

Pertanika Journal of

**SCIENCE &**

**TECHNOLOGY**

**JST**

**VOL. 32 (1) JAN. 2024**



PERTANIKA  
JOURNALS

A scientific journal published by Universiti Putra Malaysia Press

# PERTANIKA JOURNAL OF SCIENCE & TECHNOLOGY

## About the Journal

### Overview

Pertanika Journal of Science & Technology is an official journal of Universiti Putra Malaysia. It is an open-access online scientific journal. It publishes original scientific outputs. It neither accepts nor commissions third party content.

Recognised internationally as the leading peer-reviewed interdisciplinary journal devoted to the publication of original papers, it serves as a forum for practical approaches to improve quality on issues pertaining to science and engineering and its related fields.

Pertanika Journal of Science & Technology currently publishes 6 issues a year (*January, March, April, July, August, and October*). It is considered for publication of original articles as per its scope. The journal publishes in **English** and it is open for submission by authors from all over the world.

The journal is available world-wide.

### Aims and scope

Pertanika Journal of Science & Technology aims to provide a forum for high quality research related to science and engineering research. Areas relevant to the scope of the journal include: bioinformatics, bioscience, biotechnology and bio-molecular sciences, chemistry, computer science, ecology, engineering, engineering design, environmental control and management, mathematics and statistics, medicine and health sciences, nanotechnology, physics, safety and emergency management, and related fields of study.

### History

Pertanika Journal of Science & Technology was founded in 1993 and focuses on research in science and engineering and its related fields.

### Vision

To publish a journal of international repute.

### Mission

Our goal is to bring the highest quality research to the widest possible audience.

### Quality

We aim for excellence, sustained by a responsible and professional approach to journal publishing. Submissions can expect to receive a decision within 90 days. The elapsed time from submission to publication for the articles averages 180 days. We are working towards decreasing the processing time with the help of our editors and the reviewers.

### Abstracting and indexing of Pertanika

Pertanika Journal of Science & Technology is now over 27 years old; this accumulated knowledge and experience has resulted the journal being abstracted and indexed in SCOPUS (Elsevier), Clarivate Web of Science (ESCI), EBSCO, ASEAN CITATION INDEX, Microsoft Academic, Google Scholar, and MyCite.

### Citing journal articles

The abbreviation for Pertanika Journal of Science & Technology is *Pertanika J. Sci. & Technol.*

### Publication policy

*Pertanika* policy prohibits an author from submitting the same manuscript for concurrent consideration by two or more publications. It prohibits as well publication of any manuscript that has already been published either in whole or substantial part elsewhere. It also does not permit publication of manuscript that has been published in full in proceedings.

### Code of Ethics

The *Pertanika* journals and Universiti Putra Malaysia take seriously the responsibility of all of its journal publications to reflect the highest in publication ethics. Thus, all journals and journal editors are expected to abide by the journal's codes of ethics. Refer to *Pertanika's Code of Ethics* for full details, or visit the journal's web link at [http://www.pertanika.upm.edu.my/code\\_of\\_ethics.php](http://www.pertanika.upm.edu.my/code_of_ethics.php)

### **Originality**

The author must ensure that when a manuscript is submitted to *Pertanika*, the manuscript must be an original work. The author should check the manuscript for any possible plagiarism using any program such as Turn-It-In or any other software before submitting the manuscripts to the *Pertanika* Editorial Office, Journal Division.

All submitted manuscripts must be in the journal's acceptable similarity index range:  
**≤ 20% – PASS; > 20% – REJECT.**

### **International Standard Serial Number (ISSN)**

An ISSN is an 8-digit code used to identify periodicals such as journals of all kinds and on all media—print and electronic.

Pertanika Journal of Science & Technology: e-ISSN 2231-8526 (Online).

### **Lag time**

A decision on acceptance or rejection of a manuscript is reached in 90 days (average). The elapsed time from submission to publication for the articles averages 180 days.

### **Authorship**

Authors are not permitted to add or remove any names from the authorship provided at the time of initial submission without the consent of the journal's Chief Executive Editor.

### **Manuscript preparation**

For manuscript preparation, authors may refer to *Pertanika*'s **INSTRUCTION TO AUTHORS**, available on the official website of *Pertanika*.

### **Editorial process**

Authors who complete any submission are notified with an acknowledgement containing a manuscript ID on receipt of a manuscript, and upon the editorial decision regarding publication.

*Pertanika* follows a **double-blind peer-review** process. Manuscripts deemed suitable for publication are sent to reviewers. Authors are encouraged to suggest names of at least 3 potential reviewers at the time of submission of their manuscripts to *Pertanika*, but the editors will make the final selection and are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within 90 days from the receipt of manuscript. Publication of solicited manuscripts is not guaranteed. In most cases, manuscripts are accepted conditionally, pending an author's revision of the material.

### **The journal's peer review**

In the peer-review process, 2 to 3 referees independently evaluate the scientific quality of the submitted manuscripts. At least 2 referee reports are required to help make a decision.

Peer reviewers are experts chosen by journal editors to provide written assessment of the **strengths** and **weaknesses** of written research, with the aim of improving the reporting of research and identifying the most appropriate and highest quality material for the journal.

### **Operating and review process**

What happens to a manuscript once it is submitted to *Pertanika*? Typically, there are 7 steps to the editorial review process:

1. The journal's Chief Executive Editor and the Editor-in-Chief examine the paper to determine whether it is relevance to journal needs in terms of novelty, impact, design, procedure, language as well as presentation and allow it to proceed to the reviewing process. If not appropriate, the manuscript is rejected outright and the author is informed.
2. The Chief Executive Editor sends the article-identifying information having been removed, to 2 to 3 reviewers. They are specialists in the subject matter of the article. The Chief Executive Editor requests that they complete the review within 3 weeks.

Comments to authors are about the appropriateness and adequacy of the theoretical or conceptual framework, literature review, method, results and discussion, and conclusions. Reviewers often include suggestions for strengthening of the manuscript. Comments to the editor are in the nature of the significance of the work and its potential contribution to the research field.

3. The Editor-in-Chief examines the review reports and decides whether to accept or reject the manuscript, invite the authors to revise and resubmit the manuscript, or seek additional review reports. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the authors) are forwarded to the authors. If a revision is indicated, the editor provides guidelines for attending to the reviewers' suggestions and perhaps additional advice about revising the manuscript.
4. The authors decide whether and how to address the reviewers' comments and criticisms and the editor's concerns. The authors return a revised version of the paper to the Chief Executive Editor along with specific information describing how they have addressed' the concerns of the reviewers and the editor, usually in a tabular form. The authors may also submit a rebuttal if there is a need especially when the authors disagree with certain comments provided by reviewers.
5. The Chief Executive Editor sends the revised manuscript out for re-review. Typically, at least 1 of the original reviewers will be asked to examine the article.
6. When the reviewers have completed their work, the Editor-in-Chief examines their comments and decides whether the manuscript is ready to be published, needs another round of revisions, or should be rejected. If the decision is to accept, the Chief Executive Editor is notified.
7. The Chief Executive Editor reserves the final right to accept or reject any material for publication, if the processing of a particular manuscript is deemed not to be in compliance with the S.O.P. of *Pertanika*. An acceptance letter is sent to all the authors.

The editorial office ensures that the manuscript adheres to the correct style (in-text citations, the reference list, and tables are typical areas of concern, clarity, and grammar). The authors are asked to respond to any minor queries by the editorial office. Following these corrections, page proofs are mailed to the corresponding authors for their final approval. At this point, **only essential changes are accepted**. Finally, the manuscript appears in the pages of the journal and is posted on-line.



Pertanika Journal of  
**SCIENCE  
& TECHNOLOGY**

**Vol. 32 (1) Jan. 2024**



A scientific journal published by Universiti Putra Malaysia Press



## EDITOR-IN-CHIEF

**Luqman Chuah Abdullah**  
*Chemical Engineering*

## CHIEF EXECUTIVE EDITOR

**Mohd Sapuan Salit**

## UNIVERSITY PUBLICATIONS COMMITTEE

### CHAIRMAN

**Zamberi Sekawi**

### EDITORIAL STAFF

#### Journal Officers:

Ellyianur Puteri Zainal  
Kanagamalar Silvarajoo  
Siti Zuhaila Abd Wahid  
Tee Syin Ying

#### Editorial Assistants:

Ku Ida Mastura Ku Baharom  
Siti Juridah Mat Arip  
Zulinaardawati Kamarudin

#### English Editor:

Norhanizah Ismail

### PRODUCTION STAFF

#### Pre-press Officers:

Nur Farrah Dila Ismail  
Wong Lih Jiu

### WEBMASTER

#### IT Officer:

Illi Najwa Mohamad Sakri

### EDITORIAL OFFICE

#### JOURNAL DIVISION

Putra Science Park  
1<sup>st</sup> Floor, IDEA Tower II  
UPM-IMTDC Technology Centre  
Universiti Putra Malaysia  
43400 Serdang, Selangor Malaysia.

#### General Enquiry

Tel. No: +603 9769 1622 | 1616

#### E-mail:

[executive\\_editor.pertanika@upm.edu.my](mailto:executive_editor.pertanika@upm.edu.my)

URL: [www.journals-jd.upm.edu.my](http://www.journals-jd.upm.edu.my)

### PUBLISHER

#### UPM Press

Universiti Putra Malaysia  
43400 UPM, Serdang, Selangor, Malaysia.  
Tel: +603 9769 8851  
E-mail: [penerbit@putra.upm.edu.my](mailto:penerbit@putra.upm.edu.my)  
URL: <http://penerbit.upm.edu.my>



### ASSOCIATE EDITOR

#### 2021-2023

#### Adem Kilicman

Mathematical Sciences  
Universiti Putra Malaysia, Malaysia

#### Miss Laiha Mat Kiah

Security Services Sn: Digital Forensic,  
Steganography, Network Security,  
Information Security, Communication  
Protocols, Security Protocols  
Universiti Malaya, Malaysia

#### Saidur Rahman

Renewable Energy, Nanofluids, Energy  
Efficiency, Heat Transfer, Energy Policy  
Sunway University, Malaysia

### EDITORIAL BOARD

#### 2022-2024

#### Abdul Latif Ahmad

Chemical Engineering  
Universiti Sains Malaysia, Malaysia

#### Ho Yuh-Shan

Water research, Chemical Engineering  
and Environmental Studies  
Asia University, Taiwan

#### Mohd Zulkifly Abdullah

Fluid Mechanics, Heat Transfer,  
Computational Fluid Dynamics (CFD)  
Universiti Sains Malaysia, Malaysia

#### Ahmad Zaharin Aris

Hydrochemistry, Environmental  
Chemistry, Environmental Forensics,  
Heavy Metals  
Universiti Putra Malaysia, Malaysia

#### Hsiu-Po Kuo

Chemical Engineering  
National Taiwan University, Taiwan

#### Mohd. Ali Hassan

Bioprocess Engineering, Environmental  
Biotechnology  
Universiti Putra Malaysia, Malaysia

#### Azlina Harun@Kamaruddin

Enzyme Technology, Fermentation  
Technology  
Universiti Sains Malaysia, Malaysia

#### Ivan D. Rukhlenko

Nonlinear Optics, Silicon Photonics,  
Plasmonics and Nanotechnology  
The University of Sydney, Australia

#### Nor Azah Yusof

Biosensors, Chemical Sensor, Functional  
Material  
Universiti Putra Malaysia, Malaysia

#### Bassim H. Hameed

Chemical Engineering: Reaction  
Engineering, Environmental Catalysis &  
Adsorption  
Qatar University, Qatar

#### Lee Keat Teong

Energy Environment, Reaction  
Engineering, Waste Utilization,  
Renewable Energy  
Universiti Sains Malaysia, Malaysia

#### Norbahiah Misran

Communication Engineering  
Universiti Kebangsaan Malaysia,  
Malaysia

#### Biswajeet Pradhan

Digital image processing, Geographical  
Information System (GIS), Remote  
Sensing  
University of Technology Sydney,  
Australia

#### Mohamed Othman

Communication Technology and  
Network, Scientific Computing  
Universiti Putra Malaysia, Malaysia

#### Roslan Abd-Shukur

Physics & Materials Physics,  
Superconducting Materials  
Universiti Kebangsaan Malaysia,  
Malaysia

#### Daud Ahmad Israf Ali

Cell Biology, Biochemical, Pharmacology  
Universiti Putra Malaysia, Malaysia

#### Mohd Shukry Abdul Majid

Polymer Composites, Composite  
Pipes, Natural Fibre Composites,  
Biodegradable Composites, Bio-  
Composites  
Universiti Malaysia Perlis, Malaysia

#### Wing Keong Ng

Aquaculture, Aquatic Animal Nutrition,  
Aqua Feed Technology  
Universiti Sains Malaysia, Malaysia

### INTERNATIONAL ADVISORY BOARD

#### 2021-2024

#### CHUNG, Neal Tai-Shung

Polymer Science, Composite and  
Materials Science  
National University of Singapore,  
Singapore

#### Mohamed Pourkashanian

Mechanical Engineering, Energy, CFD  
and Combustion Processes  
Sheffield University, United Kingdom

#### Yulong Ding

Particle Science & Thermal Engineering  
University of Birmingham, United  
Kingdom

#### Hiroshi Uyama

Polymer Chemistry, Organic  
Compounds, Coating, Chemical  
Engineering  
Osaka University, Japan

#### Mohini Sain

Material Science, Biocomposites,  
Biomaterials  
University of Toronto, Canada

### ABSTRACTING AND INDEXING OF PERTANIKA JOURNALS

The journal is indexed in SCOPUS (Elsevier), Clarivate-Emerging Sources Citation Index (ESCI), BIOSIS, National Agricultural Science (NAL), Google Scholar, MyCite, ISC. In addition, Pertanika JSSH is recipient of "CREAM" Award conferred by Ministry of Higher Education (MoHE), Malaysia.

The publisher of Pertanika will not be responsible for the statements made by the authors in any articles published in the journal. Under no circumstances will the publisher of this publication be liable for any loss or damage caused by your reliance on the advice, opinion or information obtained either explicitly or implied through the contents of this publication.

All rights of reproduction are reserved in respect of all papers, articles, illustrations, etc., published in Pertanika. Pertanika provides free access to the full text of research articles for anyone, web-wide. It does not charge either its authors or author-institution for refereeing/publishing outgoing articles or user-institution for accessing incoming articles.

No material published in Pertanika may be reproduced or stored on microfilm or in electronic, optical or magnetic form without the written authorization of the Publisher.

Copyright © 2021 Universiti Putra Malaysia Press. All Rights Reserved.



**Pertanika Journal of Science & Technology**  
**Vol. 32 (1) Jan. 2024**

**Contents**

Foreword <i>Mohd Sapuan Salit</i>	i
Physicochemical Impacts on Bacterial Communities in Putrajaya Lake, Malaysia <i>Nurul Najwa Fariah Mat Lazim, Afiqah Mohamed, Zana Ruhaizat Zana Rudin, Fatimah Md Yusoff, Ikhsan Natrah and Shahrizim Zulkifly</i>	1
Towards Maximising Hardware Resources and Design Efficiency via High-Speed Implementation of HMAC based on SHA-256 Design <i>Shamsiah Suhaili, Norhuzaimin Julai, Rohana Sapawi and Nordiana Rajae</i>	31
Effect of Natural Ventilation on Thermal Performance of Different Residential Building Forms in the Hot-dry Climate of Jordan <i>Esraa Shehadeh Abbaas, Mazran Ismail, Ala'eddin Ahmad Saif and Muhamad Azhar Ghazali</i>	45
Hydrocyanic Acid, Protein Concentration, and Phytochemical Compounds of Pulut and White Varieties in Young and Matured Cassava ( <i>Manihot Esculenta</i> , Crantz) <i>Sarah Idris, Rosnah Shamsudin, Mohd Zuhair Mohd Nor, Mohd Noriznan Mokhtar and Siti Salwa Abd Gani</i>	67
Attention-based Spatialized Word Embedding Bi-LSTM Model for Sentiment Analysis <i>Kun Zhu and Nur Hana Samsudin</i>	79
Analysis of Pure-pursuit Algorithm Parameters for Nonholonomic Mobile Robot Navigation in Unstructured and Confined Space <i>Izzati Saleh, Azwati Azmin, Azan Yunus and Wan Rahiman</i>	99
Use of Enhanced Greedy Algorithm for Load Balancing in Cloud Computing <i>Hanaa Osman, Asma'a Yassin Hammo and Abdulnasir Younus Ahmad</i>	113

<i>Review Article</i>	
Chromatography and Spectroscopy Methods for the Analysis of Nicotine and Other Chemical Ingredients in E-Liquid Formulation: A Review <i>Mohd Rashidi Abdull Manap, Noor Hazfalinda Hamzah, Qhurratul Aina Kholili, Fatin Abu Hasan and Azhana Alhumaira</i>	127
The Mechanical Performance of Polymer Concrete Incorporating Waste Tin Fibres <i>Hosseinkhah Ashkan, Shuhairy Norhisham, Mohd Supian Abu Bakar, Agusril Syamsir, Mohammed Jalal Abdullah, Muhammad Rizal Muhammad Asyraf, Maiyozi Chairi, Mutiara Yetrina, Jihan Melasari and Muhammad Farid</i>	143
The Mechanical Performance of Polymer Concrete Incorporating Waste Tin Fibres <i>Hosseinkhah Ashkan, Shuhairy Norhisham, Mohd Supian Abu Bakar, Agusril Syamsir, Mohammed Jalal Abdullah, Muhammad Rizal Muhammad Asyraf, Maiyozi Chairi, Mutiara Yetrina, Jihan Melasari and Muhammad Farid</i>	161
Proportions of Green Area and Tree Health on University Campus: The Impact of Pavement Presence <i>Nur Ain Aiman Abd Rahim, Rahmad Zakaria, Asyraf Mansor, Mohd Ashraf Mohamad Ismail and Nik Fadzly N Rosely</i>	185
The Effect of Nutrients in Anodic Chamber to the Performance of Microbial Fuel Cell (MFC) <i>Nashley Ursula Mundi Ujai, Siti Kudnie Sahari, Marini Sawawi, Kuryati Kipli, Asmahani Awang, Mohamad Rusop Mahmood, Lilik Hasanah, Abdul Rahman Kram and Zainab Ngaini</i>	205
Investigation of Blended Seaweed Waste Recycling Using Black Soldier Fly Larvae <i>Iva Yenris Septiariva, I Wayan Koko Suryawan, Wisnu Prayogo, Sapta Suhardono and Ariyanti Sarwono</i>	217
An Attribute-based Data Privacy Classification Through the Bayesian Theorem to Raise Awareness in Public Data Sharing Activity <i>Nur Aziana Azwani Abdul Aziz, Masnida Hussin and Nur Raidah Salim</i>	235
High-performance THz Metallic Axial Mode Helix Antenna with Optimised Truncated Hollow Cone Ground Plane for 6G Wireless Communication System <i>Zahraa Raad Mayoof Hajiyat, Alyani Ismail, Aduwati Sali and Mohd. Nizar Hamidon</i>	249

Soft Set-based Parameter Reduction Algorithm Through a Discernibility Matrix and the Hybrid Approach for the Risk-Factor Prediction of Cardiovascular Diseases by Various Machine Learning Techniques <i>Menaga Anbumani and Kannan Kaniyaiah</i>	265
<i>Review Article</i>	
Understanding the Degradation of Carbofuran in Agricultural Area: A Review of Fate, Metabolites, and Toxicity <i>Nurul Syuhada Baharudin, Harlina Ahmad and Md Sohrab Hossain</i>	285
Phenolics-Enhancing <i>Piper sarmentosum</i> (Roxburgh) Extracts Pre-Treated with Supercritical Carbon Dioxide and its Correlation with Cytotoxicity and $\alpha$ -Glucosidase Inhibitory Activities <i>Mohd Shahrul Ridzuan Hamil, Fauziahanim Zakaria, Lee Chong Yew, Norzilawati Pauzi, Khamsah Suryati Mohd and Noor Hafizoh Saidan</i>	323
Estimation of Leachate Volume and Treatment Cost Avoidance Through Waste Segregation Programme in Malaysia <i>Josfirin Uding Rangga, Sharifah Norkhadijah Syed Ismail, Irniza Rasdi and Karmegam Karuppiah</i>	339
Comparison Using Intelligent Systems for Data Prediction and Near Miss Detection Techniques <i>Lek Ming Lim, Saratha Sathasivam, Mohd. Tahir Ismail, Ahmad Sufiril Azlan Mohamed, Olayemi Joshua Ibidoja and Majid Khan Majahar Ali</i>	365
Development of GIS-based Ground Flash Density and its Statistical Analysis for Lightning Performance Evaluation of Transmission Lines in Peninsular Malaysia <i>Nurzanariah Roslan, Ungku Anisa Ungku Amirulddin, Mohd Zainal Abidin Ab. Kadir and Noradlina Abdullah</i>	395
Preparation of Activated Carbon from Sugarcane Bagasse Using Microwave-assisted $ZnCl_2$ Chemical Activation: Optimization and Characterization Study <i>Atiqa Rahmawati, Fadzkurisma Robbika and Yuafni</i>	419
Conceptual Design and Selection of Natural Fibre Reinforced Composite Cyclist Helmet Liner Using an Integrated Approach <i>Nurul Ain Maidin, Mohd Sapuan Salit, Mastura Mohammad Taha and Mohd Zuhri Mohamed Yusoff</i>	

**Selected papers from the International Conference on Data Science  
and Emerging Technologies (2022)**

**Guest Editors:** Yap Bee Wah and Simon Fong

Improved GCC Technique: A Comprehensive Approach to Color  
Cast Rectification and Image Enhancement 453  
*Danny Ngo Lung Yao, Abdullah Bade, Iznora Aini Zolkifly and  
Paridah Daud*

Transfer Learning for Lung Nodules Classification with CNN and  
Random Forest 463  
*Abdulrazak Yahya Saleh, Chee Ka Chin and Ros Ameera Rosdi*



# Foreword

Welcome to the first issue of 2024 for the *Pertanika Journal of Science and Technology (PJST)*!

PJST is an open-access journal for studies in Science and Technology published by Universiti Putra Malaysia Press. It is independently owned and managed by the university for the benefit of the world-wide science community.

This issue contains 25 articles; two review articles and the rest are regular articles. The authors of these articles come from different countries namely Bangladesh, India, Indonesia, Iraq, Malaysia, Saudi Arabia, Spain and United Kingdom.

Izzati Saleh, Azwati Azmin and Azan Yunus from Malaysia and Wan Rahiman from Bangladesh presented an analysis of pure-pursuit algorithm parameters for nonholonomic mobile robot navigation in unstructured and confined space. This study's simulation-based experiment is limited to the mobile robot arrangement. The Look Ahead Distance parameter is adjusted so the mobile robot can navigate the predefined map closely following the waypoints. The optimal Look Ahead Distance value is combined with the VFH+ algorithm for obstacle avoidance. The method is enhanced by adding the  $\lambda$  weight so the robot returns to its waypoints after avoiding an obstacle. The investigation reveals that  $\lambda$  influences the mobile robot's capacity to return to its predetermined waypoints after avoiding an obstacle. The detailed information of this study is available on page 99.

An investigation on a high-performance THz metallic axial mode helix antenna with an optimised truncated hollow cone ground plane for a 6G wireless communication system was conducted by Zahraa Raad Mayoof Hajiyat et al. from Universiti Putra Malaysia. The simulation results show that the optimised copper (annealed) axial mode helix antenna performed well in the 0.52–0.98 THz frequency band with an impedance BW of 0.46 THz and FBW of 61.33%. Additionally, the highest directivity and realised gain recorded were 21.8 dBi and 21.5 dBi at 0.85 THz, respectively. The comparative analysis between the CST MWS and Ansys HFSS showed good agreement, further validating the proposed antenna design. Moreover, the performance comparison of this study shows that the proposed optimised THz antenna design offered an outstanding directivity performance compared to other available THz axial mode helix antennas. Further details of the investigation can be found on page 249.

The next article reviewed the agricultural area's carbofuran usage, toxicity, and degradation pathways. Carbofuran is a highly toxic insecticide commonly used to protect crops in agricultural areas. Exposure to carbofuran can cause harmful effects on the ecological environment and human health, particularly on non-target species such as birds and aquatic organisms. Carbofuran continues to be used, although it has been banned in some countries. Several metabolites

are formed during the breakdown of carbofuran. These include 3-hydroxy-carbofuran, 3-ketocarbofuran, carbofuran-phenol, and 3-hydroxy-5-nitrophenol. These metabolites vary in toxicity and persistence in the environment, with some being more lethal than the parent compound. Details of this study are available on page 285.

We anticipate that you will find the evidence presented in this issue to be intriguing, thought-provoking and useful in reaching new milestones in your own research. Please recommend the journal to your colleagues and students to make this endeavour meaningful.

All the papers published in this edition underwent Pertanika's stringent peer-review process involving a minimum of two reviewers comprising internal as well as external referees. This was to ensure that the quality of the papers justified the high ranking of the journal, which is renowned as a heavily-cited journal not only by authors and researchers in Malaysia but by those in other countries around the world as well.

We would also like to express our gratitude to all the contributors, namely the authors, reviewers, Editor-in-Chief and Editorial Board Members of PJST, who have made this issue possible.

PJST is currently accepting manuscripts for upcoming issues based on original qualitative or quantitative research that opens new areas of inquiry and investigation.

**Chief Executive Editor**

**Mohd Sapuan Salit**

[executive\\_editor.pertanika@upm.edu.my](mailto:executive_editor.pertanika@upm.edu.my)

## Physicochemical Impacts on Bacterial Communities in Putrajaya Lake, Malaysia

Nurul Najwa Fariah Mat Lazim<sup>1</sup>, Afiqah Mohamed<sup>1</sup>, Zana Ruhaizat Zana Rudin<sup>1</sup>, Fatimah Md Yusoff<sup>2,3</sup>, Ikhsan Natrah<sup>3</sup> and Shahrizim Zulkifly<sup>1\*</sup>

<sup>1</sup>Department of Biology, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>International Institute of Aquaculture and Aquatic Sciences, Universiti Putra Malaysia, 71050 Port Dickson, Negeri Sembilan, Malaysia

<sup>3</sup>Department of Aquaculture, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

### ABSTRACT

This study determines the associations between the bacterial communities and water physicochemical parameters in Putrajaya Lake and Putrajaya Wetlands Park, Malaysia. Bacterial communities were assessed by metagenomics of the 16S rRNA gene from lake water input, central wetlands, and primary lake area. Water samples (n=18) were collected during two different periods: post-high rainfall events (samples collected in May) and dry periods (July). The data revealed that bacterial communities of the three sites were taxonomically distinct and associated with different environmental parameters. However, no significant differences were found between the wet and dry periods. Alpha diversity analyses revealed the highest index in May 2018 in the constructed wetlands ( $H' = 5.397$ ) than those from water input or primary lake ( $p < 0.05$ ). Overall, 49 phyla, 147 classes, 284 orders, 471 families, 778 genera and 62 species of bacteria were identified.

Verrucomicrobia and Firmicutes showed a strong positive correlation with ammonia-nitrogen ( $r = 0.709$ ). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with r value ( $r = 0.673$ ) and ( $r = 0.647$ ), respectively. In this study, the metagenomics of the 16S rRNA gene amplicon by Illumina MiSeq has successfully identified the bacterial community assemblage in Putrajaya Lake and wetlands. Bacterial composition was

### ARTICLE INFO

#### Article history:

Received: 02 September 2022

Accepted: 06 March 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.01>

#### E-mail addresses:

najwafariah07@gmail.com (Nurul Najwa Fariah Mat Lazim)

haqifa.dhom@gmail.com (Afiqah Mohamed)

zruhaizatrz@gmail.com (Zana Ruhaizat Zana Rudin)

fatimamy@upm.edu.my (Fatimah Md Yusoff)

natrah@upm.edu.my (Ikhsan Natrah)

shahrizim@upm.edu.my (Shahrizim Zulkifly)

\* Corresponding author

associated with the availability of physicochemical properties of specific sites. The effectiveness of the engineered wetlands of Putrajaya in bioremediation was demonstrated by the marked decrease in certain nutrient concentrations from lake water input to the primary lake area.

*Keywords:* 16S rRNA, bacterial community, engineered wetlands, physicochemical, water quality

---

## INTRODUCTION

Environmental concerns, which include climate change, habitat degradation, and water pollution, have directly or indirectly affected aquatic ecosystems' diversity and water quality. Putrajaya Lake is one of the prides of Putrajaya and is located in Malaysia's federal government administrative centre and has a total surface area of 400 ha. It has been designed to meet Class IIB of the Malaysian National Water Quality Standard (NWQS) and is suitable for body-contact recreational activities (Suratman et al., 2016). Putrajaya Holdings is entrusted to manage and maintain Putrajaya Lake by evaluating and monitoring the water quality according to the National Water Quality Standards, which have implemented various smart initiatives, such as low-carbon initiatives, with seven focal areas, in line with the Sustainable Development Goals of the United Nations (Majizat et al., 2016).

The Putrajaya Wetlands Park is an essential part of Putrajaya Lake, filtering the water from the primary lake from two main rivers: Chuau River and Bisa River. This engineered wetland covers 200 ha and was constructed as a natural filtration and bioremediation system of water input to the lake.

In Putrajaya Wetlands Park, several different plant communities have been planted to serve as a nutrient filtration system (Newton et al., 2011). Phytoremediation is a biological process in which selected plants remove water contaminants or pollutants. The filtration system involves the accumulation of contaminants by plant roots and provides a root zone and sediment habitat for bacteria and microbes that filter and decompose contaminants (Mohamad, 2012; Sabkie et al., 2020).

Rapid population growth and urbanisation lead to inevitable repercussions on the environment, such as releasing nutrient-rich runoffs into the lake, which have been cited as the main factors affecting reduced lake quality. These continuous changes in the seasonal and physicochemical parameters have undoubtedly influenced the state of water quality in Putrajaya Lake. Many indications of pollution were found within Putrajaya Lake due to the anthropogenic activities from the inlets or surface run-off points surrounding the lake (Asmat et al., 2018). Monitoring the nutrient removal for water quality improvement has been done at a pilot scale to simulate the essential roles of inhabiting macrophytes within the constructed wetlands (Vymazal, 2007). As far as the water quality monitoring assessment of freshwater lake is concerned, the Putrajaya Corporations (PjC) has developed the Putrajaya Lake Water Quality Standards (PLWQS) (Sharip et al., 2016). A previous study which aimed to design an

algorithm prediction for water quality changes has found that ammonia concentration plays a crucial role in determining the water quality level (Najah et al., 2021).

In addition to the current monitoring of the water quality parameters, there is a need to understand Putrajaya Lake in a more holistic approach, such as in biodiversity studies. Biodiversity is one of the critical components that determine the well-being of the lake, as it can be used as an indicator to assess the lake's health.

Previous studies in Putrajaya Lake only focused on assessing water quality and diversity of phytoplankton. Research on phytoplankton community structure as an ecosystem health indicator in the management of Putrajaya Lake and Putrajaya Wetlands Park indicated that there were 148 species from 77 genera identified during a sampling period from October 2009 until September 2010 (Jamal et al., 2014). Another study from November 2017 until January 2018 on phytoplankton diversity from three sites in Putrajaya Lake and Putrajaya Wetlands Park reported the discovery of 14 genera within eight classes (Sabkie et al., 2020). There was a higher density of phytoplankton recorded in Putrajaya Lake compared to the wetland.

The bacteria population in the lake forms a vital niche in nutrient cycling, decomposition and remediation. Bacteria are essential for all life in the ecosystem as they play a role in maintaining the structure, function, and sustainability of the ecosystem (Briones & Raskin, 2003). Information on bacterial diversity has the potential to aid water quality assessments. This information is currently limited to Putrajaya Lake.

Bacteria have a significant role in regenerating and mobilising the nutrients in freshwater food webs and are important in recycling the most naturally dynamic components in the ecosystem (Newton et al., 2011). Bacteria collectively are responsible for the movement of substances and conditions of water bodies (Pernthaler & Amann, 2005) as a result of their biomass creation and trophic coupling to eukaryotic predators as well as the foremost essential degraders and converters of an organic compound into inorganic material (Cotner & Biddanda, 2002).

The identification and classification of ubiquitous prokaryotes are widely investigated (Gupta et al., 2013). Bacteria can be identified based on morphology, biochemical tests, and culturing with selective media. However, not all environmental bacteria can be cultured under artificial laboratory conditions. The most current and practical approach to bacterial identification is through molecular methods using the Next Generation Sequencing (NGS) technique.

The next-generation sequencing (NGS) is one of the new technologies in the field of genomic analysis for DNA sequencing. It is a persuasive tool for demonstrating the diversity of various samples and studying metabolic pathways (Al-Sulaiman, 2012). The application of NGS leads to a high throughput accurate sequence read length and has enabled the investigation of a vast number of samples at a greater depth (Fadrosh et al., 2014).

Metagenomics studies used the direct genetic material of samples from the environment, bypassing the culturing process. The Illumina MiSeq platform can examine the community

composition of the clinical and environmental samples (Fadrosch et al., 2014). The 16S ribosomal RNA (rRNA) gene has been previously used for the assessment of bacteria variety and structure, including soil (Liles et al., 2003), marine (Sogin et al., 2006), and freshwater environments (Mueller-Spitz et al., 2009). The 16S rRNA gene is present in all bacteria, consisting of conserved regions to allow the design of a universal primer to amplify the 16S rRNA gene via polymerase chain reaction (PCR); incorporating the hypervariable regions (V1-V9) will be used for bacteria identification (Jo et al., 2016). The characterisation of bacteria community by using 16S rDNA as a molecular marker that targets hypervariable regions (V3, V5 or V6) for amplification allows profound sequencing (Staley et al., 2013), and V3-V4 regions have been used for the identification of bacterial genera from oligotrophic freshwater reservoirs using Illumina MiSeq platform (Iliev et al., 2017).

Putrajaya experiences an equatorial climate characterised by hot and humid weather all year round. Microclimate rainfall events and anthropogenic activities surrounding Putrajaya Lake are significant contributors to the lake's water quality, which include the release of effluents from the point and non-point sources, sewage from residential houses and commercial areas, the use of fertilisers for farming and rubbish littering. Identifying the bacteria community is important as bacteria can be an environmental bioindicator to detect changes in the lake's water quality. Many factors, which include global climate change, can threaten and shift the diversity of aquatic organisms in the water body. Water temperature has the potential to influence the changes in the hydrology of a water body and the habitat suitability of species.

In this study, Illumina MiSeq Sequencing was used to identify the bacterial diversity of Putrajaya Lake and Putrajaya Wetlands Park, targeting 16S rRNA V3-V4 hypervariable regions. Identifying the bacterial community will be a useful environmental indicator to detect changes in the lake's nutrient composition and water quality.

The objectives of this research were (1) to utilise the 16S ribosomal RNA (rRNA) gene for identifying bacteria species collected from water input, water after going through natural remediation of wetlands and primary lakes using next-generation sequencing technique, and (2) to determine the correlation between water body physicochemical properties and bacterial diversity. Here, we present the implementation of amplicon sequencing, which targeted the V3-V4 region of the 16S rRNA gene and the association of the physicochemical properties for studying bacterial community composition in Putrajaya Lake and Putrajaya Wetlands Park, Malaysia.

## **MATERIALS AND METHODS**

### **Study Area and Sample Collection**

Putrajaya Lake, located in the Federal Territory of Putrajaya, Malaysia, is an artificial freshwater lake sourced from the Chuau and Bisa River systems. Water sampling was



conducted in 2018 at three sites (Figure 1). The descriptions of the chosen sites are shown in Table 1.

Five hundred (500) mL water samples were collected in triplicate at each of the three sites using a Van Dorn sampler and collected in 500 mL PET sterilised water bottles. Surface water sampling (~1m) was done in selected months beginning November 2017 to August 2018. However, for this paper, only two months of data were selected, corresponding to

Table 1  
Sampling sites

Sites	Location	GPS Coordinate	Site Description
Site 1	The point of lake water input, Upper North 8 inlet (UN8; lake inlet)	2°59'10 "N, 101°41'26 "E	The main water inlet from Chuau River, the longest arm within the constructed wetland with diverse macrophyte species
Site 2	The Central Wetland (CW; engineered wetland)	2°57'1 "N, 101°41'37 "E	Final catchment area for bioremediated water before discharged into the primary lake area
Site 3	The open water at the Primary Lake f (PLf)	2°54'19.63 "N, 101°40'21.59 "E	Centre for water sports, recreational activities, and tourism

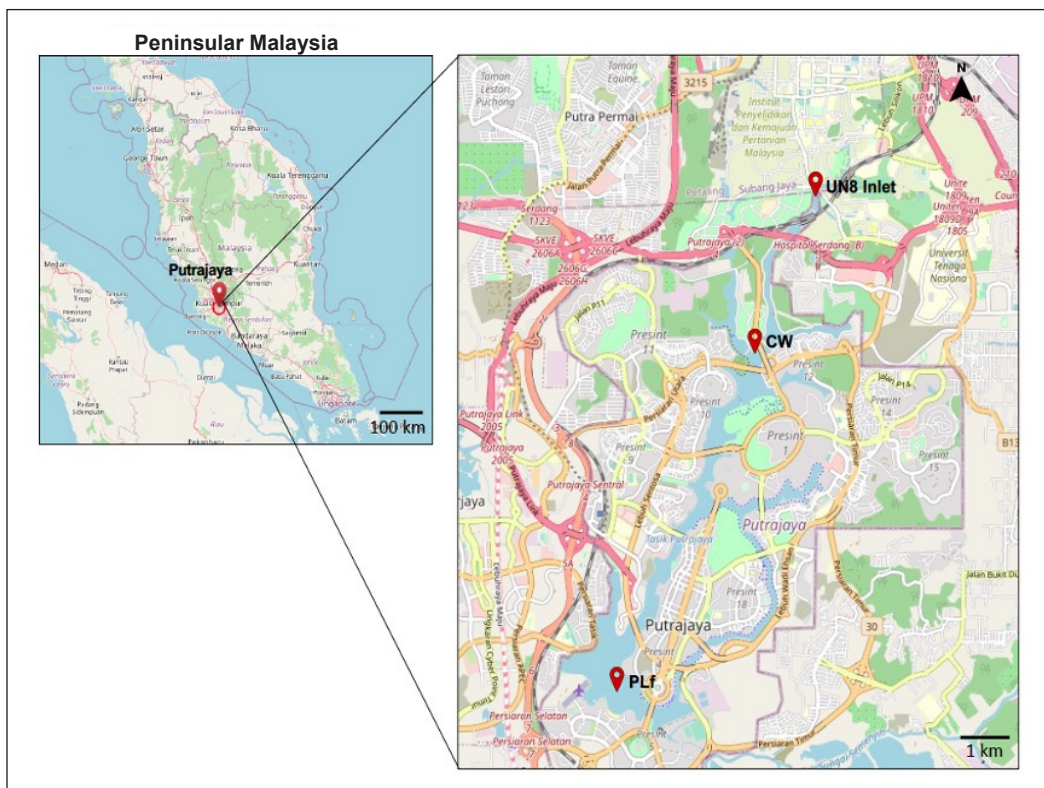


Figure 1. Three sampling sites in Putrajaya Lake and Putrajaya Wetlands Park, Site 1 (UN8: Upper North 8; lake inlet), Site 2 (CW: Central Wetland; engineered wetland), and Site 3 (PLf: Primary Lake f; open lake water)

the availability of metagenomics data and physicochemical correlations. The selection data was based primarily and limited on the samples' genomic quality and availability. In total, eighteen water samples were collected in May and July 2018 from three sources: Site 1: Upper North 8 (S1R1, S1R2, S1R3), Site 2: Central Wetland (S2R1, S2R2, S2R3), and Site 3: Primary Lake f (S3R1, S3R2, S3R3) were used in this study. Data in May 2018 was chosen to represent the wet period during the post-rainfall event, whereas data in July 2018 was chosen to represent the dry periods with less rainfall. The rainfall pattern data was provided by the Putrajaya Corporation (PjC) (Figure 2).

A total of eighteen samples collected represented three sources: Site 1 (S1R1, S1R2, S1R3), Site 2 (S2R1, S2R2, S2R3), and Site 3 (S3R1, S3R2, S3R3). All samples were immediately stored on ice in a cooler, transferred into the laboratory within 4 hours of collection, and stored at 4°C before further processing.

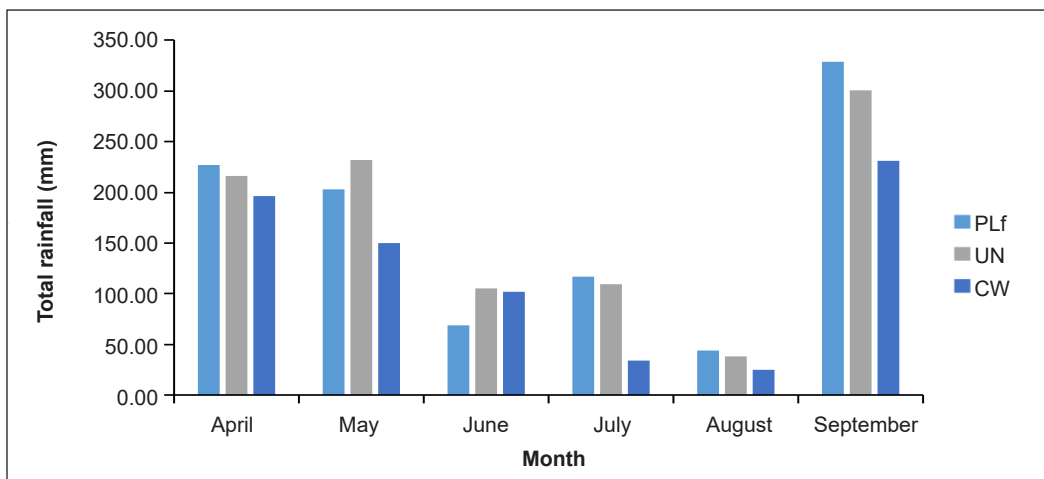


Figure 2. Rainfall pattern at three Putrajaya Lake and Putrajaya Wetlands Park sampling sites within the sampling months

### Physico-chemical Analysis

*In situ* measurements were done for all physical parameters in accordance with the Standard Methods for the Examination of Water and Wastewater (APHA, 2005). Water salinity, conductivity, and temperature were assessed by a YSI multi-parameter probe (Model: 30 SCT Handheld meter). Dissolved oxygen was measured by YSI Dissolved Oxygen (Model 58), while pH was measured by Trans Instrument WalkLAB Microprocessor pH meter (HP9000). Irradiance was determined using a Digital Light Meter (Model: LD8903) and a Secchi disk measured water transparency. Nutrients (ammonia (NH<sub>3</sub>), nitrate (NO<sub>3</sub><sup>-</sup>), phosphate (PO<sub>4</sub><sup>3-</sup>) and silica (Si)) were measured in the laboratory within the same sampling day by using Hach Multiparameter Portable Colorimeter (DR900, HACH). The chemical profiles of the water sample collected were assessed based on the instructions: nitrate



(Cadmium Reduction method: 8039), phosphate (Ascorbic acid method: 8048), ammonia (Salicylate method: 8155) and silica (Heteropoly Blue method: 8186) as described in the HACH handbook of Water Analysis (Hach, 2002). Readings for each parameter were recorded in triplicates. All equipment for *in situ* analysis was calibrated according to the manufacturer's instructions prior to sampling.

### **DNA Extraction**

Triplicate samples of 500 mL water were filtered with Whatman polyethersulfone (PES) membrane filter paper with a pore size of 0.22  $\mu\text{m}$ , 47 mm diameter, using a filtration pump. The filter papers containing the microbes were cut into small pieces using a sterile blade and stored in a  $-80^{\circ}\text{C}$  freezer prior to DNA extraction. According to the manufacturer's suggested protocol, DNA was extracted using an EZNA Soil DNA Kit (Omega Bio-Tek, Inc.). The DNA purity analysis was carried out using a Nanodrop BioSpectrometer, and all samples showed a ratio value of A260/280 nm between the range of 1.8–2.0 (Kinetic, Eppendorf, Germany).

### **PCR Amplification and Illumina MiSeq Sequencing**

The oligonucleotide primers used for PCR amplification targeting 16S V3-V4 regions in this study were (5'- CCT ACG GGN GGC WGC AG -3') and (5'-GAC TAC HVG GGT ATC TAA TCC -3') (Chen et al., 2019). The PCR was performed according to Illumina's 16S Metagenomic Sequencing Library protocol (Part # 15044223 Rev. B) (Illumina Inc., 2013). Each PCR reaction contained 25  $\mu\text{l}$  of Master Mix (5  $\mu\text{l}$  of Amplicon PCR Forward Primer (1  $\mu\text{M}$ ), 5  $\mu\text{l}$  of Amplicon PCR Reverse Primer (1  $\mu\text{M}$ ), 12.5  $\mu\text{l}$  of 2 $\times$  KAPA HiFi HotStart ReadyMix, and 2.5  $\mu\text{l}$  of microbial genomic DNA (5 ng/ $\mu\text{l}$  in 10mM Tris pH 8.5). The PCR cycle conditions used in this study were as follows:  $95^{\circ}\text{C}$  for 3 minutes; 25 cycles of  $95^{\circ}\text{C}$  for 30 seconds,  $55^{\circ}\text{C}$  for 30 seconds, followed by  $72^{\circ}\text{C}$  for 30 seconds, with a final extension of  $72^{\circ}\text{C}$  for 5 minutes, and hold at  $4^{\circ}\text{C}$ . One  $\mu\text{l}$  of PCR products was run on 1.7% TAE agarose gel at 100V with 1500 bp DNA ladder for 65 mins. The gel was then viewed under a UV light gel documentation system (Enduro<sup>TM</sup> GDS). The PCR product size was approximately 430 bp. Amplicons were sent for Illumina Library Preparation and MiSeq sequencing at Apical Scientific Sdn. Bhd., Seri Kembangan.

### **Data Analysis**

Apical Scientific Sdn. Bhd. provided all bioinformatics data analysis. The raw Illumina MiSeq sequence data were processed to remove sequence adaptors and low-quality reads using Bestus Bioinformatics Decontamination using Kmers (BBduk) of the BBTools package (<https://sourceforge.net/projects/bbmap/>). Then, the paired-end reads (forward and reverse) were merged using USEARCH software (v11.0.667) (<https://www.drive5.com/usearch/>). Sequences shorter than 150 bp or longer than 600 bp were removed from

downstream processing to reduce the presence of unwanted adaptor and primer dimers (Gane, 2018) and increase the quality of reads, respectively (Tan et al., 2019). The remaining sequences were aligned against the 16S rRNA SILVA database (Release 132) and checked for chimeric errors using VSEARCH v2.6.2. In *de novo* OTU picking, all reads were clustered at 97% using UPARSE v11.0.667. Spurious OTUs with only one read (singleton) or two reads (doubleton) were removed from downstream processing to reduce the rate of sequencing errors and increase the accuracy of diversity metrics (Edgar, 2013; Flynn et al., 2015). For each OTU, a single representative sequence was chosen randomly, and Pynast software (<https://www.ncbi.nlm.nih.gov/pubmed/19914921>) was used to align and construct a phylogenetic tree against the SILVA 132 16S rRNA database. Benchmarking taxonomic assignment of OTUs was achieved using QIIME V1.9.1 against the Silva 16S rRNA database (release 132) (Kuczynski et al., 2012).

All statistical analyses were performed in R package V3.6. Alpha-diversity indices were calculated to assess species diversity in a sample using the UPARSE pipeline (Kuczynski et al., 2012), including the observed OTU, Chao1 and Shannon indices. The Vegan package performed the beta-diversity assessment to evaluate differences in taxonomic complexity among the samples using the phylogenetic weighted Unifrac distance (Lozupone et al., 2011) and ordination techniques using Principal coordinate analysis (PCoA) (Oksanen, 2008). Linear discriminant analysis effect size (LEfSe) was performed online (<https://huttenhower.sph.harvard.edu/galaxy/>) to determine the features of the bacterial group most likely to show differences between the two conditions or among the three sites based on the OTU lists of samples (Segata et al., 2011). For LEfSe analysis, an LDA score of >4 was used as a threshold. One-way Analysis of Similarity test (ANOSIM) was performed using the statistical software package PRIMER-E to test for significant differences between bacterial communities at three sites (Clarke & Gorley, 2006).

One-way analysis of variance (ANOVA) was performed in Statistical Package for Social Science (SPSS) version 25.0 to evaluate differences in physicochemical properties among the three sampling sites at a 0.05% significance level. A post hoc test (Tukey HSD) was used to compare means significantly different from each other at a probability of 5%. Principal component analysis (PCA) was used to examine the correlation among physicochemical parameters. Canonical correspondence analysis (CCA) was also performed to explore possible correlations between bacterial community and environmental variables using Microsoft Excel XLStat (Wan et al., 2017).

## Data Availability

The raw 16S reads were deposited to NCBI Sequence Read Archive (SRA) under accession number PRJNA687158. All other data are provided in the main text or supplementary materials.

## RESULTS

### Physico-chemical Properties of the Samples

A comparison was made between the lake water input (UN) and the Central wetland (CW) due to the direct water flow from the inlet to CW for remediation. Water flowing from the wetland to the primary lake (PLf) was not significantly highlighted due to the distance (km) and various water inlets surrounding the lake, which might have an additional influence on the physicochemical parameters' readings.

Significant reduction of nutrients upon bioremediation at the wetland (CW) showed that the concentration of nitrate and silica was lower as the water passed through the Upper North (UN8) arm of the wetland (Figure 3). However, it was noted that there was no significant difference in physicochemical parameters between the dry and the wet (post-rainfall) event.

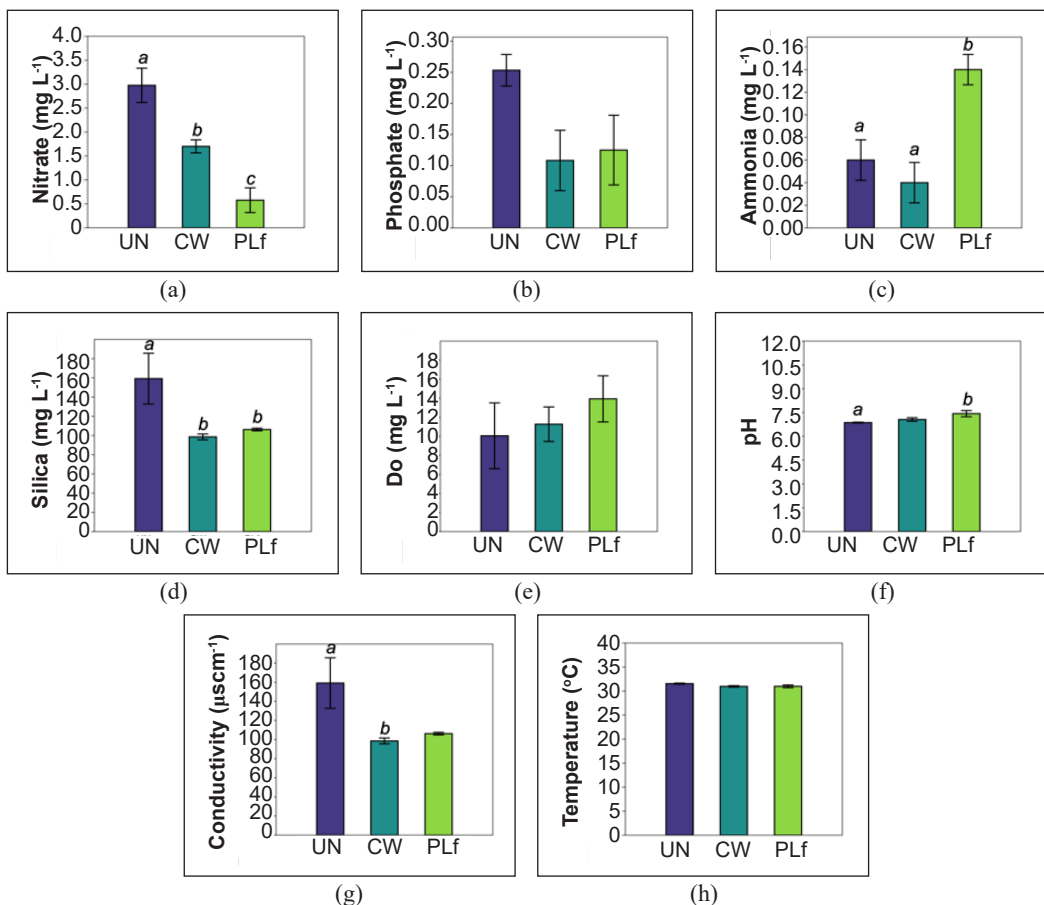


Figure 3. Mean physicochemical parameter measurements combined for May and July (A-G) in Upper North 8 (UN), Central Wetland (CW), Primary lakefront (PLf): (a) Nitrate; (b) Phosphate; (c) Ammonia; (d) Silica; (e) Dissolved Oxygen; (f) pH; (g) Conductivity; and (h) Temperature. Each bar represents the mean of triplicate data  $\pm$  SE. Bars with different letters are significantly different ( $p < 0.05$ )

### Bacterial Community Sequencing, Data Rarefaction and Diversity Indices

For biological data, a comparison was made between May and July. In total, 3,474,032 raw reads were generated by Illumina MiSeq sequencing. After quality filtering, a total of 1,910,113 16S readings were obtained. Rarefaction analysis indicated that the number of OTUs reached the plateau phase (Supplementary Figure 1), suggesting sufficient sequence numbers to represent the actual bacterial community of specific sites and conditions. The value of diversity indices (Chao1, Shannon and observed OTU) of the bacterial community at the three sites was observed (Figure 4). The OTU number at the wetland in the wet condition (May) was 1248 on average, higher than the dry weather (July). The indices value of Chao1 and Shannon at the wetland in the dry event recorded the highest value, 1439 and 5.22 on average, respectively, while the inlet in the wet season recorded the highest index, 1233 and 4.88 on average.

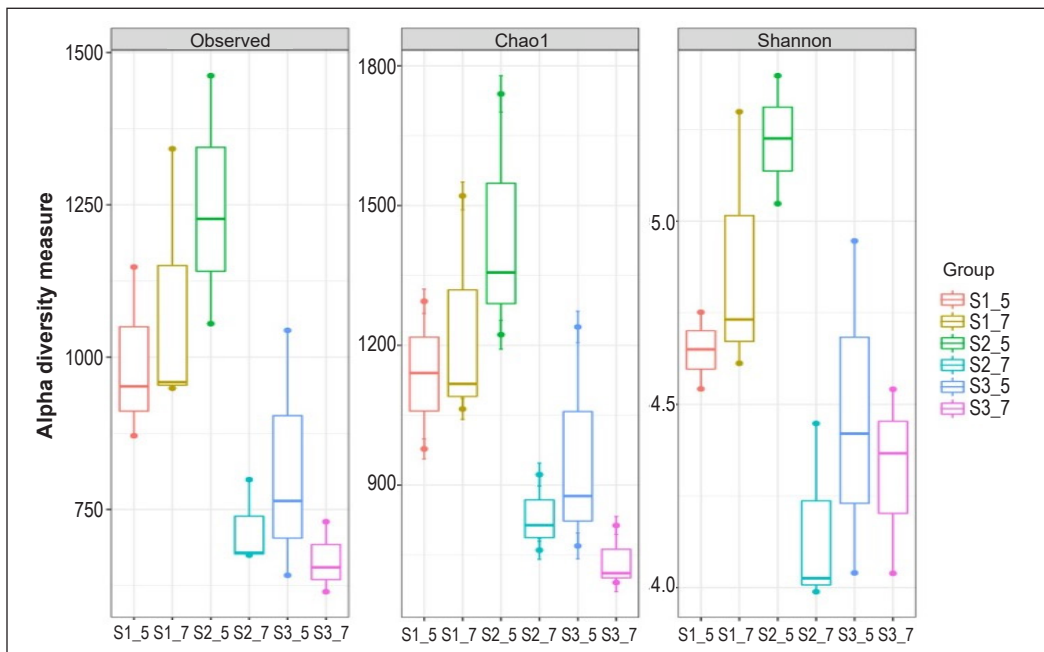


Figure 4. Comparison of diversity indices (OTU, Chao1 and Shannon). S1\_5 (Inlet, wet), S1\_7 (Inlet, dry), S2\_5 (CW, wet), S2\_7 (CW, dry), S3\_5 (Lake, wet), S3\_7 (Lake, dry)

### Beta Diversity Differences Between Sites and Months

Beta diversity analysis was performed to analyse and compare the overall bacterial diversity within the community throughout the sampling period. Principal Coordinates Analysis (PCoA) plots following the Unique Fraction (UniFrac) method (Figure 5) showed the different compositions of the bacterial diversity within the samples. PC1 and PC2 explained 45.22% of the total microbial community structure.

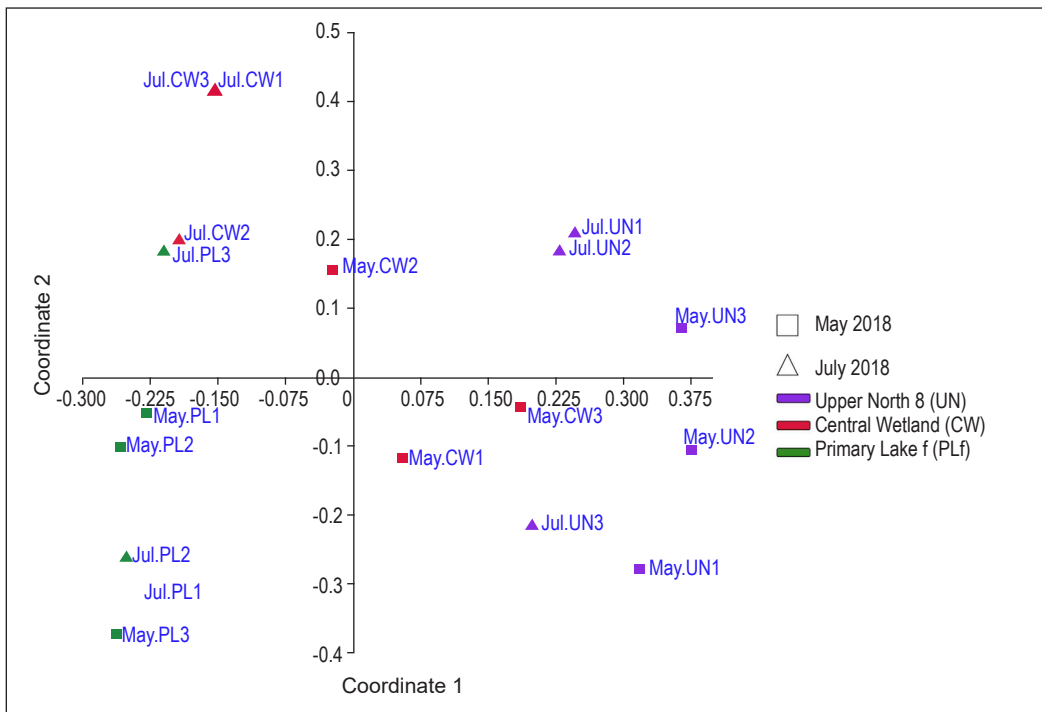


Figure 5. Principal coordinates analysis (PCoA) of bacterial communities. May: wet period, July: dry period. UN= Upper North 8, CW= Central Wetland, PL = Primary Lake f

PCoA identifies factors that differentiate the microbial communities. Clustering can be seen for PCoA ordination of samples of the primary lake (PL) communities as well as the upper North (UN) site. PCoA revealed seasonal community trends that were particularly apparent in PL and UN.

Weighted UniFrac distances were used to measure the dissimilarity coefficient between samples at different sampling sites based on the top ten taxa abundances. UniFrac distances showed that the diversity of bacterial OTUs was separated into two groups with different sampling times (Figure 6).

### Bacterial Community Composition

The phylum and genus-level bacterial diversity distributions are shown in Figures 7 and 8, respectively. Forty-six major bacteria phyla were identified, including Proteobacteria (6.49%), Actinobacteria (2.95%), Bacteroidetes (2.86%), Verrucomicrobia (1.74%), Cyanobacteria (1.48%), Planctomycetes (1.34%), Firmicutes (0.52%), Chloroflexi (0.34%), Chlamydiae (0.14%) and Acidobacteria (0.09%). LEfSe analysis was performed to identify the most differentially abundant taxa between the two conditions by having a linear discriminant analysis (LDA) score of more than 4 (Supplementary Figures 2, 3 and 4). Proteobacteria were more abundant during the dry condition, including *Pseudomonas*

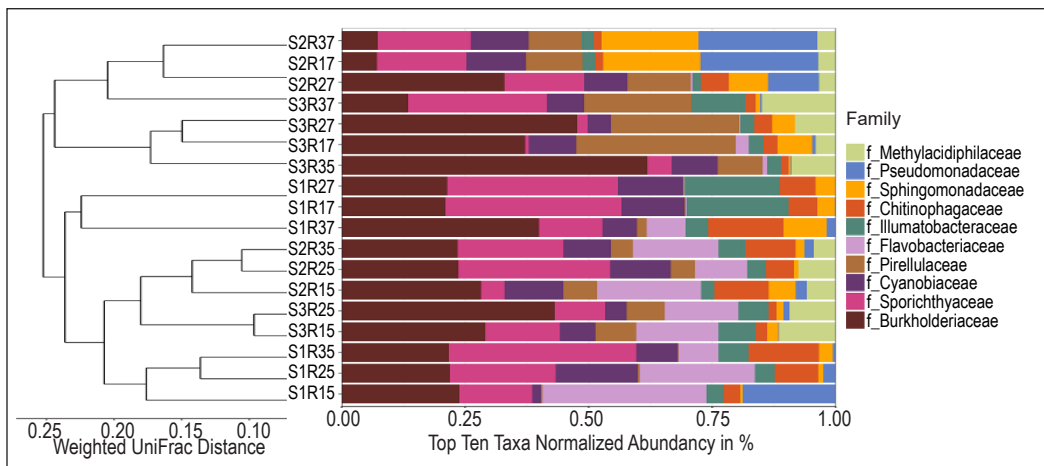


Figure 6. Weighted UniFrac dissimilarities of bacterial communities between samples at different sampling sites (S1: Inlet, S2: CW, S3: Lake, 5: Wet, 7: Dry)

sp., which were discovered in the dry condition at the wetland (0.92%) (Figures 7, 8 and Supplementary Table 2). In wet conditions, Bacteroidetes was the most common phylum (2.86%) at all three sites, with *Flavobacterium columnare* representing 2.54% of the total Bacteroidetes population. Planctomycetes were commonly found in dry conditions in the wetland, whereas the primary lake was represented by Pirellulales (Supplementary Figures 3 and 4). Phylum Firmicutes was discovered to be more common during the dry condition at both inlet and primary lake based on LEfSe analysis (Supplementary Figures 2 and 4). Upon the bioinformatics analysis against the database reference, we managed to identify 62 bacteria at the species level (Supplementary Table 1). The top 10 species and their presence in respective samples are represented in Table 2.

