optical density (OD) at 600 nm was measured and adjusted to 0.5 for use as an inoculum for ethanol fermentation.

6. Bioethanol fermentation

6.1 Bioethanol production using *S. cerevisiae* TISTR 5088

The hydrolysate obtained from enzymatic hydrolysis, adjusted to pH 5.5, was transferred into a 250 mL flask with a nutrient volume of 100 mL. It was then sterilized by autoclaving at 110°C for 20 minutes [17]. After cooling, an inoculum of *S. cerevisiae* TISTR 5088 prepared as described in section 3.5.6.1 was added at a concentration of 10% (v/v). Fermentation was conducted aerobically on a shaker at 150 rpm and 30°C for 60 hours. Samples were withdrawn every 12 hours for analysis. Subsequently, the samples were centrifuged at 4,500 rpm for 10 minutes, and the reducing sugars in the supernatant were analyzed using the DNS method. The ethanol

content was analyzed using gas chromatography (GC), along with pH measurements.

6.2 Bioethanol production using *P. stipitis* TISTR 5806

The fermentation process was conducted similarly to that described in section 6.1, utilizing *P. stipitis* TISTR 5806. Fermentation was carried out aerobically on a shaker at 150 rpm and 30°C for 60 hours. Samples were collected every 12 hours for analysis, and the analysis procedures were identical to those described above.

6.3 Bioethanol production using *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at different ratios

The hydrolysate obtained from enzymatic hydrolysis, adjusted to pH 5.5, was transferred into a 250 mL flask with a nutrient volume of 100 ml. It was then sterilized by autoclaving at 110°C for 20 minutes [18]. After cooling, inoculum of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 prepared as described in section 5 were added at various ratios (*S. cerevisiae* TISTR 5088:*P. stipitis* TISTR 5806) of 4:1, 2:1, 1:1, 1:2, and 1:4, with a final inoculum concentration of 10% (v/v). Fermentation was conducted aerobically on a shaker at 150 rpm and 30°C for 60 hours. Samples were collected every 12 hours for analysis. Subsequently, the samples were centrifuged at 4,500 rpm for 10 minutes, and the supernatant was analyzed for pH, reducing sugars and ethanol content using a pH meter, DNS method and gas chromatography (GC).

7. Study of the kinetic parameters of the fermentation process by *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806

Ethanol fermentation by *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 was conducted, and samples were taken at various time intervals for analysis. The samples were analyzed for reducing sugars, ethanol content, and pH to calculate the kinetic parameters of the fermentation process. The yield coefficient of the product per substrate ($Y_{P/S}$) and the productivity of the product were calculated in $g/g_{substrate}$ and g/L/h, respectively.

8. Ethanol quantification

Based on Jekel et al. (2005), the ethanol content of the obtained clear fraction was analyzed using a GC (Shimadzu, Japan) gas chromatography system with a DB-1 column. Helium gas was used as the carrier gas, with a column pressure of 30 kPa and an internal column temperature of 60°C. Detection was performed using a flame ionization detector (FID) at 180°C. The injector temperature was maintained at 150°C.

9. Statistical analysis

All results are presented as the mean ± standard deviation (SD). Statistical comparisons were conducted using one-way analysis of variance (ANOVA) followed by Duncan's new multiple range test (DMRT) at a significance level of $p < 0.05$. All the statistical analyses were performed using the SPSS (version 22.0) program for Windows.

III. RESULTS AND DISCUSSION

1. Effect of microwave power

The results showed that 800 Watts (W) yielded the highest reducing sugar concentration (14.96 ± 0.10 g/L), which was significantly different from that at lower power levels. This finding aligns with prior research indicating that the sugar content increases with increasing microwave power. However, excessive heat led to sugar degradation. Therefore, 800 watts was determined to be the appropriate power level for further investigation [19]. The study of Anoopkumar, Reshmy [20] revealed that microwave-assisted alkali (NaOH) pretreatment at 600 W resulted in a high quantity of monomeric sugar, specifically 0.665 g, extracted from sugarcane bagasse as the substrate. The microwave power and reducing sugar concentration are shown in Table 1.

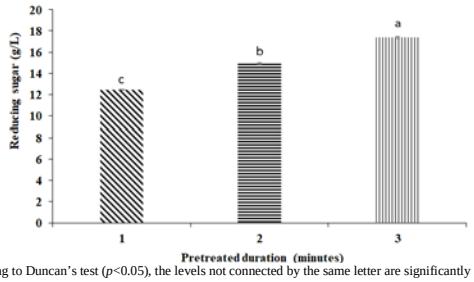
Table 1 Reducing sugar concentration from pretreated sugarcane bagasse using 2% (w/v) NaOH and different microwave powers for 2 minutes, followed by enzymatic hydrolysis with the ACCELLERASE 1500 enzyme at 50°C for 48 hours.

Microwave power (Watts; W)	Reducing sugar (g/L)
100	8.18 ± 0.12^{d}
200	10.91 ± 0.06^{c}
400	12.42 ± 0.08^{b}
800	14.96 ± 0.10^{a}

According to Duncan's test ($p<0.05$), the levels not connected by the same letter are significantly different.

2. Effect of the microwave pretreatment duration

From the study of the optimal microwave power with 2% NaOH combined with heating using an 800-watt microwave for varying durations of 1, 2, and 3 minutes, the resulting samples of treated bagasse were enzymatically hydrolyzed as previously explained, with experiments conducted in triplicate. Subsequent analysis of the reducing sugar content from enzymatic hydrolysis revealed that at 1 minute, the minimum amount of reducing sugar obtained was 12.48 ± 0.07 g/L; at 2 minutes, it was 14.95 ± 0.16 g/L; and at 3 minutes, it was 17.38 ± 0.14 g/L, as shown in Fig. 1. However, it was observed that there was charring of the sugarcane bagasse and expulsion of material from the flask during microwave irradiation. Consequently, for subsequent studies, the use of 2% (w/v) NaOH at 800 W for 2 minutes was selected, which yielded slightly lower levels of reducing sugars than did 3 minutes. According to Zhu, Rezende [19], optimizing the duration for treating lignocellulosic materials with microwave-assisted alkali solutions is vital. It enhances sugar release from cellulose and hemicellulose, boosting yields. Furthermore, a microwave power of 832.9 W, NaOH concentration of 2.7% (w/v), and a reaction time of 8.9 minutes exhibited the great. However, excessive duration risks charring and material degradation.



According to Duncan's test (p<0.05), the levels not connected by the same letter are significantly different.

Fig. 1. The effect of microwave pretreatment duration and the amount of reducing sugars.

3. Effect of the concentration of NaOH combined with microwave pretreatment

This study examined the suitable concentration of NaOH in conjunction with microwave treatment for conditioning sugarcane bagasse. Bagasse samples treated with NaOH at concentrations ranging from 1% to 10% (w/v) were subjected to enzymatic hydrolysis with the ACCELLERASE 1500 enzyme. The results showed that the highest reducing sugar content was obtained at a 10% (w/v) NaOH concentration (Table 2), with no significant difference compared to that at 8% (w/v) NaOH. Therefore, 8% (w/v) NaOH was chosen for further studies. Previous research by Xu, Zhang [21] and Zhu, Rezende [19] supported and emphasized the role of alkaline solutions in breaking down lignocellulosic materials, enhancing sugar release for efficient enzymatic digestion. Higher concentrations of NaOH were found to increase sugar yield and effectively remove lignin, optimizing enzyme performance during hydrolysis.

Table 2 Effect of NaOH concentration with microwave pretreatment.

NaOH concentration (%w/v)	Reducing sugar (g/L)
1.0	10.92 ± 0.05^{e}
2.0	14.79 ± 0.33^{d}
4.0	16.82 ± 0.05^{c}
6.0	17.81 ± 0.16^{b}
8.0	18.69 ± 0.05^{a}
10.0	18.88 ± 0.31^{a}

According to Duncan's test (p<0.05), the levels not connected by the same letter are significantly different.

4. Bioethanol production using *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806

The hydrolyzed bagasse fermented with *S. cerevisiae* TISTR 5088 initially contained 17.86 ± 0.06 g/L of reducing sugars, which rapidly decreased until 36 hours, followed by a slower decline, reaching 3.51 ± 0.02 g/L at the end of fermentation. Ethanol production increased significantly between 24 and 48 hours, reaching a peak of 3.30 ± 0.08 g/L at 48 hours. Similarly, hydrolyzed bagasse fermented with *P. stipitis* TISTR 5806 initially contained 17.15 ± 1.34 g/L reducing sugars, which decreased gradually in the first 12 hours, followed by a rapid decrease until 36 hours, and then a slight decrease until the end of fermentation, with 4.87 ± 0.05 g/L reducing sugars remaining. Ethanol production increased rapidly between 24 and 48 hours, peaking at 2.68 ± 0.06 g/L at 48 hours, followed by a slight decrease. A comparison of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 for ethanol production from hydrolyzed and enzymatically treated bagasse revealed that *S. cerevisiae* TISTR 5088 produced more ethanol than did *P. stipitis* TISTR 5806, especially at 48 hours. Previous studies by Tesfaw and Assefa [22] indicated that *S. cerevisiae* is an efficient yeast strain for ethanol production, while *P. stipitis* can utilize both C5 and C6 but has lower tolerance to high ethanol concentrations and various toxins. Additionally, research by Scanes, Hohmann [23] showed that prolonged fermentation led to increased glycerol production and yeast autolysis, releasing amino acids, which elevated the pH. The bioethanol production using *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 is shown in Fig. 2.