Table 2  
Top 10 bacteria species discovered at Putrajaya Lake and Putrajaya Wetlands Park and presence in respective samples

Bacteria	Sample					
	May 2018			July 2018		
	UN	CW	PL	UN	CW	PL
<i>Flavobacterium columnare</i>	/	/	/	/	/	X
<i>Sphingomonas changbaiensis</i>	X	/	/	/	/	/
<i>Acidovorax delafieldii</i>	/	/	/	/	/	/
<i>Flavobacterium succinicans</i>	/	/	/	/	/	X
<i>Prostheco bacter debontii</i>	/	/	/	/	/	/
<i>Candidatus Aquiluna rubra</i>	/	/	/	/	/	/
<i>Sphingomonas wittichii</i>	/	/	/	/	/	X
<i>Roseateles depolymerans</i>	/	/	/	/	/	/
<i>Sphingomonas yabuuchiae</i>	/	/	/	/	/	/
<i>Pseudomonas nitroreducens</i>	/	/	/	/	/	/

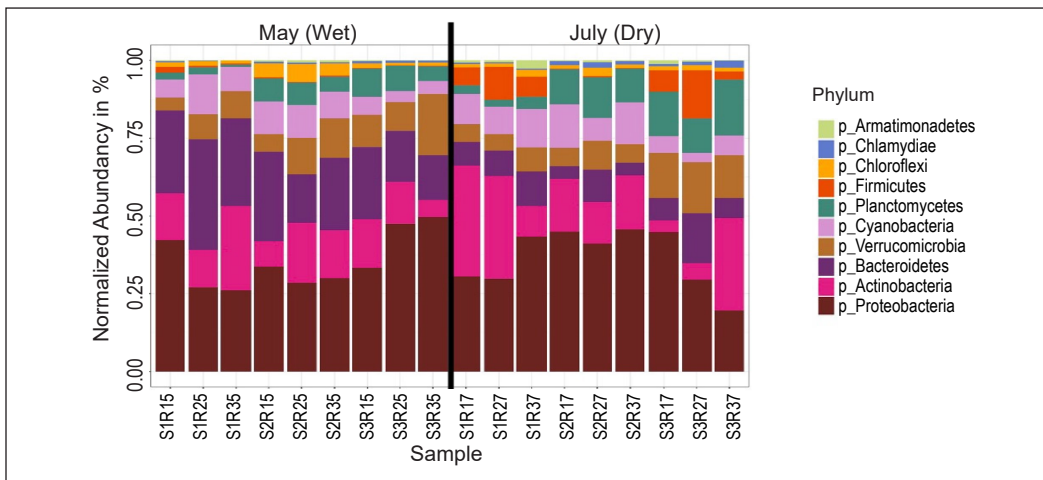


Figure 7. Phylum level distribution (%) of the bacterial community in Putrajaya Lake (S1: UN, S2: CW, S3: PLf, 5: Wet, 7: Dry)

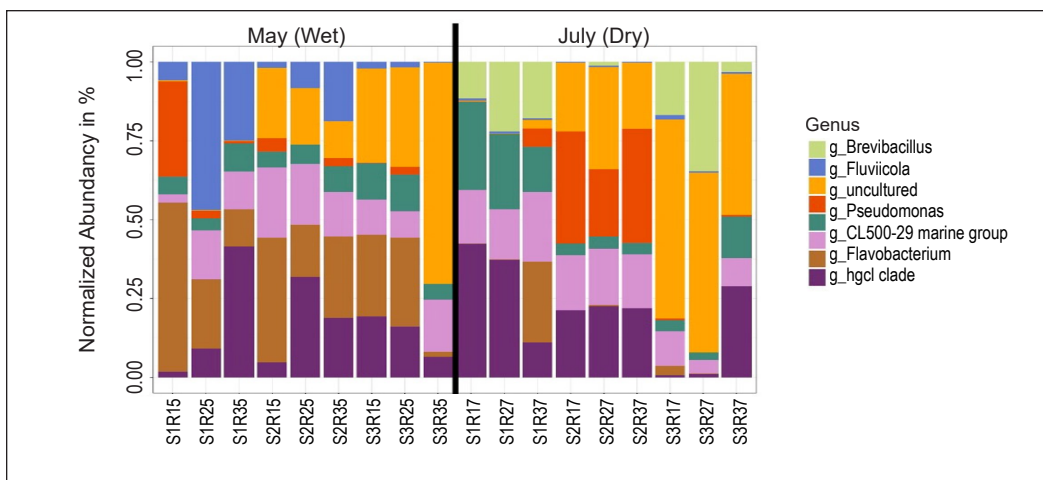


Figure 8. Genus level distribution (%) of the bacterial community in Putrajaya Lake (S1: UN, S2: CW, S3: PLf, 5: Wet, 7: Dry)

### Relationships Between Bacterial Community Composition and Physico-chemical Properties

A PCA analysis examined correlations among the physicochemical parameters and sampling sites (Figure 9). The cumulative total variability of the data was 64.36%, explained by the first and second eigenvalues. PCA analysis was performed prior to the analysis of Canonical correspondence analysis (CCA) to extract precise parameters that affect bacterial distribution in Putrajaya Lake. PCA biplot shows both active variables (physicochemical parameters) as well as active observations (months and sampling sites).

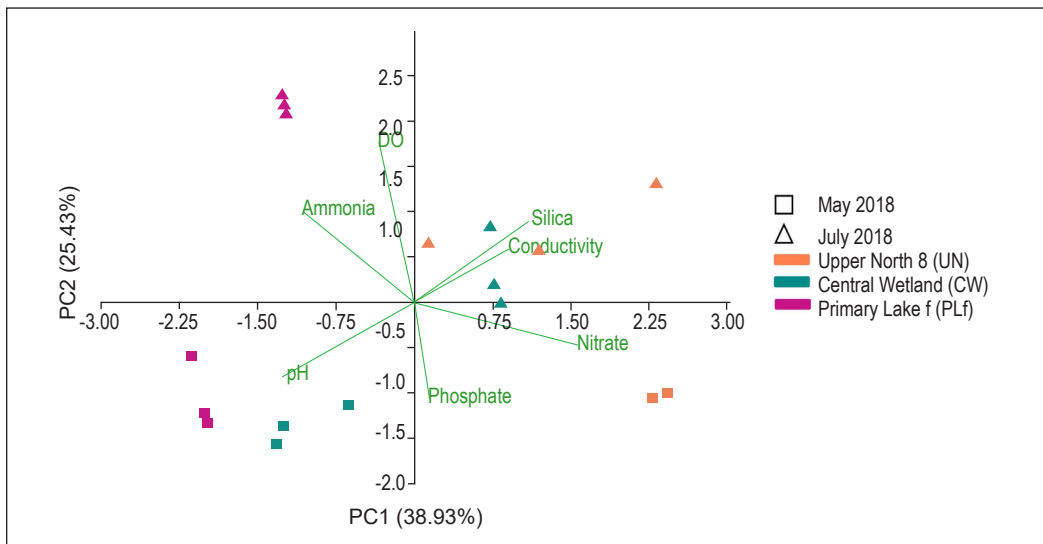


Figure 9. Principal component analysis (PCA) illustrates the correlation between variables and sites  
 Note. May: wet period. July: dry period

There was a positive correlation between nitrate and conductivity (0.52). A negative correlation was found between ammonia and nitrate (-0.72) and pH with nitrate (-0.64). pH was also negatively correlated with silica (-0.57). The post-rain fall condition (May samples) was more associated with pH, which can be observed in the central wetland and the primary lake.

ANOSIM analysis revealed significant differences in the bacterial communities among the three sites ( $R = 0.793$ ,  $P = 0.001$ ) (Figure 10). A one-way ANOSIM accepted the null hypothesis that there was a significant difference in total community structure at the level of individual sequences between sites. This finding may demonstrate a correlation between physicochemical parameters and bacterial composition.

The data in the box are the distances between and within groups R-value range (-1, 1). An R-value close to 0 represents no significant differences between and within groups, and an R-value close to 1 shows that within-group differences are greater than between-group differences. Boxes represent

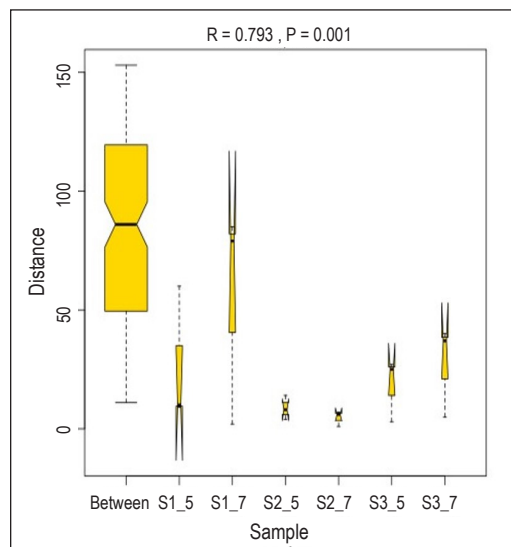


Figure 10. ANOSIM analysis results. Between represents the difference between groups (S1: UN, S2: CW, S3: PLf). Others are within groups 5: wet and 7: dry



the interquartile range between the first and third quartiles, while the line inside the box defines the median.

Pearson correlation revealed that Verrumicrobia and Firmicutes had a strong positive correlation with ammonia-nitrogen ( $r = 0.709$ ). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with  $r$  value ( $r = 0.673$ ) and ( $r = 0.647$ ), respectively, while Bacteroidetes showed a positive association with temperature ( $r = 0.64$ ) (Table 3).

CCA ordination plot (Figure 11) indicates the distribution of bacteria with the availability of physicochemical features at a site. This study’s physicochemical parameters

Table 3  
Pearson’s correlation coefficient analysis of selected physicochemical variables and bacteria diversity from metagenomics data

Variables	pH	Nitrate	Phosphate	Ammonia	Temperature
<i>Proteobacteria</i>	0.44	-0.426	-0.199	-0.311	<b>-0.15</b>
<i>Actinobacteria</i>	-0.561	<b>0.673</b>	0.443	-0.205	<b>-0.07</b>
<i>Bacteroidetes</i>	0.144	0.266	0.424	-0.254	<b>0.64</b>
<i>Verrumicrobia</i>	<b>0.597</b>	<b>-0.781</b>	-0.327	<b>0.709</b>	<b>-0.32</b>
<i>Cyanobacteria</i>	<b>-0.612</b>	<b>0.647</b>	0.006	<b>-0.788</b>	<b>-0.12</b>
<i>Planctomycetes</i>	-0.368	0.446	-0.346	-0.194	<b>-0.83</b>
<i>Firmicutes</i>	-0.324	0.028	-0.152	<b>0.709</b>	<b>-0.41</b>
<i>Chloroflexi</i>	-0.031	0.027	0.145	-0.162	<b>-0.01</b>
<i>Acidobacteria</i>	-0.113	0.325	<b>0.547</b>	-0.202	<b>0.26</b>
<i>Chlamydiae</i>	-0.174	-0.155	<b>-0.615</b>	-0.214	<b>-0.65</b>

Note. Values in bold significance level  $\alpha=0.05$

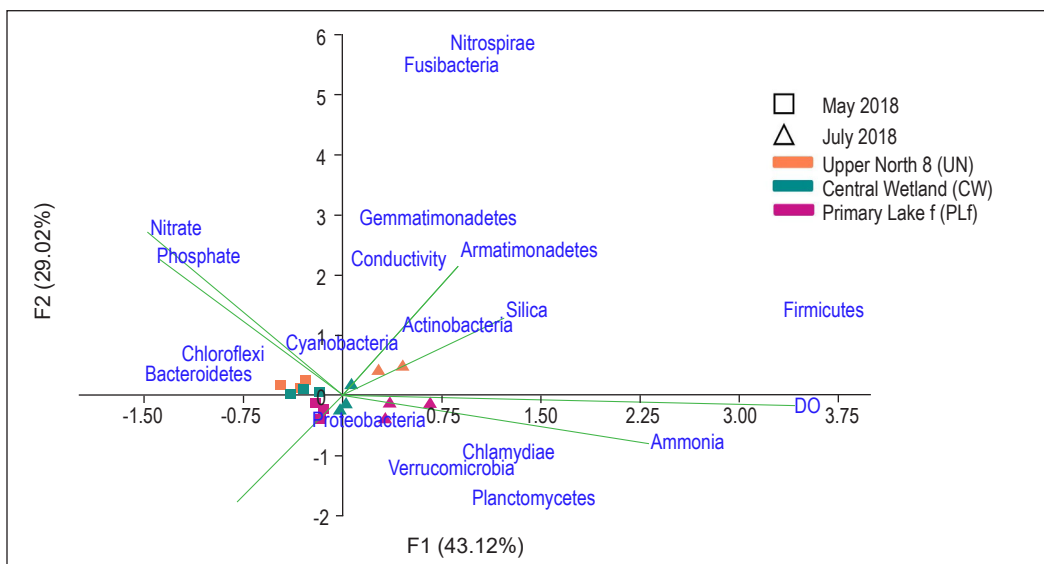


Figure 11. CCA ordination plot illustrating the correlation between variables and bacterial composition. May: wet period, July: dry period

and rainfall impacted the bacterial community composition. Overall, the first two axes explained 72.14% of the total variation. The diversity of bacteria among the samples was more clustered within the different sampling months than the sites. Samples from May were more driven by the nitrate and phosphate concentration, particularly associated with the community of Chloroflexi, Bacteroidetes and Cyanobacteria. Meanwhile, those from August were more influenced by the availability of silica and ammonia.

## DISCUSSION

The advancement in sequencing methods through Next-Generation Sequencing (NGS) has been adopted as the main standard for bacterial identification, replacing the conventional culturing method and first-generation sequencing (Sanger sequencing). Implementing NGS for bacteria community profiling of freshwater lakes and water bodies in Malaysia has become increasingly popular as it provides much information for various research fields. A previous study on a tropical man-made lake in Malaysia, Lake Temenggor, discovered the influence of aquaculture activities on water quality by discovering the functional capacity of the bacteria profile through shotgun metagenomics (Lau et al., 2019). The same method was also carried out in a study for endophytic bacteria identification in Kenyir Lake, which may provide insightful information for antimicrobial properties among medicinal plants (Abidin et al., 2020). Bacterial communities are responsible for the biogeochemical cycling of nutrients in the ecosystem, particularly in the freshwater environment. This study was the first attempt to characterise the bacterial community based on molecular data in Putrajaya Lake within the dry season of May after rainfall events (wet) and in July (dry). The wet condition showed higher proportions of bacterial diversity than the dry condition, though they were not significantly different.

In eutrophic lakes, high phytoplankton concentration produces large amounts of organic compounds through photosynthesis and releases oxygen into the water (Viet et al., 2016). The dissolved oxygen is inversely proportional to the water temperature in the lake, which might explain the lower DO readings in the water inlet. However, bacteria that decompose phytoplankton organic compounds also use oxygen, causing net respiration (Kragh et al., 2020). Thus, the respiration of phytoplankton and bacteria which consume the oxygen can cause a decrease in the dissolved oxygen concentration in eutrophic lakes. Conductivity was also commonly used as a lake water quality indicator, which is proportional to the water temperature - a higher conductivity value indicates higher ionic concentrations. The pH of Putrajaya Lake is neutral to slightly alkaline. The alluvium soil, which can be found at the lake bed, consists of silty to sandy clay, which is low in bicarbonate, which may explain the overall neutral pH (Jaapar et al., 2002). The pH can fluctuate rapidly in aquatic ecosystems. High photosynthesis removes carbon dioxide and bicarbonate from the water, increasing carbonate concentrations and pH. The bacterial community may not adapt optimally to altered pH conditions during rapid pH changes, leading to impaired functions (Viet et al., 2016).

Nitrate and phosphate are among the chemical parameters used to assess water quality. The value of nitrate and silica decreased significantly from the water inlet near the Upper North 8 site (UN) to the central wetlands (CW), which may reflect the bioremediation impact of wetlands. In the central wetlands, plants associated with microbes (holobionts) could successfully reduce nitrate and silica concentrations. Holobionts can absorb nutrients such as nitrate, which cycles through several transformation processes, including fixation, ammonification, nitrification, and denitrification conducted by bacteria (Vymazal, 2007). Holobionts are chosen based on soil type and water level in the Putrajaya wetland area. *Hanguana malayana* (Bakong) and *Phragmites karka* (Rumput gedabong) are the most suitable plants found to be grown in wetland areas from 1998 until now (Mohamad, 2012). Silica is an essential element for the growth of diatoms. Aquatic macrophytes like *Phragmites australis* (reed) can uptake, translocate and accumulate silica in wetland areas (Struyf et al., 2004). This species can temporarily store silica, making it unavailable for diatoms. Ammonia-nitrogen can be used by phytoplankton or oxidised by bacteria into nitrate or nitrite for their growth and metabolism (Glibert et al., 2016). The higher concentrations of ammonia-nitrogen were influenced by the possible discharge of domestic sewage, agricultural activities and animal waste (Huang et al., 2015).

### **Importance and Application of Bacterial Species Found in Putrajaya Lake and Putrajaya Wetlands Park**

Proteobacteria were more common in the dry condition than the wet condition based on LEfSe analysis. Soil bacteria such as Proteobacteria are important drivers of energy flow and nutrient cycling from the physical environment into living organisms in lake ecosystems (Wan et al., 2017). High abundances of Proteobacterial groups have previously been related to the higher availability of carbon (Fierer et al., 2007). Photosynthesis by aquatic plants and phytoplankton is a major carbon source in lakes. Aquatic ecosystems also get carbon sources through the runoff from the land nearby. In this study, the availability of carbon in the neutral pH range to slightly more alkaline might have contributed to the change in the relative abundance of Proteobacteria, which include *Pseudomonas* sp., which might have originated from non-point sources such as agricultural and domestic. *Pseudomonas* sp. is able to degrade organic compounds into simpler molecules in a variety of ecological conditions (Batrich et al., 2019). The availability of these microbes in the constructed wetland may be involved in nutrient cycling, filtering and removing pollutants coming into the lake.

The eutrophication of the lake enriched with sediments is preferred by nutrient-cycling bacteria such as Bacteroidetes, Firmicutes, Alphaproteobacteria and Gammaproteobacteria (Huang et al., 2017). *Sphingomonas* sp., an example of the Alphaproteobacteria, were abundant and capable of surviving in various environments. *Sphingomonas* are known for degrading hydrocarbons, which can aid in bioremediation. The *Sphingomonas* species

discovered within Putrajaya Lake and Putrajaya Wetlands Park have been previously reported to be capable of degrading hydrocarbons to some degree. It allows for the degradation of pollutants such as petrol spills and other oil-based spills that are potentially harmful to the ecosystem. *Sphingomonas changbaiensis*, as an example, is often added as parts of bacterial consortia used together with biosurfactants such as alkyl polyglycoside, which enhances the rate of degradation of total petroleum hydrocarbons (Li et al., 2020). *Sphingomonas wittichii*, specifically strain RW1, is a highly-interest bacterium due to its ability to degrade dioxin-based pollutants such as polychlorinated (Colquhoun et al., 2012).

The highest operational taxonomic unit (OTU) number assigned at the species level in this study was a Gram-negative Bacteroidetes under the class Flavobacteria. *Flavobacterium* sp. is a ubiquitous fish pathogen genus, notably within freshwater habitats. It can cause infections in different parts of fish (Loch & Faisal, 2015). The *Flavobacterium columnare* is responsible for a prevalent fish disease known as columnaris or cotton mouth disease, which has been reported to infect fishes in wild habitats and cultures. Previous studies reported infection by this bacterial pathogen at a global scale among various fish species (Davis, 2011), such as trout (Singh et al., 2021), red tilapia (Ponpukdee et al., 2021), Asian seabass (Chokmangmeepisarn et al., 2021) and catfish (Lange et al., 2021). This species has also been reported to interact with the parasitic protozoan *Chilodonella hexasticha* to cause fish death. This ciliate is correlated with the abundant presence of *F. columnare* (Gomes et al., 2019).

Another species, *F. succinicans*, is associated with gill disease among rainbow trout (Good et al., 2015). In Malaysia, the presence of *Flavobacterium* sp. is widely spread in freshwater habitats to the extent that some may have developed antibiotic resistance (Hassan et al., 2020). A study among the ornamental fish industry in Malaysia has also found the presence of this genus infecting fishes sold in local retail pet shops (Anjur et al., 2021). The treatment for *Flavobacterium* infection is currently by administering vaccines or antibiotics to the culture (Lange et al., 2019). Though these species are not a concern for human health, more strains have become resistant to the treatments (Elgendy et al., 2022), causing a massive economic loss in the aquaculture sector (LaFrentz et al., 2018).

The distribution of the *Flavobacterium* population within the aquatic habitat is reportedly associated with heterotrophic activity within the environment, favouring the abundance of available resources (Newton et al., 2011). The ability to survive and quickly adapt to high nutrient concentrations might explain the high OTUs assigned to *Flavobacterium* reported in this study across all sampling sites.

Ecology management services by the Putrajaya Corporation (PjC) encompass maintenance activities, monitoring exercises and biodiversity assessment. This research will contribute towards the monitoring exercises of the physical and chemical properties of water. The report on the diversity of bacterial communities from the water input site (Upper North: UN8), remediation site (Central Wetland: CW), as well as the main water

sports activities centre (Primary Lake f: PLf), provide specific snapshots of spatiotemporal bacterial community data which will contribute towards Putrajaya's biodiversity and ecological assessment. Water flowing from the Upper North to the primary lake area showed a marked decrease in certain nutrient concentrations, providing evidence of the effectiveness of the engineered wetlands of Putrajaya.

### **Implications for Remediation Strategies**

As a constructed lake that is continuously exposed to various anthropogenic pollutants coming from untreated water of the two main rivers, there is a need to develop efficient strategies to ensure that the Putrajaya Lake and its surrounding water inputs are of good quality, on par with the level set by the National Water Quality Standard for Malaysia (NWQSM) and maintained at pristine condition. One of the major issues in Putrajaya Lake is the accidental release of nutrients in runoffs, such as landscape fertilisers and pesticides, which cause the water clarity to decrease and result in consistent slight eutrophic conditions (Hakim et al., 2016). The lake also serves as a major tourism site for water-based activities. Hence, a strict monitoring program is essential to sustain and maintain the water quality. Pollutants come from point and non-point sources, inside and outside the boundaries of Putrajaya Lake; thus, the multiple stakeholders' management is essential, in addition to the local government. Therefore, a holistic management approach needs to be emphasised by all stakeholders.

### **CONCLUSION**

The use of the 16S rRNA gene as the genetic marker was successful in identifying bacteria species. Overall, 49 phyla, 147 classes, 284 orders, 471 families, 778 genera and 62 species of bacteria were identified. Proteobacteria was the most dominant bacteria class found in all samples, while *Flavobacterium columnare* of Bacteroidetes was the most common bacteria to be identified at the species level. Verrucomicrobia and Firmicutes showed a strong positive correlation with ammonia-nitrogen ( $r = 0.709$ ). Actinobacteria and Cyanobacteria had a moderate positive correlation with nitrate with  $r$  value ( $r = 0.673$ ) and ( $r = 0.647$ ), respectively. This report has been the first to employ the 16S rRNA gene amplicons to evaluate the bacterial community assemblages of Putrajaya Lake, Malaysia. We have found that the presence of the engineered wetland plantings was influential in nutrient removal, as shown by the significant reduction in nitrate and silica concentration, which might have impacted the diversity of bacteria found within the different sampling sites. For future research, other wetland arms, such as the Upper Bisa (UB), Upper West (UW) and Upper East (UE), will be considered for sampling sites. Engineered wetlands have benefited the natural bioremediation system of lake water inlets and could be emulated in other cities for environmentally friendly water management.

## ACKNOWLEDGEMENTS

The authors would like to acknowledge the contribution and constructive criticisms of Linda Graham of University of Wisconsin, Madison, Wisconsin, USA, in writing this paper. The authors would also like to acknowledge the Ministry of Higher Education Malaysia and Universiti Putra Malaysia for funding this project. This project was financially supported by MOHE Malaysia FRGS (FRGS/1/2020/STG03/UPM/02/9), PUTRA IPM grant (GP/2017/9568100) and UPM PUTRA IPS grant (GP-IPS/2018/9664300).

## REFERENCES

- Abidin, Z. A. Z., Zolkefli, N., Shahari, R., & Chowdhury, A. J. K. (2020). Endophytic bacteria of *Anisophyllea disticha* (Raja berangkat) from tropical lake environment in Malaysia. *Desalination and Water Treatment*, 188, 282-287. <https://doi.org/10.5004/dwt.2020.25274>
- Al-Sulaiman, S. A. A. W. (2012). Chemical safety of drinking-water: Assessing priorities for risk management. *International Journal of Environmental Studies*, 69(6), 1001. <https://doi.org/10.1080/00207233.2011.565947>
- Anjur, N., Sabran, S. F., Daud, H. M., & Othman, N. Z. (2021). An update on the ornamental fish industry in Malaysia: *Aeromonas hydrophila*-associated disease and its treatment control. *Veterinary World*, 14(5), 1143-1152. <https://doi.org/10.14202/vetworld.2021.1143-1152>
- APHA. (2005). *Standard Methods for the Examination of Water and Wastewater* (21st ed.). Washington DC: American Public Health Association.
- Asmat, A., Hazali, N. A., Nor, A. N. M., & Zuhan, F. K. (2018). Seasonal-spatial of Putrajaya Lake Water Quality Parameter (WQP) concentration using geographic information system (GIS). *International Journal of Engineering & Technology*, 7(3), 176-181. <https://doi.org/10.14419/ijet.v7i3.11.15956>
- Batrach, M., Maskeri, L., Schubert, R., Ho, B., Kohout, M., Abdeljaber, M., Abuhasna, A., Kholoki, M., Psihogios, P., Razzaq, T., Sawhney, S., Siddiqui, S., Xoubi, E., Cooper, A., Hatzopoulos, T., & Putonti, C. (2019). Pseudomonas diversity within urban freshwaters. *Frontiers in Microbiology*, 10, Article 195. <https://doi.org/10.3389/fmicb.2019.00195>
- Briones, A., & Raskin, L. (2003). Diversity and dynamics of microbial communities in engineered environments and their implications for process stability. *Current Opinion in Biotechnology*, 14(3), 270-276. [https://doi.org/10.1016/S0958-1669\(03\)00065-X](https://doi.org/10.1016/S0958-1669(03)00065-X)
- Chen, Z., Hui, P. C., Hui, M., Yeoh, Y. K., Wong, P. Y., Chan, M. C. W., Wong, M. C. S., Ng, S. C., Chan, F. K. L., & Chan, P. K. S. (2019). Impact of preservation method and 16S rRNA hypervariable region on gut microbiota profiling. *MSystems*, 4(1), 1-15. <https://doi.org/10.1128/msystems.00271-18>
- Chokmangmeepisarn, P., Thangsunan, P., Kayansamruaj, P., & Rodkhum, C. (2021). Resistome characterization of *Flavobacterium columnare* isolated from freshwater cultured Asian sea bass (*Lates calcarifer*) revealed diversity of quinolone resistance associated genes. *Aquaculture*, 544, Article 737149. <https://doi.org/10.1016/j.aquaculture.2021.737149>
- Clarke, K. R., & Gorley, R. N. (2006). *Primer v. 6: User manual/Tutorial*. Scientific Research Publishing Inc. [https://www.scirp.org/\(S\(vtj3fa45qm1ean45vfvfcz55\)\)/reference/ReferencesPapers.aspx?ReferenceID=1960070](https://www.scirp.org/(S(vtj3fa45qm1ean45vfvfcz55))/reference/ReferencesPapers.aspx?ReferenceID=1960070)



- Colquhoun, D. R., Hartmann, E. M., & Halden, R. U. (2012). Proteomic profiling of the dioxin-degrading bacterium *Sphingomonas wittichii* RW1. *Journal of Biomedicine and Biotechnology*, 2012, Article 408690. <https://doi.org/10.1155/2012/408690>
- Cotner, J. B., & Biddanda, B. A. (2002). Small players, large role: Microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems*, 5(2), 105-121. <https://doi.org/10.1007/s10021-001-0059-3>
- Davis, H. S. (2011). *A New Bacterial Disease of Fresh-Water Fishes*. US Government Printing Office. <https://doi.org/10.5962/bhl.title.49773>
- Edgar, R. C. (2013). UPARSE: Highly accurate OTU sequences from microbial amplicon reads. *Nature Methods*, 10, 996-998. <https://doi.org/10.1038/nmeth.2604>
- Elgendy, M. Y., Abdelsalam, M., Mohamed, S. A., & Ali, S. E. (2022). Molecular characterization, virulence profiling, antibiotic susceptibility, and scanning electron microscopy of *Flavobacterium columnare* isolates retrieved from Nile tilapia (*Oreochromis niloticus*). *Aquaculture International*, 30, 845-862. <https://doi.org/10.1007/s10499-021-00819-x>
- Fadrosh, D. W., Ma, B., Gajer, P., Sengamalay, N., Ott, S., Brotman, R. M., & Ravel, J. (2014). An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome*, 2(6), 1-7. <https://doi.org/10.1186/2049-2618-2-6>
- Fierer, N., Bradford, M. A., & Jackson, R. B. (2007). Toward an ecological classification of soil bacteria. *Ecology*, 88(6), 1354-1364. <https://doi.org/10.1890/05-1839>
- Flynn, J. M., Brown, E. A., Chain, F. J. J., Macisaac, H. J., & Cristescu, M. E. (2015). Toward accurate molecular identification of species in complex environmental samples: Testing the performance of sequence filtering and clustering methods. *Ecology and Evolution*, 5(11), 2252-2266. <https://doi.org/10.1002/ece3.1497>
- Gane, A. (2018). *Size selection brings better data to NGS workflows*. Cytiva. <https://www.cytivalifesciences.com/en/us/news-center/better-data-via-size-selection-10001>
- Glibert, P. M., Wilkerson, F. P., Dugdale, R. C., Raven, J. A., Dupont, C. L., Leavitt, P. R., Parker, A. E., Burkholder, J. M., & Kana, T. M. (2016). Pluses and minuses of ammonium and nitrate uptake and assimilation by phytoplankton and implications for productivity and community composition, with emphasis on nitrogen-enriched conditions. *Limnology and Oceanography*, 61(1), 165-197. <https://doi.org/10.1002/lno.10203>
- Gomes, G. B., Hutson, K. S., Domingos, J. A., Villamil, S. I., Huerlimann, R., Miller, T. L., & Jerry, D. R. (2019). Parasitic protozoan interactions with bacterial microbiome in a tropical fish farm. *Aquaculture*, 502, 196-201. <https://doi.org/10.1016/j.aquaculture.2018.12.037>
- Good, C., Davidson, J., Wiens, G. D., Welch, T. J., & Summerfelt, S. (2015). *Flavobacterium branchiophilum* and *F. succinicans* associated with bacterial gill disease in rainbow trout *Oncorhynchus mykiss* (Walbaum) in water recirculation aquaculture systems. *Journal of Fish Diseases*, 38(4), 409-413. <https://doi.org/10.1111/jfd.12249>
- Gupta, V., Ratha, S. K., Sood, A., Chaudhary, V., & Prasanna, R. (2013). New insights into the biodiversity and applications of cyanobacteria (blue-green algae) - Prospects and challenges. *Algal Research*, 2(2), 79-97. <https://doi.org/10.1016/j.algal.2013.01.006>

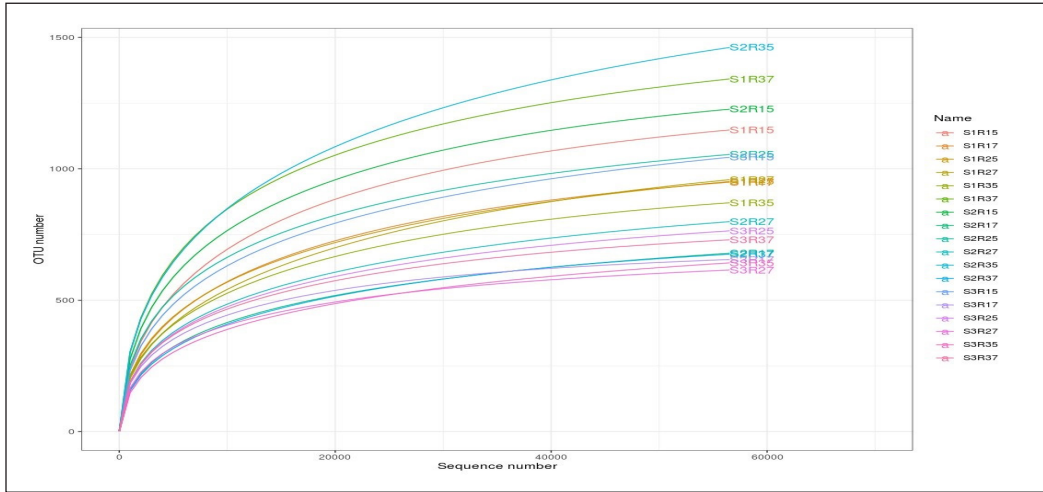
- Hach, C. (2002). *Water Analysis Handbook*. Loveland Publishers & Pioneers.
- Hakim, M., Heryansyah, A., & Ismail, T. (2016, June 7-8). Sustainability in water resources management in Putrajaya wetland: Impact assessment on new emerging pollutants. In *1st Proceeding of Civil Engineering* (pp. 301-317). Universiti Teknologi Malaysia, Malaysia.
- Hassan, N. L., Kam, K. Y., & Zain, N. A. M. (2020). Isolation of antibiotic resistant bacteria from rivers in Kelantan, Malaysia. *International Journal of Life Sciences and Biotechnology*, 3(2), 291-307. <https://doi.org/10.38001/ijlsb.712542>
- Huang, W., Chen, X., Jiang, X., & Zheng, B. (2017). Characterization of sediment bacterial communities in plain lakes with different trophic statuses. *MicrobiologyOpen*, 6(5), 1-14. <https://doi.org/10.1002/mbo3.503>
- Huang, Y. F., Ang, S. Y., Lee, K. M., & Lee, T. S. (2015). Quality of water resources in Malaysia. In *Research and Practices in Water Quality* (pp. 65-94). InTech. <https://doi.org/10.5772/58969>
- Iliev, I., Yahubyan, G., Marhova, M., Apostolova, E., Gozmanova, M., Gecheva, G., Kostadinova, S., Ivanova, A., & Baev, V. (2017). Metagenomic profiling of the microbial freshwater communities in two Bulgarian reservoirs. *Journal of Basic Microbiology*, 57(8), 669-679. <https://doi.org/10.1002/jobm.201700137>
- Jaapar, A. R., Ahba, N. A., & Hussin, A. (2002, May 26). The occurrence and classification of hard rock body in Putrajaya and its implication to construction activities. *Bulletin of the Geological Society of Malaysia*, 45, 123-128. <https://doi.org/10.7186/bgsm45200218>
- Jamal, A., Yusoff, F. M., Banerjee, S., & Shariff, M. (2014). Littoral and limnetic phytoplankton distribution and biodiversity in a tropical man-made lake, Malaysia. *Advanced Studies in Biology*, 6(4), 149-168. <http://dx.doi.org/10.12988/asb.2014.4631>
- Jo, J. H., Kennedy, E. A., & Kong, H. H. (2016). Research techniques made simple: Bacterial 16S ribosomal RNA gene sequencing in cutaneous research. *Journal of Investigative Dermatology*, 136(3), e23-e27. <https://doi.org/10.1016/j.jid.2016.01.005>
- Kragh, T., Martinsen, K. T., Kristensen, E., & Sand-Jensen, K. (2020). From drought to flood: Sudden carbon inflow causes whole-lake anoxia and massive fish kill in a large shallow lake. *Science of the Total Environment*, 739, Article 140072. <https://doi.org/10.1016/j.scitotenv.2020.140072>
- Kuczynski, J., Stombaugh, J., Walters, W. A., González, A., Caporaso, J. G., & Knight, R. (2012). Using QIIME to analyze 16s rRNA gene sequences from microbial communities. *Current Protocols in Microbiology*, 27(1), 1E.5.1-1E.5.20. <https://doi.org/10.1002/9780471729259.mc01e05s27>
- LaFrentz, B. R., García, J. C., Waldbieser, G. C., Evenhuis, J. P., Loch, T. P., Liles, M. R., Wong, F. S., & Chang, S. F. (2018). Identification of four distinct phylogenetic groups in *Flavobacterium columnare* with fish host associations. *Frontiers in Microbiology*, 9, Article 452. <https://doi.org/10.3389/fmicb.2018.00452>
- Lange, M. D., Abernathy, J., & Farmer, B. D. (2019). Evaluation of a recombinant flavobacterium columnare DnaK protein vaccine as a means of protection against columnaris disease in channel catfish (*Ictalurus punctatus*). *Frontiers in Immunology*, 10, Article 1175. <https://doi.org/10.3389/fimmu.2019.01175>
- Lange, M. D., Abernathy, J., Farmer, B. D., & Beck, B. H. (2021). Use of an immersion adjuvant with a *Flavobacterium columnare* recombinant protein vaccine in channel catfish. *Fish and Shellfish Immunology*, 117, 136-139. <https://doi.org/10.1016/j.fsi.2021.07.025>



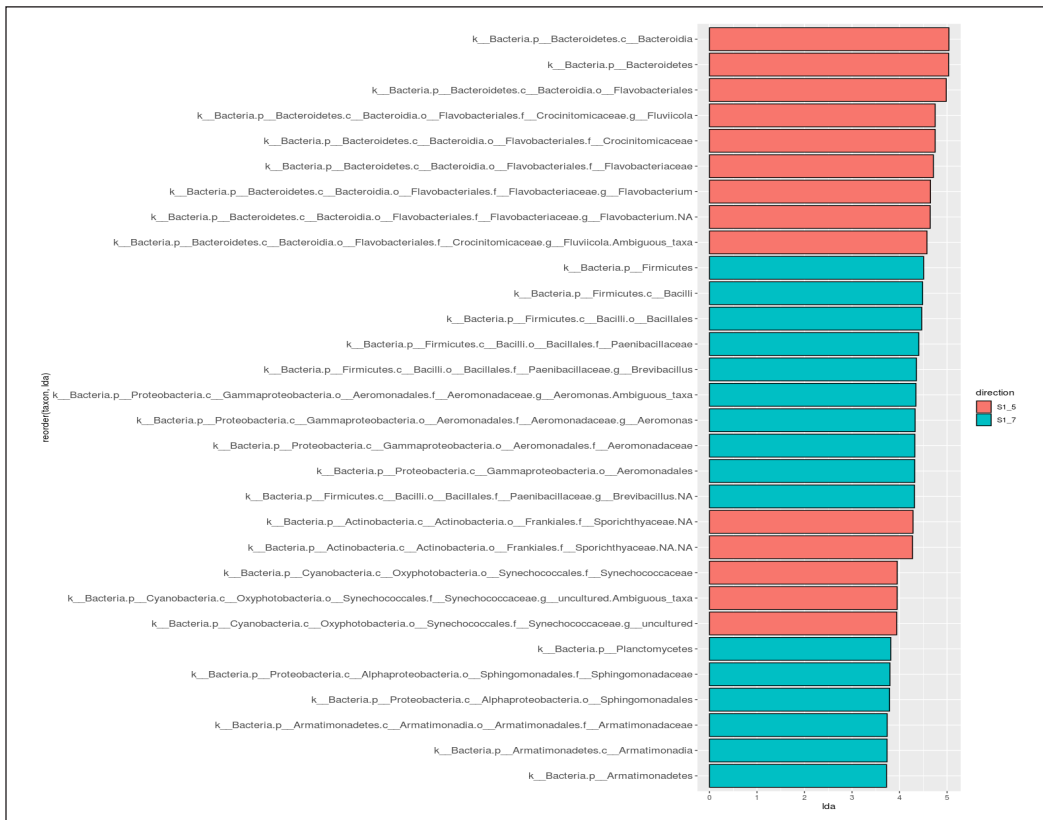
- Lau, N. S., Zarkasi, K. Z., Sah, A. S. R. M., & Shu-Chien, A. C. (2019). Diversity and coding potential of the microbiota in the photic and aphotic zones of tropical man-made lake with intensive aquaculture activities: A case study on Temengor Lake, Malaysia. *Microbial Ecology*, 78, 20-32. <https://doi.org/10.1007/s00248-018-1283-0>
- Li, Q., Huang, Y., Wen, D., Fu, R., & Feng, L. (2020). Application of alkyl polyglycosides for enhanced bioremediation of petroleum hydrocarbon-contaminated soil using *Sphingomonas changbaiensis* and *Pseudomonas stutzeri*. *Science of the Total Environment*, 719, Article 137456. <https://doi.org/10.1016/j.scitotenv.2020.137456>
- Liles, M. R., Manske, B. F., Bintrim, S. B., Handelsman, J., & Goodman, R. M. (2003). A census of rRNA genes and linked genomic sequences within a soil metagenomic library. *Applied and Environmental Microbiology*, 69(5), 2684-2691. <https://doi.org/10.1128/AEM.69.5.2684-2691.2003>
- Loch, T. P., & Faisal, M. (2015). Emerging flavobacterial infections in fish: A review. *Journal of Advanced Research*, 6(3), 283-300. <https://doi.org/10.1016/j.jare.2014.10.009>
- Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J., & Knight, R. (2011). UniFrac: An effective distance metric for microbial community comparison. *The ISME Journal*, 5, 169-172. <https://doi.org/10.1038/ismej.2010.133>
- Majizat, A., Ahmad, B., Noordin, N., & Sharip, Z. (2016). Adopting an integrated catchment-based management approach for Putrajaya Lake and Wetlands. *International Journal of Water*, 10(4), 343-358. <https://doi.org/10.1504/IJW.2016.079701>
- Mohamad, N. H. N. (2012). Ecological approaches in designing neighbourhood green spaces as urban wildlife habitat in the Klang valley, Peninsular Malaysia. *International Journal of Applied Science and Technology*, 2(3), 192-213.
- Mueller-Spitz, S. R., Goetz, G. W., & McLellan, S. L. (2009). Temporal and spatial variability in nearshore bacterioplankton communities of Lake Michigan. *FEMS Microbiology Ecology*, 67(3), 511-522. <https://doi.org/10.1111/j.1574-6941.2008.00639.x>
- Najah, A., Teo, F. Y., Chow, M. F., Huang, Y. F., Latif, S. D., Abdullah, S., Ismail, M., & El-Shafie, A. (2021). Surface water quality status and prediction during movement control operation order under COVID-19 pandemic: Case studies in Malaysia. *International Journal of Environmental Science and Technology*, 18, 1009-1018. <https://doi.org/10.1007/s13762-021-03139-y>
- Newton, R. J., Jones, S. E., Eiler, A., McMahan, K. D., & Bertilsson, S. (2011). A guide to the natural history of freshwater lake bacteria. *Microbiology and Molecular Biology Reviews*, 75(1), 14-49. <https://doi.org/10.1128/membr.00028-10>
- Oksanen, J. (2008). Vegan: An Introduction to Ordination. <https://cran.r-project.org/web/packages/vegan/vignettes/intro-vegan.pdf>
- Pernthaler, J., & Amann, R. (2005). Fate of heterotrophic microbes in pelagic habitats: Focus on populations. *Microbiology and Molecular Biology Reviews*, 69(3), 440-461. <https://doi.org/10.1128/membr.69.3.440-461.2005>
- Ponpukdee, N., Wangman, P., Rodkhum, C., Pengsuk, C., Chaivisuthangkura, P., Sithigorngul, P., & Longyant, S. (2021). Detection and identification of a fish pathogen *Flavobacterium columnare*

- using specific monoclonal antibodies. *Aquaculture*, 545, Article 737231. <https://doi.org/10.1016/j.aquaculture.2021.737231>
- Sabkie, N., Zulkifly, S., Lazim, N., Rosli, N., & Yong, C. (2020). Physico-chemical influence on the diversity of phytoplankton at Putrajaya Lake and Wetlands, Putrajaya, Malaysia. *Malayan Nature Journal*, 72(2), 153-163.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biology*, 12, Article R60. <https://doi.org/10.1186/gb-2011-12-6-r60>
- Sharip, Z., Saman, J. M., Noordin, N., Majizat, A., Suratman, S., & Shaaban, A. J. (2016). Assessing the spatial water quality dynamics in Putrajaya Lake: A modelling approach. *Modeling Earth Systems and Environment*, 2, Article 46. <https://doi.org/10.1007/s40808-016-0104-z>
- Singh, S., Mallik, S. K., Kala, K., Shahi, N., Pathak, R., Giri, A. K., Chandra, S., Pant, K., & Patiyal, R. S. (2021). Characterization of *Flavobacterium columnare* from farmed infected rainbow trout, *Oncorhynchus mykiss* (Walbaum, 1792) of Central Indian Himalayan region, India. *Aquaculture*, 544, Article 737118. <https://doi.org/10.1016/j.aquaculture.2021.737118>
- Sogin, M. L., Morrison, H. G., Huber, J. A., Welch, D. M., Huse, S. M., Neal, P. R., Arrieta, J. M., & Herndl, G. J. (2006). Microbial diversity in the deep sea and the underexplored "rare biosphere." *Proceedings of the National Academy of Sciences of the United States of America*, 103(32), 12115-12120. <https://doi.org/10.1073/pnas.0605127103>
- Staley, C., Unno, T., Gould, T. J., Jarvis, B., Phillips, J., Cotner, J. B., & Sadowsky, M. J. (2013). Application of Illumina next-generation sequencing to characterize the bacterial community of the Upper Mississippi River. *Journal of Applied Microbiology*, 115(5), 1147-1158. <https://doi.org/10.1111/jam.12323>
- Struyf, E., van Damme, S., & Meire, P. (2004). Possible effects of climate change on estuarine nutrient fluxes: A case study in the highly nutrified Schelde estuary (Belgium, the Netherlands). *Estuarine, Coastal and Shelf Science*, 60(4), 649-661. <https://doi.org/10.1016/j.ecss.2004.03.004>
- Suratman, S., Hussein, A. N. A. R., Tahir, N. M., Latif, M. T., Mostapa, R., & Weston, K. (2016). Seasonal and spatial variability of selected surface water quality parameters in Setiu Wetland, Terengganu, Malaysia. *Sains Malaysiana*, 45(4), 551-558.
- Tan, G., Opitz, L., Schlapbach, R., & Rehrauer, H. (2019). Long fragments achieve lower base quality in Illumina paired-end sequencing. *Scientific Reports*, 9, Article 2856. <https://doi.org/10.1038/s41598-019-39076-7>
- Viet, N. D., Bac, N. A., & Huong, H. T. T. (2016). Dissolved oxygen as an indicator for eutrophication in freshwater lakes. In *Proceedings of International Conference on Environmental Engineering and Management for Sustainable Development* (pp. 1-6). ResearchGate.
- Vymazal, J. (2007). Removal of nutrients in various types of constructed wetlands. *Science of the Total Environment*, 380(1-3), 48-65. <https://doi.org/10.1016/j.scitotenv.2006.09.014>
- Wan, Y., Ruan, X., Zhang, Y., & Li, R. (2017). Illumina sequencing-based analysis of sediment bacteria community in different trophic status freshwater lakes. *MicrobiologyOpen*, 6(4), Article e00450. <https://doi.org/10.1002/mbo3.450>

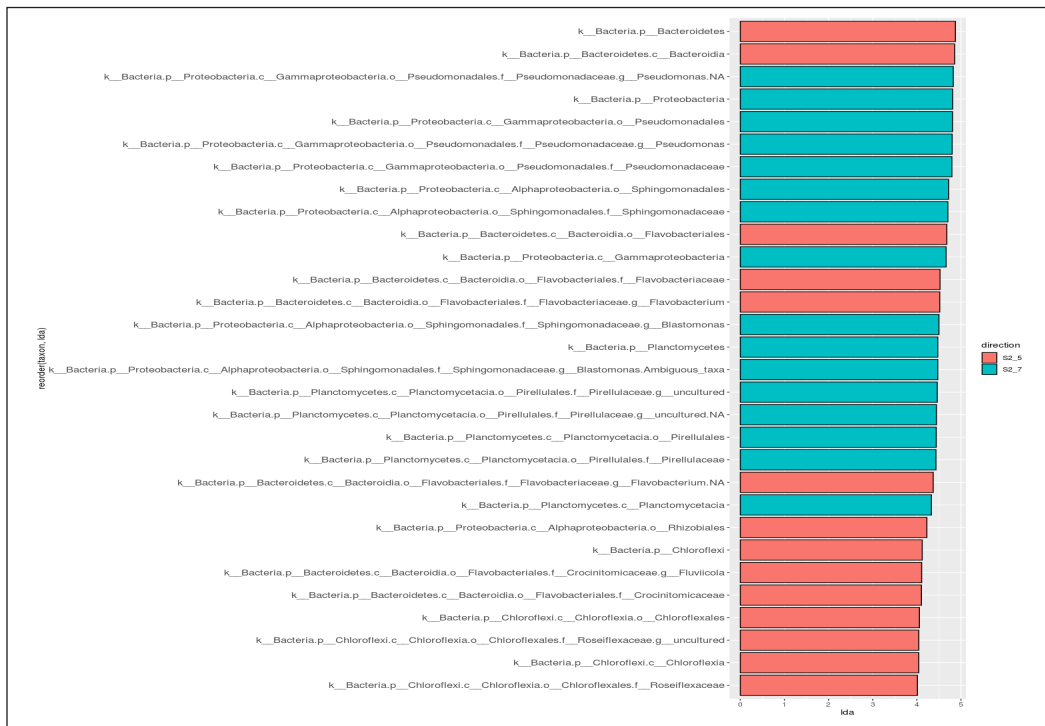
APPENDICES



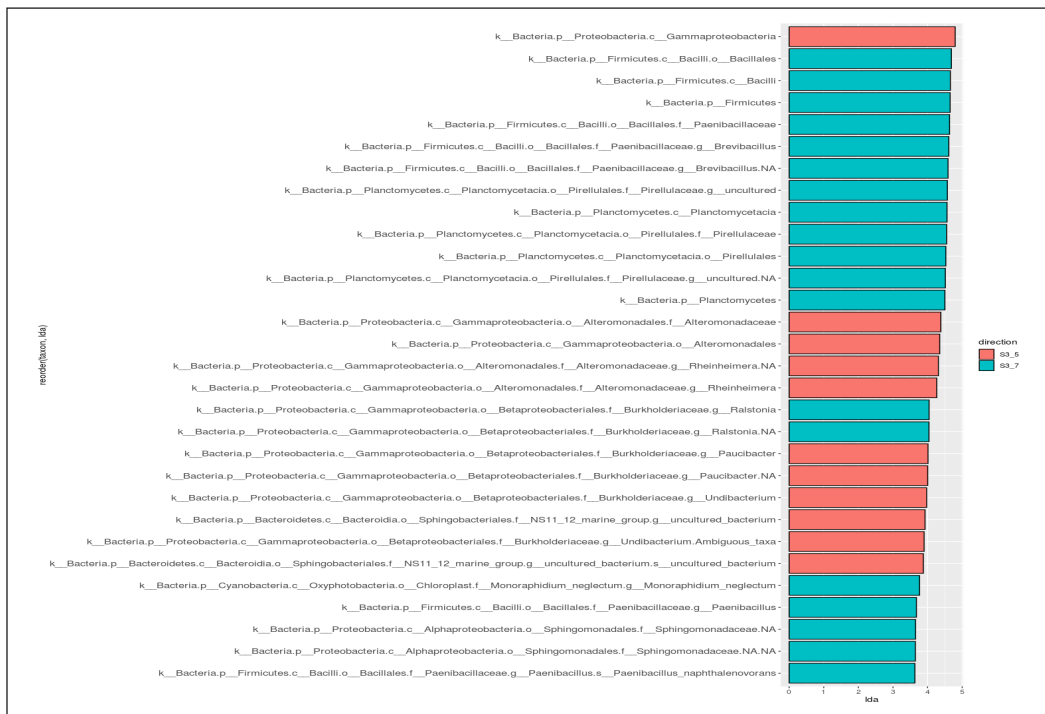
Supplementary Figure 1. Rarefaction curve. Wet season (May), Site 1 (S1R15, S1R25, S1R35); Site 2 (S2R15, S2R25, S2R35) and Site 3 (S3R15, S3R25, S3R35); dry season (July), Site 1 (S1R17, S1R27, S1R37); Site 2 (S2R17, S2R27, S2R37) and Site 3 (S3R17, S3R27, S3R37)



Supplementary Figure 2. Indicator bacterial groups at Site 1 on wet (S1\_5) and dry conditions (S1\_7), respectively



Supplementary Figure 3. Indicator bacterial groups at Site 2 on wet (S2\_5) and dry conditions (S2\_7), respectively



Supplementary Figure 4. Indicator bacterial groups at Site 3 on wet (S3\_5) and dry conditions (S3\_7), respectively

Supplementary Table 1

*List of bacteria identified at the species level*

Phylum	Class	Species
Acidobacteria	Acidobacteriia	<i>Edaphobacter modestum</i>
Actinobacteria	Actinobacteria	<i>Candidatus Aquiluna rubra</i>
		<i>Virgisporangium ochraceum</i>
		<i>Mycobacterium arupense</i>
		<i>Mycobacterium celatum</i>
		<i>Propionibacterium acnes</i>
		<i>Pseudonocardia halophobica</i>
		<i>Bifidobacterium adolescentis</i>
		<i>Bifidobacterium longum</i>
Bacteroidetes	Coriobacteriia	<i>Collinsella aerofaciens</i>
	Saprospirae	<i>Aquirestis calciphila</i>
	Bacteroidia	<i>Bacteroides caccae</i>
		<i>Bacteroides uniformis</i>
		<i>Prevotella copri</i>
		<i>Flexibacter ruber</i>
	Cytophagia	<i>Elizabethkingia meningoseptica</i>
	Flavobacteriia	<i>Flavobacterium columnare</i>
		<i>Flavobacterium succinicans</i>
	Sphingobacteriia	<i>Solitalea canadensis</i>
		<i>Sphingobacterium multivorum</i>
Cyanobacteria	Chloroplast	<i>Acutodesmus obliquus</i>
	Nostocophycideae	<i>Cylindrospermopsis raciborskii</i>
Firmicutes	Bacilli	<i>Anoxybacillus kestanbolensis</i>
		<i>Lactococcus garvieae</i>
	Clostridia	<i>Faecalibacterium prausnitzii</i>
Fusobacteria	Fusobacteriia	<i>Cetobacterium somerae</i>
Proteobacteria	Alphaproteobacteria	<i>Asticcacaulis biprosthecium</i>
		<i>Brevundimonas diminuta</i>
		<i>Ochrobactrum intermedium</i>
		<i>Hyphomicrobium sulfonivorans</i>
		<i>Methylobacterium organophilum</i>
		<i>Paracoccus marcusii</i>
		<i>Rhodovarius lipocyclicus</i>
		<i>Phaeospirillum fulvum</i>
		<i>Orientia tsutsugamushi</i>
		<i>Blastomonas natatoria</i>
		<i>Sphingomonas changbaiensis</i>
		<i>Sphingomonas echinoides</i>
		<i>Sphingomonas wittichii</i>
		<i>Sphingomonas yabuuchiae</i>
	Betaproteobacteria	<i>Salinispora tropica</i>
		<i>Acidovorax delafieldii</i>

Supplementary Table 1 (continue)

<b>Phylum</b>	<b>Class</b>	<b>Species</b>
		<i>Roseateles depolymerans</i>
		<i>Variovorax paradoxus</i>
		<i>Janthinobacterium lividum</i>
		<i>Polynucleobacter cosmopolitanus</i>
		<i>Methylothermobacter mobilis</i>
	Deltaproteobacteria	<i>Bdellovibrio bacteriovorus</i>
		<i>Desulfovibrio putealis</i>
	Epsilonproteobacteria	<i>Sulfuricurvum kujiense</i>
	Gammaproteobacteria	<i>Plesiomonas shigelloides</i>
		<i>Acinetobacter rhizosphaerae</i>
		<i>Acinetobacter schindleri</i>
		<i>Pseudomonas nitroreducens</i>
		<i>Nevskia ramosa</i>
		<i>Luteibacter rhizovicinus</i>
		<i>Lysobacter brunescens</i>
		<i>Pseudoxanthomonas mexicana</i>
Spirochaetes	Leptospirae	<i>Leptospira biflexa</i>
	Spirochaetes	<i>Spirochaeta aurantia</i>
Verrucomicrobia	Verrucomicrobiae	<i>Akkermansia muciniphila</i>
		<i>Prostheco bacter debontii</i>

Supplementary Table 2  
*Top 10 genera normalised abundance in percentage*

	S1R15	S1R17	S1R25	S1R27	S1R35	S1R37	S2R15	S2R17	S2R25	S2R27	S2R35	S2R37	S3R15	S3R17	S3R25	S3R27	S3R35	S3R37	
1	<b>g__hgcl clade</b>	0.018	0.423	0.091	0.373	0.416	0.112	0.047	0.213	0.319	0.226	0.189	0.220	0.194	0.008	0.162	0.011	0.065	0.289
2	<b>g__Flavobacterium</b>	0.536	0.001	0.220	0.001	0.118	0.255	0.396	0.000	0.165	0.005	0.257	0.000	0.259	0.029	0.281	0.001	0.016	0.000
3	<b>g__Cyanobium PCC-6307</b>	0.026	0.170	0.155	0.159	0.118	0.222	0.223	0.173	0.192	0.177	0.141	0.170	0.110	0.109	0.083	0.042	0.166	0.089
4	<b>g__uncultured</b>	0.004	0.002	0.001	0.001	0.001	0.014	0.113	0.165	0.059	0.250	0.050	0.155	0.120	0.352	0.139	0.221	0.153	0.264
5	<b>g__CL500-29 marine group</b>	0.056	0.279	0.039	0.237	0.090	0.143	0.050	0.039	0.063	0.038	0.081	0.037	0.117	0.035	0.116	0.025	0.050	0.133
6	<b>g__Pseudomonas</b>	0.301	0.002	0.024	0.001	0.007	0.057	0.043	0.354	0.000	0.213	0.028	0.361	0.001	0.007	0.025	0.000	0.000	0.003
7	<b>g__uncultured</b>	0.001	0.001	0.000	0.000	0.001	0.002	0.109	0.053	0.120	0.068	0.066	0.055	0.179	0.047	0.174	0.071	0.163	0.184
8	<b>g__Fluviicola</b>	0.057	0.008	0.470	0.007	0.250	0.004	0.019	0.003	0.083	0.005	0.189	0.002	0.021	0.013	0.018	0.005	0.002	0.006
9	<b>g__Brevibacillus</b>	0.000	0.115	0.000	0.221	0.000	0.179	0.000	0.000	0.000	0.011	0.000	0.000	0.000	0.169	0.000	0.347	0.000	0.032
10	<b>g__uncultured</b>	0.001	0.000	0.000	0.000	0.000	0.013	0.000	0.000	0.000	0.007	0.000	0.000	0.001	0.231	0.001	0.277	0.386	0.000





## Towards Maximising Hardware Resources and Design Efficiency via High-Speed Implementation of HMAC based on SHA-256 Design

Shamsiah Suhaili\*, Norhuzaimin Julai, Rohana Sapawi and Nordiana Rajae

*Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Malaysia Sarawak, 94300 UNIMAS, Kota Samarahan, Sarawak, Malaysia*

### ABSTRACT

Some applications, such as Message Authentication Code (MAC), rely on different hashing operations. There are various hash functions, including Message-Digest 5 (MD5), RACE Integrity Primitives Evaluation Message Digest 160 (RIPEMD-160), Secure Hash Algorithm 1 (SHA-1), and Secure Hash Algorithm 256 (SHA-256), among others. The network layer is the third of seven layers of the Open Systems Interconnection (OSI) concept, also known as the Internet. It handles network addressing and physical data routing. Nowadays, enhanced internet security is necessary to safeguard networks from illegal surveillance. As a result, Internet Protocol Security (IPsec) introduces secure communication across the Internet by encrypting and/or authenticating network traffic at the IP level. IPsec is an internet-based security protocol. Encapsulating Security Payload (ESP) and Authentication Header (AH) protocols are separated into two protocols. The MAC value is stored in the authentication data files of the Authentication Header and Encapsulating Security Payload. This article analyses a fast implementation of the Hash-based Message Authentication Code (HMAC), which uses its algorithm to ensure the validity and integrity of data to optimise hardware efficiency and design efficacy using the SHA-256 algorithm. During data transfer, HMAC

is critical for message authentication. It was successfully developed using Verilog Hardware Description Language (HDL) code with the implementation of a Field Programmable Gate Array (FPGA) device using the Altera Quartus II Computer-Aided Design (CAD) tool to enhance the maximum frequency of the design. The accuracy of the HMAC design, which is based on the SHA-256 design, was examined and confirmed

### ARTICLE INFO

#### *Article history:*

Received: 04 November 2022

Accepted: 10 May 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.02>

#### *E-mail addresses:*

sushamsiah@unimas.my (Shamsiah Suhaili)

jnorhuza@unimas.my (Norhuzaimin Julai)

srohana@unimas.my (Rohana Sapawi)

rnordiana@unimas.my (Nordiana Rajae)

\* Corresponding author

using ModelSim. The results indicate that the maximum frequency of the HMAC-SHA-256 design is approximately 195.16 MHz.

*Keywords:* Field Programmable Gate Array, hash function, Hash-based Message Authentication Code, Secure Hash Algorithm 256, Verilog Hardware Description Language

---

## INTRODUCTION

There are seven layers, and the internet layer is the network layer in which data transfer from one terminal to another depends on the address and routing network. Traffic networks are prone to eavesdropping and illegal access without a network-integrated security element. However, selecting a suitable encryption and authentication product for the network can solve this problem. The internet community created the Security Protocol (Randall, 1999). The third network layer of the seven-layer OSI architecture employs the IPsec protocol. The seven layers are divided into application, presentation, session, transport, network, data link, and physical layers. One of the network encryption protocols is IPsec (IP Security), the most recent IP-based technology.

The IP provides network authentication and encryption to protect the network from illegal surveillance. Because of its improved capabilities, IP Security has become a fact of life in terms of network security for Internet Protocol version 4 (IPv4) and Internet Protocol version 6 (IPv6). The IPsec is divided into two protocols: Authentication Header (AH), which examines IP packet authentication and data integrity, and Encapsulating Security Payload (ESP), which encrypts and authenticates the message. Both AH and ESP are equipped with two different modes: tunnel mode and transit mode; as a whole, the IP packet is encrypted in tunnel mode, while only the transport layer is encrypted in the latter. On the other hand, HMAC-MD5, HMAC-SHA, and HMAC-RIPEMD160 are authentication and data integrity methods. These methods may be used to safeguard all distributed applications, e-mail, file transfers, and web access.

This article focuses on computing the Hash-based Message Authentication Code (HMAC) using the MAC (Message Authentication Code) algorithm. Message Authentication Code (MAC) is used to verify the validity of a message, while HMAC is a subset of MAC that uses a cryptographic hash function and a private key for verification. It accepts arbitrary input with a specified key and produces MAC output. The authentication data element in the AH header contains this MAC value. Network transmission operations are followed using the same key to obtain the same MAC at the destination. The message is valid if the MAC value received at the destination corresponds to the one broadcasted. Similar to AH, ESP enables the use of MAC with HMAC. The encryption procedure takes place before the IP layer, which is at the IPsec layer when the application sends the message across the network. A message is routed via the network to its destination using an IP address, part of an IP layer. The router will then determine

the destination address based on the sender's IP address. The decryption of the packet is required to access the sent data.

Security has recently emerged as a hot topic among researchers. Various cryptography algorithms have been developed to enhance the effectiveness of these information-protecting processes. Message digests are generated using hash function techniques during data transmission. As a result, it becomes a crucial tool for embedding security in e-mail, Internet banking, and other applications. A hash function generates a fixed-length output from an arbitrary-length message input. The one-way nature of hash functions makes converting a hash value to a message input challenging. A hash function is a cryptography technique that does not require a key, such as MD5, RIPEMD160, or SHA-1. In this study, an SHA family was constructed and tailored to meet the performance requirements for cryptographic algorithms. There are four types of hash functions in SHA-2, which are SHA-224, SHA-256, SHA-384, and SHA-512. The length of SHA determines the output length of these hash algorithms, ranging from 256 to 512 bits. This article presents the design of the SHA-256 hash function.