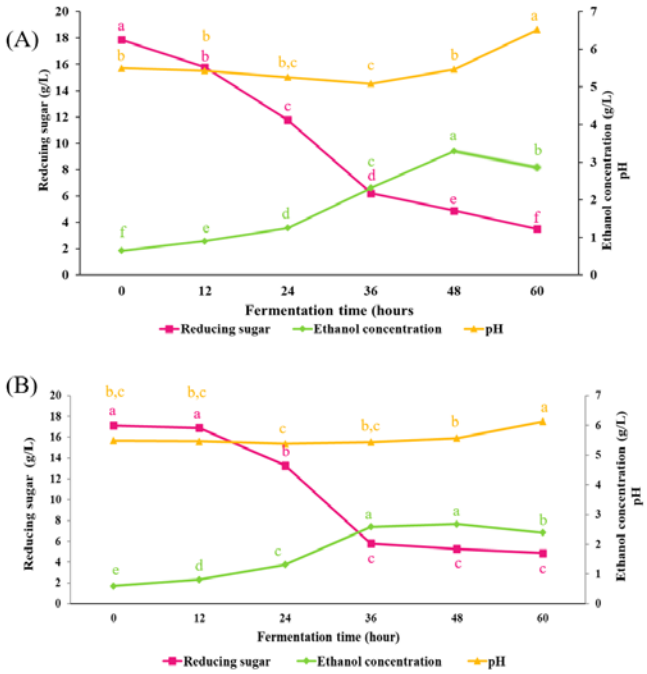


Fig. 2. Bioethanol production using *S. cerevisiae* TISTR 5088 (A) and *P. stipitis* TISTR 5086 (B). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article).

5. Bioethanol production using *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at different ratios

When a mixed culture of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 was used at ratios of 1:1, 1:2, 1:4, 2:1, and 4:1, the ethanol concentrations were 3.56 ± 0.13 g/L, 3.13 ± 0.07 g/L, 3.91 ± 0.17 g/L, 5.34 ± 0.01 g/L, and 4.68 ± 0.12 g/L, respectively. Statistical analysis revealed that the 2:1 ratio produced significantly higher ethanol concentrations than the other fermentation conditions, as shown in Table 3.

Using a mixed culture of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at a 2:1 ratio proved to be more efficient for ethanol production than other mixing ratios and single-strain fermentations. However, the ethanol yield at a 4:1 ratio was lower than that at a 2:1 ratio, possibly due to the higher abundance of *S. cerevisiae*, which was up to four times greater than that of *P. stipitis*. During fermentation, *S. cerevisiae* produced inhibitory compounds, which slowed the growth of *P. stipitis*, leading to reduced ethanol production. This finding is consistent with the research of Patel, Chapla [24], who reported that co-fermentation involving yeast strains capable of utilizing C6, such as *Saccharomyces cerevisiae*, produced growth-inhibiting compounds during fermentation, resulting in lower

efficiency than that of *Pichia stipitis*, which was less affected by growth inhibition and thus yielded less ethanol. Sjulander and Kikas [25] reported that the inhibitors in the bioethanol fermentation from the lignocellulose, such as sugar-derived aldehydes, aromatic compounds, metals, short-chain organic acids and aldehydes, and other inhibitors.

Table 3 Comparison of the maximum ethanol yields from the fermentation of hydrolyzed sugarcane bagasse obtained by pretreatment with 8% NaOH, microwave irradiation and enzymatic hydrolysis using ACCELLERASE 1500, using both single strains and mixed cultures of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at different ratios.

Fermented conditions(s)	Ethanol concentration (g/L) at 48-hour
Saccharomyces cerevisiae	3.30 ± 0.08^{e}
Pichia stipitis	2.68 ± 0.06^{f}
S. cerevisiae : *P. stipitis* (1:1)	3.56 ± 0.13^{d}
S. cerevisiae : *P. stipitis* (1:2)	3.13 ± 0.07^{e}
S. cerevisiae : *P. stipitis* (1:4)	3.91 ± 0.17^{c}
S. cerevisiae : *P. stipitis* (2:1)	5.34 ± 0.01^{a}
S. cerevisiae : *P. stipitis* (4:1)	4.68 ± 0.12^{b}

According to Duncan's test (p<0.05), the levels not connected by the same letter are significantly different.