This study optimises hardware resources and performance by utilising the hash function of SHA-256 with a Message Authentication Code. Meanwhile, IPsec and HMAC-SHA-256 are focused on several related projects to optimise hardware size, performance, and consumption. McLoone and McCanny (2002) presented IPsec hardware on a single chip that included Rijndael and HMAC-SHA-256. The wireless design raises the maximum frequency, as shown in a previous study (Selimis et al., 2003). On the other hand, the HMAC with SHA-1/MD5 was initially presented in earlier research (Wang et al., 2004), where hardware complexity was minimised, and an efficient hash function structure was devised to share hardware. Additionally, Michail et al. (2004) demonstrated the HMAC-SHA-1 implementation on an FPGA device, whereas Yiakoumis et al. (2005) showed the execution of a small-sized, high-speed HMAC-SHA-1.

Improvement methods adopted by Khan et al. (2007) used pipelining and parallelism to create the HMAC-hash unit and combine them into a single reconfigurable unit. Even though these ideas worked well, the greatest frequencies they could reach were only a few tens of MHz, or up to 111 MHz, as presented in research by Yiakoumis et al. (2005). Meanwhile, an FPGA implementation of HMAC based on SHA-256 was designed in previous studies by Juliato and Gebotys (2011) and Rubayya and Resmi (2015). The results were obtained using a Xilinx device in both designs. Furthermore, the HMAC design in Rubayya and Resmi (2015) demonstrated significant improvement. It suggests that greater performance is needed to meet the demands of current systems.

Table 1 shows the previous design of HMAC with various types of hash functions, such as SHA-1, MD5, and RIPEMD-160. Numerous studies have been conducted on HMAC, but not all of them have focused on the frequency maximum of FPGA design implementation (Choi & Seo, 2020; Oku et al., 2018; Lin et al., 2017; Ravilla & Putta, 2015a; Ravilla &

Table 1  
*Previous HMAC design*

No.	Author	FPGA Device	HMAC Design	Frequency Maximum (MHz)
1	McLoone and McCanny (2002)	Xilinx XCV1000E	HMAC (SHA-1)	50
2	Selimis et al. (2003)	V150bg352	HMAC (SHA-1)	82
3	Wang et al. (2004)	EP20K1000EBC652-IX	HMAC (SHA-1/MD5)	22.67
4	Michail et al. (2004)	Xilinx V3200efg1156	HMAC (SHA-1)	62.0
5	Yiakoumis et al. (2005)	Xilinx VirtexE-8	HMAC (MD5) HMAC (SHA-1)	55 111
6	Khan et al. (2007)	Xilinx XC2V4000	HMAC (MD5, SHA-1, RIPEMD-160)	43.47
8	Juliato and Gebotys (2011)	Altera Apex 20K, EP20K1000EBC652	HMAC (SHA-256)	35.55
	Juliato and Gebotys (2011)	Xilinx Virtex-E, XCV1600EBG1156	HMAC (SHA-256)	48.12
	Juliato and Gebotys (2011)	Xilinx Virtex-II, XCV2V4000BF957	HMAC (SHA-256)	59.66
9	Rubayya and Resmi (2015)	Xilinx Device (no mention device name)	HMAC (SHA-256)	110.009
10	Ravilla and Putta (2015a, 2015b)	No FPGA implementation	HMAC (SHA-256)	-
11	Choi and Seo (2020)	No FPGA implementation	HMAC (SHA-256)	-
12	Chen and Yuan (2012)	No FPGA implementation	HMAC (SHA-256)	-
13	Lin et al. (2017)	No FPGA implementation	HMAC (SHA-256)	-
14	Oku et al. (2018)	No FPGA implementation	HMAC (SHA-256)	-
15	Jung and Jung (2013)	No FPGA implementation	HMAC-based RFID mutual authentication	-
16	Kieu-Do-Nguyen et al. (2022)	Virtex 4/virtex 5	HMAC-SHA-256	188
17	Pham et al. (2022)	Virtex2 XC2VP20	SHA-256	165

Putta, 2015b; Jung & Jung, 2013). The reference study by Kieu-Do-Nguyen et al. (2022) combined all SHA-2 families into one core, such as HMAC-SHA2-224/256/384/512, whereas Pham et al. (2022) designed only the combination of SHA-256/512/256d hash function. This article focuses on SHA-256 because of its wide implementation in security design implementations, such as Bitcoin, also known as cryptocurrency.

## MATERIALS AND METHODS

Message authentication is the process of verifying the authenticity of messages, where two important factors must be considered in verifying the authenticity of a message, such

as a source and that the message has not been altered. The traditional encryption method authentication takes place when the sender and the receiver possess the same key throughout the transmission process (Stallings, 1996). In other words, only the genuine user has the key to decrypt the message. Figure 1 shows the block design for the complete message authentication process, which contains a secret key to generate data into the algorithm. The sender receives this message and its MAC; once the MAC has been compared to the MAC at the receiver, the receiver must determine whether it has received the same MAC using the same secret key. The message has not been changed and is deemed authentic if the output MAC matches the one transmitted.

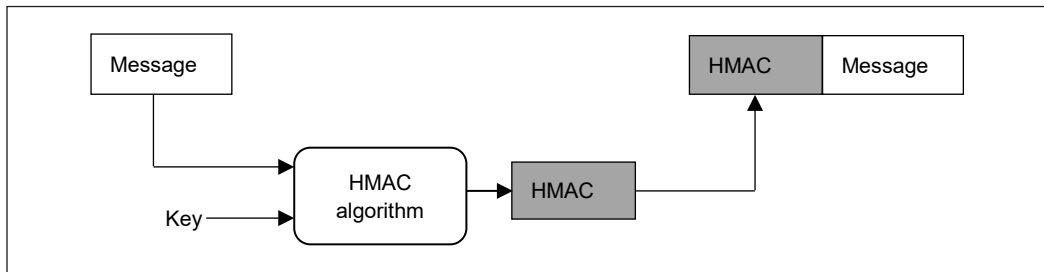


Figure 1. Message authentication using HMAC

One application that uses the hash function to verify message authentication is HMAC. An HMAC based on SHA-256 was implemented in this study. The input key is first hashed before any input text using HMAC. The input key is generated by XORing the inner pad (Ipad) with the input text. An inner pad value is 36 hexadecimal values with a 64-byte timeout, depending on the length of the key. The output of this SHA-256 hash function will be fed into the next SHA-256 to produce the HMAC design output. Before hashing the first SHA-256 output, the key inputs must be XORed with the outer pad (Opad), which has 5C values in hexadecimal (FIPS PUB 198-1, 2008; FIPS PUB 180-4, 2015). The concatenation of the 64-byte key,  $K_0$ , and Opad outputs, as well as the output of the first SHA-256, will then be hashed together to produce the HMAC output. Equation 1 shows an MAC calculated with the HMAC function over a textual representation of the data, while Figure 2 depicts the HMAC structure using SHA-256. The symbols  $\oplus$  and  $\parallel$  stand for XOR and concatenation, respectively.

$$\begin{aligned} \text{MAC}(\text{Message input})_t &= \text{HMAC}(K, \text{Message input})_t \\ &= H((K_0 \oplus \text{Opad}) \parallel H((K_0 \oplus \text{Ipad}) \parallel \text{Message input}))_t \end{aligned} \quad [1]$$

Message Authentication Code is one of the applications of the hash function. Figure 2 depicts the high-level implementation of the HMAC design, comprising an input pad (Ipad) with a fixed 36 hexadecimal value and an output pad (Opad) with a fixed 5C hexadecimal

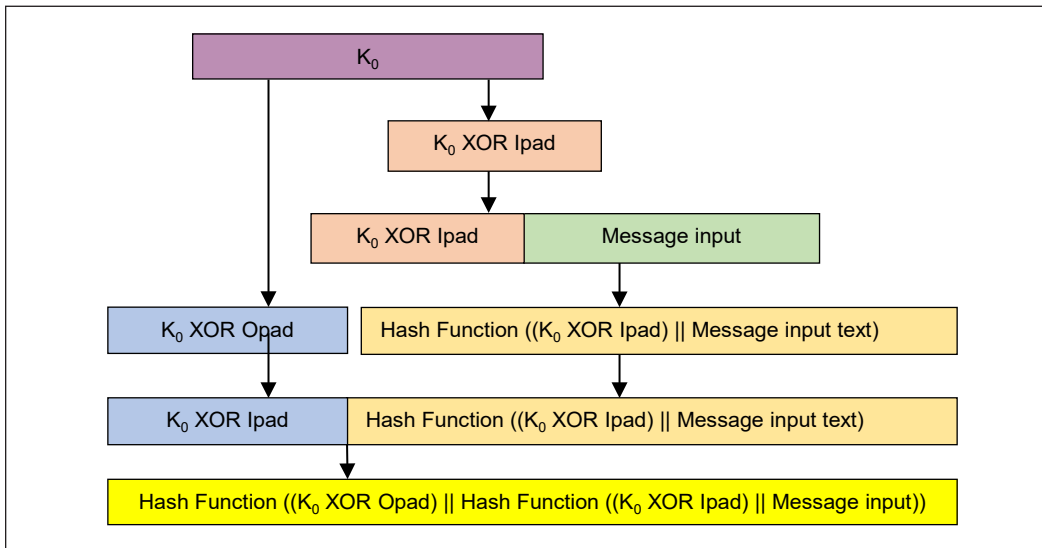


Figure 2. Illustration of HMAC construction

value. The values from the Ipad are sent into the initial SHA-256 design. The counter then causes the HMAC count to generate the output of the first SHA-256 to the second SHA-256 to obtain the overall HMAC output. HMAC is an abbreviation for Hash-based Message Authentication Code, a cryptographic authentication technique that uses a hash function and a secret key. Two identical hash function architectures are used in the HMAC design to obtain the HMAC results. Meanwhile, two hash functions must be used in the HMAC architecture to complete the HMAC structure. In this study, SHA-256 algorithms were employed in the HMAC design. Figure 3 depicts the architecture of an HMAC design with two SHA-256 hash algorithms.

Moreover, Figure 3 illustrates how the key Ipad input message was integrated with the text input. These values will be used as the SHA-256 hash function’s input message. There will be two instances of 512-bit blocks if the initialisation value and message input are used. The state machine controls the first and second SHA-256 input in the proposed design. As a result, the first SHA-256 must be executed first, and the second SHA-256 must wait for the message input from the output of the first SHA-256 design, as shown in Figure 3. HMAC uses two hash functions to operate. Two SHA-256 hash functions are employed in HMAC design. The initial SHA-256 algorithm must be executed until results are achieved. These findings serve as the input for the second SHA-256 algorithm with message input. Figure 3 depicts the concatenation of a key with input from an Ipad and SHA-256 with input from a message. Therefore, the validity of the results of initial SHA-256 remains crucial in acquiring the final HMAC based on SHA-256.

The 64 states in this design have generated the sequence input for the first and second SHA-256 algorithms. Thirty-two states were executed: 16 for the first 512 bits and another

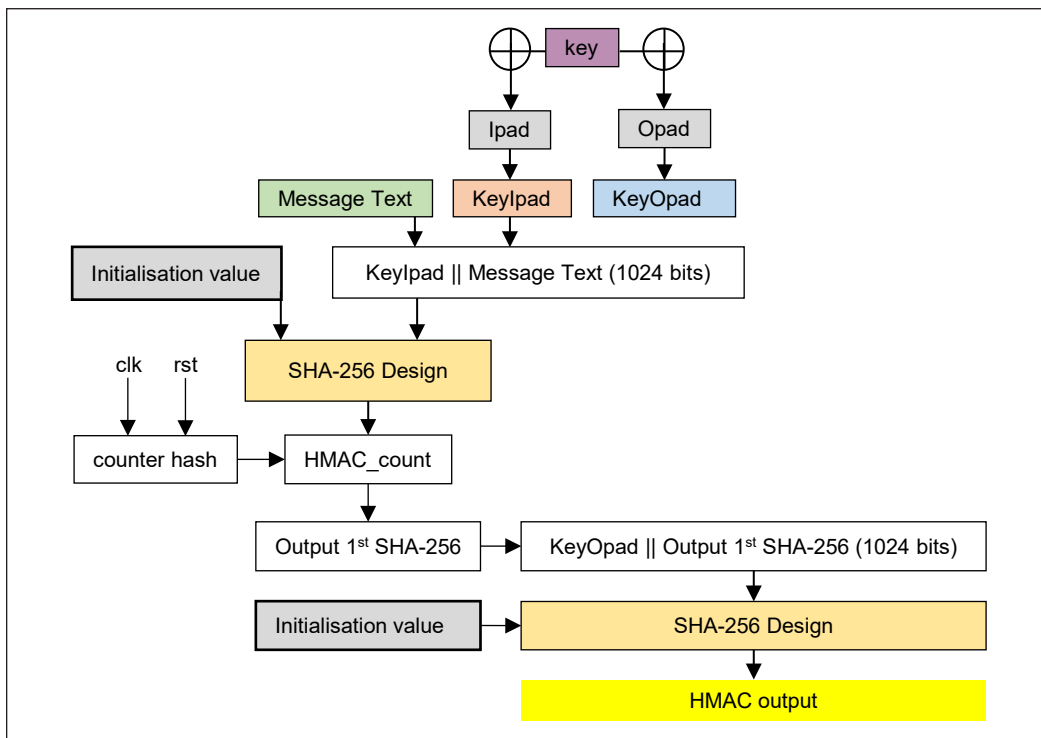


Figure 3. Implementation of proposed HMAC-SHA-256

16 for the second 512 bits SHA-256 to process the message for the first SHA-256. The message will use the output of the first 512-bit hash function to obtain the output of the second 512-bit loop of the input message. An HMAC count is used to generate the correct output SHA-256 hash function with the help of a counter. Input Ipad is the same concept as message input to the second SHA-256. The key Opad output will be combined with the first SHA-256 output. There will be two 512-bit blocks. The state begins at 33 and progresses to 48 for the first 512 bits. The remaining 512 bits will then be executed until 64 states are reached. Similar to the previous SHA-256 process, the output of the second 512-bit block of the second SHA-256 will use the output of the first 512 bits of the second hash function. Finally, this will produce the resulting output obtained in this study.

Secure Hash Functions authenticate messages; therefore, it is necessary to comply with the hash function requirement. A hash function must possess certain properties. Some of these properties include generating outputs of fixed length and the computational infeasibility of determining an  $x$  such that  $\text{Hash}(x)$  equals hash value. It is true for any given code  $h$ . It is easy to calculate the hash code, but it is impossible to recover the original message by reversing it. In addition, it is computationally impossible to determine  $y \neq x$  for any given block  $x$  for which  $\text{Hash}(x) = \text{Hash}(y)$ . In other words, identical hash codes are impossible to find, and all these characteristics constitute weak hash functions. Additionally,



no pair  $(x, y)$  can be found for which  $\text{Hash}(x) = \text{hash value}$ , computationally,  $(y)$ . If the last property of the hash function criteria is met, the function is considered “strong.” Any message less than 264 bits in length can be entered into the SHA-256 algorithm. SHA-256 processes 512-bit message inputs and 160-bit initial values to produce a 160-bit hash code output. Figure 4 depicts the 512-bit message input of the SHA-256 hash function.

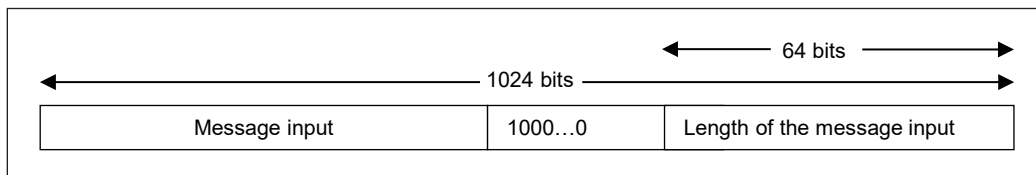


Figure 4. 512-bit message input of SHA-256

Several considerations must be made to generate output. The message must first be padded to almost exactly match 960 modulo 1024. Figure 4 shows a 1-bit input following the input message. It is then padded with 0 bits to produce the total length of the message. After the message is padded, 64-bit inputs are added to the message. Message padding consists of 512 bits and the message’s total length. HMAC-based SHA-256 uses the SHA-256 approach to

Table 2  
Buffer initialisation of SHA-256

Register	Buffer Initialisation (Hex)
A	32'h6a09e667
B	32'hbb67ae85
C	32'h3c6ef372
D	32'ha54ff53a
E	32'h510e527f
F	32'h9b05688c
G	32'h1f83d9ab
H	32'h5be0cd19

determine authentication in this design. In other words, the SHA-256 hash algorithm was utilised to construct HMAC. Based on the SHA-256 algorithm, the SHA-256 architecture has eight fixed inputs. Hence, the SHA-256 algorithm requires eight variables as initial input during hash computation. Table 2 displays the hexadecimal buffer initialisation of the SHA-256 hash function. Eight distinct input buffer initialisations exist; they are used during the first phase of execution, and their values are fixed for all SHA-256 hash functions. After the startup operation, the input message is processed in 1024-bit blocks of 32 bits each. Figure 5 displays the 64 highest-level steps of the SHA-256 message compression process.

The input message is padded during an early step of the SHA-256 hash process. Padding the message begins once the input for the message is received, and the message is completed by appending a single one-bit. Next, n zero bits will be added, and this pattern will continue until the total number of bits in the message equals 448 modulo 512. The final 64 bits are set aside specifically for use in relation to the calculation of how long the message should be. The message input capacity is 512 bits. The message scheduler computes the message,  $W_t$  of SHA-256. For  $0 \leq t \leq 15$ , a message is extracted directly from the input message,



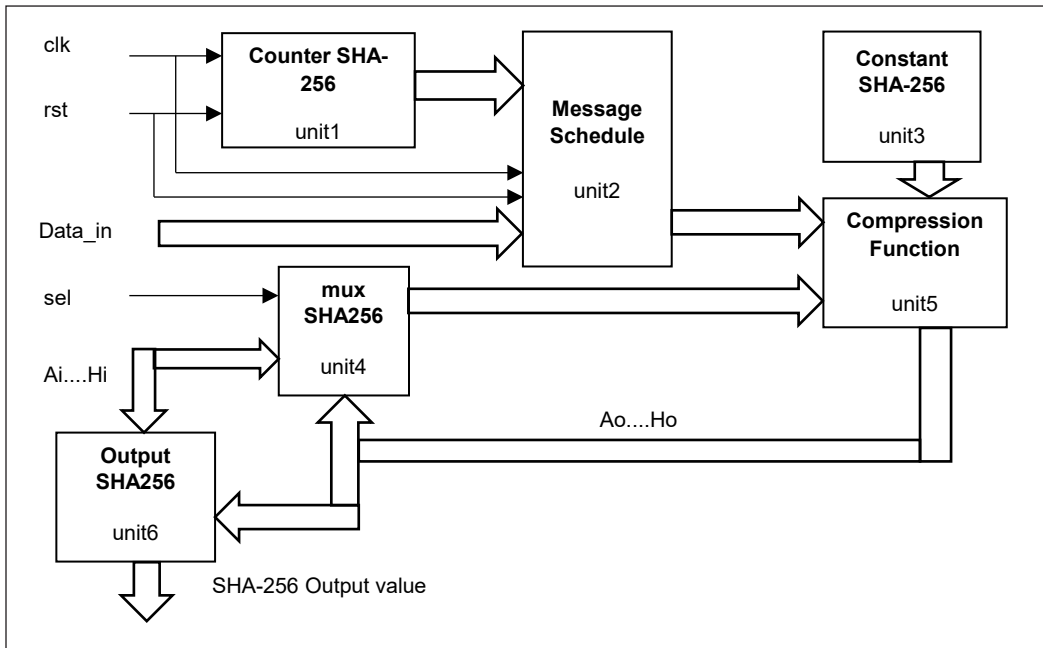


Figure 5. Top level of SHA-256 design

whereas for  $16 \leq t \leq 63$ , a message  $W_t$  is calculated using Equation 2. The value  $t$  denotes the number of transformation rounds.  $ROTR^n(x)$  is a right rotation of  $x$  by  $n$  bits, whereas  $SHR^n(x)$  is a right shift of  $x$  by  $n$  bits, as shown in Equations 3 and 4.

Message schedule SHA-256,  $W_t$

$$W_t = \text{message input} \quad 0 \leq t \leq 15$$

$$W_t = \sigma_1^{256}(W_{t-2}) + W_{t-7} + \sigma_0^{256}(W_{t-15}) + W_{t-16} \quad 16 \leq t \leq 63 \quad [2]$$

where,

$$\sigma_0^{256}(x) = ROTR^7(x) + ROTR^{18}(x) + SHR^3(x) \quad [3]$$

$$\sigma_1^{256}(x) = ROTR^{17}(x) + ROTR^{19}(x) + SHR^{10}(x) \quad [4]$$

The SHA-256 compression function is made up of four functions that round from  $t = 0$  to  $t = 63$ . The four functions are  $Ch(x,y,z)$ ,  $Maj(x,y,z)$ ,  $\Sigma_0(x)$  and  $\Sigma_1(x)$ , as shown in Equations 5, 6, 7 and 8. The symbols  $\wedge$ ,  $\neg$  and  $\oplus$  represent the logical AND gate, NOT gate, and XOR gate, respectively.

$$Ch(e, f, g) = (e \wedge f) \oplus (\neg e \wedge g) \quad [5]$$

$$Maj(a, b, c) = (a \wedge b) \oplus (a \wedge c) \oplus (b \wedge c) \quad [6]$$

$$\sum_0(a) = ROTR^2(a) + ROTR^{13}(a) + ROTR^{22}(a) \quad [7]$$

$$\sum_1(e) = ROTR^6(e) + ROTR^{11}(e) + ROTR^{25}(e) \quad [8]$$

The hash computation was used to construct eight variables with initial values to evaluate the four functions of Equations 5, 6, 7 and 8. The message input,  $W_t$ , and constant  $K_t$  form the 64 iterative operations. The output of the following Equations 9, 10 and 11 is the output of hash values.

$$Temp_1 = h + \sum_1(e) + Ch(e, f, g) + K_t + W_t \quad [9]$$

$$Temp_2 = \sum_0(a) + Maj(a, b, c) \quad [10]$$

$$h = g$$

$$g = f$$

$$f = e$$

$$e = d + Temp_1$$

$$d = c \quad [11]$$

$$c = b$$

$$b = a$$

$$a = Temp_1 + Temp_2$$

After 64 iterations, the modulo-32-bit adders calculate the hash values,  $H_0$  to  $H_7$ . The SHA-256 hash value in its final form is generated using the Big-endian format.

$$H_0 = a + H_0, H_1 = b + H_1, H_2 = c + H_2, H_3 = d + H_3$$

$$H_4 = e + H_4, H_5 = f + H_5, H_6 = g + H_6, H_7 = h + H_7$$

$$\text{Message Digest} = H_0 \parallel H_1 \parallel H_2 \parallel H_3 \parallel H_4 \parallel H_5 \parallel H_6 \parallel H_7$$

## RESULTS AND DISCUSSION

Figure 6 shows the timing simulation waveform result of the HMAC-SHA-256 design with the message input text "Sample #1". HashCalc validated the HMAC values to ensure output accuracy. Based on the simulation waveform results, the output of the HMAC value provides the correct result of the HMAC value without error, which is similar to the calculation from HashCalc software, as shown in Figure 7. FMax is the maximum

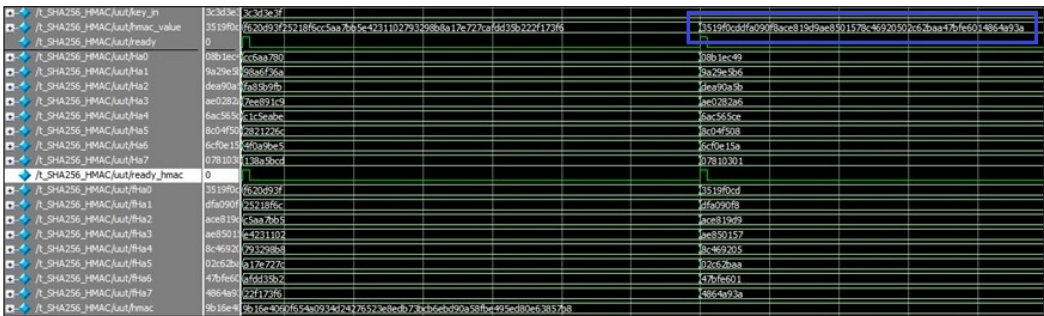


Figure 6. Simulation waveform of SHA-256 design

clock frequency of HMAC-SHA-256 that a digital design can operate at, and it improves greatly when a clock constraint is applied to the design. Figure 8 depicts the maximum clock frequency of HMAC-SHA-256 with SDC 5.3 clock limitations.

Table 3 displays the proposed HMAC-SHA-256 design and other publications utilising HMAC on various FPGA family devices. HMAC-SHA-256 was successfully designed using Altera Quartus II 15.0. ModelSim-Altera 10.3d was used for functional and timing simulation to validate

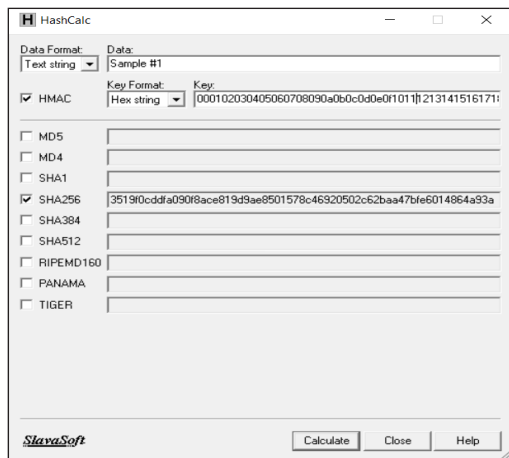


Figure 7. HMAC calculation value of SHA-256 design with specific key

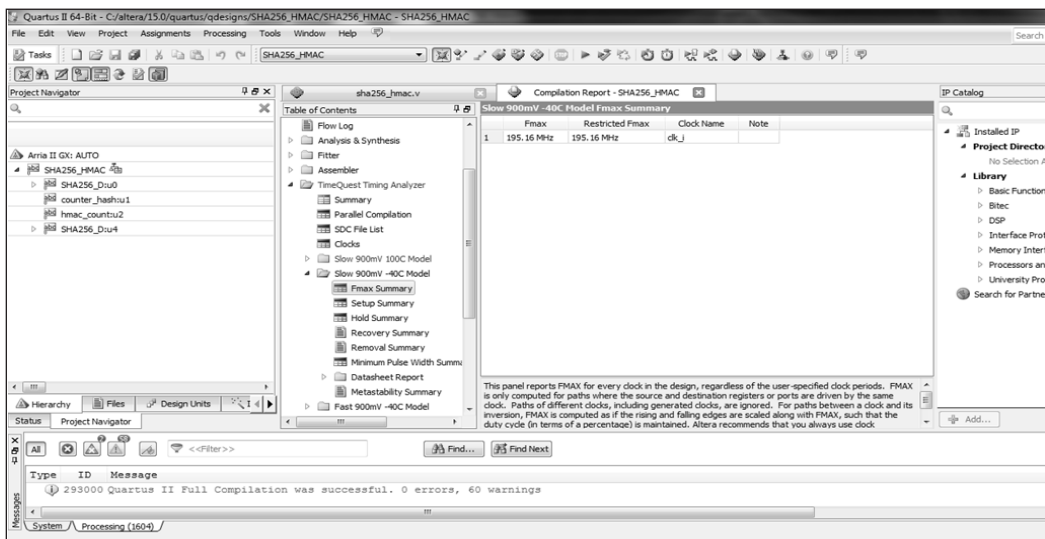


Figure 8. Maximum clock frequency (FMax) of HMAC-SHA-256

Table 3  
*FPGA-based implementation comparison of the previous HMAC design*

Authors/year	FPGA Device	Design	Maximum Frequency (MHz)	LUT/CLB /LE	Reg
Proposed Design	Altera Arria II GX	Proposed HMAC (SHA-256)	195.16	<b>3953</b>	2714
Kieu-Do-Nguyen et al. (2022)	Virtex 4/Virtex 5	HMAC (SHA-256)	188	<b>1615</b>	-
Pham et al. (2022)	Virtex2 XC2VP20	SHA-256	165	3695	-
Rubayya and Resmi (2015)	Xilinx Device (no mention device name)	HMAC (SHA-256)	110.009	6861	-
Juliato and Gebotys (2011)	Altera Apex 20K, EP20K1000EBC652	HMAC (SHA-256)	35.55	9231	-
Juliato and Gebotys (2011)	Xilinx Virtex-E, XCV1600EBG1156	HMAC (SHA-256)	48.12	3463	-
Juliato and Gebotys (2011)	Xilinx Virtex-II, XCV2V4000BF957	HMAC (SHA-256)	59.66	3608	-
Khan et al. (2007)	Xilinx XC2V4000	HMAC (MD5, SHA-1, RIPEMD-160)	43.47	7484	-
Yiakoumis et al. (2005)	Xilinx VirtexE-8	HMAC (MD5) HMAC (SHA-1)	55 111	686	-
Michail et al. (2004)	Xilinx V3200efg1156	HMAC (SHA-1)	62.0	6011	-
Wang et al. (2004)	EP20K1000EBC652-IX	HMAC (SHA-1/MD5)	22.67	-	-
Selimis et al. (2003)	V150bg352	HMAC (SHA-1)	82	1018	-
McLoone and McCanny (2002)	Xilinx XCV1000E	HMAC (SHA-1)	50	7247	-

the output results. The maximum frequency of the HMAC-SHA-256 design increases dramatically on the Arria II GX with 3953 LUT and 2714 total registers, as shown in Table 3. Based on these findings, synthesis and implementation on Arria II GX give fast speed with a maximum frequency of 195.16 MHz compared to other HMAC publications utilising various hash functions on FPGA family device types. Furthermore, with the assistance of Altera Quartus II and TimeQuest Timing Analyser advisors, HMAC-SHA-256 design results are improved ([https://www.altera.com/en\\_US/pdfs/literature/ug/ug\\_tq\\_tutorial.pdf](https://www.altera.com/en_US/pdfs/literature/ug/ug_tq_tutorial.pdf)). By applying SDC clock constraint 5.3 to the HMAC-SHA-256 design on Arria II GX, the maximum frequency of the design can be met under stable setup and hold conditions. Thus, this study proposed a high-performance and error-free HMAC-SHA-256 design with appropriate FPGA devices and clock constraints that meet the design requirement.

## CONCLUSION

The design of HMAC-SHA-256 was successful through the use and development of high-speed computing, which possessed a maximum frequency of 195.16 MHz. FPGA

implementation on the Arria II GX can offer great speed. Moreover, the design can be significantly enhanced with the assistance of advisor Altera Quartus II. Furthermore, providing the design with the proper SDC clock constraints will allow the TimeQuest timing analyser to meet the time requirements.

## ACKNOWLEDGEMENTS

The author thanks the Ministry of Higher Education, Malaysia, Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UNIMAS/02/11), Universiti Malaysia Sarawak (F02/FRGS/2035/2020) and Osaka Gas Foundation in Cultural Exchange (OGFICE) Research Grant Scheme (IG/F02/OSAKA/01/2022) for supporting this work.

## REFERENCES

- Chen, F., & Yuan J. (2012). Enhanced key derivation function of HMAC-SHA-256 algorithm in LTE network. In 2012 Fourth International Conference on Multimedia Information Networking and Security (pp. 15-18). IEEE Publishing. <https://doi.org/10.1109/MINES.2012.106>
- Choi, H., & Seo, S. C. (2020). Optimization of PBKDF2-HMAC-SHA256 and PBKDF2-HMAC-LSH256 in CPU environments. In I. You (Ed.), *Information Security Applications* (pp. 321-333). Springer Cham. [https://doi.org/10.1007/978-3-030-65299-9\\_24](https://doi.org/10.1007/978-3-030-65299-9_24)
- FIPS PUB 198-1. (2008). *Federal Information Processing Standards, The Keyed-Hash Message Authentication Code (HMAC)*. Information Technology Laboratory National Institute of Standards and Technology Gaithersburg. <https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.198-1.pdf>
- FIPS PUB 180-4. (2015). *Federal Information Processing Standards, Secure Hash Standard (SHS)*. Information Technology Laboratory National Institute of Standards and Technology Gaithersburg. <https://nvlpubs.nist.gov/nistpubs/fips/nist.fips.180-4.pdf>
- Juliato, M., & Gebotys, C. (2011). *FPGA Implementation of an HMAC Processor based on the SHA-2 Family of Hash Functions*. University of Waterloo Technical Report. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=5043ce0a65691fd16ff7a546e6c0013d9ee190ca>
- Jung, S. W., & Jung, S. (2013). HRP: A HMAC-based RFID mutual authentication protocol using PUF. In *The International Conference on Information Networking 2013 (ICOIN)* (pp. 578-582). IEEE Publishing. <https://doi.org/10.1109/ICOIN.2013.6496690>
- Khan, E., El-Kharashi, M. W., Gebali, F., & Abd-El-Barr, M. (2007). Design and performance analysis of a unified, reconfigurable HMAC-Hash unit. *IEEE Transactions on Circuits and Systems-I: Regular Papers*, 54(12), 2683-2695. <https://doi.org/10.1109/TCSI.2007.910539>
- Kieu-Do-Nguyen, B., Hoang, T. T., Tsukamoto, A., Suzuki, K., & Pham, C. K. (2022). High-performance multi-function HMAC-SHA2 FPGA implementation. In *20<sup>th</sup> IEEE International Interregional NEWCAS Conference, NEWCAS 2022* (pp. 30-34). IEEE Publishing. <https://10.1109/NEWCAS52662.2022.9842174>
- Lin, L., Chen, K., & Zhong, S. (2017). Enhancing the session security of zen cart based on HMAC-SHA256. *KSII Transactions on Internet and Information Systems*, 11(1), 466-483.

- McLoone, M., & McCanny, J. V. (2002). A single-chip IPSec cryptographic processor. In *IEEE Workshop on Signal Processing Systems* (pp. 133-138). IEEE Publishing. <https://doi.org/10.1109/SIPS.2002.1049698>
- Michail, H. E., Kakarountas, A. P., Milidonis, A., & Goutis, C. E. (2004). Efficient implementation of the keyed-hash message authentication code (HMAC) using the SHA-1 hash function. In *Proceedings of the 2004 11<sup>th</sup> IEEE International Conference on Electronics, Circuits and Systems, 2004 (ICECS 2004)* (pp. 567-570). IEEE Publishing. <https://doi.org/10.1109/ICECS.2004.1399744>
- Oku, D., Yanagisawa, M., & Togawa, N. (2018). Scan-based side-channel attack against HMAC-256 circuits based on isolating bit-transition groups using scan signatures. *IPSSJ Transactions on System LSI Design Methodology, 11*, 16-28. <https://doi.org/10.2197/ipsjtsldm.11.16>
- Pham, H. L., Tran, T. H., Duong Le, V. T., & Nakashima, Y. (2022). A high-efficiency FPGA-based multimode SHA-2 accelerator. *IEEE Access Open Access, 10*, 11830-11845. <https://doi.org/10.1109/ACCESS.2022.3146148>
- Randall, K. N. (1999). *ISCA Guide to Cryptography*. McGraw-Hill.
- Ravilla, D., & Putta, C. S. R. (2015a). Routing using trust-based system with SHA-2 authentication. *Procedia Computer Science Open Access, 46*, 1108-1115. <https://doi.org/10.1016/j.procs.2015.01.023>
- Ravilla, D., & Putta, C. S. R. (2015b). Implementation of HMAC-SHA256 algorithm for hybrid routing protocols in MANETs. In *2015 International Conference on Electronic Design, Computer Networks & Automated Verification (EDCAV)* (pp. 154-159). IEEE Publishing. <https://doi.org/10.1109/EDCAV.2015.7060558>
- Rubayya, R. S., & Resmi, R. (2015). Memory optimization of HMAC/SHA-2 encryption. In *2014 First International Conference on Computational Systems and Communications (ICCS)* (pp. 282-287). IEEE Publishing. <https://doi.org/10.1109/COMPSC.2014.7032663>
- Selimis, G., Sklavos, N., & Koufopavlou, O. (2003). VLSI implementation of the keyed-hash message authentication code for the wireless application protocol. In *10<sup>th</sup> IEEE International Conference on Electronics, Circuits and Systems, 2003 (ICECS 2003)* (Vol. 1, pp. 24-27). IEEE Publishing. <https://doi.org/10.1109/ICECS.2003.1301967>
- Stallings, W. (1996). *Data & Computer Communications* (6<sup>th</sup> ed.). Prentice Hall.
- Wang, M. Y., Su, C. P., Huang, C. T., & Wu, C. W. (2004). An HMAC processor with integrated SHA-1 and MD5 algorithm. In *ASP-DAC 2004: Asia and South Pacific Design Automation Conference 2004 (IEEE Cat. No. 04EX753)* (pp. 456-458). IEEE Publishing. <https://doi.org/10.1109/ASPDAC.2004.1337618>
- Yiakoumis, I., Papadonikolakis, M., Michail, H., Kakarountas, A. P., & Goutis, C. E. (2005). Efficient small-sized implementation of the Keyed-Hash message authentication code. In *EUROCON 2005-The International Conference on "Computer as a Tool"* (Vol. 2, pp. 1875-1878). IEEE Publishing. <https://doi.org/10.1109/EURCON.2005.1630347>

## Effect of Natural Ventilation on Thermal Performance of Different Residential Building Forms in the Hot-dry Climate of Jordan

Esraa Shehadeh Abbaas<sup>1\*</sup>, Mazran Ismail<sup>1</sup>, Ala'eddin Ahmad Saif<sup>2</sup> and Muhamad Azhar Ghazali<sup>1</sup>

<sup>1</sup>*School of Housing, Building and Planning, Universiti Sains Malaysia, 11800, Penang, Malaysia*

<sup>2</sup>*Physics department, College of Science, University of Jeddah, Jeddah, Saudi Arabia*

### ABSTRACT

This work presents a simulation study on the impact of natural ventilation on the thermal performance and thermal comfort of residential buildings of different forms in the hot-dry climate of Amman, the capital of Jordan. Three existing triple-storey residential buildings with different forms, i.e., rectangular, L-shape, and U-shape, are taken as case studies. Models with similar construction and dimensions of the buildings under investigation are designed using the OpenStudio plugin SketchUp software. Two rooms within these buildings have been considered for simulation with the aid of the EnergyPlus simulator for two cases: the basic case with no ventilation and the case with ventilation. The thermal parameters, including the air temperature, relative humidity, air speed, and mean radiant temperature of both rooms, have been extracted from the simulation. The thermal performance of these buildings is analyzed based on the indoor air temperature and mean radiant temperature, while the thermal performance is investigated via the ASHRAE-55 adaptive model. The results show that the rectangular-shaped building has the best thermal performance in unventilated conditions for the middle room on the middle floor (Room 1). In contrast, the U-shape shows better results for the west-northern room on the same floor (Room 2). On the other hand, introducing natural ventilation to the buildings reduces the indoor temperature and, subsequently, enhances the thermal performance where the buildings transform to be within the comfort zone most

of the time, according to the ASHRAE-55 adaptive model. Generally, rectangular and U-shaped buildings show comparable thermal performance, while L-shaped buildings have relatively the worst performance.

*Keywords:* Adaptive model, ASHRAE-55, building shape, EnergyPlus simulator, ventilation

### ARTICLE INFO

#### *Article history:*

Received: 09 November 2022

Accepted: 10 May 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.03>

#### *E-mail addresses:*

[esraabbas91@gmail.com](mailto:esraabbas91@gmail.com) (Esraa Shehadeh Abbaas)

[mazran@usm.my](mailto:mazran@usm.my) (Mazran Ismail)

[aasaif@uj.edu.sa](mailto:aasaif@uj.edu.sa) (Ala'eddin Ahmad Saif)

[azhar.ghazali@usm.my](mailto:azhar.ghazali@usm.my) (Muhamad Azhar Ghazali)

\* Corresponding author



## INTRODUCTION

The global increase in the energy required for human life activities, especially for the building sector, consumes around 40% of the total energy worldwide. The high cost of energy production causes a big issue for communities' development (Yu et al., 2020). The residential building sector in Jordan consumes about 24% of total energy (Almuhtady et al., 2019). The large energy consumption of the residential sector in Jordan refers to the high usage of cooling systems for hot climates in summer and heating strategies for cold weather in winter. Applying passive design approaches reduces the demands of cooling and heating loads in buildings and improves their thermal performance (Ozarisoy, 2022; Elnagar & Köhler, 2020). Many effective passive approaches, such as building form, natural ventilation, and orientation, are usually considered for buildings.

Building form and its geometrical configuration directly affect energy consumption and thermal performance by enhancing the indoor climate (Lapisa, 2019; Raof, 2017). The building shape determines the percentage of exposed surface area for the building's envelope to the outdoor ambient, indicating the amount of solar radiation the building is exposed to. Minimizing the contact area for the building with the outdoor environment can be achieved by selecting a suitable shape at the early design stage, subsequently improving the thermal comfort level (Mushtaha & Helmy, 2017). A study of different shapes of traditional residential buildings in a hot-dry climate in Diyarbakır, Turkey, showed that the building form significantly impacts indoor thermal comfort and energy consumption. A courtyard in a U-shaped building can decrease 79% of cooling loads and 63% of heating loads (Kocagil & Oral, 2015). A numerical study found that the compact form greatly decreases the energy demands for cooling in a desert area (Bekkouche et al., 2013; Deng et al., 2020). Lapisa (2019) studied the influence of two building forms in different climates, i.e., tropical, Mediterranean, and Oceanic. The results show that the square building form decreases energy consumption compared to the rectangular building form for all climates in his study due to compactness. Moreover, Ali et al. (2010) presented a general study of some buildings which are built between (1900-2000) in Karak (desert climate) and Irbid (mountainous climate), where both cities are in Jordan. It shows that the rectangular form suits desert district buildings in winter, while the compact shape is more suitable in summer.

Natural ventilation enhances the thermal performance of buildings by improving the air quality for the indoor climate in the summer season, especially at nighttime since it is pulling fresh air into the indoor zones, which leads to driving hot air to move out through openings (Raji et al., 2020; Almeida et al., 2017). This process occurs because of two mechanisms: the natural movement of outdoor air around the building and the buoyancy generated due to the difference in air temperature indoors and outdoors of the building. Therefore, the designer should consider these mechanisms during the earlier building design stage to obtain maximum natural ventilation (Krarti, 2018). It is also found that



the location of openings has a higher effect on natural ventilation efficiency scales than building orientation (Rodrigues et al., 2019; Yang et al., 2019). Besides, natural ventilation presents a healthy environment for residents and gives a satisfactory feeling, which raises their productivity and grants attractive results compared to mechanical methods (Nagy et al., 2019). For instance, Sick Building Syndrome (SBS) symptoms of infected occupants due to exposure to indoor air pollutants are higher for HVAC used by 30%–200% compared to natural ventilation (Chen et al., 2021).

Despite the importance of natural ventilation in improving indoor thermal performance, few studies explored its impact on thermal performance in different countries with a climate the same as Jordan's in summer. Al-Hemiddi and Al-Saud (2001) studied the impact of the natural ventilation in a courtyard house on internal thermal performance in a hot-aided climate in Riyadh, Saudi Arabia; they found that the cross ventilation through the courtyard contains a water pool can reduce the indoor air temperature by 5°C. Mastouri et al. (2019) examined the influence of night ventilation in a double-story house in a hot semi-arid climate in Marrakech, Morocco. They concluded that night ventilation gives high-efficiency results by reducing the ground floor temperature by 2°C and the first floor by 3°C. In addition, it can reduce the annual cooling loads by 27%. Only one study considers the impact of natural ventilation on thermal performance in Jordan, as Ma'bdeh et al. (2020) reported. They investigated the effect of introducing a wind tower, once at the north facade and another at the south facade, on interior thermal performance for a classroom in a building located at Jordan University for Science and Technology (JUST) in Irbid in the summer and winter. The results show that the wind tower improves indoor air efficiency in both cases, whereas the highest number of comfort hours is obtained when the tower is built on the southern side of the building. Therefore, the influence of natural ventilation and the impact of building form in Jordan on thermal performance needs to be investigated further. In this work, a simulation study using EnergyPlus software with the aid of the ASHRAE-55 adaptive model presents the impact of different building forms, i.e., rectangular, L-shaped, and U-shaped, in collaboration with the natural ventilation on the thermal performance and thermal comfort for existing residential buildings in the hot-dry climate in Jordan.

### **Climate Conditions and Case Studies**

Jordan lies between latitude 29°–32° north and longitude 35°–38.5° east. The climate in Jordan is mostly Mediterranean type, with cold-wet winter and hot-dry summer (Nazer, 2019). Amman, the capital of Jordan, is considered a mountain area, and it has hot-dry weather in summer with an average air temperature of 8°C–25.1°C. The highest temperature reaches in summer is 39°C in July. It has cold-rainy weather in winter, with air temperature reaching 0°C or less in January. This research will focus on the summer season (hot-dry).

Table 1 shows the monthly average temperature during the summer season. The summer wind characteristics in Amman from June to August are shown in Figure 1. It can be noticed that the wind has different speed values and directions, but mostly during the day, it has a speed of 3.3-5.5 m/s toward the west with a frequency of 11.96%, 5.5–7.9 m/s to the west with a frequency of 9.28% and 5.5–7.9 m/s with north-west direction at a frequency of 7.79%.

This work has taken three different residential building forms, i.e., rectangular, L-shape, and U-shape, located in Amman, as case studies. The selection of these buildings is based on a field survey study for the building forms at the southern side of Alqwsmeh district, where the buildings are located, in Amman, to estimate the percentage of building forms within this area. Figure 2(a) shows a satellite view of the survey area that contains 123 buildings. Figure 2(b) illustrates a solid and void analysis of the survey area. Figure 2(b) shows that the dominant building forms in this area are rectangular buildings, with a percentage of 69.1%, followed by the L-shape, 11.38%, and the U-shape, 7.31%. Additionally, it is found that other forms of

Table 1  
Monthly temperature average in the summer season in Amman, Jordan

Month	Min (°C)	Max (°C)
June	10.8	36.0
July	14.0	39.8
August	13.8	37.0

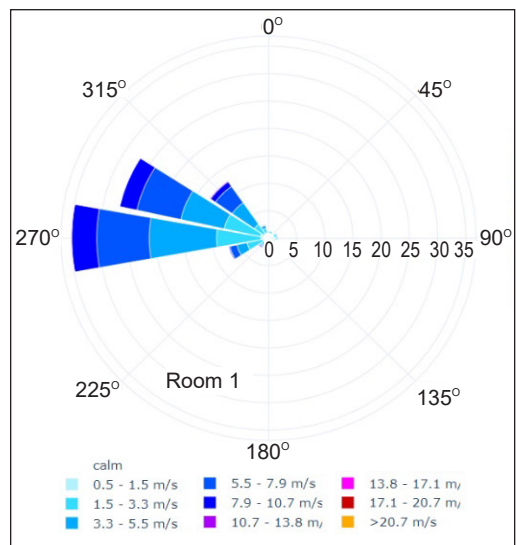
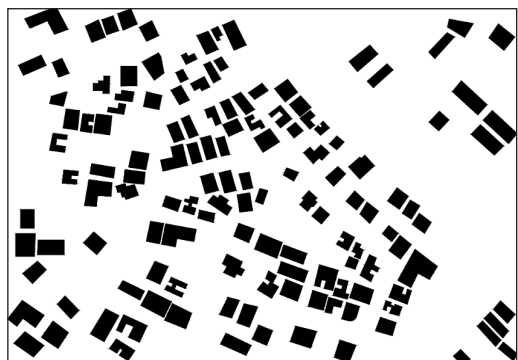


Figure 1. The wind rose from June to August for 24 hours (Betti et al., 2021)



(a)



(b)

Figure 2. (a) Satellite view and (b) solid and void analysis for field survey study area at southern of Alqwsmeh district in Amman

buildings represent only 12.19%. The selected buildings have the same construction system and consist of three floors to ensure the study's accuracy. Only two rooms on the middle floor of each building are selected to investigate the effect of building shape on thermal performance by analyzing the thermal parameters (airspeed, air temperature, mean radiant temperature, and relative humidity) with and without natural ventilation. Figure 3 shows the plan of the middle floor of the buildings, where the studied rooms are named Room 1, which has only a west window, and Room 2, which has west and north windows. Figure 4 shows the main facade of the buildings directed toward the west.

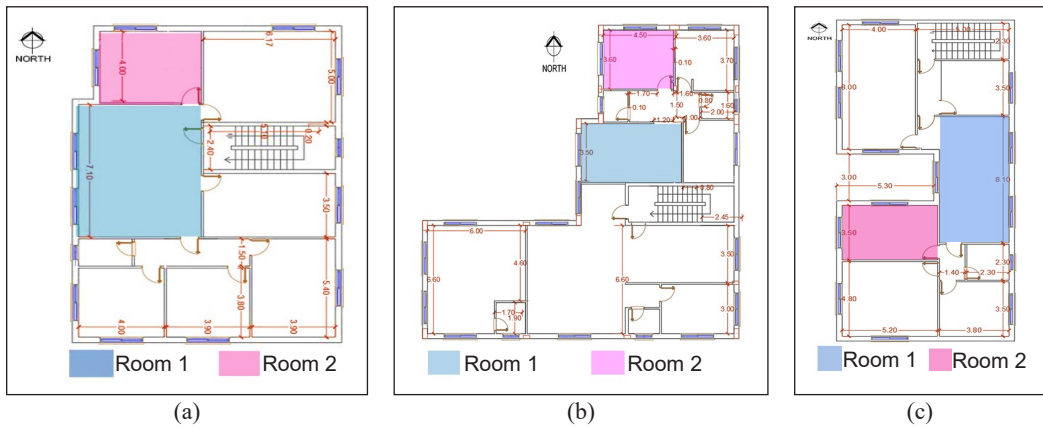


Figure 3. Plans for residential buildings with different forms studied: (a) Rectangular; (b) L-shape; and (c) U-shape



Figure 4. The main facades of the studied buildings: (a) Rectangular; (b) L-shape; and (c) U-shape

## METHODS

### Simulation Validation

EnergyPlus is one of the most extensive simulation software validated and used for building energy and thermal performance by the research community (Muslim, 2021; González et al., 2020). It is essentially the accumulation of 65 years of experience by the US Department

of Energy (DOE). It has been tested and validated under the comparative Standard Method of Test for the Evaluation of Building Energy Analysis Computer Programs BESTEST/ASHRAE Standard 140.

A model for the rectangular building is designed to ensure the validity of EnergyPlus software in this work. The details of the design and simulation are furnished in the coming subsection. Then, a site measurement for the indoor air temperature, globe temperature, and relative humidity, under the basic conditions without ventilation (where all windows are closed, and no mechanical ventilation is used) for Room 1 in the rectangular building on 21 July. This date is selected as the design day in the current study since it represents an extreme summer day in Jordan, according to the climatic data from ASHRAE (2009). The indoor air temperature, globe temperature, and relative humidity are measured using the WBGT HI meter, which has a temperature accuracy of 0.1°C and a humidity resolution of 0.1 RH. The air velocity is measured via a high-accuracy digitally certified ST6816 anemometer with a resolution of 0.01 m/s. Figure 5 illustrates the experimental set-up of devices during the measurement. The results for site measurement are compared with the simulated results on the design day and tabulated in Table 2. The table shows that the simulated results are very close to the measured values, which confirms that the EnergyPlus simulator is valid in this work.



Figure 5. Experimental set-up during data collection

Table 2  
Comparison between measured and simulated results for simulator validation

Time (H)	Measured temperature (°C)	Simulated temperature (°C)	Measured mean radiant temperature (°C)	Simulated mean radiant temperature (°C)	Measured relative humidity (%)	Simulated relative humidity (%)
1:00	32.30	32.88	33.06	33.91	35.15	35.53
2:00	32.00	32.19	33.40	33.32	36.14	36.18
3:00	31.89	31.61	33.30	32.75	36.12	36.80
4:00	31.15	31.08	32.60	32.24	37.25	37.88
5:00	30.94	30.82	32.10	31.76	37.84	38.64
6:00	30.91	30.85	31.80	31.45	39.66	39.79
7:00	31.21	30.98	31.90	31.33	41.07	41.80
8:00	31.25	31.28	31.60	31.39	42.63	42.51
9:00	31.33	31.36	31.80	31.55	43.52	42.48
10:00	31.40	31.55	31.90	31.69	43.50	42.10

Table 2 (continue)

Time (H)	Measured temperature (°C)	Simulated temperature (°C)	Measured mean radiant temperature (°C)	Simulated mean radiant temperature (°C)	Measured relative humidity (%)	Simulated relative humidity (%)
11:00	31.59	31.97	31.80	31.87	42.12	40.63
12:00	31.68	32.32	31.70	32.06	41.32	38.63
13:00	32.05	32.78	32.10	32.57	38.54	35.75
14:00	33.00	33.65	33.20	33.59	36.72	33.11
15:00	34.23	34.81	34.60	34.87	33.38	31.57
16:00	35.27	36.12	35.40	35.98	31.22	30.43
17:00	36.20	37.23	35.80	36.53	30.80	29.30
18:00	36.60	37.55	35.70	36.33	32.67	29.50
19:00	36.30	37.18	35.20	35.78	33.57	31.85
20:00	35.47	36.87	34.30	35.38	35.45	33.36
21:00	35.00	36.48	33.90	35.09	38.03	34.02
22:00	33.84	35.35	33.50	34.68	38.38	37.08
23:00	32.38	33.93	32.90	34.17	41.45	40.18
24:00	31.71	32.82	32.80	33.60	42.19	41.67

**Simulation Procedures**

Using AutoCAD, the simulation for the buildings under investigation starts with drawing two-dimensional plans for three residential buildings, i.e., rectangular, L-shaped, and U-shaped. Three-dimensional models of the buildings have been built using the OpenStudio plugin SketchUp software, as shown in Figure 6. The existing buildings' models are constructed with three layers: a 3cm-thick stone block, 10cm-thick concrete, and a 10cm-thick hollow block. After that, the loads for the electrical equipment, people, and lights are calculated and included in the simulation. The load definitions used in the simulation are summarized in Table 3. Here, every building room is set as a thermal zone to evaluate each room's internal thermal performance individually. The weather file for Amman in

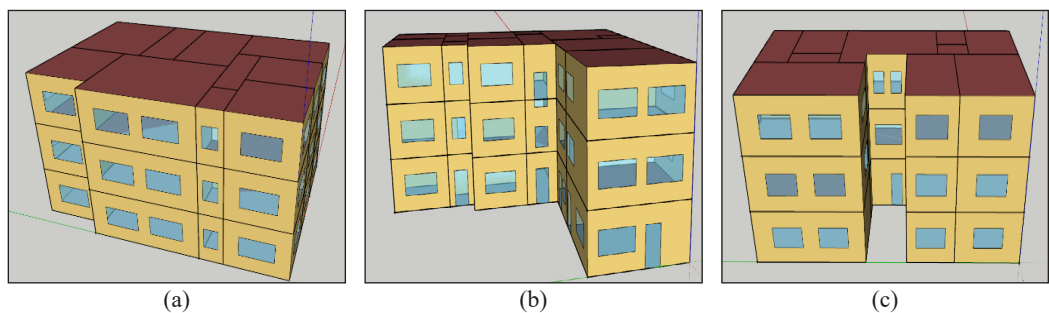


Figure 6. 3D building drawings using SketchUp Plugin OpenStudio: (a) Rectangular; (b) L-shape; and (c) U-shape



Table 3  
*Loads definition for the current work*

Definition	People	Lighting	Electrical equipment
Loads	6 persons	12 w/m <sup>2</sup>	Living room: 190 w Bedroom: 80 w Kitchen: 6010 w

EPW format is uploaded to the simulator to set the climatic parameters. In addition, the summer design day is set as 21 July 2022 and is imported in DDY format. Then, the output variables, particularly the air speed, air temperature, mean radiant temperature, and relative humidity, are measured every 15 minutes. The four readings taken per hour are averaged for high-accuracy measurement. The overall protocol for the simulation in this work is depicted in Figure 7. The thermal parameters for each building are analyzed for the unventilated case, as all openings of the middle floor are closed, and compared with output parameters for the naturally ventilated situation, as all openings of the same floor are opened. Finally, the thermal parameters for the buildings are used to analyze the results through the ASHRAE-55 adaptive model to find the comfort hours during the day.

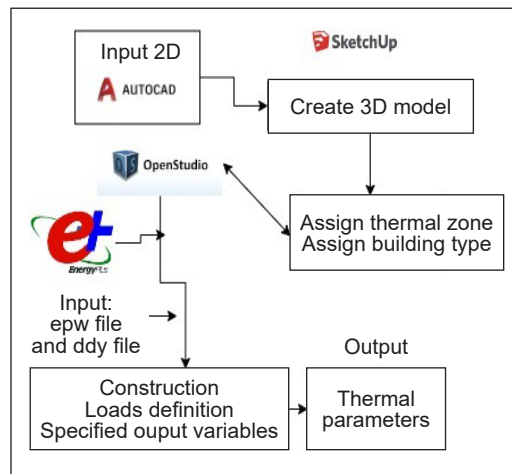


Figure 7. Summary of the simulation protocol

## RESULTS AND ANALYSIS

### Thermal Performance

Figure 8 shows the indoor air temperature comparison for the unventilated situation for the three buildings, while all openings are closed for Room 1 and Room 2. From Figure 8(a), the rectangular building has a better result in the daytime from 7:00 a.m. until 5:00 p.m., which is correlated to the compact rectangular shape; thus, it has less exposed surfaces' area to direct solar radiation. The U-shape building shows a minimum air temperature after 5:00 p.m. since it creates self-shading that blocks the direct solar radiation at the peak time for the west façade at 5:00 p.m., which reduces the heat stored inside the building during nighttime. However, the U-shape building shows the worst result between 7:00 a.m. and 5:00 p.m. since Room 1 has an additional east window that allows the solar heat to enter from early morning into the building and be stored inside. The L-shape building shows the highest indoor air temperature most of the time among the other buildings, except for the

U-shape between 7:00 a.m. to 5:00 p.m., since it has a larger surface/volume ratio compared to the rectangular, as can be seen in Table 4, and less self-shading effective compared to U-shape case. It is reported that increasing the surface/volume ratio increases solar radiation exposure while self-shading for buildings enhances thermal parameters value (Muhaisen & Abed, 2015; Mohsenzadeh et al., 2021). An almost similar trend for the mean radiant temperature for Room 1 is observed, as shown in Figure 9(a).

Figure 8(b) shows that for Room 2, the L-shape shows the highest indoor air temperature, followed by the U-shape and then the rectangular one. It can be explained using the same theory discussed in Figure 8(a), where the L-shape building has a high surface-to-volume ratio while the U-shape has self-shading, which reduces direct solar radiation. Figure 8 shows that Room 1 is hotter than Room 2 for the unventilated situation because Room 2 has two windows at the north and west facade, as shown in Figure 2. The northern window is not exposed to direct sun during daytime; thus, the outdoor temperature from the north facade is expected to be lower than the indoor temperature for Room 2. This temperature gradient may cause some loss through window glass from the room toward the outside. A similar trend for the mean radiant temperature for Room 2 is also observed, as shown in Figure 9(b).

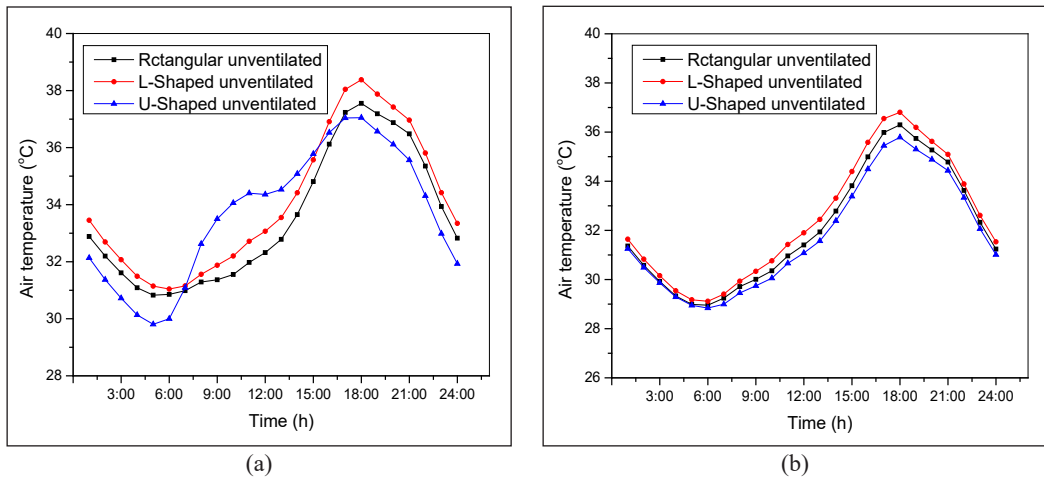


Figure 8. Air temperature comparison for the unventilated situation of different building shapes: (a) Room 1; and (b) Room 2

Table 4  
Shape parameters for the buildings in the study

	Rectangle	L-shape	U-shape
Surface area	576.00 m <sup>2</sup>	668.23 m <sup>2</sup>	582.38 m <sup>2</sup>
Volume	604.80 m <sup>3</sup>	666.51 m <sup>3</sup>	552.99 m <sup>3</sup>
S/V ratio	0.9524	1.0025	1.0531

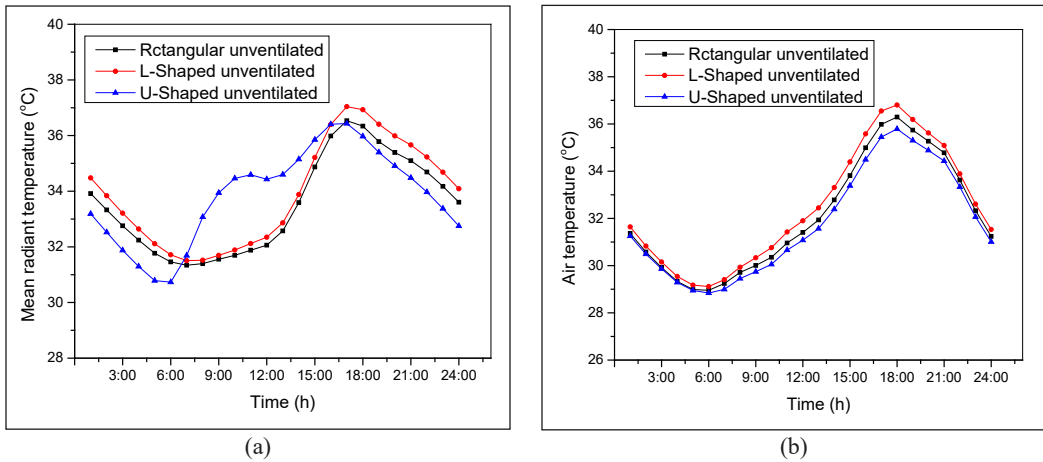


Figure 9. Mean radiant temperature comparison for the unventilated situation of different building shapes: (a) Room 1; and (b) Room 2

Figure 10 shows the indoor air temperature for the ventilated case while all windows are opened. Comparable to the unventilated case, the results show a considerable drop in indoor air temperature by almost 10°C at nighttime due to the cold outdoor climate. In comparison, it reduces by 4–8°C during the daytime. Table 5 summarizes the results of the optimum reduction in the indoor air temperature for both rooms in all building forms for the ventilation case compared to the unventilated case. This drop is attributed to the natural ventilation through the buildings that leads to renewed air and discharges the heat, which enhances indoor air quality (Hughes et al., 2012). Similar results for indoor temperature drop are reported by Yu et al. (2018), Al-Hemiddi and Al-Saud (2001), and Ma'bdeh et al. (2020). From Figure 10(a), one can notice that Room 1 in the U-shaped building shows the lowest temperature at night compared to other forms.