6. Kinetic parameters of bioethanol production by *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806

From the investigation into the kinetic properties of the fermentation process in hydrolyzed sugarcane bagasse treated with an 8% sodium hydroxide solution, along with microwave irradiation, the ethanol production rate (Q_p) and ethanol yield ($Y_{p/s}$) were compared. When fermented with *S. cerevisiae* TISTR 5088 or *P. stipitis* TISTR 5806 individually, the ethanol yields were 0.20 and 0.18 $g/_{g\ reducing\ sugars}$, respectively. When the two strains were mixed at ratios of 1:1, 1:2, 1:4, 2:1, and 4:1, the ethanol yields were 0.23, 0.20, 0.24, 0.35, and 0.31 $g/_{g\ reducing\ sugars}$, respectively.

Regarding the ethanol production rate, when fermenting with single strains of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806, the rates were 0.06 and 0.04 g/L/h, respectively. Using mixed cultures of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at ratios of 1:1, 1:2, 1:4, 2:1, and 4:1, the ethanol production rates were 0.06, 0.05, 0.07, 0.10, and 0.08 g/L/h, respectively. As shown in Table 4, *S. cerevisiae* TISTR 5088 exhibited greater ethanol production efficiency than did *P. stipitis* TISTR 5806.

Table 4 The kinetic properties of the ethanol fermentation process from hydrolyzed sugarcane bagasse modified with 8% NaOH and treated with microwave irradiation at 800 W for 2 minutes, followed by enzymatic breakdown using an ACCELLERASE 1500, with both single and mixed cultures of *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 at different ratios were examined.

Fermented conditions	Kinetic parameters of bioethanol production	
	$Y_{p/s}$	Q_p
S. cerevisiae	0.20	0.06
Pichia stipitis	0.18	0.04
S. : *P.* (1:1)	0.23	0.06
S. : *P.* (1:2)	0.20	0.05
S. : *P.* (1:4)	0.24	0.07
S. : *P.* (2:1)	0.35	0.10
S. : *P.* (4:1)	0.31	0.08

S. = *Saccharomyces cerevisiae* TISTR 5088 and P. = *Pichia stipitis* TISTR 5086.

IV. Conclusion

This study investigated the optimal conditions for ethanol fermentation from sugarcane bagasse hydrolysate achieved through conditioning and microwave-assisted breakdown. Conditions involving 800-Watt microwave power for 2 minutes combined with 8% NaOH resulted in the highest reducing sugar content. Fermentation with *S. cerevisiae* TISTR 5088 and *P. stipitis* TISTR 5806 revealed that *S. cerevisiae* TISTR 5088 outperformed *P. stipitis* TISTR 5806 in ethanol production. Mixed cultures at a 2:1 ratio of *S. cerevisiae* TISTR 5088 to *P. stipitis* TISTR 5806 exhibited the highest ethanol yield and production rate, surpassing both single and mixed cultures at other ratios. Therefore, this pretreatment process holds promise for application to other materials, and the utilization of coculture emerges as a pivotal strategy for optimizing and enhancing ethanol production.

Acknowledgments

The authors would like to acknowledge King Mongkut's Institute of Technology Ladkrabang, Thailand.

References

[1.] Demirbas, A., *Political, economic and environmental impacts of biofuels: A review.* Applied energy, 2009. **86**: p. S108-S117.

[2.] Mekonnen, M.M., et al., *Water, energy, and carbon footprints of bioethanol from the US and Brazil.* Environmental science & technology, 2018. **52**(24): p. 14508-14518.

[3.] Bashir, S. and S. Lee, *Fuel Ethanol Production from Agricultural L1GNOCELLULOSIC Feedstocks-a Review.* Fuel science & technology international, 1994. **12**(11-12): p. 1427-1473.

[4.] Chen, W.-H., Y.-J. Tu, and H.-K. Sheen, *Disruption of sugarcane bagasse lignocellulosic structure by means of dilute sulfuric acid pretreatment with microwave-assisted heating.* Applied energy, 2011. **88**(8): p. 2726-2734.

[5.] Alvira, P., et al., *Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: a review.* Bioresource technology, 2010. **101**(13): p. 4851-4861.