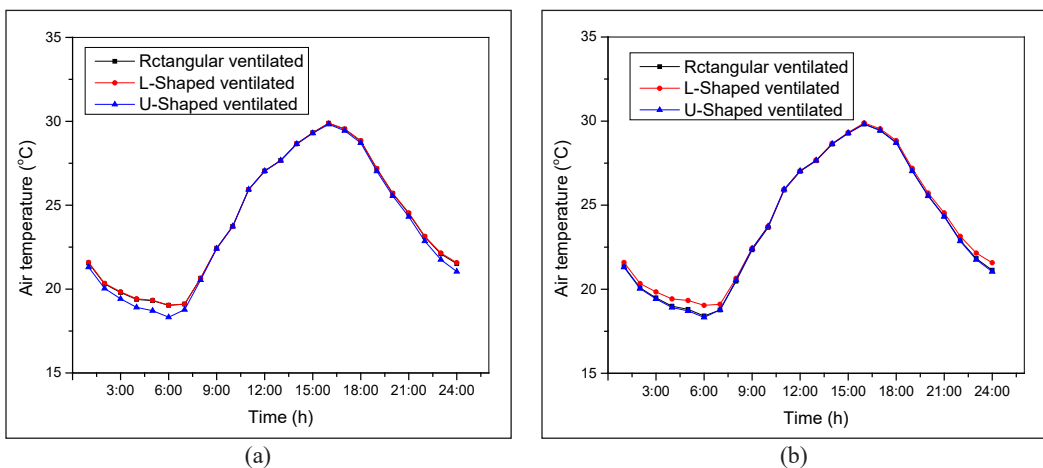


Figure 10. Air temperature comparison for different ventilated building shapes for: (a) Room 1; and (b) Room 2



Table 5

Average air temperature reduction after natural ventilation compared to the basic situation

Rectangular Shape		L-Shape		U-Shape	
Room 1	Room 2	Room 1	Room 2	Room 1	Room 2
4.99–12.24°C	4.14–10.71°C	5.70–12.60°C	4.79–10.97°C	6.43–12.31°C	3.76–10.42°C

Furthermore, from Figure 10 (b), one can notice that Room 2 in the L-shaped building shows a slightly higher temperature at nighttime than other buildings. Since the air temperature does not differ much between the buildings throughout the day, the mean radiant temperature has been plotted for Room 1 and Room 2, as shown in Figure 11. The results show good agreement with the discussion for air temperature variation with more details for the temperature throughout the day. For Room 1 in Figure 11(a), the U-shaped building shows a higher temperature from 7:00 a.m. to 5:00 p.m., possibly due to the east window's exposure to solar radiation; obviously, a higher difference is observed during early morning hours. From Figure 11(b), the L-shape building shows the highest temperature all day long for room Room 2, while the U-shaped building still has the lowest temperature most of the time among all buildings because of the self-shading effect.

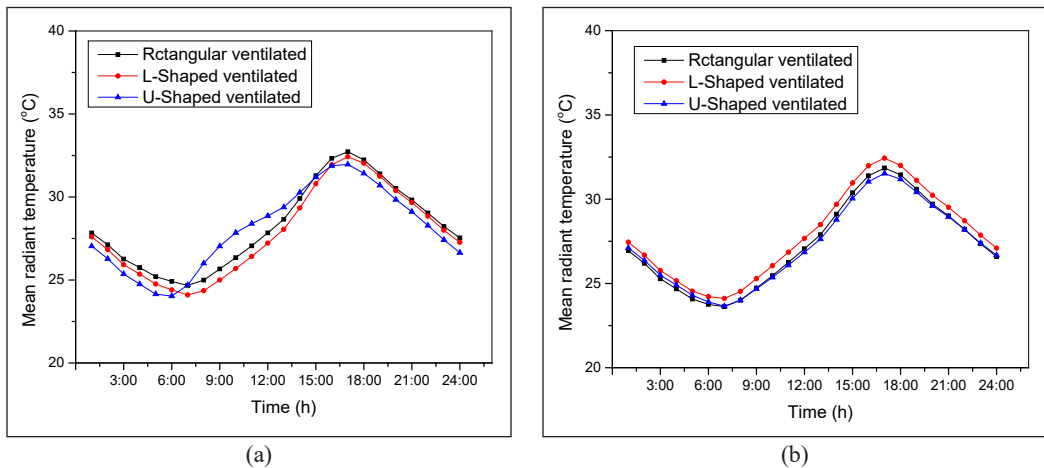


Figure 11. Mean radiant temperature comparison for the ventilated situation of different building shapes: (a) Room 1; and (b) Room 2

### Thermal Comfort

ASHRAE-55 adaptive model is a widely certified standard to evaluate thermal comfort, where there are continual revisions and updates for standard documents that mirror the latest results regarding thermal comfort zones from field experiments. ASHRAE-55 (2020) is the newest version after a long line of editions and publications from 1966 until the previous version, ASHRAE (2017). The newest version of the adaptive model shows a graphical method for Occupant-Controlled Naturally ventilated spaces, representing the

operative temperature versus the adopted prevailing mean outdoor air temperature, as shown in Figure 12(a). The prevailing mean outdoor air temperature is the arithmetic average of daily outdoor temperatures. It is usually evaluated not less than seven days and not more than 30 days before the day under investigation, and it is given by Equation 1 (Kim et al., 2019; Saif et al., 2021):

$$\overline{t_{pma(out)}} = (1 - \alpha) [t_{e(d-1)} + \alpha t_{e(d-2)} + \alpha^2 t_{e(d-3)} + \alpha^3 t_{e(d-4)} + \dots] \quad [1]$$

where  $\alpha$  is an exponential constant, its value ranging from 0 to 1, and the recommended value for  $\alpha$  is 0.8,  $t_{e(d-1)}$  is the mean external air temperature for the previous day,  $t_{e(d-2)}$  is the mean external air temperature for the previous 2 days.

The upper and lower limits for 80% acceptability limit and 90% acceptability limit for comfort area usually follow Equations 2 to 5 (Saif et al., 2021).

The upper limit of 90% acceptability:

$$= 0.31 \times \overline{t_{pma(out)}} + 20.3 \text{ [}^\circ\text{C]} \quad (10 \leq \overline{t_{pma(out)}} \leq 33.5) \quad [2]$$

The lower limit of 90% acceptability:

$$= 0.31 \times \overline{t_{pma(out)}} + 15.3 \text{ [}^\circ\text{C]} \quad (10 \leq \overline{t_{pma(out)}} \leq 33.5) \quad [3]$$

The upper limit of 80% acceptability:

$$= 0.31 \times \overline{t_{pma(out)}} + 21.3 \text{ [}^\circ\text{C]} \quad (10 \leq \overline{t_{pma(out)}} \leq 33.5) \quad [4]$$

The lower limit of 80% acceptability:

$$= 0.31 \times \overline{t_{pma(out)}} + 14.3 \text{ [}^\circ\text{C]} \quad (10 \leq \overline{t_{pma(out)}} \leq 33.5) \quad [5]$$

where  $t_{pma(out)}$  is the prevailing mean outdoor air temperature.

The increment in the average air speed increases the operative temperature, which widens the upper limit of the thermal comfort zone, as shown in Figure 12(b) (ASHRAE, 2017; ASHRAE, 2020; Bienvenido-Huertas et al., 2022). For example, as the average air speed increases from 0.3 m/s to 0.6 m/s, the operative temperature increases by 1.2°C, as the average air speed becomes 0.9 m/s, the operative temperature increases by 1.8°C, and when the average air speed reaches 1.2 m/s, the operative temperature increases by 2.2°C (ASHRAE, 2017; ASHRAE, 2020).

In order to find out the comfort zone for the rooms in all building forms, the adaptive comfort graphs have been plotted using the CBE thermal comfort tool, in which the air temperature, relative humidity, mean radiant temperature, average air speed, prevailing mean outdoor air temperature, metabolic rate, and clothing parameters are inserted to the tool for each hour individually. The metabolic rate is set as 1.2 met, and clothing is set as

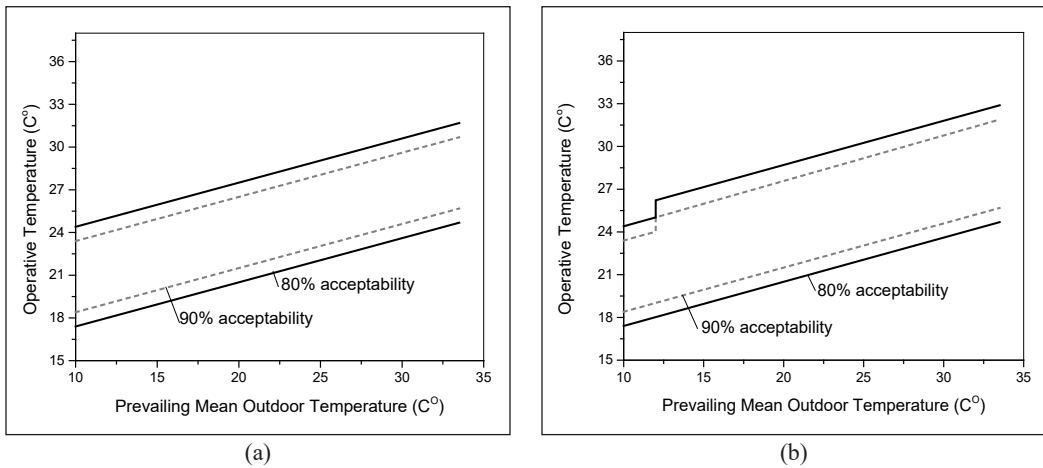


Figure 12. Modulation of the upper limit for thermal comfort zone according to the criteria established by ASHRAE-55: (a) Average air speed 0.3 m/s; and (b) average air speed 0.6 m/s

0.5 clo. Figures 13, 14, and 15 show the ASHRAE-55 adaptive models for Room 1 in all building forms with and without natural ventilation. The results show that Room 1, under basic conditions in all building forms, is out of thermal comfort zones, i.e., 80% and 90% acceptability. However, in the ventilation case, it can be noticed that a notable shift of most of the day hours to be within the comfort zone; the same observation is reported by Heracleous and Michael (2018) and Kumar et al. (2018). Room 1 in the rectangular shape building recorded 66.67% of the total hours during the day within 90% acceptability, while 16.67% of the total time within 80% acceptability and 16.67% of the day hours still within the uncomfortable status, specifically, the non-acceptable time is recorded between

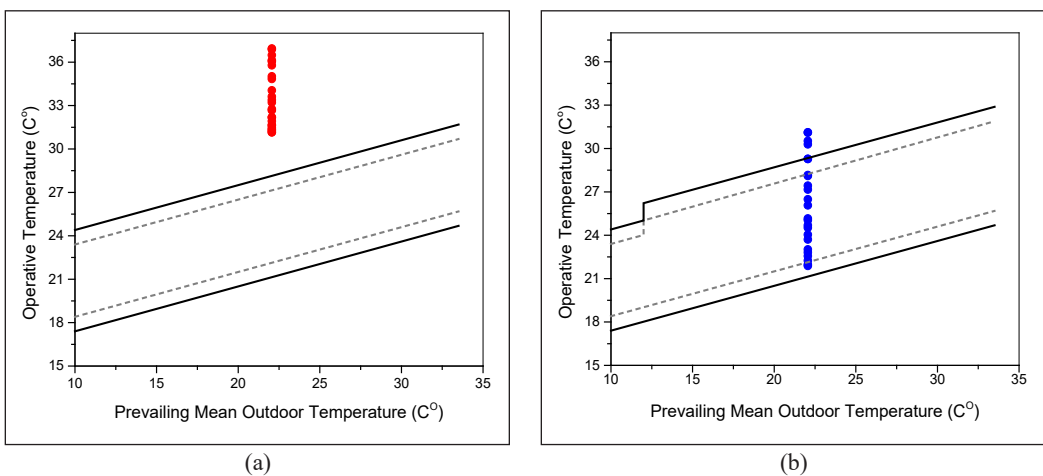


Figure 13. ASHRAE-55 adaptive model: (a) Unventilated; and (b) ventilated situation for Room 1 in the rectangular building

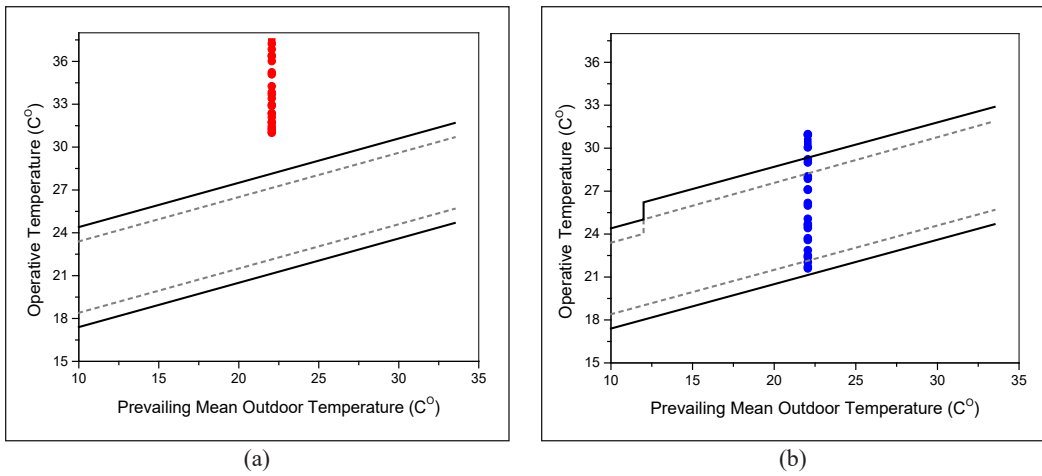


Figure 14. ASHRAE-55 adaptive model: (a) unventilated; and (b) ventilated situation for Room 1 in the L-shape building

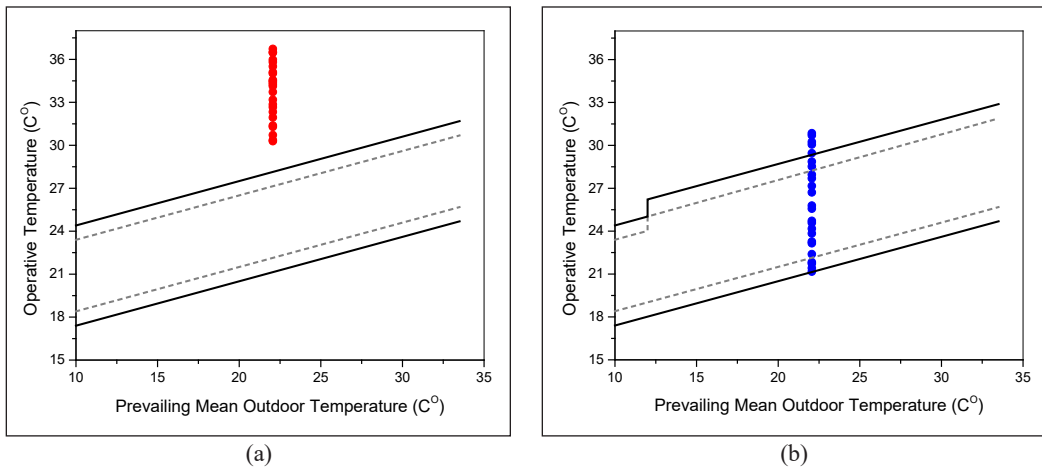


Figure 15. ASHRAE-55 adaptive model: (a) unventilated; and (b) ventilated situation for Room 1 in the U-shape building

4:00–6:00 p.m. Room 1 in the L-shape building shows that the comfort time within 90% acceptability reduced to 62.5% of total day hours. At the same time, the comfortable time for 80% acceptability increases to 20.83% of total day hours.

However, the uncomfortable status within the peak hours (4:00–6:00 p.m.) remains at 16.67% of day hours. For Room 1 in the U-shape building, the 90% acceptability, 80% acceptability, and uncomfortable status records 58.33%, 20.83%, and 20.83% of hours during the day, respectively. 83.33% of the uncomfortable time is between 2:00–6:00 p.m. and 3:00–6:00 p.m. for other buildings. Thus, it can be concluded that the rectangular-shaped building records the best performance compared to the other

building shapes since it usually presents high acceptability. At the same time, the U-shape shows a higher percentage of uncomfortable situations and less acceptability time. The uncomfortable high percentage of the U-shaped building is attributed to the eastern and western windows; thus, Room 1 will face direct sun radiation for longer, especially in the morning. Subsequently, the mean radiant temperature (MRT) for Room 1 in the U-shaped building records 1.5°C higher during morning hours than the other buildings, as illustrated in Figure 9(a). Furthermore, it is noticed that some of the 80% acceptability hours of the rectangular shape building are recorded in the early morning between 6:00–7:00 a.m. and show a slightly cool sensation, which might be shifted toward 90% acceptability by minimizing the number of opening windows at this time, thus reduces the ventilation effect.

Figures 16, 17, and 18 show the adaptive model for Room 2 performance in both cases with and without ventilation for all building forms. The results show that Room 2 in all buildings under ventilation conditions is out of the thermal comfort zone. On the other hand, for the ventilation case, Room 2 in all building forms records 58.33%, 25%, and 16.67% of day hours for the 90% acceptability, 80% acceptability, and uncomfortable status, respectively. The unacceptability time in Room 2 for all shapes has a warm sensation recorded between 3:00–6:00 p.m. Interestingly, it is noticed that the number of uncomfortable hours close to the 80% acceptability zone in the rectangular and the U-shape building relatively are more than for L-shape; this attributed to the compactness in rectangular shape and self-shading in U-shape, which indicates that their performance is relatively better than L-shaped building. Table 6 summarizes the comfort hours percentage for Room 1 and Room 2 in different building forms under ventilation conditions according to ASHRAE-55.

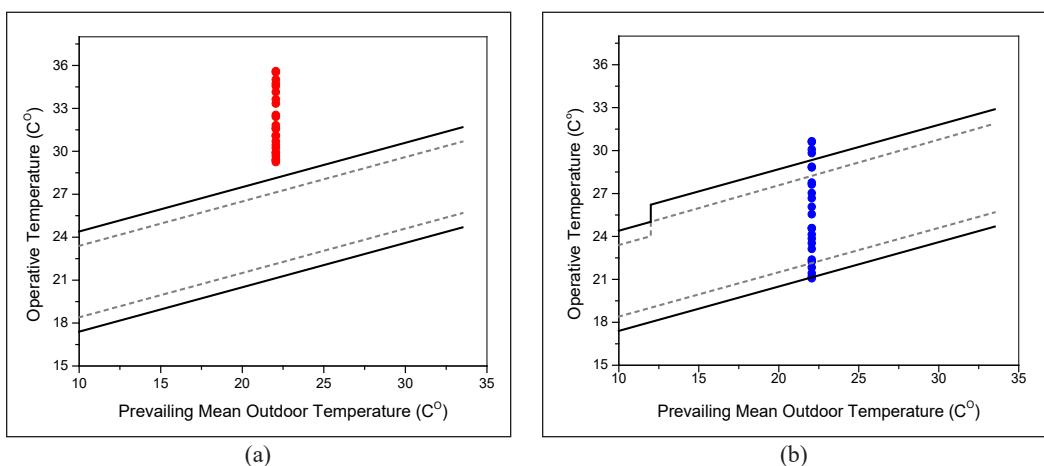


Figure 16. ASHRAE-55 adaptive model: (a) unventilated; and (b) ventilated situation for Room 2 in the rectangular building

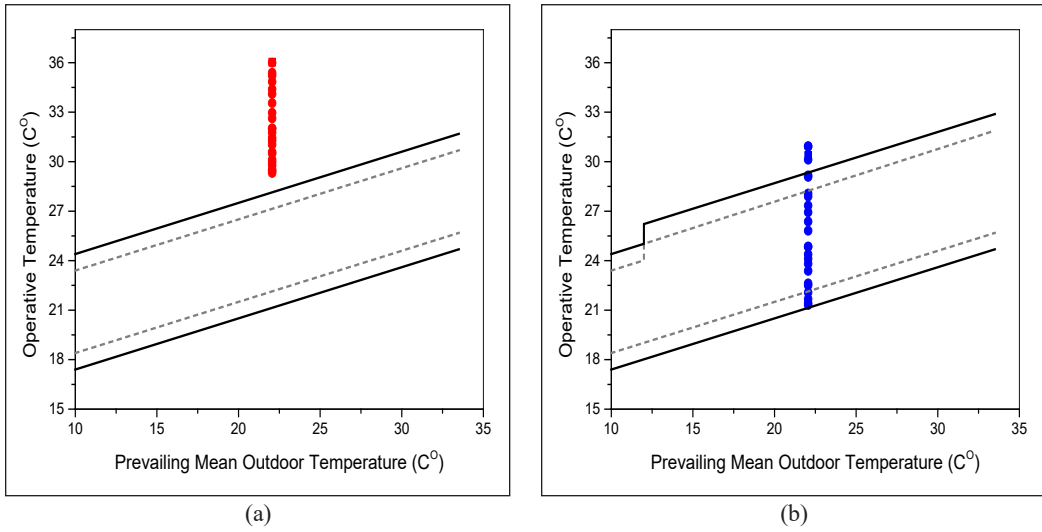


Figure 17. ASHRAE-55 adaptive model: (a) unventilated; and (b) ventilated situation for Room 2 in the L-shape building

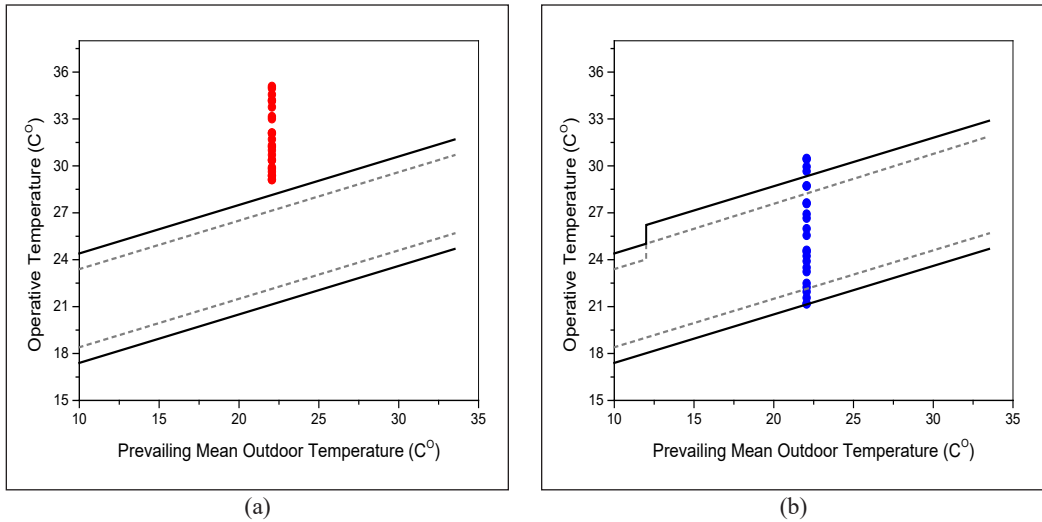


Figure 18. ASHRAE-55 adaptive model: (a) unventilated; and (b) ventilated situation for Room 2 in the U-shape building

Table 6  
Comfort hours percentage for Room 1 and Room 2 under ventilation conditions

Comfort percentage	Rectangular Shape		L-Shape		U-Shape	
	Room 1	Room 2	Room 1	Room 2	Room 1	Room 2
90% acceptability	66.67%	58.33%	62.50%	58.33%	58.33%	58.33%
80% acceptability	16.67%	25.00%	20.83%	25.00%	20.83%	25.00%
Out of ASHRAE comfort zone	16.67%	16.67%	16.67%	16.67%	20.83%	16.67%

## DISCUSSION

The impact of natural ventilation on different building forms is investigated in two situations: with and without ventilation. The thermal behavior has been examined for basic (all windows are closed) and natural ventilation cases (all windows are opened). The building form is more effective on the thermal performance when the windows are closed, with no ventilation case, since it depends on many factors such as surface-to-volume ratio, building compactness, and self-shading.

Rectangular shape building recorded the best results in the daytime for Room 1 since the compact shape minimizes exposed solar radiation. However, the L-shape shows second-ranked results because of the large surface/volume ratio; conversely, the U-shape has the worst results due to the opposite side opening. Thus, it is facing the sunrise and sunset times.

Unventilated Room 2 in the U-shape building shows better performance as compared with other forms since the maximum temperature recorded is 35°C, and the worst thermal performance for the same room is obtained for the L-shape building regarding larger surface-to-wall ratio for the building as compared to rectangular case and no self-shading as compared to U-shape case.

Natural ventilation enhances the thermal behavior of all buildings in the hot-dry climate, where it remarkably reduces the indoor air temperature and mean radiant temperature; as a result, it improves the indoor thermal performance as the buildings transform from non-comfort zone to acceptable comfort situations in most of the time according to the ASHRAE-55 adaptive model.

Room 1 in the rectangular building shows the best thermal performance according to ASHRAE-55 adaptive model as compared to the other shapes since it is recorded 66.67% of the time within 90% acceptability, 16.67% of the time within 80% acceptability, and 16.67% of the time within the uncomfortable status. For Room 2 under ventilation conditions, all building forms record the same hour percentage of comfortability according to the ASHRAE-55 adaptive model. L-shape building presents relatively less thermal performance since it records fewer hours within the uncomfortable zone, close to 80% acceptability comfort zone.

Table 7 compares natural ventilation's impact on improving indoor air temperature and comfort hours between the current work and the previous research.

Table 7

*Comparison of the impact of natural ventilation between current work and the previous research*

Reference	Climate	Approach	Findings
Al-Hemiddi and Al-Saud, 2001	Hot-arid	Experimental	Cross-ventilation provides cool indoor air.
Omrani et al., 2017	Warm-humid summers and mild to cool winters	Experimental	- Cross ventilation maintains comfortable thermal conditions 70% of the time. - Indoor condition in single-sided ventilation has an average of 3°C hotter than cross ventilation.

Table 7 (continue)

Reference	Climate	Approach	Findings
Kumar et al., 2018	Hot semi-arid	Experimental	- Natural ventilation reduces 40% and 98% of discomfort time in summer and winter, respectively.
Heracleous & Michael, 2018	Mediterranean	Simulation.	- Natural ventilation reduces operative temperature.
Mastouri et al., 2019	Hot Semi-Arid	Simulation.	- Night ventilation reduces the operative temperature of the ground floor by 2°C and 3°C on the first floor.
Ma'bdeh et al., 2020	Dry-hot	Simulation.	- Ventilation improves indoor air quality. - The comfort hours increased by 106 h during February and 170 h during August.
Current work	Hot-dry summer	Simulation	-Natural ventilation reduces indoor air temperature by almost 10°C at nighttime and 4–8°C during the daytime. - Ventilation increases comfort hours percentage (90% acceptability) from 58.33% to 66.67%.

## CONCLUSION

A simulation experiment using the EnergyPlus simulator is conducted in this paper to investigate the impact of ventilation on the thermal performance and thermal comfort of three residential buildings of different forms, namely rectangular, L-shape, and U-shape, which are in Amman, the capital of Jordan, that is characterized with a hot-dry climate in the summer season. Models for the proposed buildings according to their actual constructions are designed using the OpenStudio plugin SketchUp software. After defining the loads and the design day, they are simulated with the aid of the EnergyPlus simulator. As a preliminary step, the simulator is validated by comparing the simulated data with the site-measured results for Room 1 within the rectangular building, which shows good matching. The simulation results show that natural ventilation greatly reduces the indoor thermal temperature of the rooms under investigation. It drops by almost 10°C at nighttime due to the cold outdoor climate, while during the daytime, it reduces by 4–8°C. This reduction in air temperature is correlated to the influence of natural ventilation to refresh the indoor air and discharge the heat, subsequently improving indoor air quality. Comparably, Room 1 in the U-shaped building shows a lower temperature at night compared to other forms, and Room 2 in the L-shaped building shows a relatively higher temperature than the other buildings. The thermal comfort for the rooms has been evaluated using the ASHRAE-55 adaptive model to find out that the natural ventilation has a remarkable impact on shifting the 100% out-of-comfort rooms to less than 20% of the total design day hours. Comparably, Room 1 in the rectangular building records 66.67% of total hours during the day within the 90% comfort acceptability according to ASHRAE-55, which is the highest among the



other forms. Room 2 in all building forms records the exact acceptability percentages of 58.33%, 25%, and 16.67% of day hours for the 90% acceptability, 80% acceptability, and uncomfortable status, respectively. Based on these results, it can be concluded that natural ventilation through windows in the hot-dry climate significantly improves indoor thermal performance, especially for the rectangular and U-shape forms. For further understanding of the influence of natural ventilation, it is highly recommended to consider investigating parameters such as type of ventilation and window-to-wall ratio and correlate them to the building form.

## ACKNOWLEDGEMENT

The authors thank the houses' owners, Mr. Ahmad Saif, Mrs. Kholoud Altartir, and Mr. Shehadeh Hussein, for providing the sketch plans and facilitating the site measurements.

## REFERENCES

- Al-Hemiddi, N. A., & Al-Saud, K. A. M. (2001). The effect of a ventilated interior courtyard on the thermal performance of a house in a hot-arid region. *Renewable Energy*, 24(3-4), 581-595. [https://doi.org/10.1016/S0960-1481\(01\)00045-3](https://doi.org/10.1016/S0960-1481(01)00045-3)
- Ali, H. H., Al Zoubi, H., & Badarneh, S. (2010). Energy efficient design for thermally comforted dwelling units in hot arid zones: Case of vernacular buildings in Jordan. In *Conference on Technology & Sustainability in the Built Environment* (Vol. 1, p. 279-304). King Saud University - College of Architecture and Planning.
- Almeida, R. M. S. F., Pinto, M., Pinho, P. G., & de Lemos L. T. (2017). Natural ventilation and indoor air quality in educational buildings: Experimental assessment and improvement strategies. *Energy Efficiency*, 10, 839-854. <https://doi.org/10.1007/s12053-016-9485-0>
- Almuhtady, A., Alshwawra, A., Al Faouri, M., Al-Kouzu, W., & Al-Hinti, I. (2019). Investigation of the trends of electricity demands in Jordan and its susceptibility to the ambient air temperature towards sustainable electricity generation. *Energy, Sustainability and Society*, 9, 1-18. <https://doi.org/10.1186/s13705-019-0224-1>
- ASHRAE. (2009). *2009 ASHRAE Handbook: Fundamentals*. American Society of Heating, Refrigerating and Air-Conditioning Engineers, Atlanta, USA. <https://www.worldcat.org/title/2009-ASHRAE-handbook--fundamentals/oclc/525070649>
- ASHRAE. (2017). *Thermal Environmental Conditions for Human Occupancy*. American Society of Heating Ventilating and Air-conditioning Engineers, Atlanta, USA. [https://www.ashrae.org/file%20library/technical%20resources/standards%20and%20guidelines/standards%20addenda/55\\_2017\\_d\\_20200731.pdf](https://www.ashrae.org/file%20library/technical%20resources/standards%20and%20guidelines/standards%20addenda/55_2017_d_20200731.pdf)
- ASHRAE. (2020). *Standard 55-2020-Thermal Environmental Conditions for Human Occupancy (ANSI Approved)*. American Society of Heating Ventilating and Air-conditioning Engineers, Atlanta, USA. [https://www.techstreet.com/ashrae/standards/ashrae-55-2020?product\\_id=2207271#amendments](https://www.techstreet.com/ashrae/standards/ashrae-55-2020?product_id=2207271#amendments)
- Bekkouche, S. M. A., Benouaz, T., Cherier, M.K., & Hamdani, M. (2013). Influence of the compactness index to increase the internal temperature of a building in Saharan climate. *Energy Building*, 66, 678-687. <http://dx.doi.org/10.1016/j.enbuild.2013.07.077>

- Betti, G., Tartarini, F., Schiavon, S., & Nguyen, C. (2021). *CBE Clima Tool. Version 0.4.6*. Center for the Built Environment, University of California Berkeley. <https://clima.cbe.berkeley.edu>
- Bienvenido-Huertas, D., Sánchez-García, D., Rubio-Bellido, C., & Solís-Guzmán, J. (2022). Using adaptive strategies of natural ventilation with tolerances applied to the upper limit to improve social dwellings' thermal comfort in current and future scenarios. *Science and Technology for the Built Environment*, 28, 527-546. <https://doi.org/10.1080/23744731.2022.2040884>
- Chen, H., Du, R., Ren, W., Zhang, S., Du, P., & Zhang, Y. (2021). The microbial activity in PM2.5 in indoor air: As an index of air quality level. *Aerosol and Air Quality Research*, 21, Article 200101. <https://doi.org/10.4209/aaqr.2020.03.0101>
- Deng, X., Wang, M., Sun, D., & Fan, Z. (2020). Effect of building form on energy consumption of academic library buildings in different climate zones in China. *IOP Conference Series: Earth and Environmental Science*, 531(1), Article 012060. <https://dx.doi.org/10.1088/1755-1315/531/1/012060>
- Elnagar, E., & Köhler, B. (2020). Reduction of the energy demand with passive approaches in multifamily nearly zero-energy buildings under different climate conditions. *Frontiers in Energy Research*, 8, Article 545272. <https://doi.org/10.3389/fenrg.2020.545272>
- González, V. G., Ruiz G. R., & Bandera, C. F. (2020). Empirical and comparative validation for a building energy model calibration methodology, *Sensors*, 20, Article 5003. <https://doi.org/10.3390/s20175003>
- Heracleous, C., & Michael, A. (2018). Assessment of overheating risk and the impact of natural ventilation in educational buildings of Southern Europe under current and future climatic conditions. *Energy*, 165, 1228-1239. <https://doi.org/10.1016/j.energy.2018.10.051>
- Hughes, B. R., Calautit, J. K., & Ghani, S. A. (2012). The development of commercial wind towers for natural ventilation: A review. *Applied Energy*, 92, 606-627. <https://doi.org/10.1016/j.apenergy.2011.11.066>
- Kim, A., Wang, S., & Kim, J. (2019). Dorothy reed, indoor/outdoor environmental parameters and window-opening behavior: A structural equation modeling analysis. *Buildings*, 9(4), Article 94. <https://doi.org/10.3390/buildings9040094>
- Kocagil, I. E., & Oral, G. K. (2015). The effect of building form and settlement texture on energy efficiency for hot dry climate zone in Turkey. *Energy Procedia*, 78(4), 1835-1840. <https://doi.org/10.1016/j.egypro.2015.11.325>
- Krarti, M. (2018). *Integrated Design and Retrofit of Buildings. Optimal Design and Retrofit of Energy Efficient Buildings. Communities, and Urban Centers*. Elsevier. <https://doi.org/10.1016/C2016-0-02074-0>
- Kumar, S., Singh, M. K., Mathur, A., Mathur, S., & Mathur, J. (2018). Thermal performance and comfort potential estimation in low-rise high thermal mass naturally ventilated office buildings in India: An experimental study. *Journal of Building Engineering*, 20, 569-584. <https://doi.org/10.1016/j.jobe.2018.09.003>
- Lapisa, R. (2019). The effect of building geometric shape and orientation on its energy performance in various climate regions. *International Journal of GEOMATE*, 16(63), 113-119. <http://dx.doi.org/10.21660/2019.53.94984>

- Ma'bdeh, S. N., Al-Zghoul, A., Alradaideh, T., Bataineh, A., & Ahmad, S. (2020). Simulation study for natural ventilation retrofitting techniques in educational classrooms - A case study. *Heliyon*, 6(10), Article e0517. <https://doi.org/10.1016/j.heliyon.2020.e05171>
- Mastouri, H., Radoine, H., Bahi, H., Benhamou, B., & Hamdi, H. (2019). Effect of natural ventilation on the thermal performance of a residential building in a hot semi-arid climate. In *2019 7th International Renewable and Sustainable Energy Conference (IRSEC)* (p. 1-6). IEEE Publishing. <http://dx.doi.org/10.1109/IRSEC48032.2019.9078215>
- Mohsenzadeh, M., Marzbali, M. H., Tilaki, M. J. M., & Abdullah, A. (2021). Building form and energy efficiency in tropical climates: A case study of Penang, Malaysia. *urbe Revista Brasileira de Gestão Urbana*, 13, Article e20200280. <https://doi.org/10.1590/2175-3369.013.e20200280>
- Muhaisen, A. S., & Abed, H. M. (2015). Effect of building proportions on the thermal performance in the mediterranean climate of the Gaza strip. *Journal of Engineering Research and Technology*, 2(2), 112-121.
- Mushtaha, E., & Helmy, O. (2017). Impact of building forms on thermal performance and thermal comfort conditions in religious buildings in hot climates: A case study in Sharjah City. *International Journal of Sustainable Energy*, 36(10), 1-19. <https://doi.org/10.1080/14786451.2015.1127234>
- Muslim, S. A. (2021). EnergyPlus - Towards the selection of right simulation tool for building energy and power systems research. *Journal of Energy and Power Technology*, 3(3), Article 2103034. <http://dx.doi.org/10.21926/jept.2103034>
- Nagy, R., Mečiarová, L., Vilčeková, S., Burdová, E. K., & Košičanová, D. (2019). Investigation of a ventilation system for energy efficiency and indoor environmental quality in a renovated historical building: A case study. *International Journal of Environmental Research and Public Health*, 16(21), Article 4133. <https://doi.org/10.3390/ijerph16214133>
- Nazer, H. A. (2019). *Developing an energy benchmark for residential apartments in Amman*. Jordan Green Building Council. <https://library.fes.de/pdf-files/bueros/amman/15926.pdf>
- Omriani, S., Garcia-Hansen, V., Capra, B. R., & Drogemuller, R. (2017). Effect of natural ventilation mode on thermal comfort and ventilation performance: Full-scale measurement. *Energy and Buildings*, 156, 1-16.
- Ozarisoy, B. (2022). Energy effectiveness of passive cooling design strategies to reduce the impact of long-term heatwaves on occupants' thermal comfort in Europe: Climate change and mitigation. *Journal of Cleaner Production*, 330, Article 129675. <https://doi.org/10.1016/j.jclepro.2021.129675>
- Raji, B., Tenpierik, M. J., Bokel, R., & Dobbelsteen, A. V. D. (2020). Natural summer ventilation strategies for energy-saving in high-rise buildings: A case study in the Netherlands. *International Journal of Ventilation*, 19(1), 25-48. <https://doi.org/10.1080/14733315.2018.1524210>
- Raof, B. Y. (2017). The correlation between building shape and building energy performance. *International Journal of Advanced Research (IJAR)*, 5(5), 552-561. <http://dx.doi.org/10.21474/IJAR01/4145>
- Rodrigues, A. M., Santos, M., Gomes, M.G., & Duarte, R. (2019). Impact of natural ventilation on the thermal and energy performance of buildings in a Mediterranean climate. *Buildings*, 9(3), 1-17. <https://doi.org/10.3390/buildings9050123>

- Saif, J., Wright, A., Khattak, S., & Elfadli, K. (2021) Keeping cool in the desert: Using wind catchers for improved thermal comfort and indoor air quality at half the energy. *Buildings*, 11(3), Article 100. <https://doi.org/10.3390/buildings11030100>
- Yang, L., Liu, X., Qian, F., & Du, S. (2019). Ventilation effect on different position of classrooms in “line” type teaching building. *Journal of Cleaner Production*, 209, 886-902. <https://doi.org/10.1016/j.jclepro.2018.10.228>
- Yu, C. R., Guo, H. S., Wang, Q. C., & Chang, R. D. (2020). Revealing the impacts of passive cooling techniques on building energy performance: A residential case in Hong Kong. *Applied Sciences*, 10(12), Article 4188. <https://doi.org/10.3390/app10124188>
- Yu, J., Ye, H., Xu, X., Huang, J., Liu, Y., & Wang, J. (2018). Experimental study on the thermal performance of a hollow block ventilation wall. *Renewable Energy*, 122, 619-631. <https://doi.org/10.1016/j.renene.2018.01.126>

## Hydrocyanic Acid, Protein Concentration, and Phytochemical Compounds of Pulut and White Varieties in Young and Matured Cassava (*Manihot Esculenta*, Crantz)

Sarah Idris<sup>1</sup>, Rosnah Shamsudin<sup>1,2\*</sup>, Mohd Zuhair Mohd Nor<sup>2</sup>, Mohd Noriznan Mokhtar<sup>2</sup> and Siti Salwa Abd Gani<sup>1</sup>

<sup>1</sup>Halal Products Research Institute, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Department of Process and Food Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

### ABSTRACT

In Malaysia, there is a demand and a promising market for cassava (*Manihot esculanta*, Crantz) leaves as a supplementary animal feed because of their nutritional value and availability throughout the year. However, cyanide accumulation has been a problem due to its toxicity in animal feed; therefore, finding the best variety with low cyanide, high protein, and phytochemical content can address this issue. The hydrocyanic (HCN) contents were analyzed for two local varieties, White and Pulut cassava, distinguished from leaf shapes and the color of the leaf petioles. Young leaves were identified from the plant's top leaves, while matured leaves were defined from the plant's bottom leaves. Two-way ANOVA was conducted to determine the interaction between the maturity and variety of cassava leaves for the cyanide and protein concentrations with Tukey's multiple ranges to observe the significant difference at  $p < 0.05$ . The findings indicated significant differences in the HCN content of cassava leaves between different maturities, while other varieties significantly affected protein concentration. The maturity and variety of cassava leaves showed significant interactions with the HCN content. The young Pulut variety had the highest protein concentration and low

HCN content. Thus, it is the best option as an animal feed by reducing its HCN content and maintaining its total phenolic (TPC) and total flavonoid content (TFC). The results imply that variety, as well as maturity, have significant effects on the protein and cyanide concentration of cassava leaves.

**Keywords:** Cassava leaves, cyanide concentration, phytochemical contents, protein concentration

### ARTICLE INFO

#### Article history:

Received: 06 December 2022

Accepted: 04 April 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.04>

#### E-mail addresses:

sarahidris85@gmail.com (Sarah Idris)

rosnahs@upm.edu.my (Rosnah Shamsudin)

zuhair@upm.edu.my (Mohd Zuhair Mohd Nor)

noriznan@upm.edu.my (Mohd Noriznan Mokhtar)

ssalwaag@upm.edu.my (Siti Salwa Abd Gani)

\* Corresponding author

## INTRODUCTION

*Manihot esculenta*, Crantz (cassava), also known as yucca, manioc, tapioca and commonly known as “ubi kayu” in Malaysia, belongs to a woody shrub in the Euphorbiaceae family (Jamil & Bujang, 2016; Latif & Müller, 2015). Cassava roots have a high carbohydrate. Therefore, it is utilized to produce chips and starch. The aerial components (leaves and stems) are usually thrown as waste or used as animal feed after the roots have been harvested (Chaiareekitwat et al., 2022; Idris et al., 2021). Since cassava plants can withstand drought and grow in acidic, low-fertility soil, their leaves are widely available throughout the year (Ubwa et al., 2015). Cassava cultivation is widespread in many tropical countries, with an estimated annual production of 291.9 million tonnes (Lansche et al., 2020). In Asia, including Malaysia, Indonesia, and sub-Saharan Africa, tender leaves are edible as a traditional vegetable commonly served with rice (Latif & Müller, 2015).

It is crucial to evaluate the protein content of cassava leaves because it has been known that green vegetables are the lowest-priced source of protein. Since cassava leaf has three times more crude protein than various cassava plants' components, it can be used as a ruminant protein source (Idris et al., 2021). Another study reported that the crude protein content of cassava leaf was 17.7%–38.1% D.M. (Awoyinka et al., 1995). Crude protein levels in Malaysian cassava cultivars ranged from 21.51 to 30.31% D.M. (Jamil & Bujang, 2016). Since the root of the cassava contains low protein, the leaves should receive better attention as human food and animal feed.

Cyanogen glucosides are present throughout cassava, and the leaf is six to 20 times higher than in the root (Ekpo & Baridia, 2020; Jamil & Bujang, 2016; Eleazu & Eleazu, 2012). It determines the bitterness of the root and other parts of the plant, and it is vital as a defense mechanism against insect herbivores (da Silva Santos et al., 2020; Roslim et al., 2016). The variety, age, soil, geographic location, and environmental factors affect how much cyanogen glucosides are synthesized (Nwokoro et al., 2010; Ubwa et al., 2015). Cassava is divided into two types, sweet and bitter, based on the concentration of cyanogen glucosides in the root (CIAT, 1983). The bitter type contains 50 mg HCN/kg and above (wet basis), while the sweet type has less than this amount (Ojiambo et al., 2017; Pendak, 2011). Sweet cassava matures 6–12 months after planting (MAP) and loses quality if not harvested on time, while bitter cassava matures 10–14 MAP, becoming fibrous if not harvested on time (Pendak, 2011). Previous research has shown that HCN is synthesized at the cassava plant shoot apex and transported to the root (Chaiareekitwat et al., 2022; Latif & Müller, 2015). Therefore, the concentration of cyanogen glucosides in cassava leaves is highest at the top and decreases as it descends. Cyanogen glucosides are hydrolyzed by endogenous enzymes when the plant tissue is damaged during harvesting, by herbivore chewing, food preparation, or within the digestive system to sugar and alpha-hydroxynitrile (Ubwa et al., 2015). Alpha-hydroxynitrile has undergone an intramolecular

reaction to release the toxic HCN gas (Müller-Schwarze, 2009; Zitnak, 1973). HCN inhibits cytochrome oxidase, eventually leading to death, as shown in Figure 1. The tissues readily absorb the HCN through the bloodstream because of their tiny size and low charge density (da Silva Santos et al., 2020).

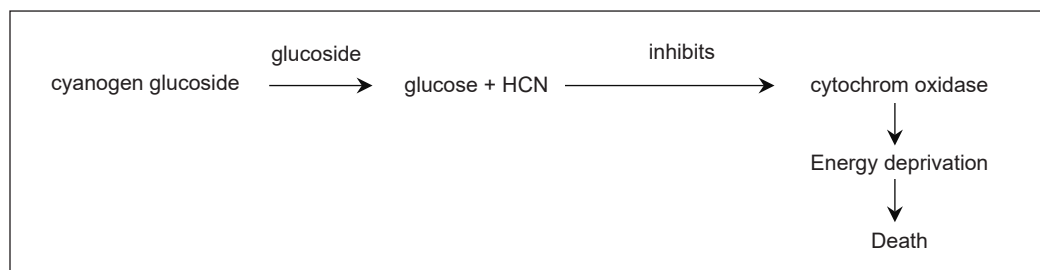


Figure 1. Mechanism of action of cyanogens according to Makkar et al. (2007)

WHO recommends that the maximum lethal dose of HCN in oral foods is 10 mg HCN/kg of body weight (Ojiambo et al., 2017; Ubwa et al., 2015; Umuhozariho et al., 2014). However, ruminants are more vulnerable to HCN poisoning than non-ruminants (Ajayi & Joseph, 2019). It has been demonstrated that an animal's lethal dose of HCN is between 2 and 2.5 mg/kg. Moreover, if the HCN concentration is over 1000 ppm, it may cause death among ruminants (Gensa, 2019).

In the last few decades, there has been much research on polyphenols, phenolics, and flavonoids as alternative bioactive compound-based functional food supplements. Phytochemicals benefit health by acting as preventatives and lowering the chance of non-communicable diseases (NCDs), but they are unnecessary for survival. In animal husbandry, polyphenol supplements are needed in feeds to promote immunity against foreign pathogens by activating signaling pathways (Dragoş et al., 2022).

Furthermore, polyphenols hold benefits such that they can induce epigenetic changes in cells, regulate intestinal immune responses, allergic diseases, and antitumor immunity. In the previous literature, cassava has been reported to have secondary metabolites that may add value to animal feed. Thus, this paper aims to determine the level of HCN concentration, protein concentration, and phytochemicals in two cassava varieties at different maturity stages.

## MATERIALS AND METHODS



### Cassava Leaf Samples

One kg of cassava leaves from two varieties, i.e., White and Pulut, as shown in Table 1, was collected from a local planter in Banting, Malaysia. The samples were collected on the same day to obtain samples with the same ages at six MAP. To describe young and mature leaves,



respectively, leaves were randomly selected from the defined heights from the plants' top (tender shoot apex) and bottom (woody stem). The leaves were harvested, bagged in plastic, and preserved in a polystyrene box with ice during transportation because the protein and cyanide content is thermolabile. Therefore, the samples must be kept between 0 and 4°C. The samples were cleaned, and 12 g of each was randomly taken over the homogenization process. Then, the sample was ground using a kitchen blender. It was packaged in zip-lock bags and kept at -80°C for further analysis (Chaiareekitwat et al., 2022).

Table 1  
Variety of cassava leaf samples

Variety	Shapes	Description
White		<ul style="list-style-type: none"> <li>The color of the leaf is green</li> <li>The color of the petiole is green</li> <li>The shape of the lobes is lanceolate</li> </ul>
Pulut		<ul style="list-style-type: none"> <li>The color of the leaf is green</li> <li>The color of the petiole is reddish-green</li> <li>The shape of lobes is linear</li> </ul>

### HCN Analysis Based on Reduction of Picrate

The HCN analysis based on the reduction of picrate is the ability to reduce sodium picrate to a red compound, as described by Makkar et al. (2007) (Figure 2). A flat-bottomed screw-capped glass bottle was filled with a 20 mg sample and 1 mL of the 0.2 M phosphate buffer (pH 8). A picrate paper attached to a plastic strip was glued on the screw cap of the bottle above the solution and immediately closed. The bottle was shaken slowly to mix the sample with the buffer. The bottle was incubated at room temperature ( $30 \pm 2^\circ\text{C}$ ) for 16 hr. Later, the picrate paper was taken out and submerged for 30 minutes in 5 mL of distilled water. The colorless filter paper was removed, and the picrate solution was boiled for 5 minutes in a water bath. The solution was centrifuged at 4000 rpm for 5 min. The absorbance of the colored solution was measured against the reagent blank (0 ml of sample, 1 ml of buffer, and 5 ml of picrate solution) at 510 nm. HCN content was expressed in ppm using the calibration curve according to the potassium cyanide (KCN) standard.

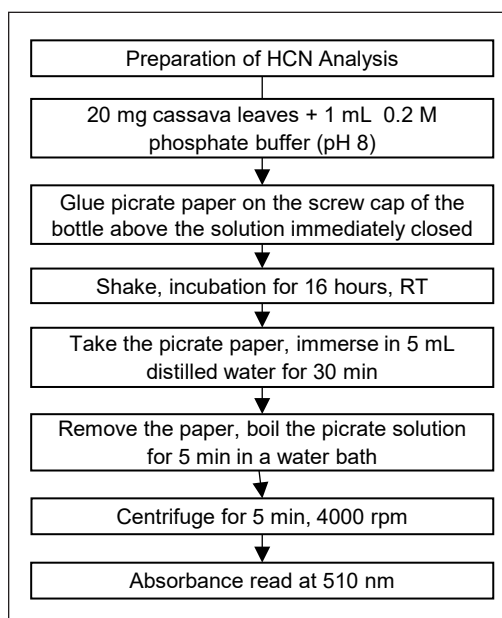


Figure 2. Flowchart representing steps of HCN determination (Makkar et al., 2007)



### Protein Analysis Based on Bradford Method

The Bradford method, which analyses the protein-dye complex of a sample, is based on the absorbance shift seen in an acidic solution of the dye Coomassie® Brilliant Blue G-250. The complexes remain dispersed in the solution for about an hour after the dye has a steady affinity for protein and starts to aggregate after one hour (Bonjoch & Tamayo, 2001). The protein concentration was measured using a Multiskan GO microplate spectrophotometer (Thermo Scientific 1510) at 595 nm. The protein concentration in each cassava leaf was determined by using the BSA (bovine serum albumin) calibration curve as a standard.

The samples were powdered using liquid nitrogen. In 1 mL of extraction buffer containing 0.05 M tris base, 0.1% ascorbic acid, 0.1% cysteine hydrochloride, 1% polyethylene glycol, 0.15% citric acid (monohydrate), and 0.008% 2-mercaptoethanol in distilled water. 0.05 g of the antioxidant polyvinyl polypyrrolidone (PVPP) is added to each sample during homogenization. Homogenates are transferred to 2 ml Eppendorf tubes and centrifuged for 20 minutes at 12000 rpm at 4°C. The supernatant was collected for protein determination using the Bradford assay described in Figure 3.

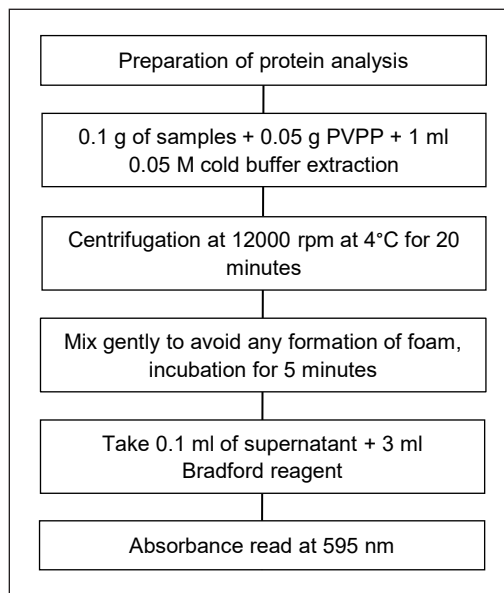


Figure 3. Flowchart representing steps in protein concentration

### Phytochemical Analysis

Total phenolic and flavonoid content extraction was analyzed using procedures described by Hosni et al. (2023) with minor modifications. Then, 0.5 g each of cassava roots and leaves was ground using mortar and pestle. Each sample of cassava leaves was extracted in 10 ml of distilled water in sealed bottles using an ultrasonic bath at 60°C for 2 hours. Then, all extracts were separated from the residues by filtering through Whatman No.1 filter paper and stored at -20°C for further quantification.

The phytochemical analysis of TPC and TFC in cassava leaves was conducted using Folin-Ciocalteu reagent and aluminum chloride assay (Samat et al., 2020), respectively, as shown in Figure 4. The TPC was analyzed with slight modifications. Firstly, 200 µL of samples were added to 5 mL of Folin-Ciocalteu reagent, allowing the mixture to react for 10 min at 25°C. Later, 4 mL of sodium carbonate was added to the solutions kept in the dark for 20 minutes at room temperature. The absorbance was quantified using a Multiskan

GO microplate spectrophotometer (Thermo Scientific 1510) at 750 nm. The TPC value was stated as mg gallic acid equivalents per 100 g of fresh weight.

To determine the TFC, 1 mL of each extract was mixed with 4 mL of distilled water. Next, 0.3 mL of a sodium nitrite ( $\text{NaNO}_2$ ) solution (1:5, w/v) was added to the flasks and left at room temperature for 6 min. Afterward, 0.3 mL of a 1:10 solution of aluminum chloride ( $\text{AlCl}_3$ ) was added, and the mixture was left at room temperature for 6 min. To the resulting solution, 2 mL of a 1 M sodium hydroxide solution was added, and the mixture was incubated for 10 minutes at 25°C. The absorbance of the mixture was measured at 510 nm using a Multiskan GO microplate spectrophotometer (Thermo Scientific 1510). The TFC was calculated using the calibration curve with the quercetin standard and expressed as mg QE/100 g fresh weight.

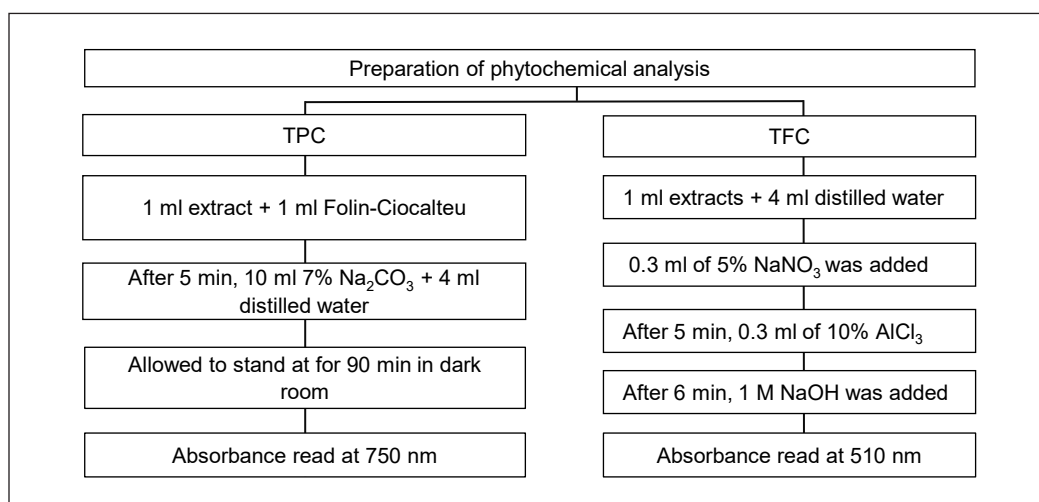


Figure 4. Flowchart representing steps in TPC and TFC determination

### Statistical Analysis

Two independent variables influenced by dependent variables were analyzed using Minitab software (Version 19, State College, Pennsylvania, USA). A two-way analysis of variance (ANOVA) at 95% ( $p < 0.05$ ) confidence interval was carried out to test differences between the means of more than two groups. Tukey’s multiple comparison tests identified the significant difference between the means and the final value as the average  $\pm$  standard deviation. The interaction between the one independent variable and the effect on the other independent variable was examined.

## RESULTS AND DISCUSSION

Table 2 demonstrates the amount of HCN, protein concentration, and phytochemical contents in cassava leaf samples according to the KCN, BSA, gallic acid, and catechin

standard curves, respectively. The cassava leaves of different maturity and variety had different HCN and protein values. The findings indicated that the HCN varied significantly ( $p < 0.05$ ) among different maturities from the White variety. The HCN value ranged from 5.72 to 8.67 mg/100 g. This finding is consistent with that of Jamil and Bujang (2016), who found that the cyanide content in cassava leaves in Malaysia was 8.867–14.72 mg/100 g. Cassava leaves from the White variety demonstrated the highest HCN, followed by the young Pulut, mature Pulut, and young White. The significant differences could be due to the different leaf positions representing leaf maturity. The youngest leaves are collected from the shoot apex (top position), and the matured leaves are located at the bottom, resulting in different amounts of HCN content. According to Ravindran (1992), the cyanide concentration in cassava leaves was highest at the top position and declined towards the lower position. Generally, young cassava leaves contain the highest level of HCN and will decrease to 50–70% in mature leaves (Nambisan & Sundaresan, 1994).

Furthermore, the results showed that HCN content in all cassava leaves exceeded the lethal dose intake of 1 mg HCN equivalent/100 g of the sample (Jamil & Bujang, 2016; Latif & Müller, 2015). Also, ruminants (based only on a studies on goats) may tolerate an intake of 0.25 ppm of HCN (Gensa, 2019). Therefore, the leaves should be detoxified before consumption because they are highly poisonous. The presence of HCN in all study leaves confirmed earlier reports that all cassava cultivars contain a cyanogenic glucoside in wide disparities according to varieties (CIAT, 1983). Hydrolysis of cyanogen glycosides at a temperature above 25.6°C will release cyanide gas, which may cause dizziness, headache, fatigue, and sometimes death (Latif & Müller, 2015).

The final value was expressed as mean  $\pm$  standard deviation;  $n=3$ . Means with the same letters (a, b) in the same column are not significantly different at  $p < 0.05$  according to the Tukey multiple range test.

Furthermore, the results showed that HCN content in all cassava leaves exceeded the lethal dose intake of 1 mg HCN equivalent/100 g of the sample (Jamil & Bujang, 2016; Latif & Müller, 2015). Additionally, ruminants (based only on studies on goats) may tolerate an intake of 0.25 ppm of HCN daily (Gensa, 2019). Therefore, the leaves should be detoxified before consumption because they are highly poisonous. The presence of HCN in all study

Table 2

*Cyanide, protein, phytochemicals concentrations of different varieties, and maturity of cassava leaves*

Samples	Maturity	HCN (mg/100 g)	Protein (mg/100 g)	Total Phenolic compound (mg GAE/100 g)	Total Flavonoid compound (mg CE/100 g)
White	young	8.67 $\pm$ 0.25 <sup>a</sup>	1.91 $\pm$ 0.69 <sup>b</sup>	242.9 $\pm$ 32.7 <sup>a</sup>	64.12 $\pm$ 0.65 <sup>b</sup>
	mature	5.72 $\pm$ 0.47 <sup>b</sup>	129.46 $\pm$ 0.2 <sup>a</sup>	305.2 $\pm$ 63.2 <sup>a</sup>	76.81 $\pm$ 3.53 <sup>a</sup>
Pulut	young	7.27 $\pm$ 0.47 <sup>a,b</sup>	233.17 $\pm$ 0.46 <sup>a</sup>	302.4 $\pm$ 61.6 <sup>a</sup>	59.73 $\pm$ 2.73 <sup>b</sup>
	mature	7.41 $\pm$ 0.41 <sup>a,b</sup>	232.7 $\pm$ 91.1 <sup>a</sup>	291.7 $\pm$ 51.8 <sup>a</sup>	88.01 $\pm$ 7.73 <sup>a</sup>

leaves confirmed earlier reports that all cassava cultivars contain a cyanogenic glucoside in wide disparities according to varieties (CIAT, 1983).

The safe level of HCN from cassava leaves varies for animals depending on the species and the animal's age. The Food and Agriculture Organization (FAO) recommends a maximum safe level of 50 mg HCN per kilogram of dry matter in the diet of ruminants and 10–20 mg HCN per kilogram of dry weight in the diet of monogastric animals (Gensa, 2019). High levels of HCN in the ruminant diet can adversely affect their health and performance.

Regarding polyphenols, not many differences were recorded for both varieties and maturities. Although no significant variances were found between flavonoid varieties, flavonoids were higher in mature leaves than in young cassava leaves. The current results agreed with the previous study conducted by Nur et al. (2013) in cultivating cassava leaves in Malaysia, where the value of TPC is five times higher than TFC—according to Fritz et al. (2001) and Ghasemzadeh et al. (2014) stated that TPC and TFC increase their production following the maturity of plant's capability to devote the resources. It explained why flavonoids as secondary metabolites and primary metabolites were accumulated in mature leaves rather than young leaves.

The results also revealed that the protein concentration demonstrated a significant difference between young White cassava leaves and others, with values ranging from 129.46 to 233.17 mg/100 g. The highest protein content was exhibited by the young Pulut (233.17 mg/100g), and the lowest was from the young White (1.91 mg/100g) cassava leaves. The protein concentration was significantly lower ( $p < 0.05$ ) for young White cassava leaves compared to the other samples. On a fresh basis, protein content was reported to be 4.0–9.6%, while dry basis content had higher crude protein at 20.6–36.4% (Latif & Müller, 2015). Therefore, to use cassava leaves as a protein source, the Pulut variety is more suitable than the White variety. Two-way ANOVA was conducted to investigate two main effects (variety and maturity) on cassava leaves' HCN and protein content (Tables 3 and 4). The two-way interaction showed that the leaf maturity and interaction (variety  $\times$  maturity) statistically significantly affected the HCN content.