[6.] Kumar, P., et al., *Methods for pretreatment of lignocellulosic biomass for efficient hydrolysis and biofuel production.* Industrial & engineering chemistry research, 2009. **48**(8): p. 3713-3729.

[7.] Esfahani, M.R. and M. Azin, *Pretreatment of sugarcane bagasse by ultrasound energy and dilute acid.* Asia-Pacific Journal of Chemical Engineering, 2012. **7**(2): p. 274-278.

[8.] Singh, A., et al., *Enhanced saccharification of rice straw and hull by microwave–alkali pretreatment and lignocellulolytic enzyme production.* Bioresource technology, 2011. **102**(2): p. 1773-1782.

[9.] Ma, H., et al., *Enhanced enzymatic saccharification of rice straw by microwave pretreatment.* Bioresource technology, 2009. **100**(3): p. 1279-1284.

[10.] Binod, P., et al., *Short duration microwave assisted pretreatment enhances the enzymatic saccharification and fermentable sugar yield from sugarcane bagasse.* Renewable Energy, 2012. **37**(1): p. 109-116.
[11.] Chen, W.H., Y.J. Tu, and H.K. Sheen, *Impact of dilute acid pretreatment on the structure of bagasse for bioethanol production.* International Journal of Energy Research, 2010. **34**(3): p. 265-274.
[12.] Chen, W.-H., S.-C. Ye, and H.-K. Sheen, *Hydrolysis characteristics of sugarcane bagasse pretreated by dilute acid solution in a microwave irradiation environment.* Applied Energy, 2012. **93**: p. 237-244.
[13.] Slininger, P.J., et al., *Nitrogen source and mineral optimization enhance D-xylose conversion to ethanol by the yeast Pichia stipitis NRRL Y-7124.* Applied microbiology and biotechnology, 2006. **72**: p. 1285-1296.
[14.] Palmqvist, E. and B. Hahn-Hägerdal, *Fermentation of lignocellulosic hydrolysates. I: inhibition and detoxification.* Bioresource technology, 2000. **74**(1): p. 17-24.
[15.] Ochaikul, D., A. Jongmeesuk, and C. Maneeruttanarungroj, *Bioethanol Fermentation From Acid/Base-Treated Water Hyacinth Biomass Using Fermentation Yeasts Saccharomyces Cerevisiae YRK 017 and Candida Shehatae ATCC 22984.* Int. J. Renew. Energy Res., 2019. **9**(3): p. 1428-1434.
[16.] Miller, G.L., *Use of dinitrosalicylic acid reagent for determination of reducing sugar.* Analytical chemistry, 1959. **31**(3): p. 426-428.
[17.] Kingkaew, E., et al., *Characterization of lactic acid bacteria from fermented fish (pla-paeng-daeng) and their cholesterol-lowering and immunomodulatory effects.* Microbes and environments, 2023. **38**(1): p. ME22044.
[18.] Kingkaew, E., et al., *Distribution, cholesterol-lowering and immunomodulation effects of lactic acid bacteria from fermented mussel (Hoi-dong).* Heliyon, 2022. **8**(12).
[19.] Zhu, Z., et al., *Efficient sugar production from sugarcane bagasse by microwave assisted acid and alkali pretreatment.* Biomass and Bioenergy, 2016. **93**: p. 269-278.
[20.] Anoopkumar, A.N., et al., *Progress and challenges of Microwave-assisted pretreatment of lignocellulosic biomass from circular bioeconomy perspectives.* Bioresource Technology, 2023. **369**: p. 128459.
[21.] Xu, L., et al., *Alkali-based pretreatment-facilitated lignin valorization: a review.* Industrial & Engineering Chemistry Research, 2020. **59**(39): p. 16923-16938.
[22.] Tesfaw, A. and F. Assefa, *Current trends in bioethanol production by Saccharomyces cerevisiae: substrate, inhibitor reduction, growth variables, coculture, and immobilization.* International Scholarly Research Notices, 2014. **2014**.
[23.] Scanes, K.T., S. Hohmann, and B.A. Prior, *Glycerol production by the yeast Saccharomyces cerevisiae and its relevance to wine: a review.* 1998.
[24.] Patel, H., D. Chapla, and A. Shah, *Bioconversion of pretreated sugarcane bagasse using enzymatic and acid followed by enzymatic hydrolysis approaches for bioethanol production.* Renewable Energy, 2017. **109**: p. 323-331.
[25.] Sjulander, N. and T. Kikas, *Origin, impact and control of lignocellulosic inhibitors in bioethanol production—A review.* Energies, 2020. **13**(18): p. 4751.

Sponsor