Table 3

*Two-way ANOVA test of the effect of variety and maturity of cassava leaf on HCN concentration*

Source of variation	df	Sum of squares	Mean square	F-value	p-value
Variety	1	0.04	0.04	0.10	0.77
Maturity	1	3.93	3.93	8.96	0.04
Variety $\times$ Maturity	1	4.75	4.74	10.81	0.03
Error	4	1.76	0.44		

*Note.*  $R$  squared = 83.24% (adjusted  $R$  squared = 70.67%)

df – degree of freedom; dependent variable –HCN concentration

Table 4

The two-way ANOVA test of the effect of variety and maturity of cassava leaf on protein concentration

Source	df	Sum of squares	Mean square	F-value	p-value
Variety	1	55941	55941	26.97	0.01
Maturity	1	8073	8073	3.89	0.12
Variety × Maturity	1	8197	8197	3.95	0.12
Error	4	8296	2074		

Note. R squared = 89.70% (adjusted R squared = 81.97%)

df – degree of freedom; dependent variable –protein concentration

## CONCLUSION

This study has shown that cassava leaves are rich in protein. However, it also contained HCN, which is harmful to humans and animals and was above the safe level recommended by WHO/FAO. There is a variation of protein and HCN content found in different varieties and maturity of the cassava leaves. The highest protein content was found in young Pulut leaves; therefore, it was selected as a potential ingredient in ruminant feed. The influence of maturity significantly affects the HCN content of cassava leaves. These results could be preliminary data on nutrient sources related to cassava leaves that can be used as a reference for future research.

## ACKNOWLEDGEMENT

The authors thank Universiti Putra Malaysia for the financial and technical support during this research. Also, they thank the 9<sup>th</sup> International Symposium on Applied Engineering and Sciences 2021 (SAES 2021) for giving them the opportunity and financial support for this publication.

## REFERENCES

- Ajayi, M. O., & Joseph, E. (2019). Characteristics of ensiled cassava leaves and maize stover as dry season feed for ruminants. *Nigerian Journal of Animal Production*, 46(4), 273-281.
- Awoyinka, A. F., Abegunde, V. O., & Adewusi, S. R. A. (1995). Nutrient content of young cassava leaves and assessment of their acceptance as a green vegetable in Nigeria. *Plant Foods for Human Nutrition*, 47(1), 21-28. <https://doi.org/10.1007/BF01088163>
- Bonjoch, N. P., & Tamayo, P. R. (2001). Protein content quantification by bradford method. In *Handbook of Plant Ecophysiology Techniques* (pp. 283-295). Kluwer Academic Publishers. [https://doi.org/10.1007/0-306-48057-3\\_19](https://doi.org/10.1007/0-306-48057-3_19)
- Chaiareekitwat, S., Latif, S., Mahayothee, B., Khuwijitjaru, P., Nagle, M., Amawan, S., & Müller, J. (2022). Protein composition, chlorophyll, carotenoids, and cyanide content of cassava leaves (*Manihot esculenta* Crantz) as influenced by cultivar, plant age, and leaf position. *Food Chemistry*, 372, Article 131173. <https://doi.org/10.1016/j.foodchem.2021.131173>

- CIAT. (1983). *Morphology of the Cassava Plant*. Centro Internacional de Agricultura Tropical. [http://ciat-library.ciat.cgiar.org/articulos\\_ciat/books/Morphology\\_of\\_the\\_casava\\_plant.pdf](http://ciat-library.ciat.cgiar.org/articulos_ciat/books/Morphology_of_the_casava_plant.pdf)
- da Silva Santos, B. R., Requião Silva, E. F., Minho, L. A. C., Brandão, G. C., Pinto dos Santos, A. M., Carvalho dos Santos, W. P., Lopes Silva, M. V., & Lopes dos Santos, W. N. (2020). Evaluation of the nutritional composition in effect of processing cassava leaves (*Manihot esculenta*) using multivariate analysis techniques. *Microchemical Journal*, 152, Article 104271. <https://doi.org/10.1016/j.microc.2019.104271>
- Dragoş, D., Petran, M., Gradinaru, T. C., & Gilca, M. (2022). Phytochemicals and inflammation: Is bitter better? *Plants*, 11(21), Article 2991. <https://doi.org/10.3390/plants11212991>
- Ekpo, U. A., & Baridia, D. F. (2020). Effect of processing on the chemical and anti-nutritional properties of cassava leaves (sweet and bitter varieties). *ARC Journal of Nutrition and Growth*, 6(2), 6-12. <https://doi.org/10.20431/2455-2550.0602002>
- Eleazu, C. O., & Eleazu, K. C. (2012). Determination of the proximate composition, total carotenoid, reducing sugars and residual cyanide levels of flours of 6 new yellow and white cassava (*Manihot esculenta* Crantz) Varieties. *American Journal of Food Technology*, 7(10), 642-649. <https://doi.org/10.3923/ajft.2012.642.649>
- Fritz, R. S., Hochwender, C. G., Lewkiewicz, D. A., Bothwell, S., & Orians, C. M. (2001). Seedling herbivory by slugs in a willow hybrid system: Developmental changes in damage, chemical defense, and plant performance. *Oecologia*, 129(1), 87-97. <https://doi.org/10.1007/s004420100703>
- Gensa, U. (2019). Review on cyanide poisoning in ruminants. *Synthesis*, 9(6), 1-12. <https://doi.org/10.7176/JBAH>
- Ghasemzadeh, A., Nasiri, A., Jaafar, H. Z. E., Baghdadi, A., & Ahmad, I. (2014). Changes in phytochemical synthesis, chalcone synthase activity and pharmaceutical qualities of sabah snake grass (*Clinacanthus nutans* L.) in relation to plant age. *Molecules*, 19(11), 17632-17648. <https://doi.org/10.3390/molecules191117632>
- Hosni, S., Gani, S. S. A., Orsat, V., Hassan, M., & Abdullah, S. (2023). Ultrasound-assisted extraction of antioxidants from *Melastoma malabathricum* Linn.: Modeling and optimization using Box–Behnken design. *Molecules*, 28(2), Article 487. <https://doi.org/10.3390/molecules28020487>
- Idris, S. B., Shamsudin, R., Nor, M. Z. M., Mokhtar, M. N., & Ghani, S. S. A. (2021). Proximate composition of different parts of white cassava (*Manihot esculenta* Crantz) plant as a ruminant feed. *Advances in Agricultural and Food Research Journal*, 2(1), 1-9. <https://doi.org/10.36877/aafjrj.a0000181>
- Jamil, S. S., & Bujang, A. (2016). Nutrient and antinutrient composition of different variety of cassava (*Manihot esculenta* Crantz) leaves. *Jurnal Teknologi*, 78(6-6), 59-63. <https://doi.org/10.11113/jt.v78.9024>
- Lansche, J., Awiszus, S., Latif, S., & Müller, J. (2020). Potential of biogas production from processing residues to reduce environmental impacts from cassava starch and crisp production - A case study from Malaysia. *Applied Sciences*, 10(8), 2975. <https://doi.org/10.3390/APP10082975>
- Latif, S., & Müller, J. (2015). Potential of cassava leaves in human nutrition: A review. *Trends in Food Science & Technology*, 44(2), 147-158. <https://doi.org/10.1016/j.tifs.2015.04.006>
- Makkar, H. P., Siddhuraju, P., & Becker, K. (2007). *Plant secondary metabolites* (Vol. 393, pp. 1-122). Humana Press.

- Müller-Schwarze, D. (2009). *Hands-on chemical ecology: Simple field and laboratory exercises*. Springer Science & Business Media.
- Nambisan, B., & Sundaresan, S. (1994). Distribution of linamarin and its metabolising enzymes in cassava tissues. *Journal of the Science of Food and Agriculture*, 66(4), 503-507.
- Nwokoro, O., Ogbonna, J. C., & Okpala, G. N. (2010). Simple picrate method for the determination of cyanide in cassava flour. *Bio-Research*, 7(2), 502-504. <https://doi.org/10.4314/br.v7i2.56582>
- Nur, F. O., Siti, A. H., & Umi, K. Y. (2013). Comparative evaluation of organic and inorganic fertilizers on total phenolic, total flavonoid, antioxidant activity and cyanogenic glycosides in cassava (*Manihot esculenta*). *African Journal of Biotechnology*, 12(18), 2414-2421.
- Ojiambo, O. C., Nawiri, M. P., & Masika, E. (2017). Reduction of cyanide levels in sweet cassava leaves grown in Busia County, Kenya based on different processing methods. *Food Research*, 1(3), 97-102.
- Pendak, Y. (2011). *Baloi: Sweet-type Tapioca Variety Suitable for Planting on Acid Sulphate Soil*. Jabatan Pertanian Sarawak. <http://myagric.upm.edu.my/id/eprint/12062>
- Ravindran, V. (1992). Preparation of cassava leaf products and their use as animal feeds. *FAO Animal Production and Health Paper*, 95, 111-125.
- Roslim, D. I., Herman, Sofyanti, N., Chaniago, M., Restiani, R., & Novita, L. (2016). Characteristics of 22 CASSAVA (*Manihot esculenta* Crantz) genotypes from Riau Province, Indonesia. *Sabrao Journal of Breeding and Genetics*, 48(2), 110-119.
- Samat, N. M. A. A., Ahmad, S., Awang, Y., Bakar, R. A. H., & Hakiman, M. (2020). Alterations in herbage yield, antioxidant activities, phytochemical contents, and bioactive compounds of Sabah snake grass (*Clinacanthus Nutans* L.) with regards to harvesting age and harvesting frequency. *Molecules*, 25(12), 1-16. <https://doi.org/10.3390/molecules25122833>
- Ubwa, S. T., Otache, M. A., Igbum, G. O., & Shambe, T. (2015). Determination of cyanide content in three sweet cassava cultivars in three local government areas of Benue State, Nigeria. *Food and Nutrition Sciences*, 06(12), 1078-1085. <https://doi.org/10.4236/fns.2015.612112>
- Umuhozariho, M. G., Shayo N, B., Msuya, J. M., & Sallah P., Y. K. (2014). Cyanide and selected nutrients content of different preparations of leaves from three cassava species. *African Journal of Food Science*, 8(3), 122-129. <https://doi.org/10.5897/ajfs2013.1100>
- Zitnak, A. (1973). Chronic cassava toxicity. In B. Nestel & R. MacIntyre (Eds.), *Proceedings of an interdisciplinary workshop* (pp. 89-96). International Development Research Centre.





## Attention-based Spatialized Word Embedding Bi-LSTM Model for Sentiment Analysis

Kun Zhu and Nur Hana Samsudin\*

*School of Computer Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*

### ABSTRACT

Movie reviews provide a medium of communication for the movie fans community. Movie reviews not only help viewers and potential viewers to obtain a general opinion about a movie but also allow the fans to construct an opinion of the movie. In this work, an analysis of over 60,000 movie reviews has been implemented to find meaningful text representation via text embedding. We improved the text embedding by proposing an attention-based Bidirectional Long-Short Term Memory (Bi-LSTM) network by using over 60,000 movie review text data as the training set and over 20,000 movie review text data as the testing set. Based on the data features, we performed a probabilistic analysis of the information related to words and phrases, combined the analysis results with text embedding, spatialized the text embedding, and compared the performance of the proposed attention-based spatialized word embedding Bi-LSTM model with several traditional machine learning models. The attention-based spatialized word embedding Bi-LSTM model proposed in this paper achieves an F1 score of 0.91 on the movie review sentiment classification dataset, with a prediction accuracy of 91%, outperforming the results of the current state-of-the-art research. The model can effectively identify the sentimental tendencies of movie reviews and use the analyzed sentimental tendencies to guide consumers in their consumption and obtain feedback on movie content.

*Keywords:* Attention-based deep neural network, data mining, deep learning, natural language processing, sentiment analysis

### ARTICLE INFO

#### *Article history:*

Received: 11 December 2022

Accepted: 14 June 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.05>

#### *E-mail addresses:*

[zhu.kun@student.usm.my](mailto:zhu.kun@student.usm.my) (Kun Zhu)

[nurhana.samsudin@usm.my](mailto:nurhana.samsudin@usm.my) (Nur Hana Samsudin)

\* Corresponding author

### INTRODUCTION

Nowadays, people are no longer limited to asking friends and family for recommendations when they want to watch a movie or buy an item, as there are many user reviews and discussions about the product online. Movie reviews are an important

piece of data on the social services platform, directly reflecting the audience's evaluation of the movie's plot, scenes, director, actors, shooting techniques, script, setting, colors, and sound effects (Zhen et al., 2018). Sentiment analysis of movie reviews in different periods has different reference and application values. Through analyzing the movie reviews, audiences who want to watch the movie can quickly obtain public opinion about a movie. From the review, film investors may also adjust their marketing strategies in time to maximize their investment return. Therefore, film review analysis has become one of a popular tool of sentiment analysis research in recent years (Wang et al., 2020).

Sentiment analysis is an important research direction in natural language processing (NLP) (Abdullah & Rusli, 2021; Fernandes & Mannepalli, 2021a). Manual analysis is unrealistic because of the massive amount of movie review data. It is necessary to use technical tools such as machine learning to perform sentiment analysis of movie reviews (Nayak et al., 2018). As one of the current popular recurrent neural networks (RNN), LSTM can mitigate the gradient disappearance and gradient explosion problems of RNN through the gating mechanism (Cheng et al., 2020; Munshi et al., 2022), so it is often used for the task of text analysis. Sentiment analysis models of LSTM tend to better capture long-term memory dependencies in textual data sequences. Meanwhile, adding an attention mechanism to LSTM networks can reduce the dimensionality of the data, reducing the computational effort of the data and giving the model the ability to capture more important inter-word information (Fernandes & Mannepalli, 2021b; Lai et al., 2015).

This paper proposes an attention-based approach to spatialize word embedding in a bidirectional LSTM model based on the above overview. Spatialization sees words as points in space and word-to-word connections as a vector. Commonly used word combinations are close, and less common word combinations are farther apart. The same vector distance is the same part of the sentence with the same combination of words. It is used to increase the efficiency and accuracy of the model. The bidirectional attentional structure can effectively capture the connections between sentences in context. The attention mechanism is used to capture the relationships between sentences. It allows sentiment analysis tasks to be performed without losing information. Based on Gu et al. (2021), the relevant data are calculated to build a matrix and combined with word embeddings to form spatially structured word embeddings. It allows for a richer input characterization and, to a certain extent, improves information extraction and avoids information loss. Then, add spatialized word embeddings in both directions of the Bi-LSTM and use the Bi-LSTM to encode and decode the text information so that sentence features and information at different locations can be concatenated. In the final part of the model, important movie review text features are re-extracted and weighted using the attention mechanism, thus reducing the difficulty of processing the data and improving the accuracy of the output results. The final result is to improve the performance of the text sentiment classification tasks.

This paper presents an experimental study of attention-based spatialized word embedding using a bidirectional LSTM model for sentiment analysis to understand the applicability and advantages of the method over baseline machine learning methods by evaluating accuracy, precision, recall, and F1 score.

## RELATED WORK

Opinions mined from user-generated content are invaluable assets (Liu, 2012). Since Nasukawa and Yi (2003) introduced the concept of sentiment analysis at the beginning of the 21<sup>st</sup> century, it has become of increasing interest (Liu, 2012). In this context, social text analytics are considered the original application of big data analytics today, as they can derive underlying opinions and sentiments from the vast amount of user-generated data shared online. Traditional sentiment analysis methods are divided into lexicon-based and machine-learning-based approaches. Machine learning methods for sentiment analysis are considered superior to lexicon-based methods because they are unsupervised techniques that rely on external resources to associate polarity scores with each feature used for opinion mining (Behera et al., 2021). Several studies on text classification as well as clustering proposed various methods that aim to enhance the performance of text classification and clustering using ensemble learning methods, clustering, and optimization algorithms (Onan, 2018a; Onan, 2018b; Onan, 2019a; Onan, Korukoğlu et al., 2017). Onan, Bulut et al. (2017), and Onan and Korukoğlu (2017) also evaluated the proposed methods on benchmark datasets and compared their performance with traditional methods. The experimental results showed that the proposed methods outperformed traditional methods in most cases. However, some proposed methods may require more computing resources or manual tuning. Overall, the studies suggest that using machine learning methods such as integrated learning, clustering, and optimization algorithms can improve the performance of text classification and clustering.

Deep learning-based approaches to emotional text analysis are considered superior to traditional sentiment analysis methods, as deep learning can extract features with very high accuracy (Li et al., 2019; Lim et al., 2021). In previous research on sentiment analysis, the most commonly used deep learning models are RNN and LSTM (Hochreiter & Schmidhuber, 1997). Word embeddings have been widely used for text-based representation and understanding since the late 1990s, while LSTM only started to flourish in 2013 when data science was developing (Jain et al., 2021). Onan (2019b) discusses the challenge of detecting irony in sentiment analysis and proposes a deep learning approach that combines word embedding techniques and feature sets. The LDA2vec algorithm outperforms other word embedding techniques, and combining word embedding-based features with traditional feature sets yields good results. In another paper, Onan (2019c) presents a method for topic extraction in scientific literature using word embedding models. The method

improves the performance of clustering algorithms, but the study may be limited to certain subject areas. In 2020, Onan proposed a text mining approach for sentiment analysis on instructor evaluation reviews, which achieved a high classification accuracy using deep learning methods with the GloVe word embedding scheme. However, the limitations of this approach are not discussed (Onan, 2020).

Despite their inconsistent emergence, LSTM and word embeddings are now well used in text-based analysis, such as sentiment analysis and opinion mining (AlKhawter & Al-Twairsh, 2021). In 2012, Socher et al. began an approach to using semantic data to improve the performance of sentiment analysis. However, traditional approaches do not take into account the semantic associations between sentences or document content. In 2015, Chen proposed the Text-CNN model, which achieved good results in sentiment analysis, and was one of the first researchers to apply CNN to sentiment analysis. Later, researchers found that deep learning-related models worked well for emotional text sequence problems (Jianqiang et al., 2018) and began to apply deep learning on a gradually larger scale in NLP (Kumar et al., 2021). In 2019, Rani and Kumar proposed a deep learning approach based on CNN.

This CNN-based deep learning model can effectively perform sentiment analysis on Hindi movie reviews. The CNN-based deep learning approach uses a hand-labeled dataset by three native Hindi experts, and the model achieves classification accuracy of up to 95%. In the same year, Shrivastava et al. (2019) proposed a method combining CNN and Attention to perform sentiment analysis on textual datasets, yielding good results. The research shows that CNN-based deep learning sentiment analysis models have the potential to perform better than machine learning solutions. On novel symbolic and picture emojis, Solanki et al. (2019) propose an opinion-mining method based on sentiment markers to identify positive, negative, and neutral opinions on sentiment topics.

Although this is a traditional method, it can be useful for the sentiment analysis of new emoticons. In 2021, Sarzynska et al. proposed a new ELMo model based on the LSTM structure, and the results were better than the original LSTM model. Onan (2021a) proposed methods for sentiment analysis on MOOC comments using machine learning, ensemble learning, and deep learning techniques. Their research showed that ensemble learning methods outperformed supervised learning methods and that deep learning architectures, particularly the LSTM network, achieved the highest predictive performance. Onan (2021b) also proposed a deep learning architecture for sentiment analysis that combined TF-IDF weighted glove word embedding with a CNN-LSTM architecture. Onan and Tocoglu (2021) developed a deep learning framework for identifying sarcasm in social media using term weighting and LSTM network structure, which outperformed traditional deep neural networks. These studies demonstrate the effectiveness of deep learning techniques and the importance of selecting appropriate text representation schemes and classification algorithms for sentiment analysis in different contexts.

However, further research is needed to explore the potential of other methods and address the limitations of current approaches, such as model ambiguity and fuzzy interference. Tan et al. (2022) built on this method by applying sliding windows to low-resource languages with very good results. Onan (2022) proposed a deep-learning framework for detecting sarcasm in text. The framework includes a three-layer bidirectional LSTM network with a weighted trigram model and a weighted word embedding model. The proposed approach outperformed traditional deep neural network architectures in prediction performance and word embedding. However, efficiency issues when processing large-scale data may be a limitation. Some of the work related to sentiment text analysis is summarized in Table 1.

Our work compares and improves the above research results and proposes an attention-based approach to spatializing word embedding in a Bi-LSTM model. The spatialized input provides more effective information to the model, such as the distance of the same word vectors in a sentence, but increases the model's running time and computational space. Based on the model's accuracy, we use the attention mechanism to assign weights to the individual features within the model, thus reducing the computational burden of processing high-dimensional input data, reducing the dimensionality of the data, and producing high-quality results. The addition of the attention mechanism increases the complexity of the model but reduces the computational effort and increases the accuracy of the model, so we believe that the complexity of the model is worth the sacrifice. This paper focuses on improving the accuracy of deep learning for sentiment analysis of movie reviews using Bi-LSTM models with spatialized inputs and attention mechanisms.

## METHODOLOGY

### Data Description and Pre-processing

The primary valid data used in this model was obtained from IMDB. The IMDB dataset was originally used by Pang et al. (2002), and then IMDB was widely used and extended as a benchmark sentiment dataset. This dataset was extracted from a large dataset of movie reviews used by the Computer Science Department of Stanford University. The IMDB dataset is a collection of 50,000 sentiment texts labeled with positive and negative polarity reviews. All these English movie review texts are balanced and contain only two sentiment categories (i.e., positive and negative sentiment). The dataset consists of a training set, a test set, and a set of unlabeled data. The total training set has 41758 data, of which 20,652 are positive and 21,106 are negative. The test set is 25,000, with 12,500 positive and 12,500 negative data. Based on the IMDB data, Easy Data Augmentation (EDA) was performed on the data related to the IMDB training set using a cloud server in this paper (Wei & Zou, 2019). We back-translated the sentiment text data from the training set while the test set was left unchanged, thus ensuring the fairness of the evaluation. The use of EDA techniques can effectively increase the amount of data for training and improve the

Table 1  
The list of some related tasks for sentiment text analysis

Authors	Method	Task	Precision	Recall	F1	Data set
Nasukawa and Yi (2003)	Sentiment keyword extraction	Sentiment Analysis	0.943	0.286	-	Benchmark Corpus
Shen and Huang (2016)	Attention-Based CNN	Semantic Relation Extraction	-	-	0.859	SemEval-2010 Task 8
Mondal et al. (2018)	Probabilistic model	Rumor identification	0.623	0.708	0.6672	Twitter
Dai et al. (2018)	LSTM	Relation Classification	-	-	0.857	Wikidata knowledge graph
Chen et al. (2018)	RNNs Autoencoders	Rumor identification	0.9249	0.8799	0.8916	Sina Weibo
Lee et al. (2019)	Bi-LSTM	Semantic Relation Classification	-	-	0.852	SemEval-2010 Task 8
Rani and Kumar (2019)	CNN	Sentiment Analysis	0.935	0.934	0.934	Hindi Movie reviews
Liu et al. (2019)	LSTM	Rumor identification	0.9475	0.9485	0.9475	Sina Weibo
Li et al. (2020)	SAMF-Bi-LSTM	Sentiment Classification	-	-	0.638	YELP3
Demotte et al. (2020)	Sentence-State LSTM	Sentiment Analysis	0.8705	0.888	0.8791	Sinhala News articles
Ranathunga and Liyanage (2021)	LSTM fastText	Sentiment Analysis	0.67	0.66	0.67	Sinhala News articles
Muhammad et al. (2021)	LSTM	Sentiment Analysis	0.825	-	-	Indonesian Hotel Reviews
Kumar et al. (2021)	Bi-LSTM-CRF	Sentiment Analysis	0.781	0.821	0.8	Reviews benchmark datasets
Rasool et al. (2021)	LSTM-CNN	Sentiment Analysis	0.87	0.88	0.881	Twitter
Asgar et al. (2021)	Bi-LSTM-CNN	Rumor identification	0.86	0.86	0.86	Twitter
Islam et al. (2021)	LSTM	Fake News Detection	0.96	0.96	0.96	COVID dataset

Note. A dash (-) represents that the item is not presented in that instance

generalization ability of the model, increase the noise data to improve the robustness of the model, and, to some extent, address the problem of insufficient data. The following multiple operations on the review texts are required to facilitate the construction of a lexicon based on a corpus of movie reviews:

- (1) Noise data removal: Remove useless symbols such as “,” “@”, “#”, “/”, “</br>” from the text, which are useless for building sentiment dictionaries.
- (2) Unification of text: Unified uppercase and lowercase, unified text length. Uniform capitalization will help feature extraction during word embedding. The unified case is because the length of text characters needs to be the same during the LSTM model training process. In this case, for texts less than the maximum characters, this paper uses ‘PAD’ characters for padding to ensure that the maximum length of the text is consistent.
- (3) Tokenization: The text is broken down into tokens, symbols, words, or unique elements with clear meaning to reduce the burden on the model.

### Word Embedding and Spatialization

This step first performs tokenization on the paragraphs that need to be analyzed. Usually, there are two ways of word segmentation: First, sentences can be dissected into individual words and expressions, allowing for a more granular analysis of language. Second, an alternative approach involves treating the entire sentence as a single entity, which can be particularly useful in certain contexts. In this step, we combine the above two methods, use regular expressions to segment words and retain certain special emotional symbols. Moreover, we use the N-gram algorithm to optimize word segmentation. The N-Gram formula is as stated in Equation 1:

$$p(Q) = p(w_1 w_2 w_3 w_4 \cdots w_n) = p(w_1^n) \quad (1)$$

$p(Q)$  is the probability of sentence  $Q$  appearing in the corpus,  $n$  is the length of sentence  $Q$ , and  $w_i (1 \leq i \leq n)$  is a word transformed by word segmentation in sentence  $Q$ . According to the chain rule and conditional probability formula, the probability of  $p(Q)$  is equal to the probability multiplication of each word  $w_i (1 \leq i \leq n)$  in  $Q$ , and the formula is as stated in Equation 2:

$$\begin{aligned} p(Q) &= p(w_1^n) \\ &= p(w_1) p(w_2 | w_1) p(w_3 | w_2 w_1) \cdots p(w_n | w_{n-1} \cdots w_2 w_1) \end{aligned} \quad (2)$$

When training an N-gram model, we also need to know the parameter estimation conditional probability of the model. Suppose  $f(w_{i-1})$  is the number of times the word



$w_{i-1}$  appears in the corpus. In that case,  $f(w_{i-1}, w_i)$  is the number of times the two-tuple  $(w_{i-1}, w_i)$  appears in the training corpus, and  $p(w_i|w_{i-1})$  does the model require the estimated conditional probability, then the formula is as stated in Equation 3:

$$p(w_i|w_{i-1}) = \frac{p(w_{i-1}, w_i)}{p(w_i)} = \frac{f(w_{i-1}, w_i)}{f(w_i)} \quad (3)$$

This research uses word embedding because if one-hot encoding is adopted, the dimensionality will be too high, and the amount of calculation will increase. Unlike one-hot encoding, word embedding uses a floating-point dense matrix to represent word segmentation. Based on the size of the dictionary, vectors are used to represent the words' dimensions. A vector represents words and sentences. Before that, we will use numbers to represent word segmentation and vectors to represent numbers: word segmentation  $\rightarrow$  number  $\rightarrow$  vector. The vectors consisting of words, sentences, and the information between them (e.g., word-to-word, sentence-to-sentence, word-to-sentence similarity, and distance) then spatialize the word embeddings to improve the model's accuracy. This paper uses a dictionary to save each word and the corresponding number in the text and implements a method to map the sentence into a list containing numbers through the dictionary. After completing this serialization, we also need to record the number of times and filter the words. For the problem of different sentence lengths, each batch of sentences cannot be constructed with the same length. For those not in the dictionary during training, This article also uses special characters to replace the words.

## Model Structure

In this paper, we first extract the text for sentence feature information, then perform word embedding and combine the results with the extracted sentence information features as the input to the neural network. Features contain information about spatialized features, i.e., the distance between the same words in a sentence. LSTM is mainly designed to extract high-level features from the text, and Bi-LSTM can get a deeper temporal relationship, enhancing the accuracy of high-level feature extraction. After completing the forward LSTM and reverse LSTM, the results are concatenated, and the results are then handed over to the Attention mechanism for processing. This layer of Attention processing gets the global information after the Bi-LSTM processing and performs weighting based on the relevant feature information. The Attention mechanism here will also get the fault information related to the model, which will also greatly improve the accuracy of the prediction. After completing these parts, it will be handed over to the fully connected layer for softmax processing and classification, with the final output being the sentiment classification of the movie text. The model structure is shown in Figure 1.

The process described above is a process of encoding. After the encoding has been completed, attention and decoder work simultaneously. The model needs to calculate the



correlation, with the input Bi-LSTM result  $l_i$  and the un-updated state  $Y_0$ . The first step is to transform linearly using the two-parameter matrices  $W_k$  and  $W_q$  to get the two vectors  $k_i$  and  $q_0$ . The third step is to do a softmax transform on these values. The correlation is expressed as  $a$ ,  $a$  is also the weight for the attention mechanism, and each  $a$  corresponds to a state  $l_i$ . We need a weighted average of the states in  $a$  and get the content vector  $C$ . Each  $C$  has a corresponding decoding state,  $Y$ . The content vector  $C_i$  formula is as stated in Equation 4:

$$C_i = a_1l_1 + a_2l_2 + \dots + a_il_i \tag{4}$$

The new state  $Y_i$  is a concat of the old state  $S_{i-1}$ , the new input function  $X$  with the content vector  $C_i$ , and then multiplied by the weight matrix plus the bias  $b$ . The formula is as stated in Equation 5:

$$Y_i = \tanh\left(S_i \cdot \begin{bmatrix} X \\ Y_{i-1} \\ C_i \end{bmatrix}\right) \tag{5}$$

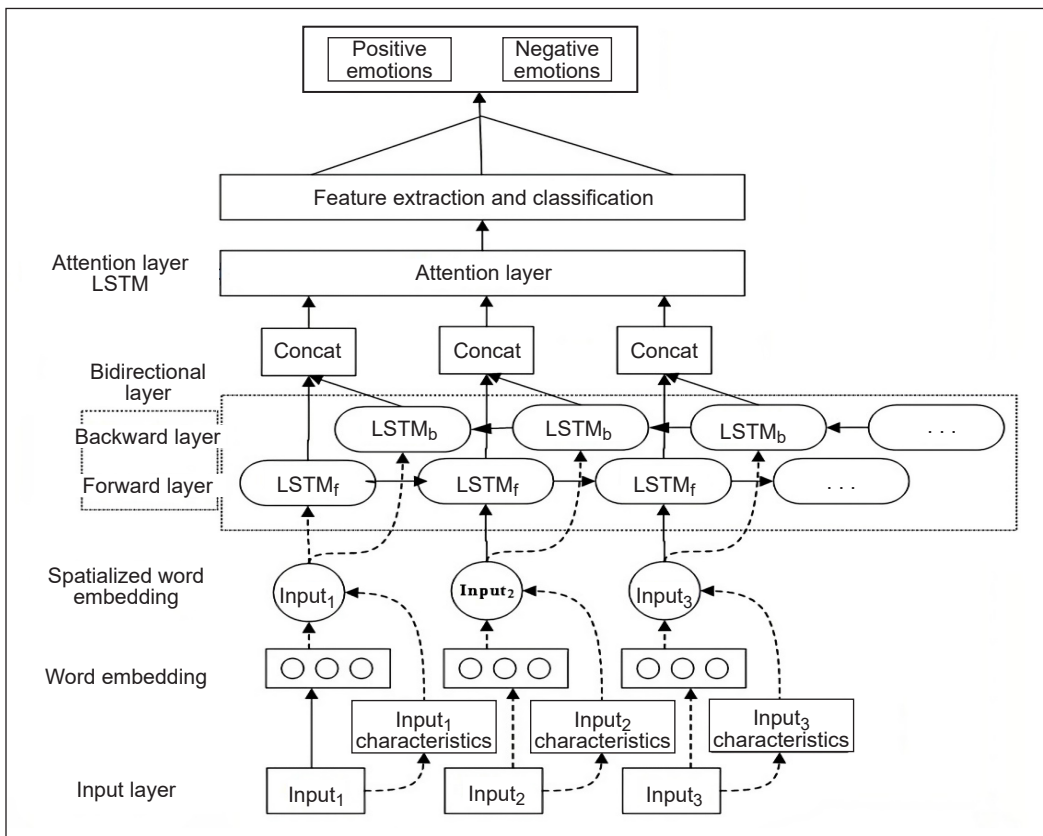


Figure 1. Model structure

## Model Training and Analysis

This paper uses torch.nn from the Pytorch library to build a Bi-LSTM model. When building the model, we need to input the necessary parameters. Table 2 shows the parameter settings for the model in this paper and the comparison model.

Table 2  
*Model parameters*

Parameters	Spatialized word embedding	Spatialized word embedding + Bi-LSTM	Spatialized word embedding + Bi-LSTM + attention
Number of layers	2-Dense	1-Bi-LSTM, 1-Dense	1-Bi-LSTM, 1-Dense
Dimension of hidden state vector	-	200, 128	200, 128
Number of neurons (Dense)	256, 2	2	2
Activation	ReLU, Softmax	Softmax	Softmax
Learning rate	0.001	0.001	0.001
Optimizer	Adam	Adam	Adam
Dropout	0.1	0.05	0.05
Sentence Max	200	200	200
Batch size	8	128	128
Epochs	30	30	30

## RESULTS AND COMPARISON

This section discusses the results of different models of traditional machine learning and deep learning for sentiment analysis attention-based spatialized word embedding LSTM. The attention-based spatialized word embedding Bi-LSTM model was constructed based on the Pytorch deep learning framework, using a Bi-LSTM +Attention deep learning network structure. In the case of deep learning using the attention-based spatialized word embedding Bi-LSTM model, the total sentiment text dataset collected was 91,758, split into a training and test set in a 3:2 ratio. In this section, we perform comparative experiments on our model and provide a detailed comparative analysis of the model's results after spatialized word embedding.

### Results

The performance of the proposed model was evaluated using the evaluation metrics of Accuracy, Precision, Recall, F1-score, and AUC.

As shown in Table 3, Spatialized word embedding methods have significant advantages over regular word embedding methods for sentiment analysis tasks, provided that the dataset remains unchanged. As shown in Table 4, among the models using only word embedding, the models using the spatialized word embedding method had a 10.12% higher prediction

rate and an 11.9% higher F1 score than those using the regular word embedding method. The attention-based Bi-LSTM model with spatialized word embedding had a 5.34% higher prediction rate and a 6.27% higher F1 score than the attention-based Bi-LSTM model with regular word embedding. The model of the spatialized word embedding method is able to capture more word and semantic information before the word is embedded, and the model can encode better based on these features.

It is one of the important reasons why the spatialized word embedding approach models work better. The Attention mechanism helps the model to reduce the computational burden of processing high-dimensional input data by structurally selecting a subset of the inputs, reducing the dimensionality of the data, and allowing the model to focus more on finding useful information in the input data that is significantly relevant to the current output, thus improving the quality of the output. As shown in Table 5, the performance of the Bi-LSTM model with the Attention mechanism was higher than that of the Bi-LSTM model without the Attention mechanism. The F1-score of the model can be improved by 7.02% after using the Attention mechanism.

Precision represents the degree of prediction precision in the sample results. Figure 2 compares the precision and recall of different models for positive and negative categories after spatialized word embedding: (a) the precision of the positive category and (b) the precision of the negative category; (c) the recall of the positive category and (d) the recall of the negative category. The combination of spatialized word embeddings, attention mechanisms, and a Bi-LSTM model has significant results for the sentiment text

Table 3  
*Comparison of spatialized word embedding*

Method	Precision	AUC	Recall	F1
Word embedding	64.63%	70.43%	62.31%	62.56%
Word embedding (spatialized)	74.75%	82.96%	74.29%	74.46%

Table 4  
*Comparison of the attention-based Bi-LSTM model with spatialized word embedding*

Method	Precision	AUC	Recall	F1
Word embedding + Bi-LSTM + attention	85.26%	88.74%	84.15%	84.23%
Word embedding (spatialized) + Bi-LSTM + attention	90.60%	95.69%	90.41%	90.50%

Table 5  
*Model comparison with an attention mechanism*

Method	Precision	AUC	Recall	F1
Word embedding (spatialized) + Bi-LSTM	84.14%	90.78%	83.65%	83.48%
Word embedding (spatialized) + Bi-LSTM +attention	90.60%	95.69%	90.41%	90.50%

classification task (Figure 2). The stability of the model and the recognition of positive and negative categories are significantly improved. It is demonstrated that the combination of spatialized word embeddings and Bi-LSTM can better capture sequence information, and the classification precision is higher with the addition of the Attention mechanism.

Figures 3 and 4 show images of the loss and accuracy processes for the three model runs after spatialization. In Figure 3, the X-axis is the epoch of the model’s run, and the Y-axis is the loss of the model. In Figure 4, the X-axis is also the epoch of the models run, and the Y-axis is the accuracy of the models. The loss of the three models decreases and gradually stabilizes as the number of training steps increases in Figure 3. In Figure 4, it can be seen from the learning curve that the models are learning from the data. As the epoch of training rounds increases, the accuracy of the models increases and stabilizes as it reaches its peak. The comparison of the different models in Figures 3 and 4 shows that the attention-based spatialized word embedding Bi-LSTM model is more stable has a faster training process, and is more accurate than the other models. The comparison of the different models in Figures 3 and 4 shows that the attention-based spatialized word embedding Bi-

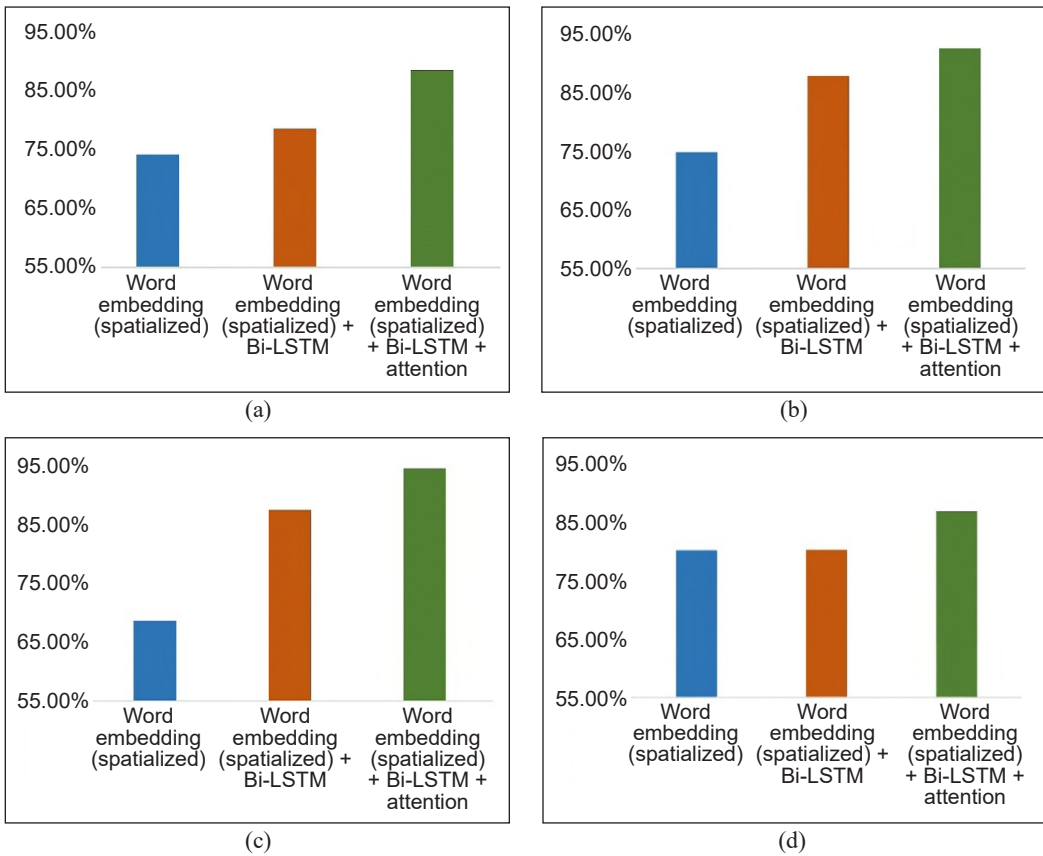
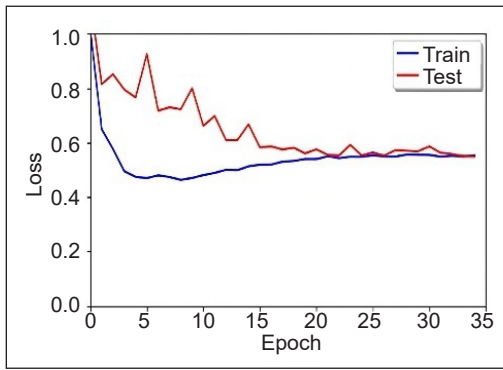
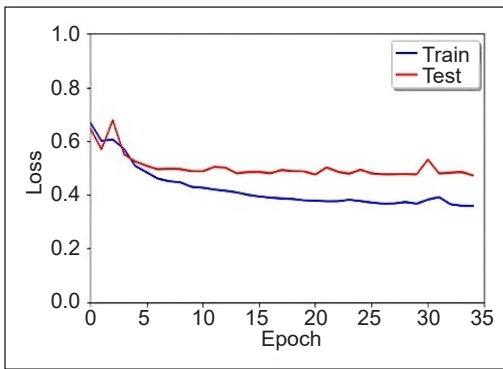


Figure 2. Precision: (a) positive category, (b) negative category; and recall: (c) positive category, (d) negative category

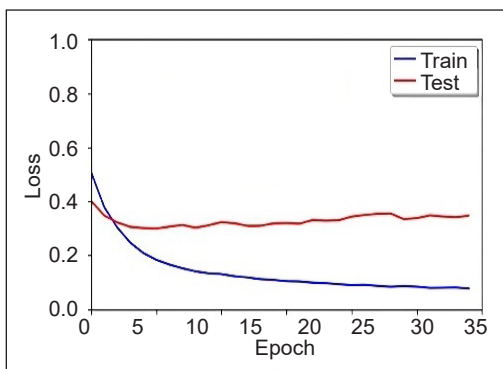
LSTM model is more stable has a faster training process, and is more accurate than the other models. It also proved that there is a gap between the traditional word embedding and deep learning models and that our Bi-LSTM model with the Attention mechanism has better results than the traditional Bi-LSTM model.



(a)

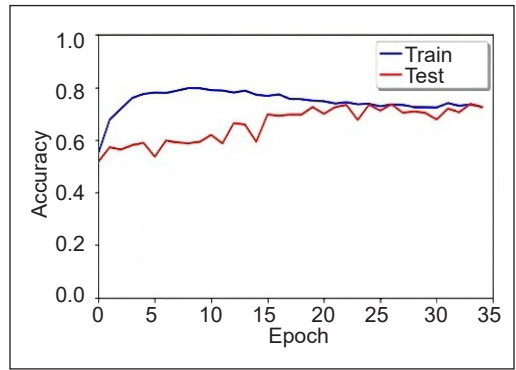


(b)

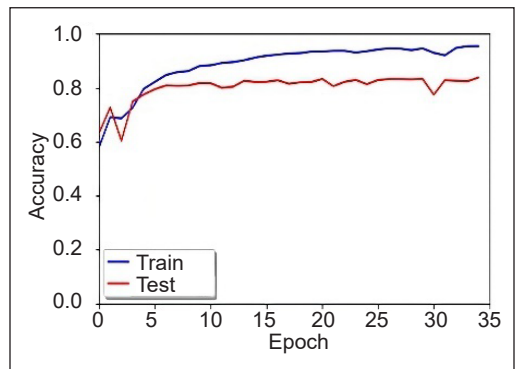


(c)

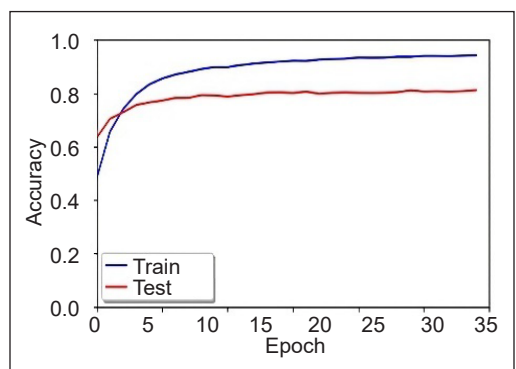
Figure 3. Loss of models: (a) word embedding (spatialized); (b) word embedding (spatialized) + Bi-LSTM; and (c) word embedding (spatialized) + Bi-LSTM + attention



(a)



(b)



(c)

Figure 4. Accuracy of models: (a) word embedding (spatialized); (b) word embedding (spatialized) + Bi-LSTM; and (c) word embedding (spatialized) + Bi-LSTM + attention

## Comparison of Related Models

In this subsection, we evaluate the performance of our attention-based spatialized word embedding Bi-LSTM model on the same test set using the work in Table 6. Traditionally, feature extraction methods were the most commonly used in the early period of targeting sentiment analysis tasks. The first three models in Table 6 mainly use traditional feature extraction methods (Yang et al., 2016), all of which have achieved good effectiveness on sentiment analysis tasks, in particular the emotions and word embeddings method proposed by Giatsoglou et al. (2017), which achieved a precision of 87.8%, which is a very good result. The spatialized word embedding method in this paper also achieved a precision of only 74.75% in a stand-alone test, but the combination of traditional feature extraction methods with deep learning in this paper achieved a precision of 90.60%. We might achieve better results if we learn from Giatsoglou et al. (2017) and add the recognition of sentiment word categories to the spatialized word embeddings. It is one of the next steps in our future work.

In the last three years, the methods used for sentiment analysis have mainly been neural networks or deep learning. Chiny et al. (2021) achieved 82.9% precision, 82.7% recall, and 83.5% F1-score using LSTM on the same test set as our paper, but were respectively 7.7%, 7.7%, and 7% lower than the model proposed in our paper. We believe that the main difference between the two models is that the LSTM is unable to obtain more complete features before the input and cannot perform importance analysis on the features after the output, which is one of the areas where we differ from most models. Based on previous work, Leng et al. (2021) and Gupta et al. (2021) added an Attention mechanism to the LSTM, which improved the precision of the model by about 5% compared to the ordinary LSTM model. A unidirectional LSTM has the limitation that it only retains information from the past and not from the future or the next state to understand the context of the sentence better.

In our paper, the Bi-LSTM model captures information from both current and previous inputs and avoids information loss, thus enabling effective prediction by retaining contextual information (current and previous) over a long period. Briskilal and Subalalitha (2022) used the BERT model to solve the sentiment analysis task with very good results, and it improved the precision value of the movie sentiment classification task to over 90%. To a certain extent, the Attention mechanism in this paper is similar to the Attention mechanism of BERT. Anisotropy in the vectors in which BERT encodes sentences can lead to the semantics of even a high-frequency word and a low-frequency word being equivalent. This paper uses a method of spatialized word embeddings that reduces this by altering the vector somewhat before input. Compared to the BERT model, the precision of the proposed model in this paper is improved by 0.2%.

Through comparative experiments, the model proposed in this paper is the best compared to other classical models on the same test set of IMDB. This paper applies the Bi-LSTM

and attention mechanism methods to the automatic codec model. It allows the automatic codec to have bi-directional computational capabilities. In addition, using spatialized word embedding, the text information can be well reconstructed using this traditional method to improve the capture of text features. It is also shown from the comparative tests that the deep learning methods are, to some extent, superior to the traditional methods and that the models with attention mechanisms outperform those without.

Table 6  
*Comparison of related models*

Authors	Method	Precision	Recall	F1
Yang et al. (2016)	Hierarchical attention networks	-	-	49.4%
Giatsoglou et al. (2017)	Emotions and word embeddings	87.8%	-	-
Kumar et al. (2019)	Hybrid feature extraction	78.3%	-	78.0%
Chiny et al. (2021)	LSTM	82.9%	82.7%	83.5%
Leng et al. (2021)	HRNaEMSA	87.5%	87.4%	87.4%
Gupta et al. (2021)	Senti_ALSTM	87.4%	-	-
Briskilal et al. (2022)	BERT	90.4%	-	89.0%
Our proposed model	Word embedding (spatialized) + Bi-LSTM + attention	90.6%	90.4%	90.5%

## CONCLUSION

In this paper, we propose an approach based on attention spatialized word embedding Bi-LSTM model to improve the accuracy of movie review sentiment analysis effectively and enrich the movie review training dataset using Easy Data Augmentation Techniques.

Firstly, Bi-LSTM can effectively capture the connections between sentences in the context. Secondly, spatialized word embeddings provide the model with more usable text data features while retaining the basic textual information features. Finally, the attention mechanism weighs the processed data to perform sentiment analysis tasks faster and better without losing information. We experimented with classifiers from a variety of traditional and deep-learning methods. They were tested on the same IMDB test set, and our results were compared with previous work. The experimental results show that classifiers from the attention-based spatialized word embedding Bi-LSTM model approach outperformed those from other approaches, achieving the best results in terms of precision (90.6%), recall (90.4%), and F1-score (90.5%).

In the future, we will extend the model to more scenarios of sentiment text analysis. In addition, we will also combine the features learned from unstructured data, such as sound and video, using multimodal joint representation to do sentiment text information retrieval and apply sentiment text information retrieval to sentiment text classification tasks, which is one of our important future works.

## ACKNOWLEDGMENTS

This manuscript publication fee was partially funded by the Universiti Sains Malaysia's Short-term Grant No.: 304/PKOMP/6315273.

## REFERENCES

- Abdullah, N. A. S., & Rusli, N. I. A. (2021). Multilingual sentiment analysis: A systematic literature review. *Pertanika Journal of Science and Technology*, 29(1), 445-470. <https://doi.org/10.47836/pjst.29.1.25>
- AlKhawter, W., & Al-Twairish, N. (2021). Part-of-speech tagging for Arabic tweets using CRF and Bi-LSTM. *Computer Speech and Language*, 65, Article 101138. <https://doi.org/10.1016/j.csl.2020.101138>
- Asghar, M. Z., Habib, A., Habib, A., Khan, A., Ali, R., & Khattak, A. (2021). Exploring deep neural networks for rumor detection. *Journal of Ambient Intelligence and Humanized Computing*, 12(4), 4315-4333. <https://doi.org/10.1007/s12652-019-01527-4>
- Behera, R. K., Jena, M., Rath, S. K., & Misra, S. (2021). Co-LSTM: Convolutional LSTM model for sentiment analysis in social big data. *Information Processing and Management*, 58(1), Article 102435. <https://doi.org/10.1016/j.ipm.2020.102435>
- Briskilal, J., & Subalalitha, C. N. (2022). An ensemble model for classifying idioms and literal texts using BERT and RoBERTa. *Information Processing and Management*, 59(1), Article 102756. <https://doi.org/10.1016/j.ipm.2021.102756>
- Chen, W., Zhang, Y., Yeo, C. K., Lau, C. T., & Lee, B. S. (2018). Unsupervised rumor detection based on users' behaviors using neural networks. *Pattern Recognition Letters*, 105, 226-233. <https://doi.org/10.1016/j.patrec.2017.10.014>
- Chen, Y. (2015). *Convolutional Neural Network for Sentence Classification* (Unpublished Master's thesis). University of Waterloo, Canada. <https://uwspace.uwaterloo.ca/handle/10012/9592>
- Cheng, Y., Yao, L., Xiang, G., Zhang, G., Tang, T., & Zhong, L. (2020). Text sentiment orientation analysis based on multi-channel CNN and bidirectional GRU with attention mechanism. *IEEE Access*, 8, 134964-134975. <https://doi.org/10.1109/ACCESS.2020.3005823>
- Chiny, M., Chihab, M., Chihab, Y., & Bencharef, O. (2021). LSTM, VADER and TF-IDF based hybrid sentiment analysis model. *International Journal of Advanced Computer Science and Applications*, 12(7), 265-275. <https://doi.org/10.14569/IJACSA.2021.0120730>
- Dai, Y., Guo, W., Chen, X., & Zhang, Z. (2018). Relation classification via LSTMs based on sequence and tree structure. *IEEE Access*, 6, 64927-64937. <https://doi.org/10.1109/ACCESS.2018.2877934>
- Demotte, P., Senevirathne, L., Karunanayake, B., Munasinghe, U., & Ranathunga, S. (2020). Sentiment analysis of Sinhala news comments using sentence-state LSTM networks. In *MERCCon 2020 - 6th International Multidisciplinary Moratuwa Engineering Research Conference* (pp. 283-288). IEEE Publishing. <https://doi.org/10.1109/MERCCon50084.2020.9185327>
- Fernandes, B., & Mannepalli, K. (2021a). An analysis of emotional speech recognition for tamil language using deep learning gate recurrent unit. *Pertanika Journal of Science and Technology*, 29(3), 1937-1961. <https://doi.org/10.47836/pjst.29.3.37>



- Fernandes, B., & Manneppalli, K. (2021b). Speech emotion recognition using deep learning LSTM for tamil language. *Pertanika Journal of Science and Technology*, 29(3), 1915-1936. <https://doi.org/10.47836/pjst.29.3.33>
- Giatsoglou, M., Vozalis, M. G., Diamantaras, K., Vakali, A., Sarigiannidis, G., & Chatzisavvas, K. C. (2017). Sentiment analysis leveraging emotions and word embeddings. *Expert Systems with Applications*, 69, 214-224. <https://doi.org/10.1016/j.eswa.2016.10.043>
- Gu, W., Tandon, A., Ahn, Y. Y., & Radicchi, F. (2021). Principled approach to the selection of the embedding dimension of networks. *Nature Communications*, 12(1), 1-10. <https://doi.org/10.1038/s41467-021-23795-5>
- Gupta, C., Chawla, G., Rawlley, K., Bisht, K., & Sharma, M. (2021). Senti\_ALSTM: Sentiment analysis of movie reviews using attention-based-LSTM. In *Proceedings of 3rd International Conference on Computing Informatics and Networks: ICCIN 2020* (pp. 211-219). Springer. [https://doi.org/10.1007/978-981-15-9712-1\\_18](https://doi.org/10.1007/978-981-15-9712-1_18)
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735-1780. [https://doi.org/10.1007/978-1-4757-5388-2\\_2](https://doi.org/10.1007/978-1-4757-5388-2_2)
- Islam, M. U., Hossain, M. M., & Kashem, M. A. (2021). COVFake: A word embedding coupled with LSTM approach for COVID related fake news detection. *International Journal of Computer Applications*, 174(10), 1-5. <https://doi.org/10.5120/ijca2021920977>
- Jain, P. K., Saravanan, V., & Pamula, R. (2021). A hybrid CNN-LSTM: A deep learning approach for consumer sentiment analysis using qualitative user-generated contents. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 20(5), Article 84. <https://doi.org/10.1145/3457206>
- Jianqiang, Z., Xiaolin, G., & Xuejun, Z. (2018). Deep convolution neural networks for twitter sentiment analysis. *IEEE Access*, 6, 23253-23260. <https://doi.org/10.1109/ACCESS.2017.2776930>
- Kumar, A., Verma, S., & Sharan, A. (2021). ATE-SPD: Simultaneous extraction of aspect-term and aspect sentiment polarity using Bi-LSTM-CRF neural network. *Journal of Experimental and Theoretical Artificial Intelligence*, 33(3), 487-508. <https://doi.org/10.1080/0952813X.2020.1764632>
- Kumar, K., Harish, B. S., & Darshan, H. K. (2019). Sentiment analysis on IMDb movie reviews using hybrid feature extraction method. *International Journal of Interactive Multimedia and Artificial Intelligence*, 5(5), Article 109. <https://doi.org/10.9781/ijimai.2018.12.005>
- Lai, S., Xu, L., Liu, K., & Zhao, J. (2015). Recurrent convolutional neural networks for text classification. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 29, No. 1). AAAI Press. <https://doi.org/10.1609/aaai.v29i1.9513>
- Lee, J., Seo, S., & Choi, Y. S. (2019). Semantic relation classification via bidirectional LSTM networks with entity-aware attention using latent entity typing. *Symmetry*, 11(6), Article 785. <https://doi.org/10.3390/sym11060785>
- Leng, X. L., Miao, X. A., & Liu, T. (2021). Using recurrent neural network structure with enhanced multi-head self-attention for sentiment analysis. *Multimedia Tools and Applications*, 80(8), 12581-12600. <https://doi.org/10.1007/s11042-020-10336-3>

- Li, W., Liu, P., Zhang, Q., & Liu, W. (2019). An improved approach for text sentiment classification based on a deep neural network via a sentiment attention mechanism. *Future Internet*, 11(4), Article 96. <https://doi.org/10.3390/FI11040096>
- Li, W., Qi, F., Tang, M., & Yu, Z. (2020). Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*, 387, 63-77. <https://doi.org/10.1016/j.neucom.2020.01.006>
- Lim, C. T., Bong, C. H., Wong, W. S., & Lee, N. K. (2021). A comprehensive review of automated essay scoring (AES) research and development. *Pertanika Journal of Science and Technology*, 29(3), 1875-1899. <https://doi.org/10.47836/pjst.29.3.27>
- Liu, B. (2012). Sentiment analysis and opinion mining. *Synthesis Lectures on Human Language Technologies*, 5(1), 1-167. [https://doi.org/10.1142/9789813100459\\_0007](https://doi.org/10.1142/9789813100459_0007)
- Liu, Y., Jin, X., & Shen, H. (2019). Towards early identification of online rumors based on long short-term memory networks. *Information Processing and Management*, 56(4), 1457-1467. <https://doi.org/10.1016/j.ipm.2018.11.003>
- Mondal, T., Pramanik, P., Bhattacharya, I., Boral, N., & Ghosh, S. (2018). Analysis and early detection of rumors in a post disaster scenario. *Information Systems Frontiers*, 20(5), 961-979. <https://doi.org/10.1007/s10796-018-9837-8>
- Muhammad, P. F., Kusumaningrum, R., & Wibowo, A. (2021). Sentiment analysis using Word2vec and long short-term memory (LSTM) for Indonesian hotel reviews. *Procedia Computer Science*, 179, 728-735. <https://doi.org/10.1016/j.procs.2021.01.061>
- Munshi, A. A., AlSabban, W. H., Farag, A. T., Rakha, O. E., Al Sallab, A., & Alotaibi, M. (2022). Automated Islamic jurisprudential legal opinions generation using artificial intelligence. *Pertanika Journal of Science and Technology*, 30(2), 1135-1156. <https://doi.org/10.47836/pjst.30.2.16>
- Nasukawa, T., & Yi, J. (2003). Sentiment analysis: Capturing favorability using natural language processing. In *Proceedings of the 2nd International Conference on Knowledge Capture* (pp. 70-77). John Wiley & Sons. <https://doi.org/10.1111/j.1469-185X.1956.tb01550.x>
- Nayak, S. K., Rout, P. K., Jagadev, A. K., & Swarnkar, T. (2018). Elitism-based multi-objective differential evolution with extreme learning machine for feature selection: A novel searching technique. *Connection Science*, 30(4), 362-387. <https://doi.org/10.1080/09540091.2018.1487384>
- Onan, A. (2018a). An ensemble scheme based on language function analysis and feature engineering for text genre classification. *Journal of Information Science*, 44(1), 28-47. <https://doi.org/10.1177/0165551516677911>
- Onan, A. (2018b). Biomedical text categorization based on ensemble pruning and optimized topic modelling. *Computational and Mathematical Methods in Medicine*, 2018, Article 2497471. <https://doi.org/10.1155/2018/2497471>
- Onan, A. (2019a). Consensus clustering-based undersampling approach to imbalanced learning. *Scientific Programming*, 2019, Article 5901087. <https://doi.org/10.1155/2019/5901087>

- Onan, A. (2019b). Topic-enriched word embeddings for sarcasm identification. In *Software Engineering Methods in Intelligent Algorithms: Proceedings of 8th Computer Science Online Conference* (pp. 293-304). Springer. [https://doi.org/10.1007/978-3-030-19807-7\\_29](https://doi.org/10.1007/978-3-030-19807-7_29)
- Onan, A. (2019c). Two-stage topic extraction model for bibliometric data analysis based on word embeddings and clustering. *IEEE Access*, 7, 145614-145633. <https://doi.org/10.1109/ACCESS.2019.2945911>
- Onan, A. (2020). Mining opinions from instructor evaluation reviews: A deep learning approach. *Computer Applications in Engineering Education*, 28(1), 117-138. <https://doi.org/10.1002/cae.22179>
- Onan, A. (2021a). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. *Computer Applications in Engineering Education*, 29(3), 572-589. <https://doi.org/10.1002/cae.22253>
- Onan, A. (2021b). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. *Concurrency and Computation: Practice and Experience*, 33(23), Article e5909. <https://doi.org/10.1002/cpe.5909>
- Onan, A. (2022). Bidirectional convolutional recurrent neural network architecture with group-wise enhancement mechanism for text sentiment classification. *Journal of King Saud University - Computer and Information Sciences*, 34(5), 2098-2117. <https://doi.org/10.1016/j.jksuci.2022.02.025>
- Onan, A., Bulut, H., & Korukoğlu, S. (2017). An improved ant algorithm with LDA-based representation for text document clustering. *Journal of Information Science*, 43(2), 275-292. <https://doi.org/10.1177/0165551516638784>
- Onan, A., & Korukoğlu, S. (2017). A feature selection model based on genetic rank aggregation for text sentiment classification. *Journal of Information Science*, 43(1), 25-38. <https://doi.org/10.1177/0165551515613226>
- Onan, A., Korukoğlu, S., & Bulut, H. (2017). A hybrid ensemble pruning approach based on consensus clustering and multi-objective evolutionary algorithm for sentiment classification. *Information Processing and Management*, 53(4), 814-833. <https://doi.org/10.1016/j.ipm.2017.02.008>
- Onan, A., & Tocoglu, M. A. (2021). A term weighted neural language model and stacked bidirectional LSTM based framework for sarcasm identification. *IEEE Access*, 9, 7701-7722. <https://doi.org/10.1109/ACCESS.2021.3049734>
- Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. In *EMNLP '02: Proceedings of the ACL-02 conference on Empirical methods in natural language processing* (pp. 79-86). ACM Publishing. <https://doi.org/10.3115/1118693.1118704>
- Ranathunga, S., & Liyanage, I. U. (2021). Sentiment analysis of Sinhala news comments. *ACM Transactions on Asian and Low-Resource Language Information Processing*, 20(4), Article 59. <https://doi.org/10.1145/3445035>
- Rani, S., & Kumar, P. (2019). Deep learning based sentiment analysis using convolution neural network. *Arabian Journal for Science and Engineering*, 44(4), 3305-3314. <https://doi.org/10.1007/s13369-018-3500-z>
- Rasool, A., Jiang, Q., Qu, Q., & Ji, C. (2021). WRS: A novel word-embedding method for real-time sentiment with integrated LSTM-CNN model. In *2021 IEEE International Conference on Real-Time Computing and Robotics (RCAR)* (pp. 590-595). IEEE Publishing. <https://doi.org/10.1109/RCAR52367.2021.9517671>

- Sarzynska-Wawer, J., Wawer, A., Pawlak, A., Szymanowska, J., Stefaniak, I., Jarkiewicz, M., & Okruszek, L. (2021). Detecting formal thought disorder by deep contextualized word representations. *Psychiatry Research*, 304, Article 114135. <https://doi.org/10.1016/j.psychres.2021.114135>
- Shen, Y., & Huang, X. J. (2016). Attention-based convolutional neural network for semantic relation extraction. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers* (pp. 2526-2536). The COLING 2016 Organizing Committee.
- Shrivastava, K., Kumar, S., & Jain, D. K. (2019). An effective approach for emotion detection in multimedia text data using sequence based convolutional neural network. *Multimedia Tools and Applications*, 78(20), 29607-29639. <https://doi.org/10.1007/s11042-019-07813-9>
- Socher, R., Huval, B., Manning, C. D., & Ng, A. Y. (2012). Semantic compositionality through recursive matrix-vector spaces. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning* (pp. 1201-1211). Association for Computational Linguistics.
- Solanki, V. K., Cuong, N. H. H., & Lu, Z. J. (2019). Opinion mining: using machine learning techniques. In *Extracting Knowledge from Opinion Mining* (pp. 66-82). IGI Global. <https://doi.org/10.4018/978-1-5225-6117-0.ch004>
- Tan, T. P., Lim, C. K., & Rahman, W. R. E. A. (2022). Sliding window and parallel LSTM with attention and CNN for sentence alignment on low-resource languages. *Pertanika Journal of Science and Technology*, 30(1), 97-121. <https://doi.org/10.47836/pjst.30.1.06>
- Wang, Q., Zhu, G., Zhang, S., Li, K., Chen, X., & Xu, H. (2020). Extending emotional lexicon for improving the classification accuracy of Chinese film reviews. *Connection Science*, 33(2), 153-172. <https://doi.org/10.1080/09540091.2020.1782839>
- Wei, J., & Zou, K. (2019). *EDA: Easy data augmentation techniques for boosting performance on text classification tasks*. ArXiv Preprint. <https://doi.org/10.18653/v1/d19-1670>
- Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016). Hierarchical attention networks for document classification. In *Proceedings of the 2016 conference of the North American chapter of the association for computational linguistics: human language technologies* (pp. 1480-1489). Association for Computational Linguistics.
- Zhen, F., Yi, G., Zhenhao, Z., & Meiqi, H. (2018). Sentiment analysis of movie reviews based on dictionary and weak tagging information. *Journal of Computer Applications*, 38(11), 3084-3088.

## Analysis of Pure-pursuit Algorithm Parameters for Nonholonomic Mobile Robot Navigation in Unstructured and Confined Space

Izzati Saleh<sup>1</sup>, Azwati Azmin<sup>2</sup>, Azan Yunus<sup>3</sup> and Wan Rahiman<sup>1,4,5\*</sup>

<sup>1</sup>*School of Electric & Electronic Engineering, Universiti Sains Malaysia, Engineering Campus, 14300 USM, Nibong Tebal, Pulau Pinang, Malaysia*

<sup>2</sup>*Faculty of Electrical Engineering, Universiti Teknologi Mara, Cawangan Pulau Pinang, 13500 UiTM, Permatang Pauh, Malaysia*

<sup>3</sup>*CL Solution Sdn Bhd., Wisma Rampai, 59200 Kuala Lumpur, Malaysia*

<sup>4</sup>*Cluster of Smart Port and Logistics Technology (COSPALT), Universiti Sains Malaysia, Engineering Campus, 14300 USM, Nibong Tebal, Pulau Pinang, Malaysia*

<sup>5</sup>*Daffodil Robotics Lab, Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh*

### ABSTRACT

This research analyses Pure-pursuit algorithm parameters for nonholonomic mobile robot navigation in unstructured and constrained space. The simulation-based experiment is limited to the mobile robot arrangement. The Look Ahead Distance parameter is adjusted so the mobile robot can navigate the predefined map closely following the waypoints. The optimal Look Ahead Distance value is combined with the VFH+ algorithm for obstacle avoidance. The method is enhanced by adding the  $\lambda$  weight so the robot returns to its waypoints after avoiding an obstacle. The investigation reveals that  $\lambda$  influences the mobile robot's capacity to return to its predetermined waypoints after avoiding an obstacle. Based on the simulation experiment, the optimal LAD value is 0.2 m, and the optimal  $\lambda$  value is 0.8.

*Keywords:* Mobile robot navigation, obstacle detection, path following, VFH+ algorithm

### ARTICLE INFO

#### *Article history:*

Received: 29 December 2022

Accepted: 20 March 2023

Published: 06 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.06>

#### *E-mail addresses:*

[izzatisaleh@student.usm.my](mailto:izzatisaleh@student.usm.my) (Izzati Saleh)

[azwati3025@uitm.edu.my](mailto:azwati3025@uitm.edu.my) (Azwati Azmin)

[mkay@clsolution.com.my](mailto:mkay@clsolution.com.my) (Azan Yunus)

[wanrahiman@usm.my](mailto:wanrahiman@usm.my) (Wan Rahiman)

\* Corresponding author

### INTRODUCTION

Typical mobile robot navigation starts with the mobile robot taking input from sensors from the surroundings. From these data, the mobile robot can localise itself and generate a feasible path to navigate itself. The navigation path could be planned based on a predefined global path or from a local

planner. Based on the generated path information, the mobile robot will send signals to the microcontroller to move the robot.

Path-following is defined as a vehicle following a globally determined geometric path using steering motions to direct it along that path (Snider, 2009). Several path-following methods for ground vehicles have been adapted to mobile robots. However, a simpler yet effective mobile robot path-following method would be the geometric path-following method. One of the most used path-following methods would be the pure pursuit algorithm (Pure-pursuit).

The Pure-pursuit was initially used for missiles to pursue its target (hence the name). The method was then translated to the application for a ground vehicle by Coulter (1992), and in recent years the application expanded to nonholonomic mobile robot and autonomous vehicle applications (Ahn et al., 2021; Girbés et al., 2011; Huang et al., 2020; Li et al., 2019; Qinpeng et al., 2019; Shan et al., 2015; Wang et al., 2019; Wang et al., 2017; Yang et al., 2022). Since then, several improvements and modifications of Pure-Pursuit have been undertaken based on specific configurations of their respective mobile robot or autonomous vehicles. Research by Girbés et al. (2011) proposed a multi-level control scheme by considering different dynamics with different sampling frequencies, Shan et al. (2015) replaced circles on traditional Pure-Pursuit with clothoid C curve to reduce fitting error and Wang et al. (2017) who calibrated the heading and steering angle of the vehicle and reduced lateral error when the vehicle was following an ideal path with an improved accuracy by 54.54%.

Both research by Chen et al. (2018) and Li et al. (2019) proposed to include a PI (Proportional Integral) controller to address tracking errors in extreme drive conditions. Chen et al. (2018) also paired the algorithm with a low-pass filter to smooth the final output steering angle. Meanwhile, Li et al. (2019) used PID (Proportional Integral Derivative) to facilitate steering angle calculation when using the Pure-Pursuit approach.

Research by Wang et al. (2019) proposed a dual-stage fuzzy logic controller to adjust the mobile robot speed and Look-ahead distance in Pure-Pursuit to ensure the robustness and stability of the system. On top of that, Ahn et al. (2021) proposed a method of selection of Pure-Pursuit Look-ahead point heuristically based on the relationship between the vehicle and the path. Finally, Yang et al. (2022) proposed an algorithm that deduced the Look-ahead behaviour and scanned the area for the ideal goal point based on the evaluation function. The research objective is to minimise lateral and heading errors to achieve adaptive optimisation of the target location.

## **Paper Objective**

This paper aims to determine the optimal value of parameters affecting the trajectory of a differential drive mobile robot (later referred to as a mobile robot) in an unstructured and

confined space. The first parameter to be tuned is the Look-Ahead distance from Pure-Pursuit. The best value is then tested on the pure pursuit controller and integrated with an obstacle avoidance algorithm, Vector Field Histogram (VFH+).

The second parameter,  $\lambda$ , is introduced to ensure the mobile robot converges into the designated waypoints immediately after avoiding an obstacle. The integration of these two parameters is further tuned, and the best parameter is deduced. The tuning of the parameters is done in simulation and is limited to the mobile robot configuration and inside a pre-set map.

## MOBILE ROBOT NAVIGATION

### Pure-pursuit Algorithm

Pure-pursuit is a vehicle tracking algorithm that measures the curvature that drives a vehicle from its current location to a target position. The pure pursuit algorithm geometrically evaluates the curvature that will move the mobile robot to a target point. Figure 1 shows the geometry diagram of a Pure-pursuit adapted from Coulter (1992).

From Figure 1, by assuming the current location of the mobile robot in the global coordinate system is in origin  $(x_{robot}, y_{robot}) = (0,0)$ , let  $l$  be the Look-ahead Distance and  $(l_x, l_y)$  be the current Look-ahead point, whereby  $l$  is the hypotenuse of the right-angled triangle (Equation 1). The relationship between the radius of the arc  $r$  that joins the current location of the mobile robot  $(x_{robot}, y_{robot})$  with the Look-Ahead Distance point was explained in Equation 2, whereby  $d$  is represented by Equation 3.

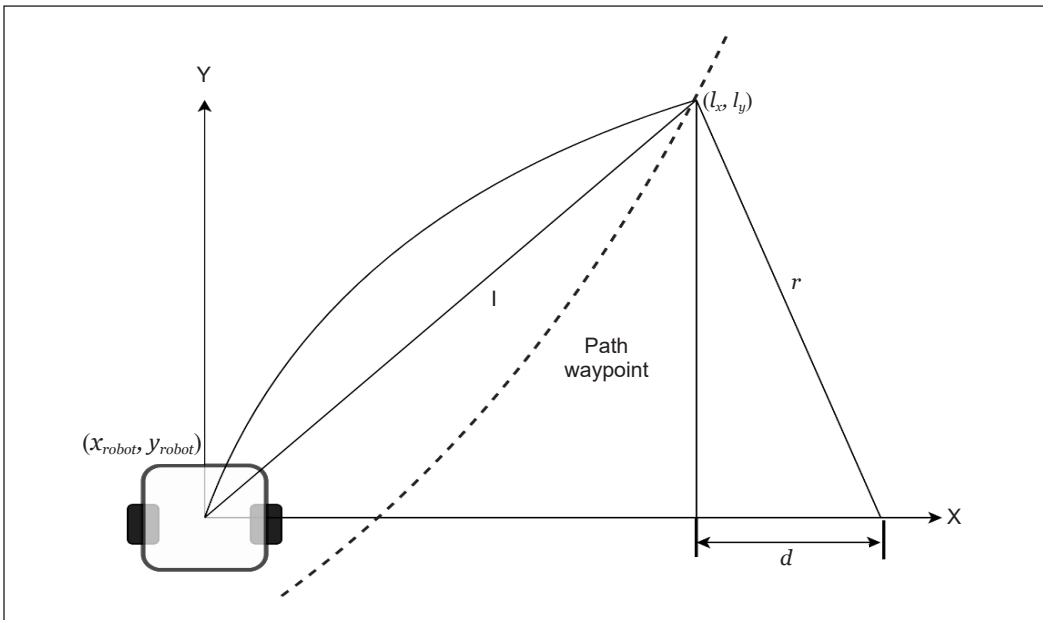


Figure 1. Geometry diagram of Pure-pursuit



$$l_x^2 + l_y^2 = l^2 \quad [1]$$

$$d^2 + l_y^2 = r^2 \quad [2]$$

$$l_x + d = r \quad [3]$$

By substituting  $d$  from Equation 3 to Equation 2, we can obtain arc radius  $r$  (Equations 4, 5 and 6):

$$(r - l_x)^2 + l_y^2 = r^2 \quad [4]$$

$$r^2 - 2rl_x + l_x^2 + l_y^2 = r^2 \quad [5]$$

$$r = l^2/2l_x \quad [6]$$

The  $r$  value determines the actual arc radius the vehicle will follow. The curvature of that radius is its reciprocal value  $\left(\frac{1}{r}\right)$ .

Figure 2 shows the Pure-Pursuit. Firstly, the waypoints of the mobile robot are established. The mobile robot will localise itself inside a global map. During every loop, the algorithm iterates to find the current position of the mobile robot, a new look-ahead point, and, consequently, the current arc radius. The algorithm will transform the goal point  $(l_x, l_y)$  to the mobile robot coordinate and consequently calculate the steering angle for the mobile robot to steer back into its path.

```

Data: Waypoints
Result: Robot Trajectory
Initialization;
while distanceToGoal > goalRadius do
    find current robot position ( $x_{robot}$ ,  $y_{robot}$ );
    find path point closest to robot position;
    find lookahead point ( $l_x$ ,  $l_y$ );
    transform goal point into robot coordinates;
    calculate the nagular velocity ( $\omega$ ) to steer back into path;
    update new robot position;
    update distanceToGoal;
    if distanceToGoal  $\leq$  goalRadius then
        | end search;
    else
        | go back to the beginning of loop;
    end
end

```

Figure 2. Pure-pursuit algorithm



## Look Ahead Distance

There are two major goals when the mobile robot tracks the path: regaining the mobile robot's position to the designated path and maintaining its position in the path. In the pure pursuit algorithm, one crucial parameter needs to be tuned: Look-ahead Distance (LAD). This parameter affects how far the mobile robot perceives the sets of waypoints. A small LAD is used for the mobile robot to follow the path of the waypoints closely. However, when the LAD is too small, the robot will overshoot the path and oscillate along the desired path. A larger LAD can be chosen for the mobile robot to converge to produce a smoother path gradually, but the robot might have difficulty manoeuvring into a small area due to larger curvatures near the corners.

## Vector Field Histogram (VFH+)

Vector Field Histogram (VFH) is an algorithm that calculates a mobile robot's obstacle-free steering direction (Bolbhat et al., 2020; Díaz & Marín, 2020; Dong et al., 2021; Pappas et al., 2020; Ulrich & Borenstein, 1998; Ulrich & Borenstein, 2000). To identify the location and proximity of obstacles, range sensor readings are used to compute polar density histograms. Unlike VFH, which is very goal-oriented and provides only one solution of steering direction, VFH+ determines a set of possible candidate directions based on all openings in the masked polar histogram. There is another extension of VFH+, which is VFH\*, that plans the waypoints based on the A\* approach; however, since, in this case, the waypoints were predefined, the former one was used instead. These candidate directions are then subjected to a cost function considering more than just the difference between the candidate and target directions. An opening is considered wide if the difference between its borders is larger than the maximum number of sectors  $s_{max}$ . For a narrow opening, there is only one candidate direction,  $c_n$ , and this can be represented by Equation 7:

$$c_n = \frac{k_r + k_l}{2} \quad [7]$$

There are two candidate directions for a wide opening: either on the left side  $c_l$  or on the right side  $c_r$ . Should the target direction lie between these two candidates, it can also be considered the third candidate  $c_t$  (Equation 8).

$$\begin{cases} c_r = k_r + s_{max}/2 \\ c_l = k_l - s_{max}/2 \\ c_t = k_t, k_t \in [c_r, c_l] \end{cases} \quad [8]$$

The cost function for a candidate  $g(c)$  can be represented with Equation 9:

$$g(c) = \mu_1 \cdot \Delta(c, k_1) + \mu_2 \cdot \Delta\left(c, \frac{\theta_n}{\alpha}\right) + \mu_3 \cdot \Delta(c, k_{d,n-1}) \quad [9]$$

and  $\Delta(c_1, c_2)$  is a function that computes the absolute angle difference between two sectors,  $c_1$  and  $c_2$  (Equation 10):

$$\Delta(c_1, c_2) = \min\{|c_1 - c_2|, |c_1 - c_2 - 360^\circ/\alpha|, |c_1 - c_2 + 360^\circ/\alpha|\} \quad [10]$$

where  $\alpha$  is the angular resolution of the histogram,  $\theta_n$  is the current orientation,  $k_t$  is the target direction divided by  $\alpha$  and  $k_{d,n-1}$  is the previously selected direction of motion/ $\alpha$ .

The terms  $\mu_1$ ,  $\mu_2$ , and  $\mu_3$  are accountable for how the mobile robot directs when facing an obstacle. The higher the  $\mu_1$ , the closer the mobile robot's steering direction to the goal point. Also, a higher  $\mu_2$  value produces an efficient path, while a higher  $\mu_3$  value ensures less oscillation in steering movements. The following condition must be satisfied to ensure the steering direction follows the goal direction (Equation 11):

$$\mu_1 > \mu_2 + \mu_3 \quad [11]$$

## METHODOLOGY

### Differential Drive Mobile Robot Kinematics

The simulation was coded using MATLAB software. To simulate a simplified vehicle model of a differential-drive mobile robot, a *differentialDriveKinematics* object (Figure 3) creates a differential-drive vehicle model. The model approximates a vehicle with a single fixed axle and wheels separated by a specified track width. For differential drive, the wheels were controlled independently. The speed and heading are defined from the axle centre. The vehicle state is defined as  $[x_{robot} \ y_{robot} \ \theta_{robot}]$ , the global coordinate inside a map measured in metres, whereas the heading was measured in radians.

### Simulation Experiment

Table 1 shows the Mobile Robot Hardware configurations. These parameters were input into the algorithm to simulate the behaviour of the differential drive robot.

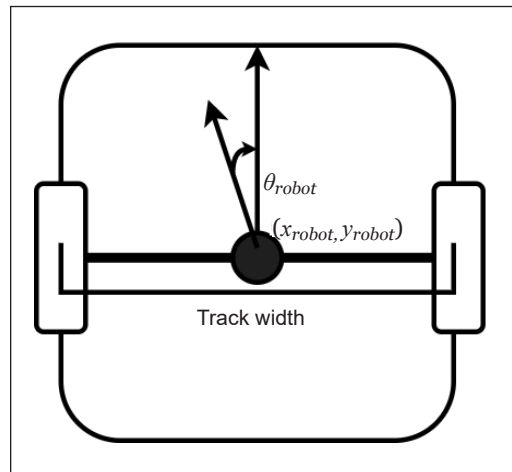


Figure 3. Differential drive robot kinematics

Table 1  
Mobile robot hardware configurations

Parameter	Value
Mobile robot radius	0.2 m
Track width	0.3 m
Minimum turning radius	0.15 m
Maximum angular velocity, $v$	1.82 rad/s
Maximum translational velocity, $v$	0.26 m/s
Range sensor angle range	0-360°
Range sensor max range	1.5m

**Test Environment.** Figure 4 shows the occupancy map of the test environment. The selected environment layout is unstructured, with non-symmetrical occupied space and a narrow pathway.

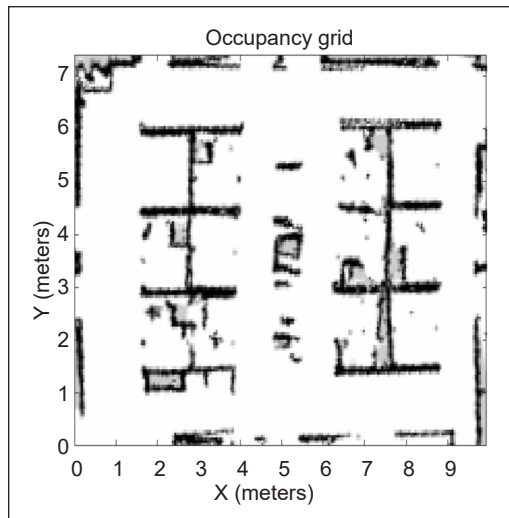


Figure 4. Map of test environment

**Path Following.** Table 2 shows pure pursuit algorithm parameters set constant in the experiment. The maximum angular velocity  $\omega$  is capped at 1 rad/s, and a set of waypoints were defined. These points were set to pass through wide and narrow gaps on the map. It should be noted that the initial waypoints did not intersect with any of the occupied spaces. Table 3 shows the simulation parameters to be tested to observe the effect of LAD and translational velocity,  $v$ , on the trajectory of the mobile robot. Since the simulated hardware’s maximum translational velocity,  $v$  is 0.26 m/s, the test parameters are capped at a translational velocity,  $v$ , of 0.2 m/s.

Table 2  
Pure pursuit algorithms parameters

Parameter	Value
Maximum angular velocity, $\omega$	1 rad/s
Initial waypoints (m)	[(1,6), (1,0.6), (4.5,0.6), (4.5,6), (5.8,6), (5.8,0.6), (9.2,0.60), (9.2,6)]

Table 3  
Look Ahead Distance (LAD)

Translational Velocity, $v$ (m/s)	LAD (m)
0.1	0.2
	0.4
	0.6
	0.8
	1.0
	2.0
0.2	0.2
	0.4
	0.6
	0.8
	1.0
	2.0

**Obstacle Avoidance.** Table 4 shows the selected parameters for the VFH+ algorithm. Safety Distance is the parameter to ensure safe navigation of the mobile robot. The obstacle-free space calculation considers the mobile robot’s radius with an added Safety Distance value. Histogram Threshold was used to compute binary histograms from the polar obstacle density. Any values higher than the upper threshold are considered occupied (1), whereas values smaller than the lower threshold are considered free space (0). Values between the threshold limit are set to follow the previous binary histogram with the initial value of free space (0).

The third parameter, Distance Limit, was set to consider only meaningful readings from the range sensor. The lower limit prevents false positive readings, while obstacles beyond the upper limit are ignored to accelerate computation. The consequent parameters  $\mu_1$ ,  $\mu_2$ , and  $\mu_3$  values were taken from Ulrich and Borenstein (1998, 2000), and the  $\alpha$  value is based on the property of the range sensor.

Table 4  
VFH+ parameters

Parameter	Value
Safety Distance	0.1 m
Histogram Threshold	[3,10]
Distance Limits	[0.05 m, 1.5 m]
$\mu_1$	5
$\mu_2$	2
$\mu_3$	2
$\alpha$	360°

**Heuristic Function.** A heuristic function is proposed to ensure the robot returns to its designated waypoints after avoiding obstacles (Equation 12). The angular velocity  $\omega$  will consider the magnitude of the calculated angular velocity from pure pursuit algorithm  $\omega_{path}$  and the calculated angular velocity from VFH+. A weight parameter  $\lambda$  is introduced to prevent the mobile robot from steering too far from the waypoints, specifically waypoints with sharp turns. Table 5 shows the selection values of  $\lambda$  to be tested.

Table 5  
 $\lambda$  parameters

Parameter	Value
$\lambda$	0.5
	0.6
	0.7
	0.8
	0.9
	1.0

$$\omega = \begin{cases} \lambda * \omega_{path} - \omega_{VFH}, & \text{if } \omega_{path} < 0 \\ \lambda * \omega_{path} + \omega_{VFH}, & \text{if } \omega_{path} > 0 \\ \omega_{VFH}, & \text{if } \omega_{path} = 0 \end{cases} \quad [12]$$

## RESULTS AND DISCUSSION

### Path Following

Figures 5 and 6 show plots of the trajectory of different LADs, where the translational velocity is fixed at 0.1 m/s and 0.2 m/s, respectively. In this simulation, mobile robot movement is based solely on waypoints, and the range sensor is disabled. From the plot, it could be observed that as the LAD value increases, the mobile robot will have a larger curvature. At LAD = 2.0 m, the curvature is too large that it collides with occupied space. While the larger LAD produces a smoother transition, the robot does not closely follow the designated waypoints. It could also be observed in Figure 7 that due to an increase in velocity, at LAD = 0.2 m, there is slight oscillation when the mobile robot turns at sharp corners. Based on these two plots, it could be concluded that the best LAD parameter for

the mobile robot configuration would be at 0.2 m for the velocity of 0.1 m/s and 0.4 m for 0.2 m/s (refer to marked points in Figures 5 and 6). Of the two, the best overall path following trajectory parameters is at a translational velocity of 0.1m/s and LAD 0.2 m. It is due to its ability to follow the waypoints at a sharp corner closely.

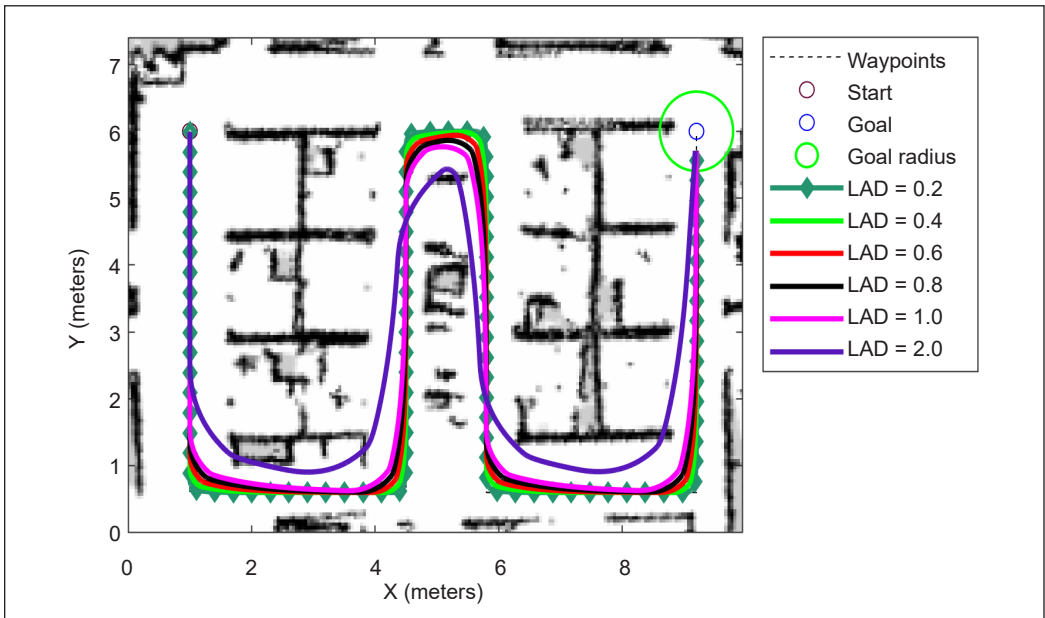


Figure 5. Plot of waypoints vs trajectory of different LAD (velocity = 0.1 m/s)

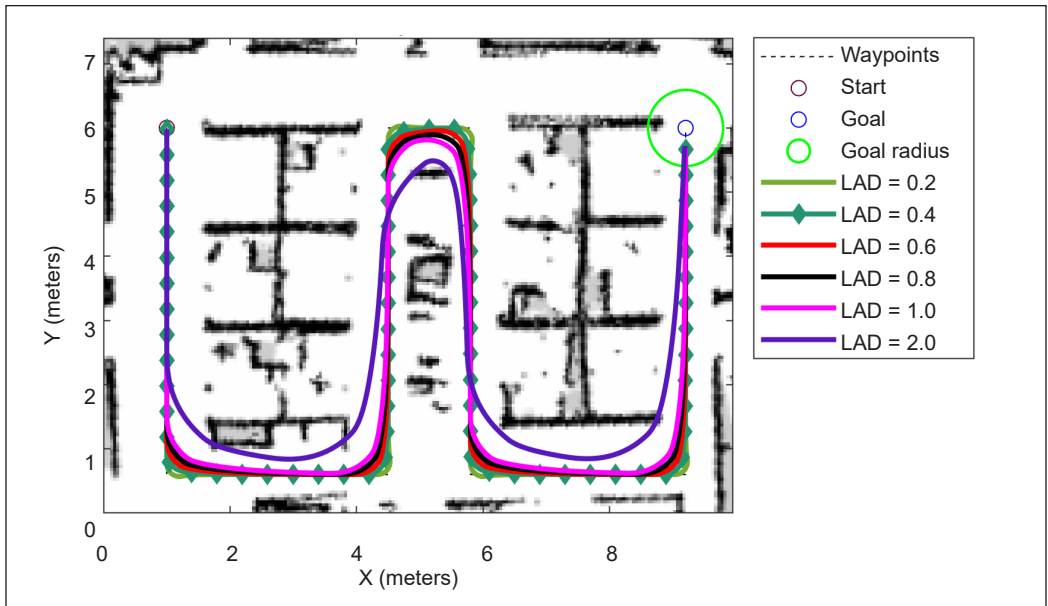


Figure 6. Plot of waypoints vs trajectory of different LAD (velocity = 0.2 m/s)



on the simulation data, the best value of  $\lambda$  is 0.8, as the mobile robot managed to avoid the obstacle and safely return to its designated waypoints. The trajectory of  $\lambda = 0.8$  is also relatively smoother compared to  $\lambda = 0.7$ .

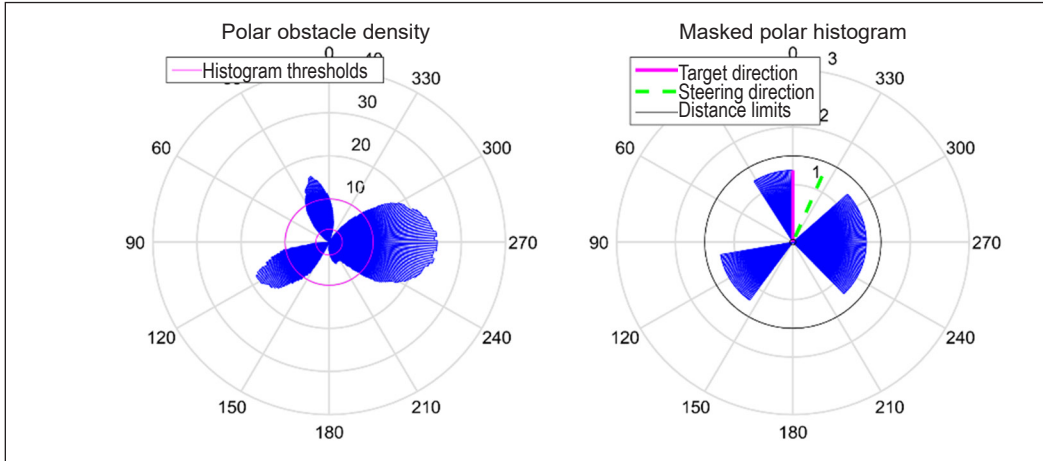


Figure 9. VFH+ histogram

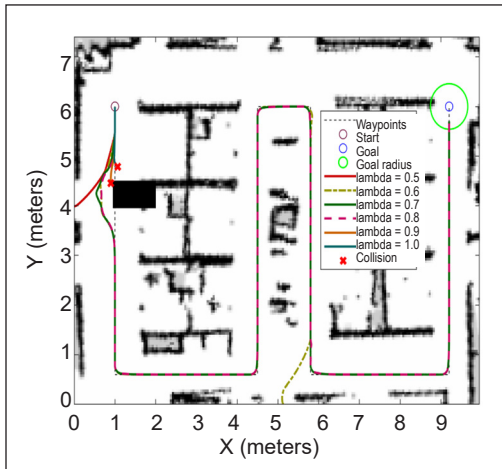


Figure 10. Plot of trajectory using Pure-pursuit and VFH+

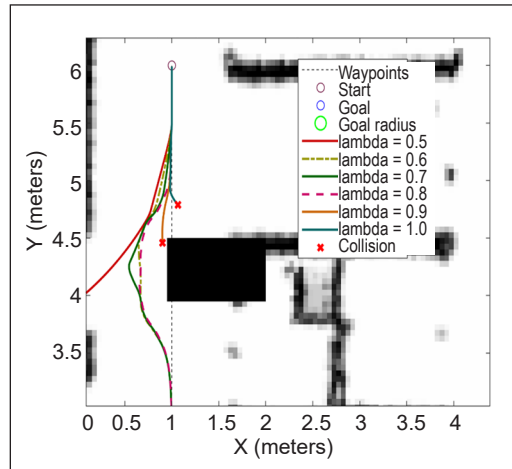


Figure 11. Plot of trajectory using Pure-pursuit and VFH+ (zoomed)

## CONCLUSION

In conclusion, Pure-pursuit is an effective geometric path-following algorithm. The simulation shows that by tuning the value of LAD, the mobile robot can navigate closely to its waypoints inside an unstructured and confined space. By integrating the VFH+ algorithm as an obstacle avoidance method and introducing a weight parameter  $\lambda$ , the mobile robot can avoid an obstacle and return to its designated waypoints.

## FUTURE WORKS

As for future works, the author plans to expand the work by applying the navigation concepts to industrial hardware and drones (Ibrahim et al., 2017). The work can also be significantly improved using metaheuristic optimisation, as proposed by Wang et al. (2020).

## ACKNOWLEDGEMENT

This work was supported by Collaborative Research in Engineering, Science, and Technology (CREST), Malaysia, with grant no. 304/PELECT/6050423/C121.

## REFERENCES

- Ahn, J., Shin, S., Kim, M., & Park, J. (2021). Accurate path tracking by adjusting look-ahead point in pure pursuit method. *International Journal of Automotive Technology*, 22(1), 119-129. <https://doi.org/10.1007/s12239-021-0013-7>
- Bolbhat, S. S., Bhosale, A. S., Sakthivel, G., Saravanakumar, D., Sivakumar, R., & Lakshmipathi, J. (2020). Intelligent obstacle avoiding agv using vector field histogram and supervisory control. In *Journal of Physics: Conference Series* (Vol. 1716, No. 1, p. 012030). IOP Publishing. <https://doi.org/10.1088/1742-6596/1716/1/012030>
- Chen, Y., Shan, Y., Chen, L., Huang, K., & Cao, D. (2018). Optimization of pure pursuit controller based on pid controller and low-pass filter. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)* (pp. 3294-3299). IEEE Publishing. <https://doi.org/10.1109/ITSC.2018.8569416>
- Coulter, R. C. (1992). *Implementation of the Pure Pursuit Path Tracking Algorithm*. Defense Technical Information Center.
- Díaz, D., & Marín, L. (2020). VFH+ D: An improvement on the VFH+ algorithm for dynamic obstacle avoidance and local planning. *IFAC-PapersOnLine*, 53(2), 9590-9595. <https://doi.org/10.1016/j.ifacol.2020.12.2450>
- Dong, X., Wang, M., Kang, C., Zhang, Y., Pan, W., & Gao, S. (2021). Improved auto-regulation VFH algorithm for obstacle avoidance of unmanned vehicles. In *2021 China Automation Congress (CAC)* (pp. 6467-6472). IEEE Publishing. <https://doi.org/10.1109/CAC53003.2021.9727888>
- Girbés, V., Armesto, L., Tornero, J., & Solanes, J. E. (2011). Smooth kinematic controller vs. pure-pursuit for non-holonomic vehicles. In *Towards Autonomous Robotic Systems: 12th Annual Conference, TAROS 2011* (pp. 277-288). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-23232-9\\_25](https://doi.org/10.1007/978-3-642-23232-9_25)
- Huang, Y., Tian, Z., Jiang, Q., & Xu, J. (2020). Path tracking based on improved pure pursuit model and pid. In *2020 IEEE 2nd International Conference on Civil Aviation Safety and Information Technology (ICCASIT)* (pp. 359-364). IEEE Publishing. <https://doi.org/10.1109/ICCASIT50869.2020.9368694>
- Ibrahim, I. N., Izhevsk, K., Pavol, B., Aiman, A. A. M., Izhevsk, K., & Karam, A. (2017). Navigation control and stability investigation of a hexacopter equipped with an aerial manipulator. In *Proceedings of the 2017 21st International Conference on Process Control, PC 2017* (pp. 204-209). IEEE Publishing. <https://doi.org/10.1109/PC.2017.7976214>



- Li, H., Luo, J., Yan, S., Zhu, M., Hu, Q., & Liu, Z. (2019). Research on parking control of bus based on improved pure pursuit algorithms. In *2019 18th International Symposium on Distributed Computing and Applications for Business Engineering and Science (DCABES)* (pp. 21-26). IEEE Publishing. <https://doi.org/10.1109/DCABES48411.2019.00013>
- Pappas, P., Chiou, M., Epsimos, G.-T., Nikolaou, G., & Stolk, R. (2020). Vfh+ based shared control for remotely operated mobile robots. In *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)* (pp. 366-373). IEEE Publishing. <https://doi.org/10.1109/SSRR50563.2020.9292585>
- Qinpeng, S., Zhonghua, W., Meng, L., Bin, L., Jin, C., & Jiayang, T. (2019). Path tracking control of wheeled mobile robot based on improved pure pursuit algorithm. In *2019 Chinese Automation Congress (CAC)* (pp. 4239-4244). IEEE Publishing. <https://doi.org/10.1109/CAC48633.2019.8997258>
- Shan, Y., Yang, W., Chen, C., Zhou, J., Zheng, L., & Li, B. (2015). CF-Pursuit: A pursuit method with a clothoid fitting and a fuzzy controller for autonomous vehicles. *International Journal of Advanced Robotic Systems*, 12(9), Article 134. <https://doi.org/10.5772/61391>
- Snider, J. M. (2009). *Automatic Steering Methods for Autonomous Automobile Path Tracking*. Carnegie Mellon University.
- Ulrich, I., & Borenstein, J. (1998). VFH+: Reliable obstacle avoidance for fast mobile robots. *IEEE International Conference on Robotics and Automation*, 2, 1572-1577. <https://doi.org/10.1109/ROBOT.1998.677362>
- Ulrich, I., & Borenstein, J. (2000). VFH\*: Local obstacle avoidance with look-ahead verification. *Proceedings-IEEE International Conference on Robotics and Automation*, 3, 2505-2511. <https://doi.org/10.1109/ROBOT.2000.846405>
- Wang, H., Chen, X., Chen, Y., Li, B., & Miao, Z. (2019). Trajectory tracking and speed control of cleaning vehicle based on improved pure pursuit algorithm. In *2019 Chinese Control Conference (CCC)* (pp. 4348-4353). IEEE Publishing. <https://doi.org/10.23919/ChiCC.2019.8865255>
- Wang, R., Li, Y., Fan, J., Wang, T., & Chen, X. (2020). A novel pure pursuit algorithm for autonomous vehicles based on salp swarm algorithm and velocity controller. *IEEE Access*, 8, 166525-166540. <https://doi.org/10.1109/ACCESS.2020.3023071>
- Wang, W. J., Hsu, T. M., & Wu, T. S. (2017). The improved pure pursuit algorithm for autonomous driving advanced system. In *2017 IEEE 10th International Workshop on Computational Intelligence and Applications (IWCIA)* (pp. 33-38). IEEE Publishing. <https://doi.org/10.1109/IWCIA.2017.8203557>
- Yang, Y., Li, Y., Wen, X., Zhang, G., Ma, Q., Cheng, S., Qi, J., Xu, L., & Chen, L. (2022). An optimal goal point determination algorithm for automatic navigation of agricultural machinery: Improving the tracking accuracy of the pure pursuit algorithm. *Computers and Electronics in Agriculture*, 194, Article 106760. <https://doi.org/10.1016/j.compag.2022.106760>



## Use of Enhanced Greedy Algorithm for Load Balancing in Cloud Computing

Hanaa Osman<sup>1</sup>, Asma'a Yassin Hammo<sup>1</sup> and Abdulnasir Younus Ahmad<sup>2\*</sup>

<sup>1</sup>Computer Science and Mathematics, University of Mosul, 259 AL-Majmoaa St., Mosul, Nineveh, Iraq

<sup>2</sup>Education for Pure Science, University of Mosul, 259 AL-Majmoaa St., Mosul, Nineveh, Iraq

### ABSTRACT

Because of the Internet's phenomenal growth in recent years, computing resources are now more widely available. It led to the development of a new computing concept known as Cloud Computing, allowing users to share resources such as networks, servers, storage, applications, services, software, and data across multiple devices on demand for economical and fast. Load balancing is an important branch of cloud computing as it optimizes machine utilization by distributing tasks equally over resources. It occurs among physical hosts or Virtual Machines in a cloud environment. Round robin is a commonly used algorithm in load balancing. RR gives a time quantum for each task and is in circular order. It is noted that it suffers from many problems, such as the waste of time and the high cost. In the present study, the greedy algorithm was enhanced and implemented to allocate and schedule tasks that come to the cloud on Virtual Machines in balance. The task with the longest execution time is given to the virtual machine with the least load using an improved greedy algorithm. The outcomes demonstrate that the suggested algorithm outperformed round robin in makespan. Also, all Virtual Machines in the proposed algorithm finish their work simultaneously, whereas round robin is unbalanced.

*Keywords:* Cloud computing, cloudSim, greedy algorithm, load balancing, makespan, round robin, Virtual Machine

### ARTICLE INFO

*Article history:*

Received: 25 October 2022

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.07>

*E-mail addresses:*

hanaaosman@uomosul.edu.iq (Hanaa Osman)

asmahammo@uomosul.edu.iq (Asma'a Yassin Hammo)

abdulnasir.younus@uomosul.edu.iq (Abdulnasir Younus Ahmad)

\* Corresponding author

### INTRODUCTION

Computing resources are now more widely available, allowing for the development of a new computing concept known as Cloud computing because of the Internet's phenomenal growth in recent years (Fatima et al., 2019). Cloud computing is a kind of Internet-based computing that enables

clients to share resources, software, and data across various devices on demand (Mishra et al., 2018). Generally, the cloud computing platform has three major issues: virtualization, distributed framework, and load balance (Kumar et al., 2020).

Virtualization technology improves cloud resource use by making various resources available online for customers to purchase on a pay-per-use basis. Physical servers (maybe even only one) can be set up and divided into numerous unconnected “virtual servers” (PS), each of which functions independently and appears to the user to be a single physical device. PSs can be moved in any direction and scaled up or down without causing harm to the end user because they are not physically tied (Fatima et al., 2019).

A distributed system is made up of numerous independent computers that communicate with one another via a computer network. Computers communicate to reach a common aim (Nerkar, 2012). A distributed cloud that connects several geographically dispersed and smaller data centers can be a compelling alternative to today’s big, centralized data centers. A distributed cloud can reduce communication overheads, prices, and latency by providing adjacent processing and storage resources. Improved data locality can also help with privacy. Many smaller data centers are installed closer to consumers to complement or enhance the bigger mega-data centers under the distributed cloud deployment paradigm, and the smaller data centers are administered as a single pooled resource (Coady et al., 2015).

When certain Virtual Machines (VMs) are overloaded with processing duties while others are underloaded, the load must be balanced to achieve optimal machine utilization (Babu & Krishna, 2013; Kumar & Kumar, 2019). Typically, web traffic, application access, databases, and other entities with large loads can use load balancer software to support uninterrupted service to Clients.

Load balancing may occur among physical hosts or Virtual Machines in a cloudy medium (Kumar & Kumar, 2019).

There are two basic challenges with load balancing:

- Task allocation refers to the distribution of a fixed number of tasks over a large number of Physical Machines (PMs), followed by VMs concerning the PMs.
- VMs Relocation Management: The process of moving virtual machines from one PMs to another to increase the data center’s resource consumption is called Migration (Kumar & Kumar, 2019).

Millions of users share cloud resources by submitting their computing tasks to the cloud system. The cloud computing environment faces a hurdle in scheduling these millions of tasks. The scheduling of cloud services is divided into two categories: user and system. Scheduling at the user level addresses issues arising from service supply between providers and customers. Within the data center, resource management is handled through system-level scheduling (Tawfeek et al., 2013).

Some of the most typical goals of adopting load balancing techniques are reducing waiting time, reducing the response time, increasing the utilization of resources, improving

reliability, increasing throughput, minimizing turn-around-time, and minimizing makespan (Dave & Maheta, 2014).

One of the most common load balancing algorithms used is round robin (RR). Although its technique is straightforward and convenient, it has several flaws, including time loss and a high cost. It is primarily due to the job being assigned to the incorrect virtual machine, which does not consider the task's size or the requirements of the virtual machine to which it is attached. This work aims to enhance and implement a greedy algorithm to allocate and schedule tasks to the cloud on balance. The algorithm is compared with what is called the RR algorithm.

### **Related Works**

Caragiannis et al. (2011) investigate two scenarios: selfish load balancing and online load balancing. The researchers describe the impact of selfishness and greediness on load balancing. They show that anarchy and stability of selfish load balancing have been enhanced and tightened. They also constrain the greedy algorithm's competitiveness for online load balancing, where the goal is to reduce latency for all clients on servers. Babu and Krishna (2013) proposed an algorithm named honey bee behavior-inspired load balancing (HBB-LB). The proposed algorithm balances the priority of tasks on the machines with the waiting time. When compared to existing algorithms, they found that the algorithm is efficient. The method considerably reduces average execution time and task queue waiting time. Ramezani et al. (2014) proposed a "Task-Based System Load Balancing with Particle Swarm Optimization" approach for achieving a system load balancing by relocating surplus tasks from an overloaded VM instead of migrating the entire overloaded VM. This approach significantly reduces the load-balancing process time compared to traditional load-balancing systems. It reduced VM downtime and the risk of losing a customer's most recent activity while improving cloud customers' Quality of Service. A unique greedy algorithm was given in Paduraru (2014). The researchers concluded that the suggested algorithm outperformed traditional approaches for load balancing algorithms, but it had a higher overhead than other well-known methods. Kapoor and Dabas (2015) suggested a load-balancing technique based on clusters. The algorithm was tested against current and modified throttled algorithms and found superior execution time, response time, throughput, and turnaround time.

The ant colony optimization algorithm was compared to alternative cloud scheduling algorithms FCFS and round-robin. According to the researchers, the ant colony optimization surpassed the FCFS and round-robin algorithms. Awad et al. (2015) announced that their proposed Load Balancing Mutation a Particle Swarm Optimization (LBMP SO) approach was compared to three algorithms: random algorithm, standard PSO, and Longest Cloudlet to Fastest Processor (LCFP). According to them, the comparison shows that LBMP SO

outperformed those algorithms in terms of execution time, makespan, transmission cost, and round trip time. Devi and Uthariaraj (2016) proposed the firefly algorithm, which reduced the time it took to execute the submitted jobs. The proposed approach takes less time to execute than First Come First Served (FCFS). Lu et al. (2016) introduced a hybrid scheduling algorithm DCBT algorithm for load balancing in a distributed environment by merging the methodologies of “Divide-and-Conquer” and “Throttled” algorithms. The researchers defined two situations. The DCBT method was utilized in both situations to schedule incoming client requests to the available virtual machines based on the load on each machine.

The suggested DCBT uses virtual machines better while lowering task execution time. In Mohanty et al. (2017), researchers proposed a metaheuristic load-balancing approach employing Particle Swarm Optimization (MPSO). The proposed method tries to reduce task overhead while maximizing resource utilization. Performance comparisons are done with the genetic algorithm and other algorithms on multiple parameters, including makespan calculation and resource utilization. The proposed method outperforms previous approaches. The imbalance of the load in System Wide Information Management (SWIM) task scheduling is the focus of Li and Wu (2019). A load balancing-based SWIM ant colony task scheduling method (ACTS-LB) is given in that study. The ACTS-LB algorithm outperforms the standard min-min strategy, the ACO algorithm, and the particle swarm optimization (PSO) algorithm in experimental simulations. It not only speeds up job execution and makes better use of system resources, but it also keeps SWIM in a more load-balanced state. According to Kruekaew and Kimpan (2020), virtual machine scheduling management using the “artificial bee colony” (ABC) algorithm and the largest job first (HABC LJF) surpassed ACO, PSO, and IPSO. The results demonstrate that the HABC with the Largest Job First (HABC LJF) heuristic method performs the best in scheduling and load balancing. Sinha and Sinha (2020) developed a load-balancing algorithm for effective resource usage and compared the performance of the proposed algorithms to those of well-known load-balancing algorithms. Compared to the RR, Throttled, ACO, and Hybrid approaches, the EWRR method has less response time.

Singh et al. (2021) employed the Crow intelligent algorithm to distribute tasks among VMs. They evaluated the effectiveness of their method using the CloudSim simulator. 32 GB of Memory and a 1 TB hard drive were employed in 16 virtual machines. They discovered that their system reached an ideal state after a limited number of rounds. Kruekaew and Kimpan (2022) employed the ABC algorithm to improve load balancing and resource utilization. For their performance analysis, they employed CloudSim. Using CloudSim, they compare their results utilizing FCFS, MOPSO (Multi-Objective Particle Swarm Optimization), and MOCS with random data sets. Replication improved the performance of cloud services, according to Javadpour et al. (2023), on one or more virtual machines, they found a solution by choosing and eliminating the VMs whose storage capacity was overcrowded when a VM was requested by more than one user yet needed storage space.

The CloudSim simulator was employed. A greedy algorithm was used for balancing, but the sorting in the algorithm needs time, especially when the job list is very long. So, an enhanced version of Greedy is used in this research.

## METHOD

“CloudSim” is a cloud computing simulation program developed in the CLOUDS laboratory at the “University of Melbourne.” It allows users to assess individual system concerns without taking into account the low-level details of cloud-based infrastructures and services (Ahmad & Khan, 2019).

CloudSim is an open-source platform for simulating cloud computing services and infrastructure. It enables users to repeatedly test applications under their control, detect system bottlenecks without using real clouds, and experiment with different configurations to build adaptive provisioning ways.

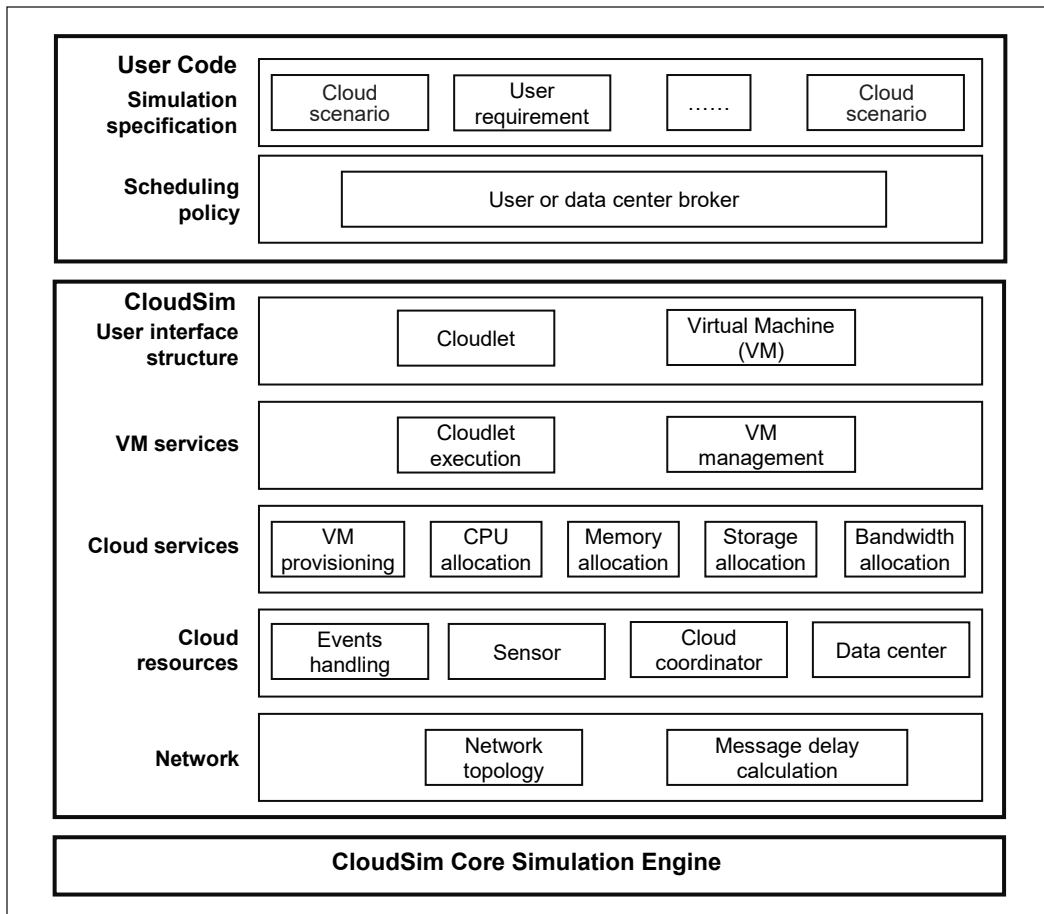


Figure 1. The architecture of CloudSim (Goyal et al., 2012). Adapted from <https://doi.org/10.1016/j.proeng.2012.06.412>

Version 3.0.1 is utilized in this study. Figure 1 depicts the CloudSim Simulation architecture (Goyal et al., 2012). The simulation layer includes specialized management interfaces for virtual machines, memory, storage, and bandwidth, as well as support for modeling and simulation of virtualized Cloud-based data center infrastructures.

This layer deals with the basics, such as assigning hosts to virtual machines, managing application execution, and keeping track of the system's dynamics (Wang & Wu, 2009). The "User Code" is the top layer of the CloudSim simulation toolkit, and it is the major interface for configuring simulation specifications and characteristics, including the number of machines, apps, jobs, users, and scheduling regulations, as well as their basic structure (Sinha & Shekhar, 2015).

Figure 2 depicts the basic scenario of the CloudSim simulation. Each component of CloudSim is described in depth in this scenario: Data centers (DC) offer resources such as multiple hosts. A host is a hardware machine that can host many virtual machines. A virtual machine (VM) is a piece of logical hardware on which the cloudlet will run. Broker has data center capabilities, which allows it to send virtual machines to the appropriate host. The "Cloud Information Service (CIS)" is in charge of locating resources, indexing them, and calculating the efficiency of data centers (Ahmad & Khan, 2019).

Results can be presented in figures, graphs, tables, and others that make the reader understand easily (Kapoor & Dabas, 2015; Saura et al., 2019). The discussion can be made in several sub-sections.

Two scheduling policies are available in CloudSim: Cloudlet scheduler policies and VMs scheduler policies (Ahmad & Khan, 2019). A greedy algorithm is used in this research to enhance the cloudlet scheduler and obtain a balanced load.

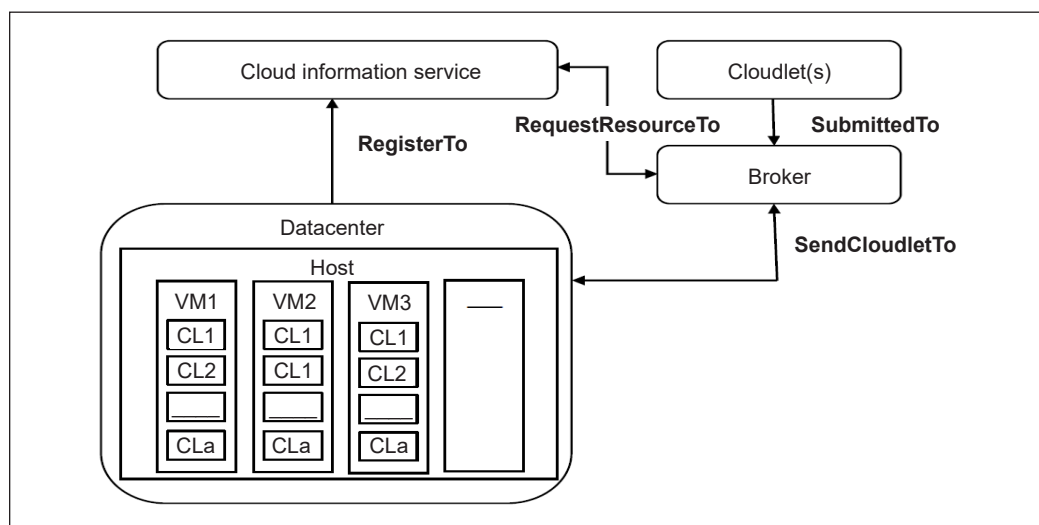


Figure 2. Scenario of CloudSim simulation (Ahmad & Khan, 2019). Adapted from <https://doi.org/10.35940/ijrte.B3669.078219>



As mentioned before, the current version of the Greedy algorithm needs sorting. It is better to take only the MAX and Min. It will decrease the overhead of the algorithm. The Greedy method proposes allocating the most time-consuming and difficult work to the more efficient virtual machine. i.e., VM with the least load, to rapidly do complicated and tough jobs and reduce the overall system's execution time and makespan.

Makespan refers to the duration of a user's task scheduling. Its value is expressed by the time from the start of the first task to the end of the last task's execution. The performance criteria values were calculated using Equations 1, 2, and 3.

- Makespan for the entire system

$$\text{Makespan} = \text{Max} (FT_{ij}) \quad [1]$$

Such as  $FT_{ij}$  finishing (ending) time for task  $T_i$ . Consider the start time is 0.

- The total capacity of all Virtual machines:

$$C = \sum_{j=1}^m C_{vmj} \quad [2]$$

Where  $C$  represents the combined capacity of all VMs, the capacity of each virtual machine, or  $C_{vm}$ , is determined using Equation 3.

$$C_{vm} = \text{Penum} * \text{Pemips} \quad [3]$$

The number of processing elements (Pe) in VM is called Penum.

The million instructions per second that Pe can execute are known as Pemips.

Following is the *Pseudo code of the greedy algorithm*:

Input: No. of tasks, No. of VMs.

Output: assigning tasks to VMs, system makespan.

Begin

While the tasks' queue is not empty Do

Choose Task T(i) that has the biggest execution time

Choose VM (j) with Minimum load

Allocate the task T(i) to Vm(j)

End While

Calculate the finishing time for all VMs

Calculate the makespan for all the systems.

End

## RESULTS AND DISCUSSION

It is contrasted to the popular RR method, which is typically employed for load balancing to improve the performance of the proposed algorithm. The mentioned algorithms in

the related work are not available to us. So, it is impossible to compare with them. The lengths of tasks generated by the random built-in function are presented in Figure 3. The random is chosen because the jobs could arrive at the cloud in any size. The job size is considered in seconds. Cloudsim is used for the simulation of the cloud. The Cloudsim configurations listed below are considered: four VMs are employed for the experiments. The specs are the same for all VMs. As a result, the experiment concentrated on the algorithm rather than the VM requirements. Each virtual machine operates on 250 MIPS, 512 MB of RAM, and 1000 Mbps (Million Byte per second) of bandwidth to run. Indeed, the experiments to test the suggested algorithm use this setting as an example. Ten times, starting at 50 tasks and going up by 50 tasks each time, the experiment is run until it reaches 500 tasks. It will change the system's load and allow researchers to assess how it affects its makespan. Table 1 shows how jobs were completed using the suggested Greedy method. The beginning and ending times of each task, as well as the virtual machine on which it is running. It is independent of task sequencing and instead uses task length to place it on a virtual machine with minimal load.

For each number of tasks, the finishing time of every VM and the overall makespan of the system is calculated. The results of both algorithms are shown in Table 2. As seen in the suggested algorithm, every VM

Table 1  
A snapshot of greedy output result for 4 VMs

Cloudlet ID	VM ID	Time of task (sec)	Start time (sec)	Finish time (sec)
6	1	8	0.1	8.1
3	2	12	0.1	12.1
8	1	24	8.1	32.1
0	0	40	0.1	40.1
2	0	8	40.1	48.1
1	3	60	0.1	60.1
5	2	52	12.1	64.1
7	0	36	48.1	84.1
4	3	32	60.1	92.1
11	0	12	84.1	96.1
9	3	4	92.1	96.1

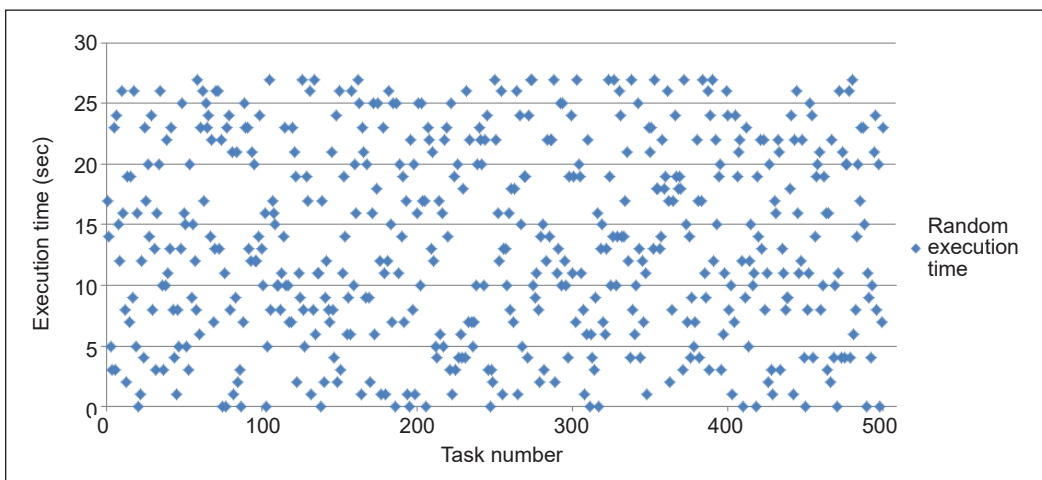


Figure 3. Random execution time

completes its work in a balanced manner (with more or less). However, the difference in the VMs' completion times is evident in RR.

A comparison of the makespan of all systems utilizing RR and the suggested technique is shown in Figure 4. The system's makespan is shorter when using the suggested approach than when utilizing RR. The reason for that is that when the number of tasks increases, there will be a higher probability of the arrival of different sizes of tasks, which in turn improves the selection and gives a better chance to achieve convergent makespan. This behavior differs from traditional RR, which makes no selection but chooses the task from the front of the task queue.

Table 2  
VMs finishing times and makespan for both algorithms

No. of Task	Greedy VMs finishing time (sec)				RR VMs finishing time (sec)				Makespan (sec)	
	VM 1	VM 2	VM 3	VM 4	VM 1	VM 2	VM 3	VM 4	Makespan G	Makespan R.R.
50	596.1	596.1	596.1	600.1	652.1	604.1	656.1	476.1	600.1	656.1
100	1408.1	1408.1	1412.1	1412.1	1616.1	1448.1	1204.1	1372.1	1412.1	1616.1
150	2156.1	2156.1	2156.1	2156.1	2120.1	2160.1	1948.1	2396.1	2156.1	2396.1
200	2920.1	2920.1	2920.1	2920.1	2940.1	2864.1	3156.1	2720.1	2920.1	3156.1
250	3672.1	3672.1	3672.1	3672.1	3604.1	3568.1	3856.1	3660.1	3672.1	3856.1
300	4072.1	4076.1	4076.1	4076.1	3964.1	4188.1	4040.1	4108.1	4076.1	4188.1
350	5172.1	5172.1	5172.1	5172.1	5228.1	5260.1	5276.1	4924.1	5172.1	5276.1
400	5824.1	5824.1	5828.1	5828.1	6276.1	5548.1	5720.1	5760.1	5828.1	6276.1
450	6324.1	6328.1	6328.1	6328.1	6208.1	6344.1	6432.1	6324.1	6328.1	6432.1
500	7324.1	7324.1	7328.1	7328.1	6912.1	7888.1	7344.1	7160.1	7328.1	7888.1

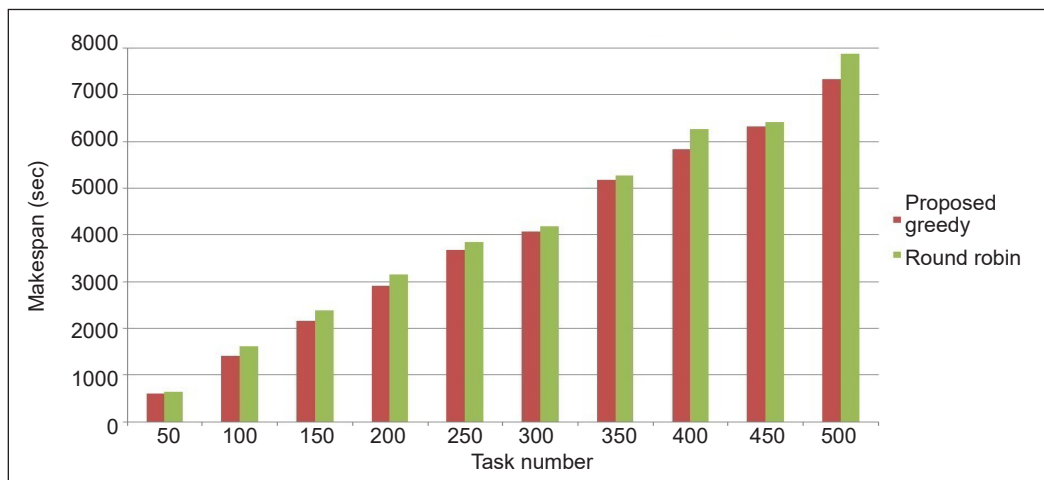


Figure 4. Makespan for different no. of tasks for round robin and proposed greedy algorithm

Figure 5 shows that all VMs in RR finish their jobs at various times. Whereas in greedy, VMs finish jobs nearly simultaneously (Figure 6). The greedy allocator takes the initial step in load balancing by assigning the task to the VM with the lowest load.

According to Table 3 and Figure 7, the improved greedy algorithm's standard deviation of the finishing values spans from 0 to 2.309401, whereas the RR algorithm's standard deviation ranges from 84.0555372 to 414.3750314. They also demonstrate no correlation between the number of tasks and the standard deviation of the VMs' completion values.

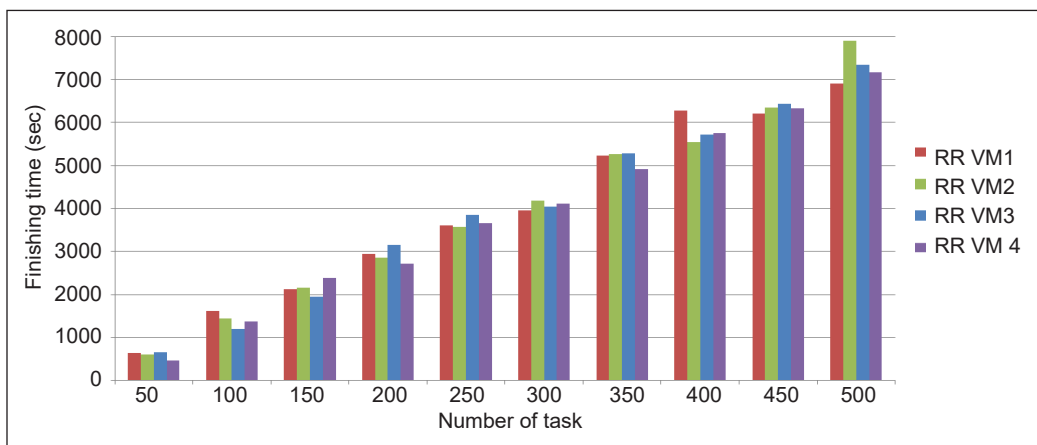


Figure 5. Unbalanced Virtual Machines' (VM) finishing time in round robin (RR)

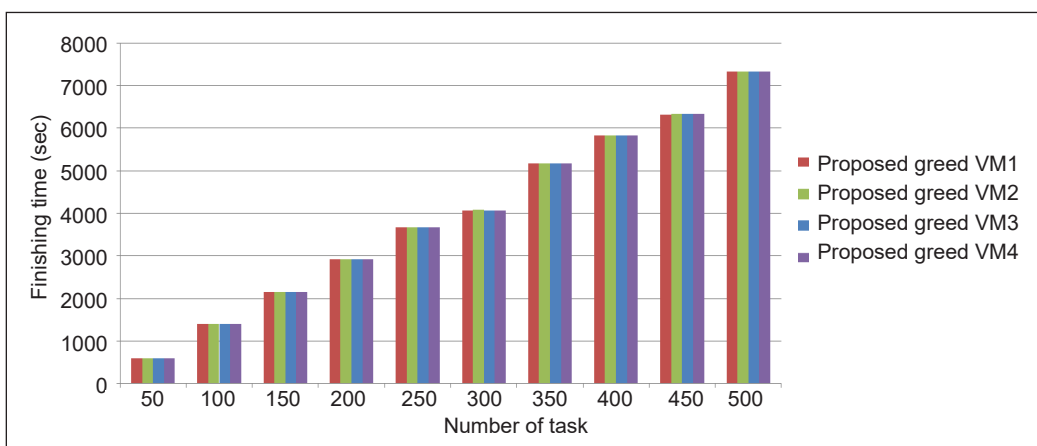


Figure 6. Balanced Virtual Machines' (VM) finishing time in greedy algorithm

Table 3  
Enhanced greedy and round robin algorithm standard deviation

No. of tasks	50	100	150	200	250	300	350	400	450	500
STDEV Greedy	2.00	2.31	0.00	0.00	0.00	2.00	0.00	2.31	2.00	2.31
STDEV RR	84.06	171.04	184.55	181.87	128.37	95.58	166.53	313.78	92.17	414.38

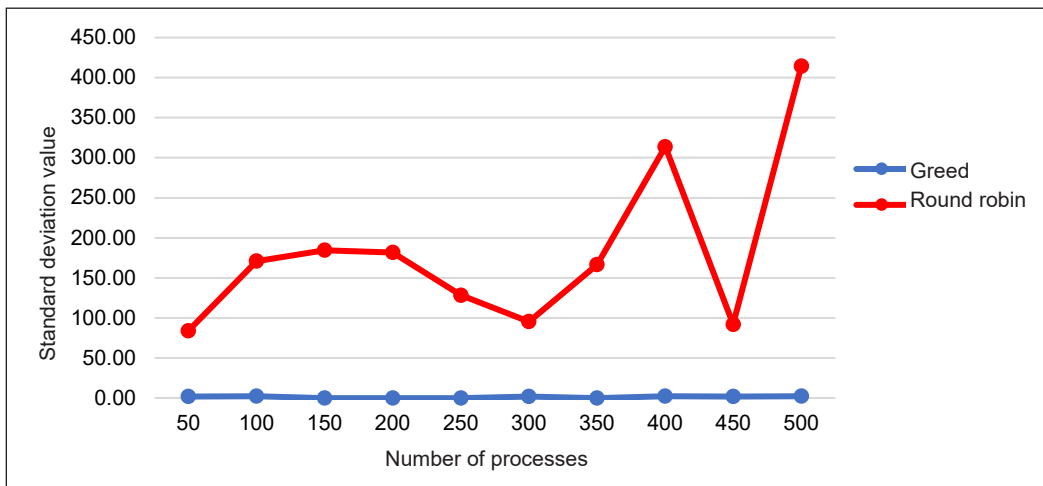


Figure 7. Standard deviation of makespans at different loads

## CONCLUSION

The present research proposes an enhanced version of a Greedy algorithm to allocate tasks in the cloud. When the request arrives at the data center, deciding which VM is the most suitable to execute it is necessary. A greedy algorithm will help to make this choice to achieve load balancing. There is no need for migration since the chosen VM will satisfy its constraints. The algorithm was compared to RR, and the results showed that the makespan of the system using Greedy is less than its equivalent using RR. Also, all VMs in Greedy finish their work simultaneously, whereas RR VMs vary in finishing time, leading to an unbalanced system load.

## ACKNOWLEDGMENTS

The authors appreciate the cooperation of the University of Mosul, Iraq, in completing the experimental work.

## REFERENCES

- Ahmad, M. O., & Khan, R. Z. (2019). Cloud computing modeling and simulation using cloudsim environment. *International Journal of Recent Technology and Engineering*, 8(2), 5439-5445. <https://doi.org/10.35940/ijrte.B3669.078219>
- Awad, A. I., El-Hefnawy, N. A., & Abdel-Kader, H. M. (2015). Enhanced particle swarm optimization for task scheduling in cloud computing environments. *Procedia Computer Science*, 65, 920-929. <https://doi.org/10.1016/j.procs.2015.09.064>
- Babu, L. D. D., & Krishna, P. V. (2013). Honey bee behavior inspired load balancing of tasks in cloud computing environments. *Applied Soft Computing Journal*, 13(5), 2292-2303. <https://doi.org/10.1016/j.asoc.2013.01.025>

- Caragiannis, I., Flammini, M., Kaklamanis, C., Kanellopoulos, P., & Moscardelli, L. (2011). Tight bounds for selfish and greedy load balancing. *Algorithmica*, 61, 606-637. <https://doi.org/10.1007/s00453-010-9427-8>
- Coady, Y., Hohlfeld, O., Kempf, J., McGeer, R., & Schmid, S. (2015). Distributed cloud computing: Applications, status quo, and challenges. *Computer Communication Review*, 45(2), 38-43. <https://doi.org/10.1145/2766330.2766337>
- Dave, S., & Maheta, P. (2014). Utilizing round robin concept for load balancing algorithm at virtual machine level in cloud environment. *International Journal of Computer Applications*, 94(4), 23-29. <https://doi.org/10.5120/16332-5612>
- Devi, D. C., & Uthariaraj, V. R. (2016). Load balancing in cloud computing environment using improved weighted round robin algorithm for nonpreemptive dependent tasks. *Scientific World Journal*, 2016, Article 3896065. <https://doi.org/10.1155/2016/3896065>
- Fatima, S. G., Fatima, S. K., Sattar, S. A., Khan, N. A., & Adil, S. (2019). Cloud computing and load balancing. *International Journal of Advanced Research in Engineering and Technology*, 10(2), 189-209. <https://doi.org/10.34218/IJARET.10.2.2019.019>
- Goyal, T., Singh, A., & Agrawa, A. (2012). Cloudsim: Simulator for cloud computing infrastructure and modeling. *Procedia Engineering*, 38, 3566-3572. <https://doi.org/10.1016/j.proeng.2012.06.412>
- Javadpour, A., Sangaiah, A. K., Pinto, P., Ja'fari, F., Zhang, W., Abadi, A. M. H., & Ahmadi, H. R. (2023). An energy-optimized embedded load balancing using DVFS computing in cloud data centers. *Computer Communications*, 197, 255-266. <https://doi.org/10.1016/j.comcom.2022.10.019>
- Kapoor, S., & Dabas, C. (2015). Cluster based load balancing in cloud computing. In *2015 8th International Conference on Contemporary Computing, IC3 2015* (pp. 76-81). IEEE Publishing. <https://doi.org/10.1109/IC3.2015.7346656>
- Kruekaew, B., & Kimpan, W. (2020). Enhancing of artificial bee colony algorithm for virtual machine scheduling and load balancing problem in cloud computing. *International Journal of Computational Intelligence Systems*, 13(1), 496-510. <https://doi.org/10.2991/ijcis.d.200410.002>
- Kruekaew, B., & Kimpan, W. (2022). Multi-objective task scheduling optimization for load balancing in cloud computing environment using hybrid artificial bee colony algorithm with reinforcement learning. *IEEE Access*, 10, 17803-17818. <https://doi.org/10.1109/ACCESS.2022.3149955>
- Kumar, K. P., Ragunathan, T., Vasumathi, D., & Prasad, P. K. (2020). An efficient load balancing technique based on cuckoo search and firefly algorithm in cloud. *International Journal of Intelligent Engineering and Systems*, 13(3), 422-432. <https://doi.org/10.22266/IJIES2020.0630.38>
- Kumar, P., & Kumar, R. (2019). Issues and challenges of load balancing techniques in cloud computing: A survey. *ACM Computing Surveys*, 51(6), Article 120. <https://doi.org/10.1145/3281010>
- Li, G., & Wu, Z. (2019). Ant colony optimization task scheduling algorithm for SWIM based on load balancing. *Future Internet*, 11(4), Article 90. <https://doi.org/10.3390/fi11040090>
- Lu, Y., Zhang, J., Wu, S., Zhang, S., Zhang, Y., Li, Y., Ghosh, S., Banerjee, C., Kulkarni, A. K., Annappa, B., Domanal, S. G., Reddy, G. R. M., Komarasamy, D., & Muthuswamy, V. (2016). Load balancing in cloud environment using a novel hybrid scheduling algorithm. In *2015 IEEE International Conference on Cloud*

- Computing in Emerging Markets, CCEM 2015* (pp. 37-42). IEEE Publishing. <https://doi.org/10.1109/CCEM.2015.31>
- Mishra, S. K., Sahoo, B., & Parida, P. P. (2018). Load balancing in cloud computing: A big picture. *Journal of King Saud University - Computer and Information Sciences*, 32(2), 149-158. <https://doi.org/10.1016/j.jksuci.2018.01.003>
- Mohanty, S., Patra, P. K., Ray, M., & Mohapatra, S. (2017). A novel meta-heuristic approach for load balancing in cloud computing. *International Journal of Knowledge-Based Organizations*, 8(1), 29-49. <https://doi.org/10.4018/ijkbo.2018010103>
- Nerkar, M. H. (2012). Cloud computing in distributed system. *International Journal of Computer Science and Informatics*, 1(10), 97-101. <https://doi.org/10.47893/ijcsi.2012.1072>
- Paduraru, C. I. (2014). A greedy algorithm for load balancing jobs with deadlines in a distributed network. *International Journal of Advanced Computer Science and Applications*, 5(2), 56-59. <https://doi.org/10.14569/ijacsa.2014.050209>
- Ramezani, F., Lu, J., & Hussain, F. K. (2014). Task-based system load balancing in cloud computing using particle swarm optimization. *International Journal of Parallel Programming*, 42(5), 739-754. <https://doi.org/10.1007/s10766-013-0275-4>
- Saura, J. R., Herraiez, B. R., & Reyes-Menendez, A. (2019). Comparing a traditional approach for financial brand communication analysis with a big data analytics technique. *IEEE Access*, 7, 37100-37108. <https://doi.org/10.1109/ACCESS.2019.2905301>
- Singh, H., Tyagi, S., & Kumar, P. (2021). Cloud resource mapping through crow search inspired metaheuristic load balancing technique. *Computers and Electrical Engineering*, 93, Article 107221. <https://doi.org/10.1016/j.compeleceng.2021.107221>
- Sinha, G., & Sinha, D. (2020). Enhanced weighted round robin algorithm to balance the load for effective utilization of resource in cloud environment. *EAI Endorsed Transactions on Cloud Systems*, 6(18), Article 166284. <https://doi.org/10.4108/eai.7-9-2020.166284>
- Sinha, U., & Shekhar, M. (2015). Comparison of various cloud simulation tools available in cloud computing. *International Journal of Advanced Research in Computer and Communication Engineering*, 4(3), 171-176. <https://doi.org/10.17148/ijarccc.2015.4342>
- Tawfeek, M. A., El-Sisi, A., Keshk, A. E., & Torkey, F. A. (2013). Cloud task scheduling based on ant colony optimization. In *2013 8th International Conference on Computer Engineering & Systems (ICCES)* (pp. 64-69). IEEE Publishing. <https://doi.org/10.1109/ICCES.2013.6707172>
- Wang, Y. H., & Wu, I. C. (2009). Achieving high and consistent rendering performance of java AWT/Swing on multiple platforms. *Software - Practice and Experience*, 39(7), 701-736. <https://doi.org/10.1002/spe>





*Review Article*

## Chromatography and Spectroscopy Methods for the Analysis of Nicotine and Other Chemical Ingredients in E-Liquid Formulation: A Review

Mohd Rashidi Abdull Manap<sup>1\*</sup>, Noor Hazfalinda Hamzah<sup>2</sup>, Qhurratul Aina Kholili<sup>1</sup>, Fatin Abu Hasan<sup>1</sup> and Azhana Alhumaira<sup>1</sup>

<sup>1</sup>Department of Chemistry, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor, Malaysia

### ABSTRACT

Spectroscopy and chromatography methods have become the most preferred techniques for detecting ingredients in e-liquid analysis. Both methods are widely used for separating and quantifying volatile compounds in a sample, providing individual chemical information in complex mixtures. This paper aims to review the detection and quantification of nicotine and other chemical ingredients in e-liquid used in e-cigarettes. E-cigarettes use the evaporation-condensation principle of aerosolization to produce an inhaled vapor containing nicotine, excipients, and flavoring agents. This review covers sample preparation, identification, and quantification of nicotine and other ingredients using chromatography and spectroscopy analysis. The spectroscopy methods are useful for quickly identifying and quantifying volatile compounds, including propylene glycol (PG), vegetable glycol (VG), and nicotine, while spectroscopic methods, particularly the Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) method, have lower analytical performance compared to chromatography methods in detecting nicotine and other chemical ingredients. Based on

the review, chromatographic methods are the primary option for analyzing nicotine in all e-liquid samples, offering a better alternative for a future study with the presence of multiple detectors.

**Keywords:** Aerosolization, chromatography, e-cigarettes, e-liquid, ingredients, nicotine, quantification, spectroscopy

### ARTICLE INFO

*Article history:*

Received: 31 December 2022

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.08>

*E-mail addresses:*

rashidichemistry@upm.edu.my (Mohd Rashidi Abdull Manap)

raviera@yahoo.com (Noor Hazfalinda Hamzah)

210208@student.upm.edu.my (Qhurratul Aina Kholili)

209995@student.upm.edu.my (Fatin Abu Hasan)

200672@student.upm.edu.my (Azhana Alhumaira)

\* Corresponding author

## INTRODUCTION

Nicotine is an organic compound that belongs to the pyridine alkaloids. According to IUPAC nomenclature, nicotine is known as 3-(1-methyl-2-pyrrolidinyl) pyridine, with the chemical formula  $C_{10}H_{14}N_2$ . Nicotine is a naturally produced dinitrogen alkaloid present in dried leaves of Solanaceae, specifically the tobacco plant, with concentrations as high as 3% (*Nicotiana tabacum*) (Kimbrough, 2019). The molecular structure of nicotine and other common materials in e-liquid are shown in Figure 1.

Nicotine exists in a variety of forms, including freebase, non-protonated Form 1, and protonated Forms 2 and 3 in solution. Typically, nicotine appears as a colorless or pale-yellow liquid that is very hygroscopic and oily, with an uncomfortable strong odor. It is the main addictive component in tobacco products and e-liquids.

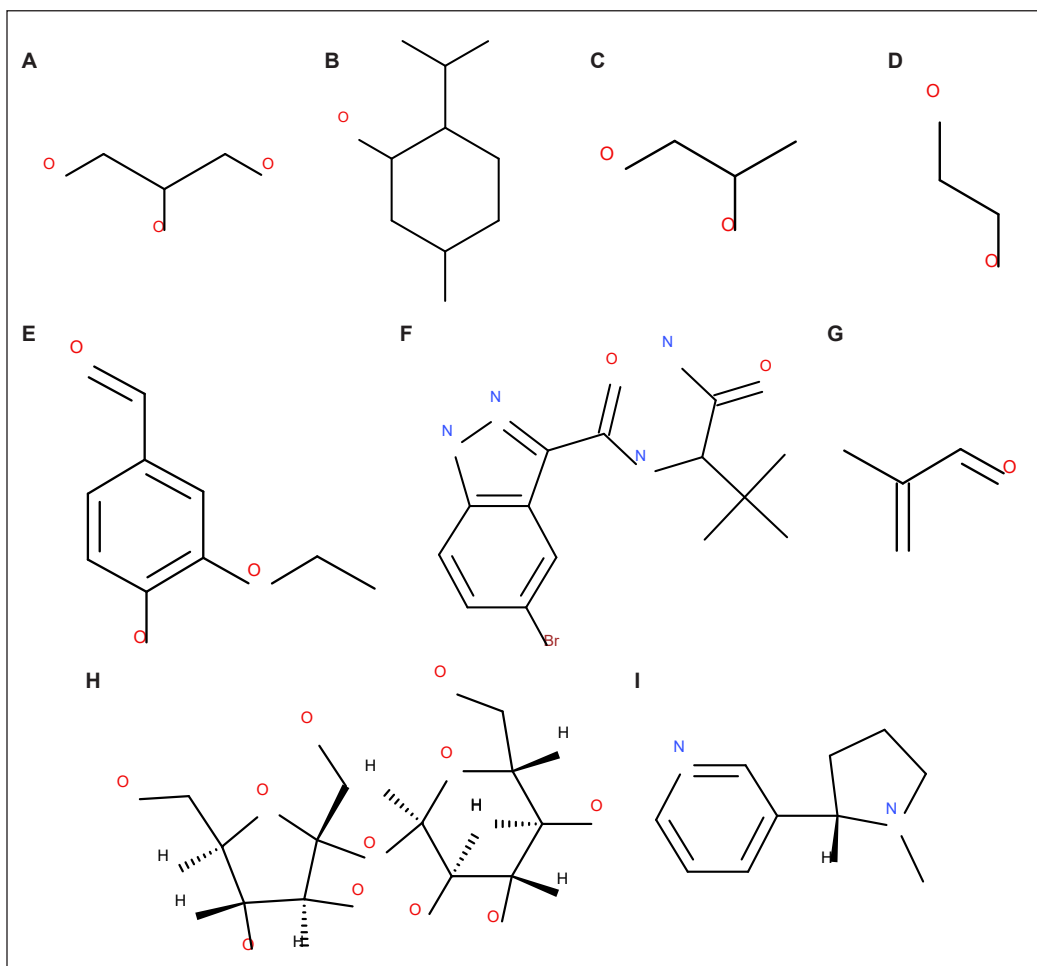


Figure 1. Molecular structure: (A) Glycerol; (B) Menthol; (C) Propylene glycol; (D) Ethylene glycol; (E) Ethyl vanillin; (F) N-(1-amino-3,3-dimethyl-1-oxobutan-2-yl)-5-bromo-1H-indazole-3-carboxamide (ADB-BRINACA); (G) methacrylaldehyde; (H) sucrose; and (I) nicotine

E-liquids in electronic cigarettes (e-cigarettes) come in various packaging types and formulations. Generally, the main ingredients found in e-liquid formulations are vegetable glycerin, propylene glycol, flavoring liquids, as well as nicotine (Cao et al., 2021; Heldt et al., 2021). With the new generation of e-cigarettes, users can customize e-liquids based on their preferences (Vivarelli et al., 2022). Keeping up with the new generation of e-liquid formulations is a significant challenge for analysts because nicotine may be difficult to identify and detect. Moreover, the use of prohibited chemicals in e-liquid formulations raises significant concerns for consumers because currently, there are no restrictions or laws for users to customize their e-liquids, and electronic cigarettes and their e-liquid formulations are virtually unregulated in some countries (Giroud et al., 2015).

Nicotine is a central nervous system (CNS) stimulant that increases alertness and can reduce appetite. Consumption of nicotine can result in different sensations compared to other drugs like amphetamines, heroin, cocaine, and crack, as the main symptoms may include dizziness, nausea, or vomiting. Although the harmful effects of nicotine may not be immediately observable, nicotine and additives are toxic chemicals that should be controlled and monitored, as they are classified as reproductive or developmental toxicants (RDT) and addictive (AD) (Patel et al., 2021).

Table 1 compares the legislative status of nicotine in e-liquids between Southeast Asia and other countries. The Nicotine Concentration in Vaping Products Regulations is a Canadian regulation that sets maximum nicotine concentration limits for vaping products sold in Canada (Canada.ca, 2021). According to this regulation, the maximum allowable nicotine concentration for vaping products sold in Canada is 20 mg/ml, irrespective of whether they are locally produced or imported. UAE.S Standard (2019) is a standard for electronic cigarettes and e-liquids issued by the Emirates Authority for Standardization and Metrology (ESMA) in the United Arab Emirates (UAE). The nicotine content in the electronic liquid must be less than or equal to 20 mg/ml as regulated by this standard. In New Zealand, under the Smokefree Environments and Regulated Products (Vaping) Amendment Act 2020, the maximum allowable nicotine concentration in e-liquids is also

Table 1  
*Legislation of nicotine concentration level in e-liquids in different countries*

Country	*(mg/ml)	Reference
Canada	less than 20	Canada.ca, 2021
Saudi Arabia	less or equal to 20	UAE.S Standard, 2019
New Zealand	less than 20	New Zealand Legislation, 2020
Philippines	NA	NA
Singapore	NA	NA
Malaysia	NA	NA

\*Maximum nicotine concentration level in e-liquid. NA indicates not available

20 mg/ml. Therefore, any e-liquid sold or supplied in New Zealand must not contain more than 20 mg of nicotine per milliliter of liquid.

The sale, importation, and distribution of e-liquids containing nicotine are prohibited in Singapore, the Philippines, and Malaysia until March 2023, resulting in the absence of regulated nicotine concentrations in e-liquids. It is important to note that the regulation of e-cigarettes and vaping products in Southeast Asian countries is still evolving, and there may be updates to the guidelines in the future. Therefore, individuals and businesses involved in the vaping industry are advised to stay updated on the latest regulations and guidelines from their respective countries' authorities.

In Malaysia, nicotine is only covered by the main law, the Malaysian Poison Act 1952. Only authorized persons, such as registered doctors and pharmacists, can handle nicotine and nicotine-containing products. However, nicotine in tobacco is exempted from this classification. To fully address the issue of legalizing e-liquid in Malaysia, it is essential to go beyond simply excluding e-liquid from the Poisons Act 1952. Proper regulation is also needed to ensure that sales are restricted to those over 18, sales and marketing are properly regulated, and the ingredients are closely monitored.

After April 2023, e-liquid is legal, but there are no regulations on the ingredients and prohibited compounds in Malaysia. Implementing regulations is crucial to controlling nicotine concentration and hazardous flavorings containing diacetyl. Failure to do so could result in unsafe nicotine concentration levels, as demonstrated by the 20mg/ml limit enforced in Canada, Saudi Arabia, and New Zealand. A flat tax rate per ml should be established to avoid the difficulties associated with determining nicotine concentration for taxation purposes. It would discourage misdeclaration and ensure a fair and easy taxation system.

Gas chromatography (GC) and liquid chromatography (LC) have been commonly used for determining nicotine levels in e-liquids in the presence of different detectors. However, new methods have also been proposed, including spectroscopic techniques like nuclear magnetic resonance (NMR), infrared (IR) spectroscopy, and Raman spectroscopy. While GC has traditionally been the preferred method, the spectroscopy techniques offer numerous advantages as a qualitative method.

The recommended method of ISO 20714 (2019) for determining nicotine in e-liquids, according to the International Organization for Standardization (ISO), is classified as a quantitative method.

This review aims to identify and quantify the presence of nicotine in commercially available e-liquid formulations used in e-cigarettes using several analytical methods. Consequently, determining the nicotine concentration may contribute to establishing any requirements regarding national regulations.

## METHODOLOGY

Scientific articles were identified and selected from several sites such as ScienceDirect and PubMed until December 2022 using the keywords: “e-cigarette,” “electronic cigarette device,” “vape,” “e-liquids,” “toxic compounds,” “nicotine,” “toxic,” “quantification,” “identification,” “chromatography,” and “spectroscopy.” A total of 31 references were used, including papers related to the scientific research with specific criteria set up to decide each paper. The criteria required must be (1) an original science-related research article, (2) an original science-related review article, or (3) related to the subject of chemistry. All the scientific articles were carefully chosen to match the main purposes of the review.

## RESULTS AND DISCUSSION

### Sample Preparation

“E-liquids” are typically sold as viscous liquid solutions in bottles in many countries (Almazrouei et al., 2022; Augustini et al., 2021; Ketonen & Malik, 2020; Laestadius et al., 2019; Patel et al., 2021). They come in a variety of complex formulations that include a mixture of hundreds of volatile and non-volatile compounds (Eddingsaas et al., 2018; Li et al., 2021; Patel et al., 2021). They also come in two types: flavored e-liquids and non-flavored e-liquids (Augustini et al., 2021). Prior to chromatography analysis, e-liquid samples need to be removed from their respective bottle packaging (Almazrouei et al., 2022). Before gas chromatography analysis, sample preparation, such as adding an internal standard (IS), dissolution, dilution, homogenization, and pre-concentration. Adding an internal standard to the sample is sometimes required for specific purposes, such as data analysis (Patel et al., 2021; Qin et al., 2022). Common GC solvents for the dissolution of e-liquids include methanol, dichloromethane, and hexane. The common sample-to-solvent ratio for dilution is 1:4, 1:9, and 1:10 (Almazrouei et al., 2022; Augustini et al., 2021; Cowan et al., 2022). Recently, an innovative method, such as QUECHERS extraction, has been applied, which can successfully reduce the matrix effect in sample formulations (Almazrouei et al., 2022). These laboratory techniques may result in simple or complicated chromatograms or time-consuming calibrations. The chromatogram will help to understand the formulation of e-liquids, whereas incompatible sample preparation processes could reduce the identification and quantification of nicotine and other ingredients in e-liquids.

### Current Analytical Methods for e-liquid Samples

**Gas Chromatography-Mass Spectrometry (GC-MS).** GC-MS is widely recognized as the gold standard technique for analyzing e-liquids due to its ability to provide highly specific spectral data on individual compounds in a complex mixture while requiring minimal sample separation. When dealing with viscous liquid solutions, sample preparation

typically involves a series of steps, including sample removal, dissolution, extraction, and clean-up. However, the derivatization of the sample is not reported during the GC analysis. Each step in the pre-treatment of the sample before analysis is crucial for enhancing the analytes' recovery and the method's sensitivity. However, as the number of pre-treatment steps increases, the method becomes more complex, and the analysis time also increases.

The analysis of e-liquids has long benefited from the Gas Chromatography (GC) method. GC can provide an easy way to analyze qualitatively and quantitatively, making it a highly advanced method for nicotine analysis. In identifying the volatile compounds in e-liquids, gas chromatography coupled with mass spectrometry (GC-MS) is one of the most widely used analytical methods in understanding the complexity and the mass-to-charge ratio of molecules in the formulations. Gas chromatography with flame-ionization detection (GC-FID) method is rarely used in the study of e-liquids except in the 2018 study, which was used to quantify levels of major components in e-liquids (Dai et al., 2018).

GC characterizes chemical compounds in e-liquids due to their volatile properties. Although e-liquids are a complex formulation, GC can quickly separate many compounds. For example, nicotine was identified after 10 minutes of injection into the GC column and other compounds such as PG and glycerol (Almazrouei et al., 2022). DB-5MS capillary column and HP-5MI Ultra Inert column are frequently used in GC methods to analyze e-liquids (Almazrouei et al., 2022; Cowan et al., 2022). For 2D chromatography, Rxi®-624Sil MS column and Stabilwax® column were used (Patel et al., 2021). These non-polar columns separate many volatile compounds in e-liquids, including nicotine.

The GC-GC-TOF method used headspace analysis to identify an extensive range of toxic compounds in e-liquids of six common flavors obtained online in Australia (Patel et al., 2021). Nicotine was one of the 1064 volatile compounds identified in all tested samples in a 2021 study. Similarly, headspace analysis of blood orange-flavored e-liquids using GC-IMS and GC-MS in Germany and Poland also found nicotine among the 37 volatile compounds identified (Augustini et al., 2021). GC-IMS used a retention index and was validated using GC-MS data. Figure 2 displays an example of an e-liquid sample chromatogram demonstrating the separation of multiple components at different concentration ratios. This particular e-liquid sample contains drugs of abuse, such as tetrahydrocannabinol (THC) and cannabidiol (CBD), alongside nicotine, PG, and glycerol (Almazrouei et al., 2022).

A powerful MS method with chromatographic capability is essential for targeted screening. In a study of 27 e-liquid samples for screening humectant and active ingredients, a Direct Analysis in Real Time™ ionization source coupled to a JEOL JMS-T100LC AccuTOF™ mass spectrometer (JEOL USA, Inc., Peabody, MA) (DART-MS) was used (Peace et al., 2016). For example, the DART-MS spectrum of the Cherry flavor e-liquid sample contains the protonated molecular ion of nicotine, glycol, and several flavor additives within 5 mmu, as shown in Figure 3.

However, several low-intensity peaks that should correspond to components of the formulation were left unidentified using this method. Despite its efficiency, rapidity, and lack of sample preparation requirements, this targeted screening method can be expensive due to the need for reference standards, high temperatures, and a large helium flow rate (2 L/min).

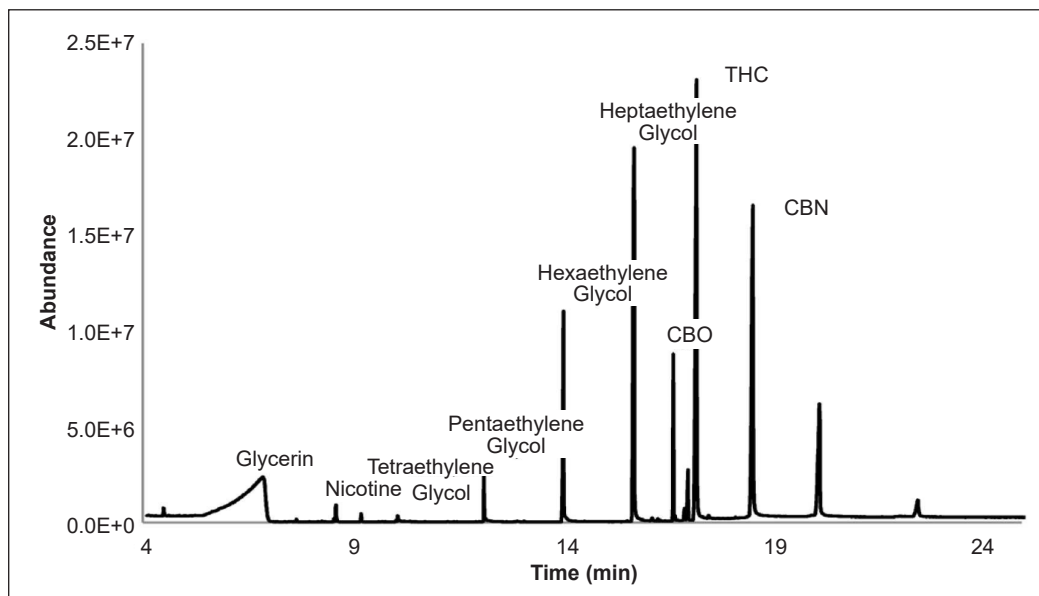


Figure 2. An example of an e-liquid sample chromatogram by GC-MS (Almazrouei et al., 2022)

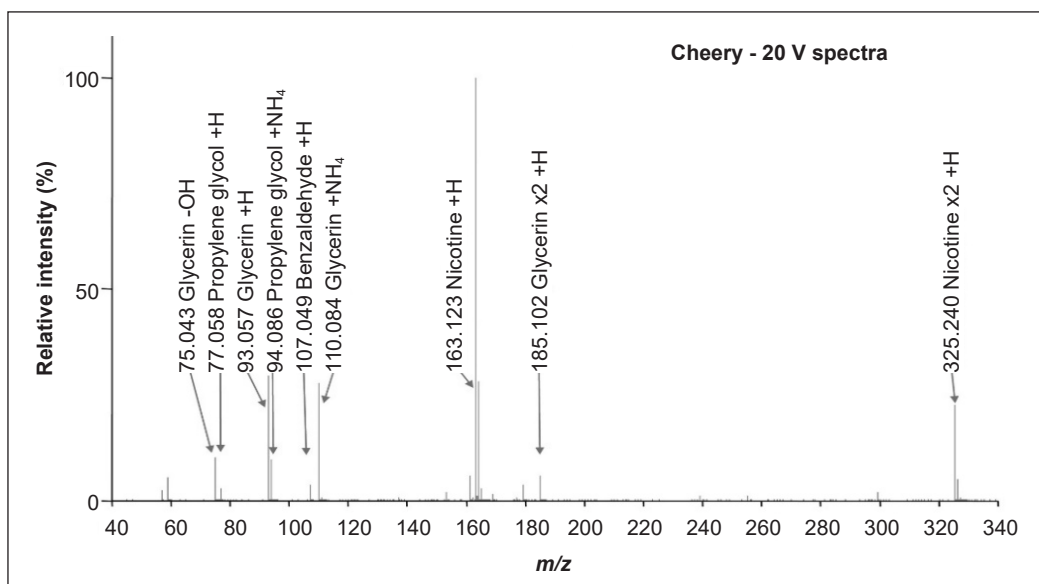


Figure 3. DART-MS 20 V spectra of the e-cigarette liquid formulations named Cheery (Peace et al., 2016)



**High-Performance Liquid Chromatography (HPLC).** The current discussion focuses on the sample preparation, screening, and identification of nicotine in e-liquids using high-performance liquid chromatography (HPLC). The discussion evaluates the capabilities of several liquid chromatography methods coupled with different types of detectors.

In the HPLC method, sample treatment is required before injecting the sample into the chromatographic system. The liquid samples need to be diluted 10,000 fold with ACN:H<sub>2</sub>O (1:9, v/v) before evaluating the concentration of nicotine (Aszyk et al., 2018). After injecting a 5 µL sample into liquid chromatography-tandem mass spectrometry with electrospray ionization (HPLC-ESI-MS/MS), nicotine and 42 other compounds were detected. In addition to dilution, the samples need to be prepared under acidic and basic conditions before separating the components in the e-liquid samples obtained in Poland (Aszyk et al., 2018).

Since the first detection of nicotine in 2015, the compositions of e-liquid formulations have changed over time. Liquid chromatography coupled to diode-array detection (LC-DA) has proven efficient for detecting nicotine concentrations in various nicotine-labeled and unlabeled e-liquid products (Davis et al., 2015). A single wavelength of detector at 260 nm is typically used for nicotine analysis in e-liquids. Similarly, after undergoing the electromembrane extraction (EME) process, nicotine was detected in aerosol matrices from e-cigarettes at a similar wavelength (259 nm). In contrast, a recent research study 2019 used liquid chromatography coupled to photodiode-array detection (LC-PDA) to detect nicotine and multiple nicotine-related alkaloids (Palazzolo et al., 2019). The average nicotine concentration ranged from 18 to 25 mg/mL, and multiple nicotine-related alkaloids were also detected using LC-PDA. The UV wavelengths of the PDA detector were set between 230 nm and 300 nm to detect multiple nicotine-related alkaloids. In addition to the HPLC-DA and PDA methods, a 3200 Q Trap (Applied Biosystems, Foster City, CA) attached to an SCL HPLC system (Shimadzu, Columbia, MD) high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS-MS) system was used to confirm and quantify nicotine in 27 e-liquids (Peace et al., 2016). HPLC methods have many advantages, including high specificity, high separation of nicotine and other components, and the availability of multiple detectors on the market. However, it is less likely to be chosen as it is only available in well-known university laboratories.

LC-DA proved efficient for detecting nicotine concentration in various nicotine-labeled and unlabeled e-liquid products (Davis et al., 2015). A single wavelength detector at 260 nm is normally utilized in nicotine analysis in e-liquids. On the other hand, nicotine was detected in matrices of aerosols from e-cigarettes at a similar wavelength (259 nm) after undergoing the electromembrane extraction (EME) process. In contrast, a recent research study 2019 used liquid chromatography coupled to photodiode-array detection (LC-PDA) to detect nicotine and multiple nicotine-related alkaloids (Palazzolo et al., 2019). The



average concentration of nicotine was found in a range between 18 to 25 mg/mL. Nicotine combined with multiple nicotine-related alkaloids was also detected using LC-PDA. The UV wavelengths of the PDA detector were set between 230 nm and 300 nm to detect multiple nicotine-related alkaloids.

Besides HPLC-DA and PDA method, a 3200 Q Trap (Applied Biosystems, Foster City, CA) attached to an SCL HPLC system (Shimadzu, Columbia, MD) high-performance liquid chromatography–tandem mass spectrometry (HPLC–MS-MS) system was used for the confirmation and quantitation of nicotine in 27 e-liquids (Peace et al., 2016).

HPLC method has many advantages, including high specificity, high separation of nicotine and other components, and availability of multiple detectors in the market, but it is less likely to be chosen as it is only available in well-known university laboratories.

**Nuclear Magnetic Resonance (NMR) Spectroscopy.** NMR is one of the spectroscopy methods used to identify nicotine and other compounds, including those not mentioned in the formulation. In the absence of nicotine in the formulation, NMR was applied for the analysis of e-liquids containing synthetic cannabinoids with different types of magnetic strength: high field (HF) and low field (LF) (Wu et al., 2021). HF method is useful compared to LF because the complexity of  $^1\text{H}$  NMR signals can be enhanced, and using  $^{19}\text{F}$  NMR can confirm the presence of fluorine in cannabinoids.

The NMR method can also be implemented to observe e-liquid signals containing nicotine. As seen in Figure 4, the aromatic proton signal of the e-liquids (for instance, from nicotine and ethyl vanillin) can be observed at low field region between  $\delta$  6 and 10 ppm, while the mid-field region contains signals associated with solvents such as glycerol, propylene glycol, and ethylene glycol (Hahn et al., 2014). Despite distinctive H signals at the low field region and using a 400 MHz spectrometer,  $^1\text{H}$  NMR is unsuitable for quantifying nicotine because they showed strong overlap signals with other matrix compounds at the middle to high field region.

**Infrared (IR) Spectroscopy.** IR spectroscopy has several advantages, including its simplicity, which allows non-spectroscopists to perform the analysis. Additionally, vibrational spectroscopy methods provide rapid analysis, with results obtained within seconds, and it is non-destructive, meaning that it does not require sample preparation or the use of chemical

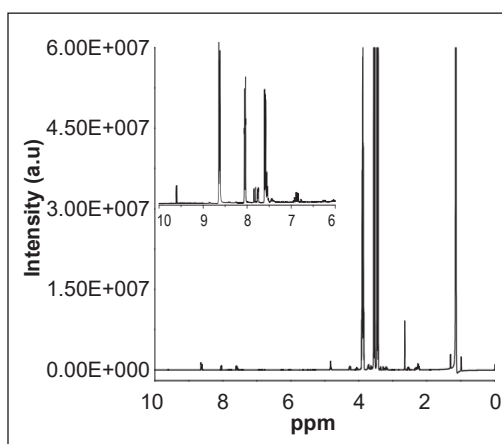
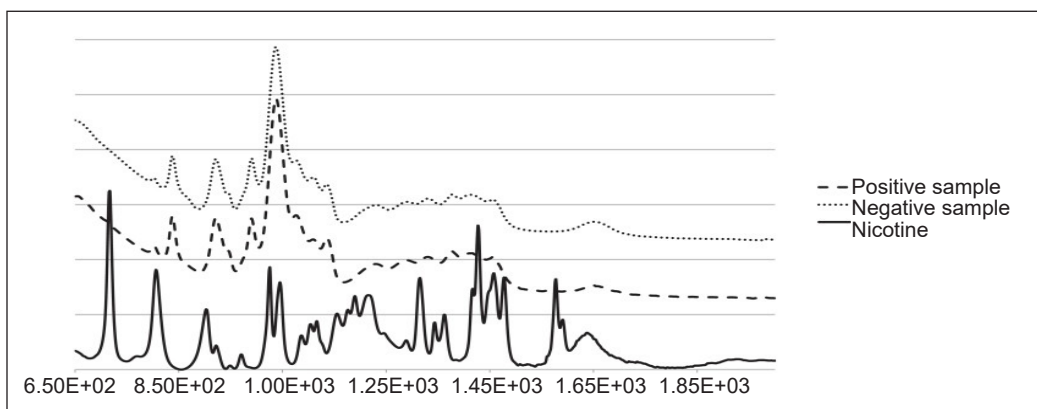
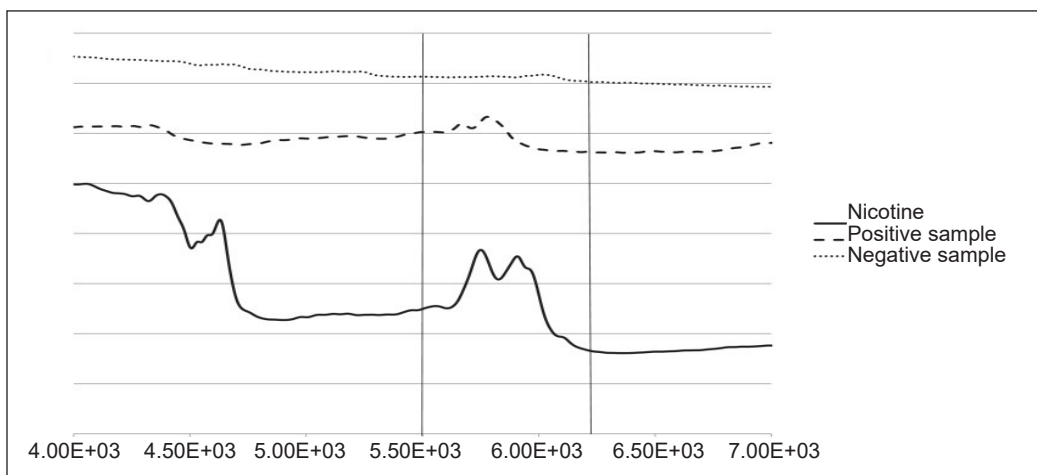


Figure 4.  $^1\text{H}$  NMR spectrum of the aqueous sample of a typical e-cigarette liquid sample (Hahn et al., 2014)

reagents. Before acquiring the IR spectra, e-liquid samples must be taken out of their plastic containers, as the IR source cannot penetrate the sample in the plastic. In a large study of e-liquid samples (n=68, containing both nicotine and non-nicotine samples), the Infrared spectroscopy (IR) method was used to detect nicotine concentration, which was acquired from Belgium. The study used a Nicolet iS10 FT-IR (ThermoFisher Scientific, Waltham, USA) spectrometer equipped with a Smart iTR accessory and a deuterated triglycine sulfate (DTGS) detector (Deconinck et al., 2016). As seen in Figure 5, Attenuated Total Reflectance-Fourier Transform Infrared Spectroscopy (ATR-FTIR) and the Near Infrared Spectroscopy (NIR) method were used to observe the difference between the spectrum of the samples and the spectrum of pure nicotine (reference standard). However, it was impossible to identify their spectral differences. Using a low-resolution ( $4\text{ cm}^{-1}$ ) spectrometer in this experiment hindered the vibrational assignment for several weak IR



(a)



(b)

Figure 5. (a) ATR-IR spectra obtained for nicotine, a positive and a negative sample; and (b) NIR spectra obtained for nicotine, a positive and negative (Deconinck et al., 2016).

peaks below the spectral region of  $1000\text{ cm}^{-1}$ . However, many strong nicotine signals were observed within the  $2000\text{--}650\text{ cm}^{-1}$  spectral region. With the use of NIR, only limited peaks of nicotine were observed.

In addition to e-liquids, a study was conducted on a small-scale sample of the aerosol phase of e-liquid using the FTIR method, which successfully provided an emission profile of volatile organic compounds (VOCs). However, this method is unable to detect nicotine as it only detects smaller molecules such as  $\text{CO}_2$ , CO, methane, ethane, ethylene, methanol, ethanol, and formaldehyde emitted from the e-cigarette device analyzed with FTIR (Ooi et al., 2019). Traditionally, cigarette smoke has been studied using a high-resolution spectrometer to detect stretching frequencies of a variety of gaseous components, including hydrocarbons, nitrogen, and carbon oxides (Bacsik et al., 2007).

### Raman Spectroscopy and Surface-Enhanced Raman Spectroscopy (SERS).

The Raman spectroscopy method can also be used for the identification of nicotine. This method requires little to no sample preparation for analysis and is known for its non-invasive approach (Poulsen et al., 2022). As depicted in Figure 6, the Raman spectrum of nicotine polacrilex resin (NPR) from the gum ingredients displays unique nicotine-specific peaks at  $1033$  and  $1052\text{ cm}^{-1}$  (Poulsen et al., 2022).

Apart from the traditional Raman spectroscopy method, the modern method of Surface-Enhanced Raman Spectroscopy (SERS) is also used for the identification of nicotine along with other related compounds. Similarly, two main peaks of nicotine in an aqueous solution can be distinctly observed at  $1032$  and  $1052\text{ cm}^{-1}$ , as shown in Figure 7 (Mamián-López & Poppi, 2013). Anabasine and cotinine also have the same spectral features. The overlapping peaks between the same functional groups are expected to hinder the identification of nicotine in the presence of anabasine and cotinine.

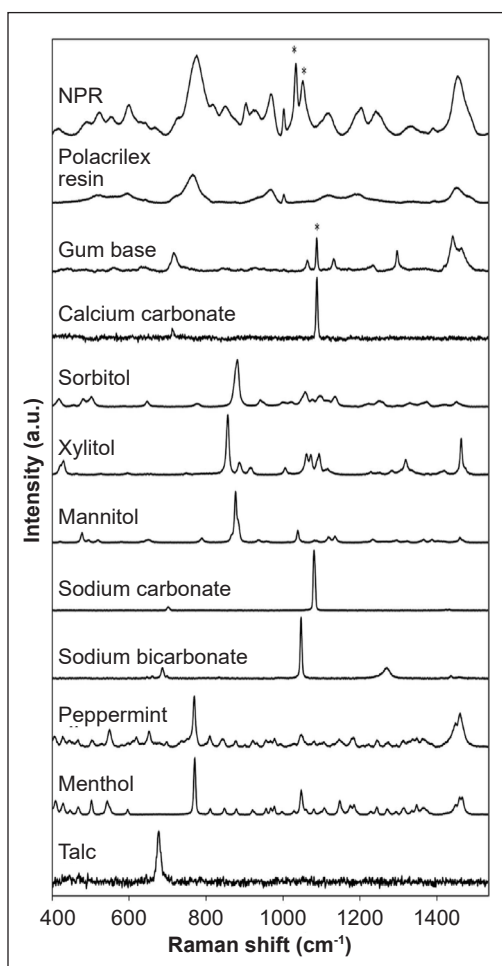


Figure 6. Raman spectra of the individual gum ingredients (Poulsen et al., 2022)

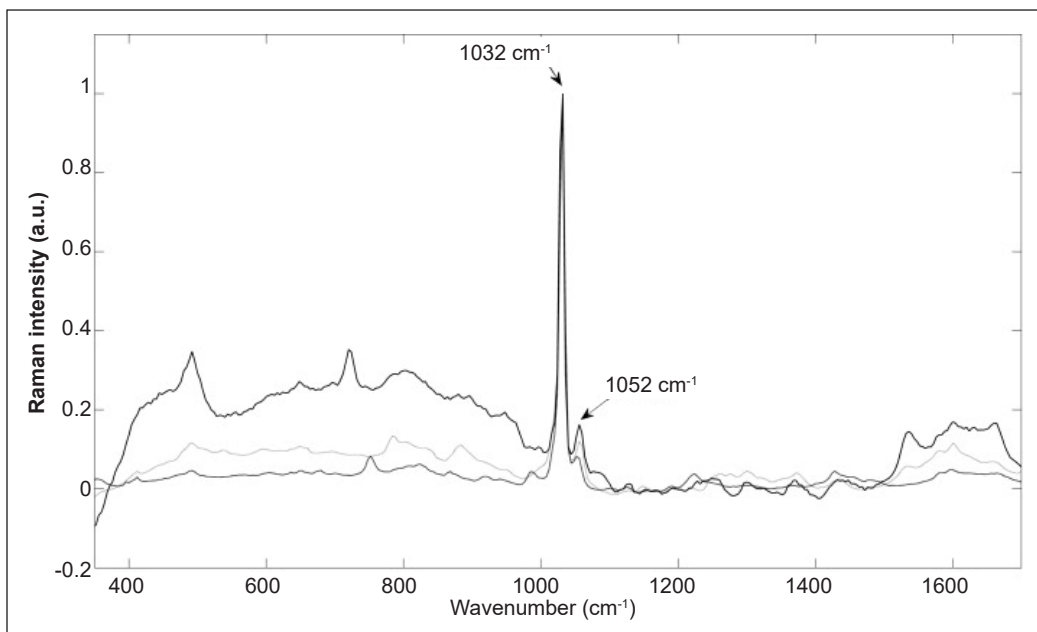


Figure 7. SERS spectra (normalized) of nicotine (solid line), anabasine (dotted line) and cotinine (dashed line) in aqueous solution (Mamián-López & Poppi, 2013)

## CONCLUSION

In recent years, analytical methods used in analyzing e-liquids can be categorized into chromatographic and spectroscopic methods, commonly used in this field of study. Among chromatographic methods, the GC-MS technique is the standard method for analyzing volatile compounds in e-liquid samples, including PG, VG, and nicotine. Additionally, HPLC is a well-known technique for identifying and quantifying all soluble compounds, including PG, VG, and nicotine, in a short total analysis time. Several spectroscopic methodologies, such as NMR, FTIR, and Raman analysis, are also available for identification and quantification. However, based on the studied articles, chromatographic methods are the primary option for analyzing nicotine in all e-liquid samples. Limited studies have applied the ATR-FTIR method for identification and quantification, especially in e-liquids found in e-cigarettes. This method has lower analytical performance compared to chromatography methods in detecting nicotine. The modernization of spectrometer design will likely involve the development of new complex statistical models.

## ACKNOWLEDGEMENT

This study was supported by the Research Management Centre (RMC) Universiti Putra Malaysia for the Inisiatif Putra Muda (GP-IPM) research grant of GP-IPM/2023/9741500 and Universiti Kebangsaan Malaysia grant GUP-2020-052.

## REFERENCES

- Almazrouci, E. S., Bintamim, A. A., Khalil, S. E. A., Alremeithi, R., & Gewily, S. (2022). The identification of drugs of abuse in e-cigarette samples seized in Dubai between 2016 and 2020. *Forensic Science International*, 333, Article 111233. <https://doi.org/10.1016/J.FORSCIINT.2022.111233>
- Aszyk, J., Kubica, P., Namieśnik, J., Kot-Wasik, A., & Wasik, A. (2018). New approach for e-cigarette aerosol collection by an original automatic aerosol generator utilizing melt-blown non-woven fabric. *Analytica Chimica Acta*, 1038, 67-78. <https://doi.org/https://doi.org/10.1016/J.ACA.2018.08.015>
- Augustini, A. L. R. M., Sielemann, S., & Telgheder, U. (2021). Strategy for the identification of flavor compounds in e-liquids by correlating the analysis of GCxIMS and GC-MS. *Talanta*, 230, Article 122318. <https://doi.org/10.1016/J.TALANTA.2021.122318>
- Bacsik, Z., McGregor, J., & Mink, J. (2007). FTIR analysis of gaseous compounds in the mainstream smoke of regular and light cigarettes. *Food and Chemical Toxicology*, 45(2), 266-271. <https://doi.org/10.1016/J.FCT.2006.08.018>
- Canada.ca. (2021). *Nicotine Concentration in Vaping Products Regulations*. <https://laws.justice.gc.ca/eng/regulations/SOR-2021-123/>
- Cao, Y., Wu, D., Ma, Y., Ma, X., Wang, S., Li, F., Li, M., & Zhang, T. (2021). Toxicity of electronic cigarettes: A general review of the origins, health hazards, and toxicity mechanisms. *Science of The Total Environment*, 772, Article 145475. <https://doi.org/10.1016/J.SCITOTENV.2021.145475>
- Cowan, E. A., Tran, H., Gray, N., Perez, J. J., Watson, C., Blount, B. C., & Valentín-Blasini, L. (2022). A gas chromatography-mass spectrometry method for quantifying squalane and squalene in aerosol emissions of electronic cigarette, or vaping, products. *Talanta*, 238, Article 122985. <https://doi.org/10.1016/J.TALANTA.2021.122985>
- Dai, J., Kim, K. H., Szulejko, J. E., Jo, S. H., Kwon, K., & Choi, D. W. (2018). Quantification of nicotine and major solvents in retail electronic cigarette fluids and vaped aerosols. *Microchemical Journal*, 140, 262-268. <https://doi.org/https://doi.org/https://doi.org/10.1016/j.microc.2018.04.028>
- Davis, B., Dang, M., Kim, J., & Talbot, P. (2015). Nicotine concentrations in electronic cigarette refill and do-it-yourself fluids. *Nicotine & Tobacco Research*, 17(2), Article 134. <https://doi.org/10.1093/NTR/NTU080>
- Deconinck, E., Bothy, J. L., Barhdadi, S., & Courselle, P. (2016). Discriminating nicotine and non-nicotine containing e-liquids using infrared spectroscopy. *Journal of Pharmaceutical and Biomedical Analysis*, 120, 333-341. <https://doi.org/10.1016/J.JPBA.2015.12.054>
- Eddingsaas, N., Pagano, T., Cummings, C., Rahman, I., Robinson, R., & Hensel, E. (2018). Qualitative analysis of e-liquid emissions as a function of flavor additives using two aerosol capture methods. *International Journal of Environmental Research and Public Health*, 15(2), Article 323. <https://doi.org/10.3390/IJERPH15020323>
- Giroud, C., de Cesare, M., Berthet, A., Varlet, V., Concha-Lozano, N., & Favrat, B. (2015). E-cigarettes: A review of new trends in cannabis use. *International Journal of Environmental Research and Public Health*, 12(8), 9988-10008. <https://doi.org/10.3390/IJERPH120809988>

- Hahn, J., Monakhova, Y. B., Hengen, J., Kohl-Himmelseher, M., Schössler, J., Hahn, H., Kuballa, T., & Lachenmeier, D. W. (2014). Electronic cigarettes: Overview of chemical composition and exposure estimation. *Tobacco Induced Diseases*, 12(1), Article 23. <https://doi.org/10.1186/S12971-014-0023-6>
- Heldt, N. A., Reichenbach, N., McGary, H. M., & Persidsky, Y. (2021). Effects of electronic nicotine delivery systems and cigarettes on systemic circulation and blood-brain barrier: Implications for cognitive decline. *The American Journal of Pathology*, 191(2), 243-255. <https://doi.org/10.1016/J.AJPATH.2020.11.007>
- ISO 20714. (2019). *E-Liquid - Determination of Nicotine, Propylene Glycol and Glycerol in Liquids used in Electronic Nicotine Delivery Devices - Gas Chromatographic Method*. <https://www.iso.org/standard/68905.html>
- Ketonen, V., & Malik, A. (2020). Characterizing vaping posts on Instagram by using unsupervised machine learning. *International Journal of Medical Informatics*, 141, Article 104223. <https://doi.org/10.1016/J.IJMEDINF.2020.104223>
- Kimbrough, D. (2019). *Vaping: What you need to know*. American Chemical Society. <https://www.acs.org/education/resources/highschool/chemmatters/past-issues/2019-2020/dec-2019/vaping.html>
- Laestadius, L. I., Wahl, M. M., Pokhrel, P., & Cho, Y. I. (2019). From apple to werewolf: A content analysis of marketing for e-liquids on Instagram. *Addictive Behaviors*, 91, 119-127. <https://doi.org/10.1016/J.ADDBEH.2018.09.008>
- Li, Y., Burns, A. E., Tran, L. N., Abellar, K. A., Poindexter, M., Li, X., Madl, A. K., Pinkerton, K. E., & Nguyen, T. B. (2021). Impact of e-liquid composition, coil temperature, and puff topography on the aerosol chemistry of electronic cigarettes. *Chemical Research in Toxicology*, 34(6), 1640-1654. [https://doi.org/10.1021/ACS.CHEMRESTOX.1C00070/SUPPL\\_FILE/TX1C00070\\_SI\\_001.PDF](https://doi.org/10.1021/ACS.CHEMRESTOX.1C00070/SUPPL_FILE/TX1C00070_SI_001.PDF)
- Mamián-López, M. B., & Poppi, R. J. (2013). Standard addition method applied to the urinary quantification of nicotine in the presence of cotinine and anabasine using surface enhanced Raman spectroscopy and multivariate curve resolution. *Analytica Chimica Acta*, 760, 53-59. <https://doi.org/10.1016/J.ACA.2012.11.023>
- New Zealand Legislation. (2020). *Smokefree Environments and Regulated Products (Vaping) Amendment Act 2020 No 62, Public Act*. <https://www.legislation.govt.nz/act/public/2020/0062/40.0/whole.html>
- Ooi, B. G., Dutta, D., Kazipeta, K., & Chong, N. S. (2019). Influence of the e-cigarette emission profile by the ratio of glycerol to propylene glycol in e-liquid composition. *ACS Omega*, 4(8), 13338-13348. [https://doi.org/10.1021/ACSOMEGA.9B01504/ASSET/IMAGES/LARGE/AO9B01504\\_0004.JPEG](https://doi.org/10.1021/ACSOMEGA.9B01504/ASSET/IMAGES/LARGE/AO9B01504_0004.JPEG)
- Palazzolo, D., Nelson, J. M., & Hudson, Z. (2019). The use of HPLC-PDA in determining nicotine and nicotine-related alkaloids from e-liquids: A comparison of five e-liquid brands purchased locally. *International Journal of Environmental Research and Public Health*, 16(17), Article 3015. <https://doi.org/10.3390/IJERPH16173015>
- Patel, D., Taudte, R. V., Nizio, K., Herok, G., Cranfield, C., & Shimmon, R. (2021). Headspace analysis of e-cigarette fluids using comprehensive two dimensional GC×GC-TOF-MS reveals the presence of volatile and toxic compounds. *Journal of Pharmaceutical and Biomedical Analysis*, 196, Article 113930. <https://doi.org/10.1016/J.JPBA.2021.113930>

- Peace, M. R., Baird, T. R., Smith, N., Wolf, C. E., Poklis, J. L., & Poklis, A. (2016). Concentration of nicotine and glycols in 27 electronic cigarette formulations. *Journal of Analytical Toxicology*, 40(6), 403-407. <https://doi.org/10.1093/JAT/BKW037>
- Poulsen, J., Nielsen, K. A., & Bauer-Brandl, A. (2022). Raman imaging as a powerful tool to elucidate chemical processes in a matrix: Medicated chewing gums with nicotine. *Journal of Pharmaceutical and Biomedical Analysis*, 209, Article 114519. <https://doi.org/10.1016/J.JPBA.2021.114519>
- Qin, J., Dou, Y., Wu, F., Yao, Y., Andersen, H. R., Hélix-Nielsen, C., Lim, S. Y., & Zhang, W. (2022). In-situ formation of Ag<sub>2</sub>O in metal-organic framework for light-driven upcycling of microplastics coupled with hydrogen production. *Applied Catalysis B: Environmental*, 319, Article 121940. <https://doi.org/https://doi.org/10.1016/j.apcatb.2022.121940>
- UAE.S Standard. (2019). *Electronic Nicotine Products (Equivalents of Traditional Tobacco Products)*. Emirates Authority for Standardization & Metrology (ESMA). <https://d3vqfzrx1ccvd.cloudfront.net/uploads/legislation/United%20Arab%20Emirates/United-Arab-Emirates-UAE.S-50302018.pdf>
- Vivarelli, F., Granata, S., Rullo, L., Mussoni, M., Candeletti, S., Romualdi, P., Fimognari, C., Cruz-Chamorro, I., Carrillo-Vico, A., Paolini, M., & Canistro, D. (2022). On the toxicity of e-cigarettes consumption: Focus on pathological cellular mechanisms. *Pharmacological Research*, 182, Article 106315. <https://doi.org/10.1016/J.PHRS.2022.106315>
- Wu, N., Danoun, S., Balayssac, S., Malet-Martino, M., Lamoureux, C., & Gilard, V. (2021). Synthetic cannabinoids in e-liquids: A proton and fluorine NMR analysis from a conventional spectrometer to a compact one. *Forensic Science International*, 324, Article 110813. <https://doi.org/10.1016/J.FORSCIINT.2021.110813>





## The Mechanical Performance of Polymer Concrete Incorporating Waste Tin Fibres

Hosseinkhah Ashkan<sup>1</sup>, Shuhairy Norhisham<sup>1</sup>, Mohd Supian Abu Bakar<sup>2</sup>, Agusril Syamsir<sup>1,2\*</sup>, Mohammed Jalal Abdullah<sup>1</sup>, Muhammad Rizal Muhammad Asyraf<sup>3</sup>, Maiyozi Chairi<sup>4</sup>, Mutiara Yetrina<sup>4</sup>, Jihan Melasari<sup>4</sup> and Muhammad Farid<sup>4</sup>

<sup>1</sup>Department of Civil Engineering, College of Engineering, Universiti Tenaga Nasional, Jalan IKRAM - UNITEN, 43000 Kajang, Selangor, Malaysia

<sup>2</sup>Institute of Energy Infrastructure, Universiti Tenaga Nasional, Jalan IKRAM - UNITEN, 43000 Kajang, Selangor, Malaysia

<sup>3</sup>Centre for Advanced Composite Materials (CACM), Universiti Teknologi Malaysia, 81310 UTM, Johor Bahru, Johor, Malaysia

<sup>4</sup>Faculty of Engineering, Universitas Putra Indonesia YPTK, Padang, Indonesia

### ABSTRACT

Concrete is the most widely used construction material in the world. It is now possible to construct structures out of concrete because this durable compound that consists of water, aggregate, and Portland cement not only gives us many scopes of design but also has a very high compressive strength at a low cost. This paper deals with alternative materials for the most common construction material, cement-based concrete and polymer concrete (PC), containing waste tin fibres. The study covers the fabrication of polymer concrete and the execution of three tests: compressive strength, flexural tensile, and splitting

tensile. Tests were conducted to determine the mechanical properties of the PC, and the results were analysed and evaluated on several PC specimens with different ratios of waste tin fibre. The results showed that using waste tin as fibre reinforcement in PC would substantially enhance the overall mechanical performance. Specifically, the optimum amount of waste tin as reinforcement in PC was 0.16% for compressive and splitting tensile strengths, while 0.20% was the optimum fibre loading for the flexural tensile strength. In this case, a positive outcome

### ARTICLE INFO

#### Article history:

Received: 31 December 2022

Accepted: 20 July 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.09>

#### E-mail addresses:

ashkan\_hosseinkhah@yahoo.de (Hosseinkhah Ashkan)

Shuhairy@uniten.edu.my (Shuhairy Norhisham)

mohdsupian7779@gmail.com (Mohd Supian Abu Bakar)

Agusril@uniten.edu.my (Agusril Syamsir)

eng.mj96@hotmail.com (Mohammed Jalal Abdullah)

muhammadasyraf.mr@utm.my (Muhammad Rizal Muhammad Asyraf)

maiyozi@upiyptk.ac.id (Maiyozi Chairi)

mutiarayetrina@upiyptk.ac.id (Mutiara Yetrina)

jihan\_melasari@upiyptk.ac.id (Jihan Melasari)

farid2500@gmail.com (Muhammad Farid)

\* Corresponding author

was found at a constant resin-to-filler ratio of 40:60 by volume and a matrix-to-aggregate ratio of 1:1.35 by weight.

*Keywords:* Compressive strength, flexural tensile strength, polymer concrete, splitting tensile strength, waste tin

---

## INTRODUCTION

Polymer concrete is a composite material of polymeric resins that serve as binder materials for aggregates and micro fillers, resulting in a hardened composite when catalysts and accelerators are added. A polymer resin must be used for concrete to be poured, trowelled, and subsequently cured. The most common polymer resins used in concrete are polyester, vinyl ester, and ordinary epoxy (Afroughsabet & Ozbakkaloglu, 2015; Ahmadi et al., 2021; Martínez & López et al., 2021; Muda, Kamal et al., 2016; Oyebisi et al., 2020; Supian et al., 2016). The polymer material causes a chemical reaction that leads to hardening and curing. Using different resins leads to different benefits, e.g., acrylic binders set very quickly and thus lead to higher weathering resistance, while epoxies create a very strong material that shrinks very little as it cures. It must be considered that the mixing of polymer concrete should be done very precisely and very thoroughly (Al-Nini et al., 2020; Ismail et al., 2022; Reddy & Santhosha, 2018; Yan et al., 2016; Zabihi et al., 2017). It cannot be mixed beforehand and kept turning (to avoid curing like cement-based concrete). The chemical reaction happens as soon as the mixture begins (Mohamad, 2014; Supian et al., 2018).

Research has shown that polymer concrete, compared to cement-based concrete, is stronger, more durable, and has lower maintenance requirements (Afroughsabet & Ozbakkaloglu, 2015; Kamal et al., 2019; Martínez-López et al., 2021; Muda, Alam et al., 2016; Parikh & Modhera, 2012; Reddy & Santhosha, 2018). Polymer concrete's mechanical strength can reach four to five times higher than cement-based concrete. Furthermore, it has a high resistance against chemicals and corrosive agents, resistance to frost, good abrasion, fast curing times, excellent durability, and water permeability (ACI 318-11, 2011). Therefore, polymer concrete is widely used in civil engineering applications, such as underground pipes, trench lines, overlays of bridges, building and highway repairs, and swimming pools (Ali & Ansari, 2013). Nevertheless, producing polymer concrete requires attention during the casting process and to curing temperature, composition, and selecting resins and additional particles carefully. Furthermore, because PC is about 5–10 times more expensive than regular concrete, its use is now confined to constructions where the increased cost justifies the improved performance.

Polymer concrete is composed of various aspects, including the type and amount of resin and filler, curing process, curing temperature, humidity, and matrix-to-aggregate ratio (Atiqah et al., 2021; Esfahani et al., 2016; Frigione, 2013; Kumlutaş et al., 2003; Mohamad et al., 2022; Sakai et al., 2005; Sosoi et al., 2018; Vaggar et al., 2021). Various studies have

shown that unsaturated polyester (UP) resin is a good mix of mechanical, electrical, and chemical properties that enhance splitting tensile and flexural strength by a maximum of 30% and a minimum of 20% compared to Portland cement concrete. It is also used as a cement substitute in 20%, 25%, and 30% concentrations (Gao et al., 2019; Hashemi et al., 2018; Yeon et al., 2020). Meanwhile, the studies by Bulut and Şahin (2017) and Hameed and Hamza (2019) investigated the mechanical and physical properties of polymer concrete using waste components and natural and river sand, revealing that increasing the amount of polymer resin in the compressive strength test increased its strength. A mixture of 70% ceramic waste and 30% polyester resin gave a maximum strength of 132 MPa, while a mixture of 80% concrete waste and 20% plastic resin gave a minimum strength of 28 MPa. The increasing polymer resin also enhances the concrete's splitting tensile strength and flexural strength.

Furthermore, the study by Sosoi et al. (2018) on the workability of both substitution mixes of fresh concrete improved when the PET quantity was raised and decreased as the sawdust amount was increased. Besides, the results showed that the chipped PET as a replacement had a greater compressive strength value than sawdust blends, with 25-50% sawdust and 50% and 75% chopped PET being the most common results. Polymer concrete with waste substitution developed fractures gradually until they were destroyed. Otherwise, the study by Niaki et al. (2018) has defined basalt fibre-reinforced polymer concrete (BFRPC) and examined the effect of temperature on the mechanical properties of a polymer concrete sample with 2% basalt fibre and one without basalt fibre, which was subjected to compressive, three-point bending, and splitting tensile tests. Based on the results, plain PC has a lower residual strength than basalt fibre polymer concrete at elevated temperatures.

Besides, many studies have conducted various works on waste material as filler material in PC (Hameed & Hamza, 2019; Sonebi et al., 2022), waste materials such as waste glass (Ramadan et al., 2023; Saribiyik et al., 2013), electronic plastic waste (Bulut & Şahin, 2017) and PET bottle (Asdollah-Tabar et al., 2021), which is used as a reinforcement material in polymer composites.

Hence, this research paper presents an innovative investigation of the mechanical performance of polymer concrete using wasted tin fibres. The study intends to analyse the impact of these common fibres on improving the mechanical properties of polymer concrete by investigating their utilisation. Using scrap tin fibres in this context is a distinctive and creative strategy that advances the growing areas of sustainable building materials and waste utilisation. The results of this study will provide important information for potential advances in composite materials, and sustainable engineering practises by providing insight into the feasibility and possible advantages of incorporating scrap tin fibres into polymer concrete. At the end of this research, potential applications and insights into waste tin fibre-reinforced polymer concrete were highlighted and recommended.

## METHODOLOGY

### Materials

Polymer concrete (PC) is a composite that uses liquid resins as a binder instead of the traditional cement-hydrate binders found in ordinary concrete. Resin, hardener, fly ash, aggregates, and scrap tin fibre are all used to make polymer concrete.

**Resin and Hardener.** An unsaturated polyester resin, Reversol P 9509, was combined with PC samples to create a stiff and low-reactive composition. Unsaturated polyester resin is less expensive and more commercially accessible than epoxy resin and has superior mechanical and chemical properties. Meanwhile, the hardener, combined with the resin, creates a chemical reaction that changes from a liquid to a solid state. Therefore, it is also called a catalyst that causes the transformation, referred to as “curing” or “polymerisation.” As shown in Table 1, the chemical hardener used with polyester resin is called Methyl Ethyl Ketone Peroxide ([http://kawamia.com/files/Mepoxe\\_new.pdf](http://kawamia.com/files/Mepoxe_new.pdf)). The specifications and properties of the resin are shown in Table 1.

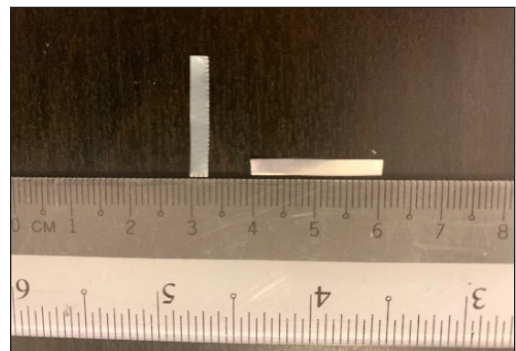
**Fly Ash.** Fly ash is a fine powder that is a by-product of burning pulverised coal in electric power plants (Figure 1a). Mixing fly ash with resin forms a compound similar to Portland cement, which makes fly ash a prime material in PC. Fly ash can be used as a replacement for river sand and as a fine

Table 1  
*Polyester resin properties of Reversol P 9509 (Luxchem, 2018)*

Specifications of Reversol P 9509	
Appearances	Hazy, pinkish
Non-volatile, %	56–59
Viscosity @ 25°C, cps	
- Brookfield, #3/60	450–600
- Brookfield, #3/6	900–1350
Thixotropic index	1.5–2.8
Gel time @ 25°C, mixture	18–23
- 1% MEKP	
Acid value, mgKOH/g	29–34
Solid resin	
Typical properties	
Specific gravity	1.12
Volumetric shrinkage, %	8



(a)



(b)

Figure 1. (a) Sample of fly ash; and (b) Sample of 5 mm aggregates

aggregate material. There are two different classes of fly ash: class C and class F. This study used class C fly ash, which generally has a high calcium ratio.

**Aggregates.** Coarse aggregates were sieved, and only aggregates with a size of 5 mm were separated and used for the mixture (Figure 1b). This size of aggregates guarantees a low void and an equal distribution of the aggregates while leaving enough space for the waste in fibre. In order to achieve good bond strength between the polymer binder and the aggregates, they should be free of dust and debris, clean, and dry.

**Waste Tin Fibre.** Aluminium cans were cut into 20 mm × 3 mm strips and twisted spirally to reinforce and to produce a sustainable product. This size ensures the resin fully covers the tin fibres while well distributed between the aggregate gaps (Figures 2a and 2b).

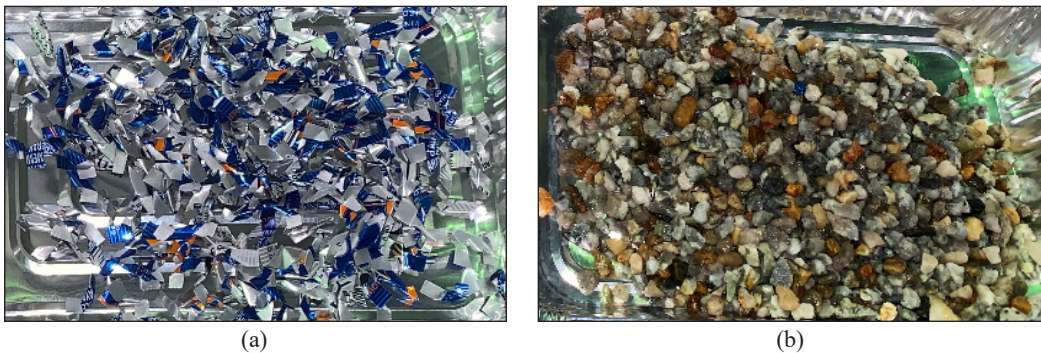


Figure 2. (a) Sample of cut waste tin fibre; and (b) Sample of twisted tin fibre

### Mixture Composition

Based on previous research, mainly referred to by Ferdous et al. (2016), where the effect of resin-to-filler ratio and matrix-to-aggregate ratio on mechanical and durability properties was investigated, the optimum resin-to-filler ratio of 60:40 and the optimum matrix-to-aggregate ratio of 1:1.35 were chosen for the preparation of polymer concrete. These ratios ensure a uniform distribution of aggregates and are highly recommended to balance mechanical performance and cost. This study used a volume ratio 60:40:135 for resin/hardener, fly ash, and coarse aggregates. Additionally, to increase the mechanical properties, waste tin fibre with percentages of 0%, 0.12%, 0.16%, and 0.2% was added to the polymer concrete composite as reinforcement.

Basically, from the volume of the cylindrical and prismatic moulds, the density of the materials, and the ratio given, the mass of each material for the different specimens was calculated. The amount of hardener used depends on the mass of the resin. Finally, waste tin fibre was calculated based on the percentage used and the mould volume. Table 2 shows the calculation for each material used.



Table 2

Total material amount for each sample

Mixes	Tin Fibre ratio (%)	Aggregate (kg/m <sup>3</sup> )	Resin (kg/m <sup>3</sup> )	Filler (kg/m <sup>3</sup> )	Waste Tin Fibre (kg/m <sup>3</sup> )
SP0	0	2929	1095	2006	2700
SP12	0.12	2929	1095	2006	2700
SP16	0.16	2929	1095	2006	2700
SP20	0.2	2929	1095	2006	2700

### Specimen Preparation

Thirty-six specimens were prepared and subjected to three tests To investigate the mechanical properties of polymer concrete containing waste tin fibre: the compressive strength test, the splitting tensile test, and the flexural tensile test. Twelve (12) cylindrical samples with a diameter of 50 mm and a height of 100 mm were cast in PVC pipes (Figure 3). The flexural strength test requires 12 prismatic samples from 25 mm and 250 mm, while the sieving of coarse aggregates with a maximum size of 5 mm was used to create a mixture that guarantees a low void and an equal distribution of the aggregates while leaving enough space for the waste gin fibre. Polymer concrete composite was cut into 3 mm × 20 mm stripes and added to the polymer concrete composite at different percentages (0%, 0.12%, 0.16%, and 0.25%). A constant ratio of resin, aggregate, and fly ash was tested three times, requiring 12 specimens for each test.

The composite mixing began with polyester resin with hardener at a resin-to-hardener weight ratio of 100:32. The exact amount and weight of the materials for each specimen were calculated (Table 2). The mixture was then poured over the aggregates and mixed for three minutes until the amount of waste tin fibre strips were added and stirred for another minute until a brownish, viscous, and homogeneous mixture was observed. The aggregates must be of good quality, free of dust and other debris, and dry, and the materials must be soaked in polymer.

The mixture was poured into the moulds in two to three layers and was subjected to vibration for two minutes after each layer. The samples were demoulded 24 hours later and cured at room temperature. Due to the closure of the laboratory caused by the pandemic in 2020 and because polymer



Figure 3. Example of cylindrical and prismatic mould

concrete reached 90% of its 28-day strength in only seven days, the specimens were tested after seven days of curing. Due to the small specimens used in this study, all mixing was done by hand. All steps are previewed in Figure 4, which is the production procedure for the specimens for this study.

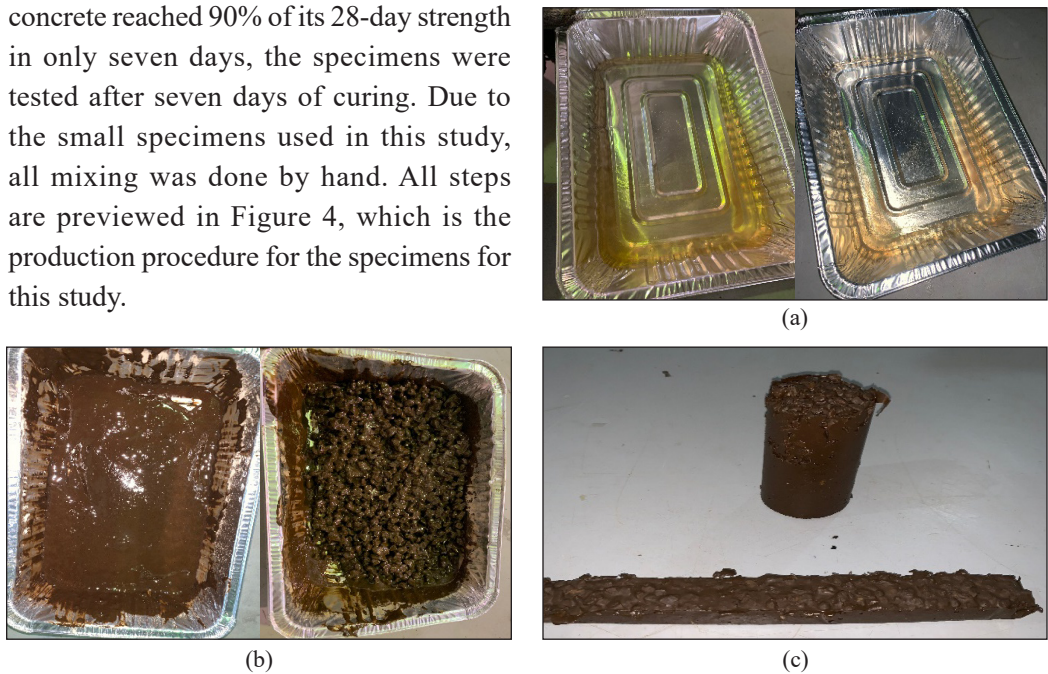


Figure 4. Fabrication process of PC specimens: (a) Mixing process – Resin and hardener; (b) Mixing process – Resin mixture with fly ash and aggregates; and (c) Demoulded cylindrical and prismatic specimens

## Experimental Design

All tests were done in the laboratory at Universiti Tenaga Nasional with the specific standards of compression strength, flexural tensile, and splitting tensile tests.

**Compressive Strength Test.** The compressive strength test is the most popular test performed on concrete, as it gives a general idea of the characteristics of concrete (Figure 5). Based on the test, the use of the concrete can be determined. The ASTM C39/C39M-01 (2001) was used in this study. After a curing period of seven days, the polymer concrete specimens were tested by the compression testing machine. The cylindrical specimens were placed vertically in the appropriate locations in the test machine. The specimens were aligned centrally on the base plate of the



Figure 5. Compressive strength test

machine. The movable part of the machine touches the specimen's top surface, and the load is applied to the face perpendicular to the casting direction. A 0.4 N/mm<sup>2</sup>/sec load was applied gradually until the specimen failed. It must be secured to avoid the load from being applied with shock. Three specimens for each ratio of waste tin fibre (a total of 12 specimens) were tested, and the maximum load of each implementation was recorded. Meanwhile, the  $F_{ck}$  (N/mm<sup>2</sup>) is the characteristic compressive strength of each grade of concrete. The average of the maximum load of the specimens (N) divided by the area of the specimen under load (mm<sup>2</sup>) will give this value in Equation 1.

$$\text{Compressive Strength} = \frac{\text{Load}}{\text{Cross - Section Area}} \quad (1)$$

**Flexural Tensile Test.** The flexural test was a method to evaluate concrete's tensile strength indirectly. Thereby, unreinforced concrete's ability to withstand bending failure was tested. The modulus of concrete rupture was the result gained through the flexural test. There were two ways to conduct the flexural test: the three-point load test (ASTM C78/C78M-18, 2016) or the centre point load test (ASTM C293/C293M-16, 2016). This study used ASTM C78/C78M-18 since the laboratory provides the three-point load test.

Specimens should be tested immediately after being taken out of the curing condition to prevent surface drying to avoid reducing flexural strength (Figure 6). The specimens were placed on two circular rollers manufactured of steel, which will support the specimen from the bottom left and bottom right sides. The length of the rollers was at least 10 mm longer than the width of the specimen. One other roller provided the load from the top side of the specimen. The specimen was placed in the machine correctly centred, with the longitudinal axis of the specimen orthogonal to the rollers. Any gap greater than 0.1 mm between the specimen and the rollers was eliminated. The load applied to the specimen had a rate of 0.2 MPa/s as per ASTM C78/C78M-18 and should not be applied with shock. The modulus of rupture was estimated using the following Equation 2:

$$MR = \frac{3PL}{2bd^2} \quad (2)$$

where, MR: modulus of rupture, P: ultimate applied load indicated by testing machine, L: span length, b: average width of the specimen, d: average depth of the specimen.

**Splitting Tensile Test.** Due to its brittle nature, concrete is weak when it comes to tensile force. When the tensile forces exceed

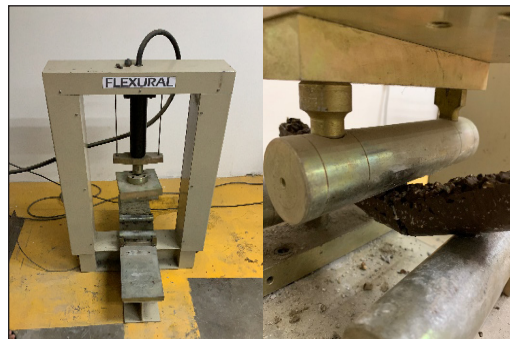


Figure 6. Flexural strength test



the concrete's tensile strength, it develops cracks. Therefore, it is necessary to determine the load at which the concrete may crack. In order to achieve this information, the splitting tensile test was implemented after the 7-day curing period. Diametrical lines were drawn on the specimen's two ends to ensure they were on the same axial plane (Figure 7). The weight and dimension of the specimens were recorded, and the compression testing machine was set for the required range. The specimen was placed between two plates. The load will be applied continuously without shock at a 0.4 N/mm<sup>2</sup>/sec rate from the upper plate until failure. The maximum load was noted at the end. The splitting tensile strength was calculated as in Equation 3:

$$T = 2P/\pi LD \quad (3)$$

where, P = Maximum load, L = Length, D = Diameter



Figure 7. Splitting tensile test

## RESULTS AND DISCUSSION

Experiments were carried out for 7 days to determine the axial compressive, flexural tensile, and splitting tensile strengths of the specimens of PC containing waste tin fibre in ratios of 0%, 0.12%, 0.16%, and 0.20%. Three specimens for each test and each ratio were produced. The experimental results were obtained through the testing machines and calculated with the designated formula for each test.

### Compressive Strength Test

Twelve cylinders with a diameter of 50 mm and a length of 100 mm, containing different amounts of waste in fibre, were tested, and the results are in Table 3. The values of the maximum load that differ by more than 25% were initially ignored when evaluating the results (Table 4). Wrong results could be caused by possible errors occurring during the mixing procedure or by not perfectly compacting the mixture due to the small specimen size. From the other values, the average was calculated as the maximum load divided by the cross-sectional area. The samples with a 0.20% waste tin fibre ratio show the lowest compressive strength, 6.94 N/mm<sup>2</sup>, while the highest compressive strength observed is 10.85 N/mm<sup>2</sup> and belongs to the samples with a 0.16% waste tin fibre ratio. Samples with 0.12% and 0% waste tin fibre show a compressive strength of 10.64 and 7.73, respectively (Figure 8).

The results show that the compressive strength of the PC in this study is low compared to other studies. The reason for that is the use of different resins. This study used polyester resin, unlike most studies that use epoxy resin. Epoxy resin is generally better than polyester resin due to its higher strength and greater shelf life. Epoxy resin is also odourless and non-flammable, which makes it safer. However, epoxy resin is more expensive and much harder to obtain. If it is possible to overcome these issues, it is recommended to use epoxy resin to produce PCs.

Table 3  
*Compressive test results for various sample ratios*

No.	Sample	Length (mm)	Diameter (mm)	Area (mm <sup>2</sup> )	Elastic modulus (MPa)	Stress (N/mm <sup>2</sup> )	Max. Load (kN)
1	SP0				3.95	0.09	1.45
2	SP0				3.94	0.07	1.10
3	SP0				2.11	0.09	1.42
4	SP12				2.96	0.13	1.99
5	SP12				1.94	0.13	2.11
6	SP12	100	50	15708	1.92	0.14	2.17
7	SP16				-	0.03	0.51
8	SP16				2.25	0.12	1.93
9	SP16				2.61	0.15	2.33
10	SP20				2.82	0.08	1.30
11	SP20				2.19	0.10	1.60
12	SP20				2.27	0.08	1.19

Table 4  
*Compressive strength test result of various tin fibre percentage*

Samples	Compressive test			
	Tin fibre amount	Force in N	Average Force in N	F/A (N/mm <sup>2</sup> )
1	0 %	1450	1435	7.73
		1100		
		1420		
2	0.12 %	1990	2090	10.64
		2110		
		2170		
3	0.16 %	510	2130	10.85
		1930		
		2330		
4	0.20 %	1300	1363	6.94
		1600		
		1190		

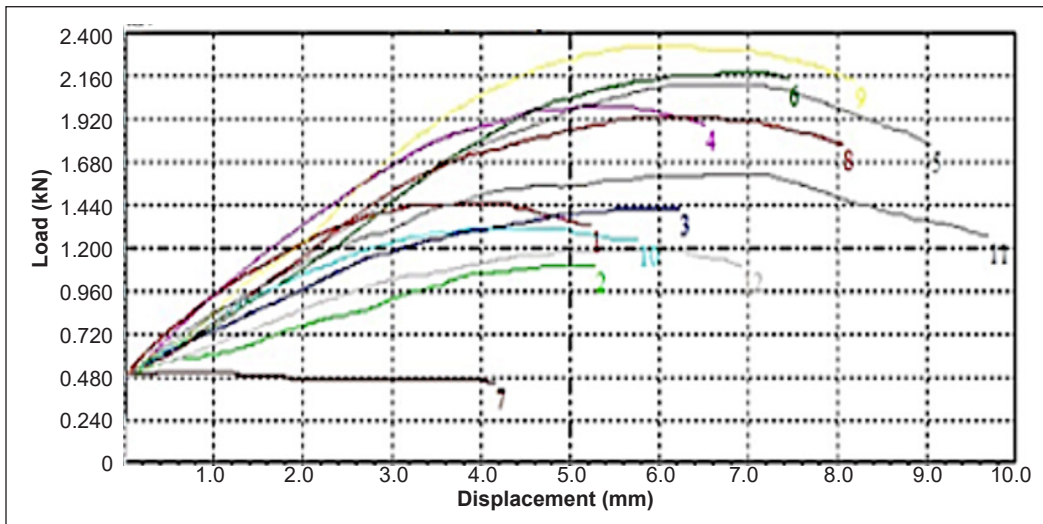


Figure 8. Compressive strength versus displacement of various sample ratios

### Splitting Tensile Test

Table 5 shows the results of the splitting tensile test done on 12 cylindrical samples in the same way as the compressive test. Furthermore, values differing by more than 25% are not considered in the result evaluation. After analysing the results presented in Table 6, it was found that the highest splitting tensile strength of 295 N/mm<sup>2</sup> was obtained from the samples containing 0.16% of waste tin fibre, which also had the highest average maximum load. In descending order, the splitting tensile strengths are 271 N/mm<sup>2</sup> (SP12), 257 N/mm<sup>2</sup> (SP0), and 241 N/mm<sup>2</sup> for the samples with 0.2% of waste tin fibre (Figure 9).

Table 5  
Splitting tensile test results for various sample ratios

No	Sample	Length (mm)	Diameter (mm)	Area (mm <sup>2</sup> )	Elastic modulus (MPa)	Stress (N/mm <sup>2</sup> )	Max. Load (kN)
1	SP0				9.36	0.13	2.12
2	SP0				7.27	0.11	1.77
3	SP0				8.65	0.14	2.18
4	SP12				7.98	0.16	2.45
5	SP12				5.42	0.11	1.80
6	SP12				8.17	0.14	2.14
7	SP16	100	50	15708	8.51	0.15	2.32
8	SP16				-	0.06	0.98
9	SP16				5.73	0.09	1.41
10	SP20				8.58	0.07	1.17
11	SP20				5.30	0.12	1.89
12	SP20				7.07	0.18	2.82

Table 6  
Splitting tensile test results of various tin fibre percentage

Samples	Splitting tensile			
	Tin fibre amount	Force in N	Average Force in N	$T = 2P/\pi L D$ (N/mm <sup>2</sup> )
1	0%	2120	2023	257
		1770		
		2180		
2	0.12%	2450	2130	271
		1800		
		2140		
3	0.16%	2320	2320	295
		980		
		1410		
4	0.2%	1170	1890	241
		1890		
		2820		

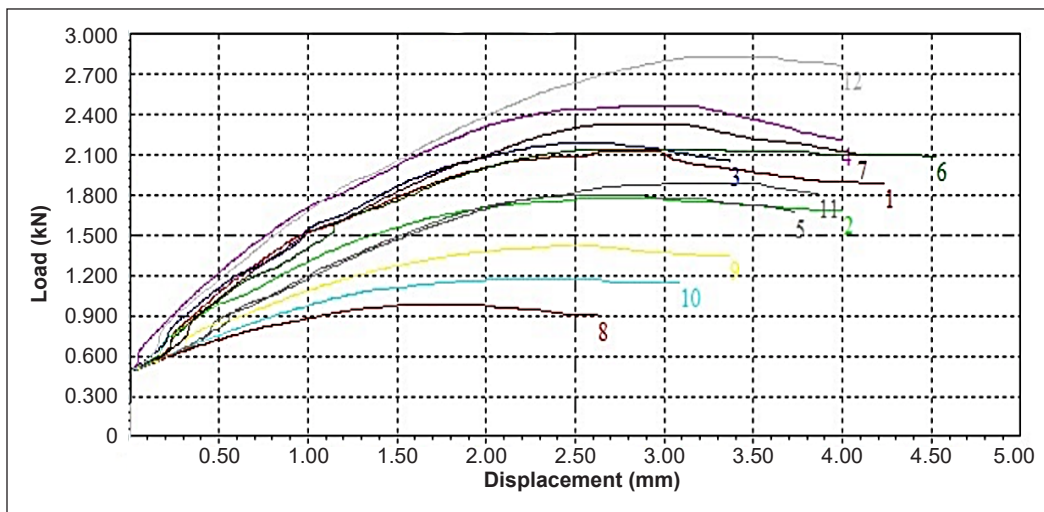


Figure 9. Splitting tensile strength versus displacement of various sample ratios

Compared to cement-based concrete, polymer concrete has a special characteristic: higher bending ability. The bending of the PC before failure causes very high flexural strength. This study shows that the flexural strength of PC is around two times higher than its splitting tensile strength. Increasing the waste tin fibre to a specific amount ensures the adherence between the resin and the tin fibres due to the resin filling the gaps in the desired way. However, increasing the waste tin ratio more than the specific amount decreases the adherence between the resins due to the even surface of aluminium, which leads to a decrement in the mechanical properties (Ferdous et al., 2016; Mohammed, 2018).

### Flexural Tensile Test Results

For the flexural tensile test, 12 prismatic specimens with 25 mm × 25 mm × 250 mm were tested. For each sample listed in Table 7, the maximum load displayed on the testing device’s panel was manually recorded (Figure 10). As shown in Table 8, the average maximum flexural tensile strength was determined from the test results and the flawed values were removed.

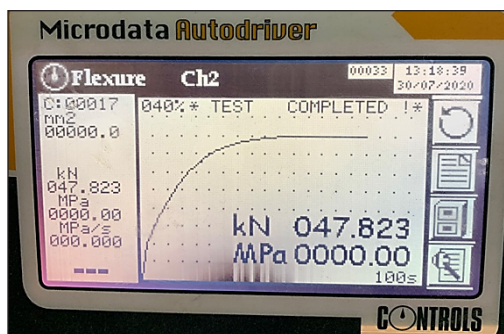


Figure 10. Flexural tensile machine (Microdata control flexure)

Table 7  
Flexural tensile test results for various sample ratios

No.	Sample	Length (mm)	Depth (mm)	Width (mm)	Max. Load (kN)
1	SP0				7.98
2	SP0				24.52
3	SP0				36.17
4	SP12				32.22
5	SP12				20.83
6	SP12				20.74
7	SP16	250	25	25	25.59
8	SP16				23.51
9	SP16				24.44
10	SP20				51.61
11	SP20				42.82
12	SP20				47.82

Table 8  
Flexural tensile test results for various sample ratios

Samples	Tin fibre amount	Flexural tensile test		
		Force in N	Average force in N	$\sigma_f = 3FL/2bd^2$ (N/mm <sup>2</sup> )
1	0 %	7980	24520	588.48
		24520		
		36170		
		32220		
2	0.12 %	20830	20785	498.84
		20740		
		25590		
3	0.16 %	23510	24510	588.24
		24440		
		51610		
4	0.20 %	42820	47416	1137.984
		47820		

PC's flexural and splitting tensile strengths are much higher than conventional Portland cement concrete (1.5-7 MPa) because of the better bonding characteristics between the matrix and aggregates. Also, using small aggregates (5 mm) leads to fewer voids in the composite, so the composite is denser and gains higher mechanical properties. These results are consistent with studies by Berrocal et al. (2018) and Pil et al. (2016).

## CONCLUSION

This study demonstrates that employing waste tin fibre as reinforcement in PC enhances its overall mechanical qualities. The effect of waste tin fibre on the mechanical properties of concrete was determined; the ideal waste tin fibre ratio for mechanical performance is 0.16%, while the optimal ratio for flexural tensile strength is 0.20%. From the experimental results, the compressive strength improved by 40% and the split tensile strength improved by 15% when a 0.16% waste tin fibre ratio was employed instead of a 0% waste tin fibre ratio. Meanwhile, a 20% waste tin fibre ratio in concrete resulted in a significant 93% improvement in flexural tensile strength. As a result of the findings, it has been demonstrated through the performance of polymer concrete, including waste tin fibres, that adding polymer during the formation of polymer concrete improves the mechanical strength of concrete by altering its properties. In addition, it has been demonstrated that polymers have enhanced the mechanical characteristics of concrete, including increased compressive strength, splitting tensile strength, and flexural strength, as well as good performance in durability and the growth and decrease of landfill wastes.

## ACKNOWLEDGEMENTS

The authors thank Universiti Tenaga Nasional (UNITEN), Malaysia, for supporting this research under BOLD 2022 through Project No: J510050002/2022004. A special thanks to those who contributed to this project directly or indirectly.

## REFERENCES

- ACI 318-11. (2011). *Building code requirements for structural concrete (ACI 318-11) and commentary*. American Concrete Institute. [https://www.pecivilexam.com/Study\\_Documents/Struc-Materials-Online/ACI\\_318-Building-Code-2011.pdf](https://www.pecivilexam.com/Study_Documents/Struc-Materials-Online/ACI_318-Building-Code-2011.pdf)
- Afroughsabet, V., & Ozbakkaloglu, T. (2015). Mechanical and durability properties of high-strength concrete containing steel and polypropylene fibers. *Construction and Building Materials*, *94*, 73-82. <https://doi.org/10.1016/j.conbuildmat.2015.06.051>
- Ahmadi, M., Shafabakhsh, G. A., Di Mascio, P., & Hassani, A. (2021). Failure behavior of functionally graded roller compacted concrete pavement under mode I and III fracture. *Construction and Building Materials*, *307*, Article 124942. <https://doi.org/10.1016/j.conbuildmat.2021.124942>



- Al-Nini, A., Nikbakht, E., Syamsir, A., Shafiq, N., Mohammed, B. S., Al-Fakih, A., Al-Nini, W., & Amran, Y. H. M. (2020). Flexural behavior of double-skin steel tube beams filled with fiber-reinforced cementitious composite and strengthened with CFRP sheets. *Materials*, *13*(14), Article 3064. <https://doi.org/10.3390/ma13143064>
- Ali, A., & Ansari, A. A. (2013, April 12-13). Polymer concrete as innovative material for development of sustainable architecture. In *2nd International Conference on Emerging Trends in Engineering & Technology* (pp. 1-4). Teerthanker Mahaveer University, India. <https://doi.org/10.13140/2.1.4761.5040>
- Asdollah-Tabar, M., Heidari-Rarani, M., & Aliha, M. R. M. (2021). The effect of recycled PET bottles on the fracture toughness of polymer concrete. *Composites Communications*, *25*, Article 100684. <https://doi.org/10.1016/j.coco.2021.100684>
- ASTM C39/C39M-01. (2001). *Standard test method for compressive strength of cylindrical concrete specimens*. ASTM International. <https://normanray.files.wordpress.com/2010/10/kuliah-7b-beton-segar-astm-c39.pdf>
- ASTM C78/C78M-18. (2016). *Standard test method for flexural strength of concrete (using simple beam with third-point loading)*. ASTM International. <https://www.scribd.com/document/426461328/ASTM-C78-Flexural-Strength-of-Concrete>
- ASTM C293/C293M-16. (2016). *Standard test method for flexural strength of concrete (using simple beam with center-point loading)*. ASTM International. <https://webstore.ansi.org/standards/astm/astmc293c293m16>
- Atiqah, A., Ismail, N., Lim, K. K., Jalar, A., Bakar, M. A., Maleque, M. A., Ilyas, R. A., & Supian, A. B. M. (2021). Properties of recycled metal matrix composites. In R. A. Ilyas, S. M. Sapuan, & E. Bayraktar (Eds.), *Recycling of Plastics, Metals, and Their Composites* (pp. 93-108). CRC press. <https://doi.org/https://doi.org/10.1201/9781003148760>
- Berrocal, C. G., Löfgren, I., & Lundgren, K. (2018). The effect of fibres on steel bar corrosion and flexural behaviour of corroded RC beams. *Engineering Structures*, *163*, 409-425. <https://doi.org/10.1016/j.engstruct.2018.02.068>
- Bulut, H. A., & Şahin, R. (2017). A study on mechanical properties of polymer concrete containing electronic plastic waste. *Composite Structures*, *178*, 50-62. <https://doi.org/10.1016/j.compstruct.2017.06.058>
- Esfahani, M. S., Janbaz, S., & Mirmazhari, S. (2016). Effect of various types of fillers on mechanical properties of concrete. *International Journal of Civil Engineering, Construction and Estate Management*, *4*(3), 20-28.
- Ferdous, W., Manalo, A., Aravinthan, T., & Erp, G. V. (2016). Properties of epoxy polymer concrete matrix: Effect of resin-to-filler ratio and determination of optimal mix for composite railway sleepers. *Construction and Building Materials*, *124*, 287-300. <https://doi.org/10.1016/j.conbuildmat.2016.07.111>
- Frigione, M. (2013). Concrete with polymers. In *Eco-Efficient Concrete* (pp. 386-436). Woodhead Publishing Limited. <https://doi.org/10.1533/9780857098993.3.386>
- Gao, Y., romero, P., Zhang, H., Huang, M., & Lai, F. (2019). Unsaturated polyester resin concrete: A review. *Construction and Building Materials*, *228*, Article 116709. <https://doi.org/10.1016/j.conbuildmat.2019.116709>
- Hameed, A. M., & Hamza, M. T. (2019). Characteristics of polymer concrete produced from wasted construction materials. *Energy Procedia*, *157*(2018), 43-50. <https://doi.org/10.1016/j.egypro.2018.11.162>



- Hashemi, M. J., Jamshidi, M., & Aghdam, J. H. (2018). Investigating fracture mechanics and flexural properties of unsaturated polyester polymer concrete (UP-PC). *Construction and Building Materials*, 163, 767-775. <https://doi.org/10.1016/j.conbuildmat.2017.12.115>
- Ismail, A. S., Jawaid, M., Hamid, N. H., Yahaya, R., Hassan, A., Asim, M., & Supian, A. B. M. (2022). Effect of curing temperature on mechanical properties of bio-phenolic/epoxy polymer blends. *Journal of Polymers and the Environment*, 30, 878-885. <https://doi.org/10.1007/s10924-021-02244-w>
- Kamal, N. L. M., Beddu, S., Syamsir, A., Mohammad, D., Itam, Z., Hamid, Z. A. A., & Manan, T. S. A. (2019). Immobilization of heavy metals for building materials in the construction industry - An overview. *Materials Today: Proceedings*, 17, 787-791. <https://doi.org/10.1016/j.matpr.2019.06.363>
- Kumlutaş, D., Tavman, I. H., & Turhan Çoban, M. (2003). Thermal conductivity of particle filled polyethylene composite materials. *Composites Science and Technology*, 63(1), 113-117. [https://doi.org/10.1016/S0266-3538\(02\)00194-X](https://doi.org/10.1016/S0266-3538(02)00194-X)
- Luxchem, P. I. (2018). *Material Safety Data Sheet - Unsaturated Polyester Resins*. LUXCHEM. [http://www.luxchem.com.my/img/pdf/reversol\\_resin.pdf](http://www.luxchem.com.my/img/pdf/reversol_resin.pdf)
- Martínez-López, M., Martínez-Barrera, G., Salgado-Delgado, R., & Gencel, O. (2021). Recycling polypropylene and polyethylene wastes in production of polyester based polymer mortars. *Construction and Building Materials*, 274, Article 121487. <https://doi.org/10.1016/j.conbuildmat.2020.121487>
- Mohamad, N. J. (2014). *Timber Reinforcement in Reinforced Concrete Beam for Light Structure* (Doctoral dissertation). University Malaysia Pahang, Malaysia. <https://core.ac.uk/download/pdf/159184667.pdf>
- Mohamad, N., Muthusamy, K., Embong, R., Kusbiantoro, A., Hashim, M. H., & Hanafi, M. (2022). Environmental impact of cement production and solutions : A review. *Materials Today: Proceedings*, 48, 741-746. <https://doi.org/10.1016/j.matpr.2021.02.212>
- Mohammed, A. A. (2018). Mechanical strength of concrete with PVC aggregates. In *Use of Recycled Plastics in Eco-efficient Concrete* (pp. 115-135). Elsevier Ltd. <https://doi.org/10.1016/B978-0-08-102676-2.00006-2>
- Muda, Z. C., Alam, M. A., Syamsir, A., Sulleman, S., Beddu, S., Mustapha, K. N., Thiruchelvam, S., Ismail, F. B., Usman, F., Kamal, N. L. M., Birima, A. H., Itam, Z., & Zaroog, O. S. (2016). The effect of thickness and mesh spacing on the impact resistance of ferrocement slab. *IOP Conference Series: Earth and Environmental Science*, 32, Article 012027. <https://doi.org/10.1088/1755-1315/32/1/012027>
- Muda, Z. C., Kamal, N. L. M., Syamsir, A., Sheng, C. Y., Beddu, S., Mustapha, K. N., Thiruchelvam, S., Usman, F., Alam, M. A., Birima, A. H., & Zaroog, O. S. (2016). Impact resistance performance of kenaf fibre reinforced concrete. *IOP Conference Series: Earth and Environmental Science*, 32, Article 012019. <https://doi.org/10.1088/1755-1315/32/1/012019>
- Niaki, M. H., Fereidoon, A., & Ahangari, M. G. (2018). Mechanical properties of epoxy/basalt polymer concrete: Experimental and analytical study. *Structural Concrete*, 19(2), 366-373. <https://doi.org/10.1002/suco.201700003>
- Oyebisi, S., Ede, A., Olutoge, F., & Omole, D. (2020). Geopolymer concrete incorporating agro-industrial wastes: Effects on mechanical properties, microstructural behaviour and mineralogical phases. *Construction and Building Materials*, 256, Article 119390. <https://doi.org/10.1016/j.conbuildmat.2020.119390>

- Parikh, K., & Modhera, C. D. (2012). Application of GFRP on preloaded retrofitted beam for enhancement in flexural strength. *International Journal of Civil and Structural Engineering*, 2(4), 1070-1080.
- Pil, L., Bensadoun, F., Pariset, J., & Verpoest, I. (2016). Why are designers fascinated by flax and hemp fibre composites? *Composites Part A: Applied Science and Manufacturing*, 83, 193-205. <https://doi.org/10.1016/j.compositesa.2015.11.004>
- Reddy, S. V. B., & Santhosha, V. (2018). Experimental study on fibre reinforced polymer concrete. *International Journal of Applied Engineering Research*, 13(15), 11844-11856.
- Ramadan, R., Jahami, A., Khatib, J., El-Hassan, H., & Elkordi, A. (2023). Improving structural performance of reinforced concrete beams with phragmites australis fiber and waste glass additives. *Applied Sciences*, 13(7), Article 4206. <https://doi.org/10.3390/app13074206>
- Sakai, E., Miyahara, S., Ohsawa, S., Lee, S. H., & Daimon, M. (2005). Hydration of fly ash cement. *Cement and Concrete Research*, 35(6), 1135-1140. <https://doi.org/10.1016/j.cemconres.2004.09.008>
- Saribiyik, M., Piskin, A., & Saribiyik, A. (2013). The effects of waste glass powder usage on polymer concrete properties. *Construction and Building Materials*, 47, 840-844. <https://doi.org/10.1016/j.conbuildmat.2013.05.023>
- Sonebi, M., Abdalqader, A., Fayyad, T., Amziane, S., & El-Khatib, J. (2022). Effect of fly ash and metakaolin on the properties of fiber-reinforced cementitious composites: A factorial design approach. *Computers and Concrete*, 29(5), 347-360. <https://doi.org/10.12989/cac.2022.29.5.347>
- Sosoi, G., Barbuta, M., Serbanoiu, A. A., Babor, D., & Burlacu, A. (2018). Wastes as aggregate substitution in polymer concrete. *Procedia Manufacturing*, 22, 347-351. <https://doi.org/10.1016/j.promfg.2018.03.052>
- Supian, A. B. M., Sapuan, S. M., Zuhri, M. Y. M., & Ya, H. H. (2016, December 28). A review of the development hybrid nature-synthetic fibre-reinforced composite in energy absorption tube applications. In *Proceedings of 5th Postgraduate Seminar on Natural Fiber Composites* (pp. 1-6). Universiti Putra Malaysia, Malaysia.
- Supian, A. B. M., Sapuan, S. M., Zuhri, M. Y. M., Zainudin, E. S., & Ya, H. H. (2018). Hybrid reinforced thermoset polymer composite in energy absorption tube application: A review. *Defence Technology*, 14(4), 291-305. <https://doi.org/10.1016/j.dt.2018.04.004>
- Vaggar, G. B., Sirimani, V. B., Sataraddi, D. P., Hiremath, N. M., & Bhajantri, F. (2021). Effect of filler materials on thermal properties of polymer composite materials: A review. *International Journal of Engineering Research & Technology (IJERT)*, 10(08), 1-5.
- Yan, L., Kasal, B., & Huang, L. (2016). A review of recent research on the use of cellulosic fibres, their fibre fabric reinforced cementitious, geo-polymer and polymer composites in civil engineering. *Composites Part B: Engineering*, 92, 94-132. <https://doi.org/10.1016/j.compositesb.2016.02.002>
- Yeon, J. H., Lee, H. J., & Yeon, J. (2020). Deformability in unsaturated polyester resin-based concrete: Effects of the concentration of shrinkage-reducing agent and type of filler. *Materials*, 13(3), 1-16. <https://doi.org/10.3390/ma13030727>
- Zabihi, O., Ahmadi, M., Nikafshar, S., Preyeswary, K. C., & Naebe, M. (2017). A technical review on epoxy-clay nanocomposites: Structure, properties, and their applications in fiber reinforced composites. *Composites Part B: Engineering*, 135, 1-24. <https://doi.org/10.1016/j.compositesb.2017.09.066>



## Health Risk Assessment of Pollutant Emissions from Coal-fired Power Plant: A Case Study in Malaysia

Mohd Shahril Kamarudin<sup>1</sup>, Abderrahim Zermene<sup>1,2</sup>, Nur Aliah Fatin Mohd Nizam Ong<sup>1,3</sup>, Noorazmin Ab Rasid<sup>1</sup>, Shafizah Masuri<sup>1</sup> and Mohd Zahirasri Mohd Tohir<sup>1,4\*</sup>

<sup>1</sup>Department of Chemical and Environmental Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Department of Civil, Environmental and Geomatic Engineering, University College London, Gower Street, London, WC1E 6BT, United Kingdom

<sup>3</sup>PETRONAS Research Sdn Bhd, Kawasan Institusi Bangi, 43000 Kajang, Selangor, Malaysia

<sup>4</sup>Department of Construction, Building Services and Structures, Universidad de Navarra, Pamplona, Spain

### ABSTRACT

Coal-fired power plants (CFPPs) are Malaysia's primary electricity source, but their emissions adversely affect human health, organism growth, climate change, and the environment. The carbon, hydrogen, and sulphur content of coal make it a viable option for electricity generation. However, the by-products from leaching, volatilisation, melting, decomposition, oxidation, hydration, and other chemical reactions significantly negatively impact the environment and human health. This study aims to quantify the emissions from a coal-fired power plant, investigate the interplay between different emissions, simulate the dispersion of emissions, and assess their health impact through a health risk assessment. The results indicate that SO<sub>2</sub> is the primary contributor to emissions and its impact on human health is a concern. The health effects, both chronic and acute, are more pronounced in children than in adults. This study combines real-time emissions data and simulations to assess emissions' health impact, raising awareness about the emissions from coal-fired power plants. Furthermore, the findings can potentially enhance working conditions for employees and promote environmental health.

### ARTICLE INFO

#### Article history:

Received: 03 January 2023

Accepted: 17 August 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.10>

#### E-mail addresses:

gs55394@student.upm.edu.my (Mohd Shahril Kamarudin)

zermene.abderrahim@gmail.com (Abderrahim Zermene)

nuraliahfatin@gmail.com (Nur Aliah Fatin Mohd Nizam Ong)

ech3501.kka@gmail.com (Noorazmin Ab Rasid)

shafizah\_m@upm.edu.my (Shafizah Masuri)

zahirasri@upm.edu.my (Mohd Zahirasri Mohd Tohir)

\* Corresponding author

*Keywords:* CFPP emissions, air pollution, health risk assessment, hazard Quotient, hazard index

### INTRODUCTION

Coal is the most abundant energy source on Earth and is extensively utilised for power generation in numerous countries (Munawar, 2018). In Malaysia, coal-fired power plants (CFPPs) contribute the largest

share of electricity, accounting for approximately 64%, while gas and hydro technologies contribute 32% and 4%, respectively (Ranjan et al., 2019). However, the utilisation of fossil fuels in thermal power plants has significantly contributed to air pollution, with primary pollutants such as carbon monoxide (CO), sulphur dioxide (SO<sub>2</sub>), carbon dioxide (CO<sub>2</sub>), and nitrogen oxides (NO<sub>x</sub>) (Shekarchian et al., 2011). The adverse effects of these pollutants, particularly CO<sub>2</sub> emissions from coal combustion, have been linked to climate change and related phenomena, resulting in an increased risk of malaria and its associated mortality (Gething et al., 2010).

It is essential to examine the emission levels and composition of various pollutants to provide a comprehensive understanding of the environmental impact of CFPPs. According to a study, CFPPs release significant quantities of carbon dioxide (CO<sub>2</sub>), with an annual rate of 17,739 tonnes per thousand households (The IBR Asia Group Sdn. Bhd., 2019). Moreover, it was found that CFPPs emit approximately 5.7 tonnes of NO<sub>x</sub>, 1.2 tonnes of particulate matter (PM), and 5.9 metric tonnes of SO<sub>2</sub> per year (The IBR Asia Group Sdn. Bhd., 2019). Additionally, CFPPs release trace amounts of heavy metals like cadmium, arsenic, mercury, and lead (Pb) (The IBR Asia Group Sdn. Bhd., 2019). These emissions contribute to the overall air pollution and environmental degradation of CFPPs. As reported by the Malaysia Ministry of Natural Resources and Environment (2015), CO<sub>2</sub> accounted for 73%, 76%, and 72% of all greenhouse gas emissions in 2000, 2005, and 2011, respectively, with methane (CH<sub>4</sub>) and nitrous oxide (N<sub>2</sub>O) also contributing to the overall greenhouse gas emissions. Nitrous oxide in the atmosphere rose from 4% in 2000 to 5% in 2011 (Babatunde et al., 2018).

The release of these pollutants from CFPPs poses significant risks to human health and other organisms, as highlighted by reports from the United Nations Environment Program (UNEP). CFPPs are estimated to be responsible for approximately 1400 additional deaths annually due to the lethal doses of air pollutants they produce (Yakubu, 2017). Furthermore, the World Health Organization (WHO) has identified air pollution as a leading cause of approximately 4.2 million premature deaths each year, associating it with various health issues such as cardiovascular diseases, cancer, respiratory illnesses, and neurological disorders. Therefore, this study is crucial for estimating the level of health risk in different age and gender groups (Combes & Franchineau, 2019). CFPPs contribute to 23% of Malaysia's air pollution, making them the second-largest source after transportation (Zubir et al., 2017). Considering these air pollutants' physical and chemical characteristics, they can spread over vast distances, posing significant risks to the environment and human health (Zubir et al., 2017).

Coal combustion yields several harmful gases, including CO<sub>2</sub>, CO, SO<sub>2</sub>, sulphur trioxide (SO<sub>3</sub>), nitrogen dioxide (NO<sub>2</sub>) and nitric oxide (NO), which have been associated with various health issues (Badman & Jaffé, 1996; Kelsall et al., 1997; Munawar, 2018). Even in residential settings, coal combustion for heating contributes to environmental and

health problems. Recent studies have shown that coal's chemical processing releases two to four times more CO<sub>2</sub> compared to oil, resulting in adverse environmental effects such as global warming and the greenhouse effect (Ren & Patel, 2009). The health problems associated with these gases range from malaria and cardiovascular disease to asthma and other respiratory ailments. Approximately 90% of all global CO<sub>2</sub> emissions in 2011 were attributed to burning fossil fuels, emphasising the significance of addressing emissions from coal combustion (Munawar, 2018).

In addition to CO<sub>2</sub>, coal combustion releases sulphur into the environment, leading to air, water, and land pollution. Unregulated coal power plants release twice as much sulphur oxides and particulate matter into the atmosphere annually compared to vehicles, trucks, and factories. The resulting SO<sub>x</sub> and PM travel long distances from a power station and decompose into sulphuric acid (H<sub>2</sub>SO<sub>4</sub>), a key component of acid rain. Besides, inhaling SO<sub>x</sub> can also negatively impact human health. The inhalation of SO<sub>x</sub> air pollution has been linked to disrupted cardiac rhythms and an increased risk of heart attacks (Peters et al., 1999). Proximity to power plants and high exposure to SO<sub>2</sub> have also been associated with respiratory issues such as suffocation, wheezing, coughing, and decreased lung function (Munawar, 2018).

Moreover, NO<sub>2</sub> generated during coal combustion accumulates in the air and causes cumulative damage to the environment and human health due to its corrosive nature and strong oxidising properties (Levy et al., 1999). Exposure to high levels of NO<sub>2</sub> (>1500 mg/m<sup>3</sup>) has been linked to decreased lung function, asthma attacks, and other respiratory problems. The impact of NO<sub>2</sub> on individuals varies, with some being vulnerable even to lower pollutant quantities (Munawar, 2018). Furthermore, coal burning releases millions of tonnes of coal fly ash (CFA) and coal dust annually, which serve as precursors to PM and pose severe health risks (Clancy et al., 2002; Miller & Sullivan, 2007). Air pollution, particularly PM, has been associated with various developmental abnormalities, including congenital malformations, adverse pregnancy outcomes, infant mortality, and genetic anomalies. The detrimental effects of PM on human health include the development of cancers, cardiovascular diseases, and reproductive abnormalities (Munawar, 2018).

According to Mahlia (2002), electricity power plants in Malaysia emit significant amounts of pollutants into the atmosphere, with a considerable portion of the chemical energy being converted into heat inefficiently. While previous studies have shed light on the air emissions from CFPPs, the effects of these emissions on health, particularly in the Johor region, remain understudied. Therefore, this study aims to evaluate the pollution emissions and assess the associated health risks among populations residing near CFPPs due to exposure to air pollutants. The findings of this research will contribute to a better understanding of the health risks faced by different populations in proximity to CFPPs. Not only that, but the research will also serve as a basis for raising awareness among the public and policymakers about the dangers associated with CFPP pollutant emissions and facilitate prompt responses to potential emergencies involving hazardous gas releases.



## MATERIALS AND METHODS

### Data Collection

The lab equipment within the CFPP offered two sets of historical data: a quarterly year interval historical data taken at four different sampling points from May 2016 to June 2019. The second consists of 30 min intervals of historical data taken at the main stack of the CFPP from January 1, 2016, to December 31, 2018.

### Target Population

Multiple neighbourhoods can be found within 10 km of the CFPP site, as Google Earth Pro 7.3 shows. Hazard quotient assessment threat locations were Kampung Sungai Boh and Kampung Sungai Chengkeh for three main reasons: (1) the nearest population area to the CFPP location, (2) the high-density population, and (3) prevailing wind direction.

Google Earth Pro 7.3's ruler tools measure the distance with the camera pointed directly north (Table 1). Meanwhile, the geography of Malaysia's southwestern region is viewed on a topographic map courtesy of Worldwide Elevation Finder/Topographic map/Altitude map (Figure 1).

Table 1

Location coordinates and distance from the CFPP (source: Google Earth Pro 7.3)

Location	Coordinates	Distance (km)
Coal Fired Power Plant (Emission source)	1°19'58.97"N, 103°32'9.61"E	-
AAA sampling point in Kg. Chokoh Besar	1°18'58.30"N, 103°30'9.70"E	4.19
Kampung Sungai Boh	1°20'38.81"N, 103°30'54.16"E	2.63
Kampung Sungai Chengkeh	1°20'39.25"N, 103°31'13.72"E	2.10

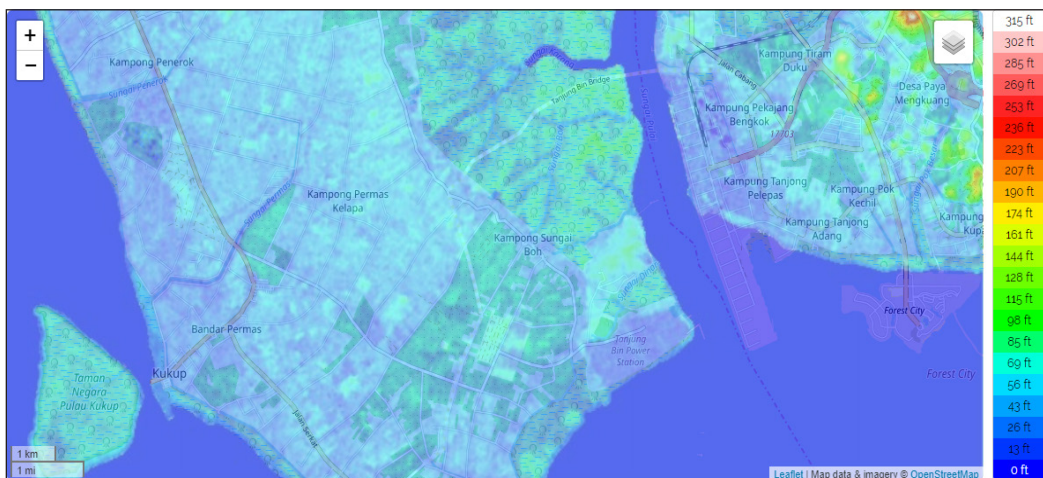


Figure 1. Topography map of southern-western Johor (source: Worldwide Elevation Finder/ Topographic map)



## Sampling and Analysing Methods

The CFPP follows the guidelines established by the Joint Standards Australia/New Zealand Standards Committee EV-007, which include the adoption of the Australian/New Zealand Standard (AS/NZS 3580.9.6:2003 and AS/NZS 3580.9.14:2013) for PM sampling and analysis (Standard Australia Committee and Standard New Zealand Committee, 2015). ISC Method (Inter-Society Committee Methods of Air Sampling and Analysis) is utilised for SO<sub>2</sub> and NO<sub>2</sub> analysis and sampling (Table 2). Sampling and testing for CO concentrations have been done using a technique different from the United States Environmental Protection Agency (USEPA) Method 10. Table 2 provides a summary of the sampling and analysis techniques that were employed.

Table 2  
*Sampling and analysing methods*

Type of gases	Sampling and analysing methods used
Particulate matter (PM <sub>10</sub> )	AS/NZS 3580.9.6:2003
Particulate matter (PM <sub>2.5</sub> )	AS/NZS 3580.9.14:2013
Sulphur dioxide (SO <sub>2</sub> )	ISC Method 704
Nitrogen dioxide (NO <sub>2</sub> )	ISC Method 408
Carbon monoxide (CO)	USEPA Method 10

## Real Time-based Emissions and Regression

In this study, the emissions were recorded over three years with a 1 min interval between each entry. These results were collected and recorded, and the overall emissions were extracted. Moreover, linear regression was used to determine how some emissions impact each other.

## Atmospheric Dispersion Model Simulation

The dispersion model is simulated using a free copy of ALOHA, version 5.4.7, downloaded from the EPA's website. The Gaussian Plume dispersion model in ALOHA software solidified the author's decision to utilise it. ALOHA dispersion models have previously been widely used in the field of consequence emissions (Khalid et al., 2019; Malik et al., 2021, 2023).

## Health Risk Assessment

After determining the population's exposure to individual air contaminants using air quality modelling, the next step was using air pollution health risk assessment (AP-HRA). Since non-carcinogenic gases are more abundant than carcinogenic gases, Hazard Quotient (HQ) is evaluated here. Acute and chronic AP-HRA estimates have been calculated for

four groups, male adult, female adult, male child, and female child, to differentiate non-chemical variables, including respiration and body weight (Das et al., 2018).

$$ADD = \frac{Ca \times IR \times ET \times EF \times ED}{BW \times AT} \quad [1]$$

$$fD = RfC \times \frac{20m^3/day}{70kg} \quad [2]$$

$$HQ = \frac{ADD}{RfD} \quad [3]$$

Equations 1, 2 and 3 are the formulas used to calculate HQ. For non-cancer assessment, the average time equals the Exposure duration (ED), converted into days (Das et al., 2018). Reference concentration (RfC) values used for health risk assessment are shown in Table 3, adapted from Kenessary et al. (2019). Maximum single concentrations of the pollutants are used for calculating HQ acute exposure (Hurt et al., 2001).

Table 3  
Reference concentration

Parameter	Adult (59 years old)		Children (11 years old)	
	Male	Female	Male	Female
Body weight, BW (kg)	68	60	32.6	32
Inhalation rate (m <sup>3</sup> /h)	0.63	0.47	0.58	0.54
Exposure duration, ED (year)	59		11	
Exposure time, ET (h/day)			24	
Exposure frequency, EF (day/year)			350	

Non-chemical factor values used in this study are summarised as displayed in Table 3. Adult Body weight (BW) is adapted based on the mean weight of Malaysia's South socio-demographic (Azmi et al., 2009). Meanwhile, children's BW is assumed based on the BW status of urban Malay primary school children (Yang et al., 2017). Inhalation rate (IR) parameters for children and adults are assumed from existing default values and recommendations for the European population's exposure assessment (Höglund et al., 2012). These default values are used because of the comprehensiveness of the respective age group and gender of the studied group. Also, the IR values for the respective studied groups are not stated in the guidelines by the Department of Environment (DoE) Malaysia (Hashim & Hashim, 2010). Values of Hazard Index (HI) less than 1 are equivalent to HQ, indicating that exposure is unlikely to cause negative consequences. This Equation 4 is used to determine HI values:

$$HI = \sum HQ_i \quad [4]$$

## RESULTS AND DISCUSSION

### Overall Emissions from 2016 to 2018

The data presented in Figure 2 indicates that the total emissions production during the period from 2016 to 2018 amounted to approximately 242 million mg/Nm<sup>3</sup> of SO<sub>2</sub>, 126 million mg/Nm<sup>3</sup> of NO<sub>2</sub>, and 91 million mg/Nm<sup>3</sup> of CO. These emission levels are substantial for a single CFPP and a relatively small developing country like Malaysia. It is noteworthy that the emissions exhibited a fluctuating trend over the studied period. Specifically, there was a decrease in emissions in 2017, but they subsequently increased again in 2018. This variation in emission levels suggests the presence of certain factors or events that influenced the emissions output during those specific years.

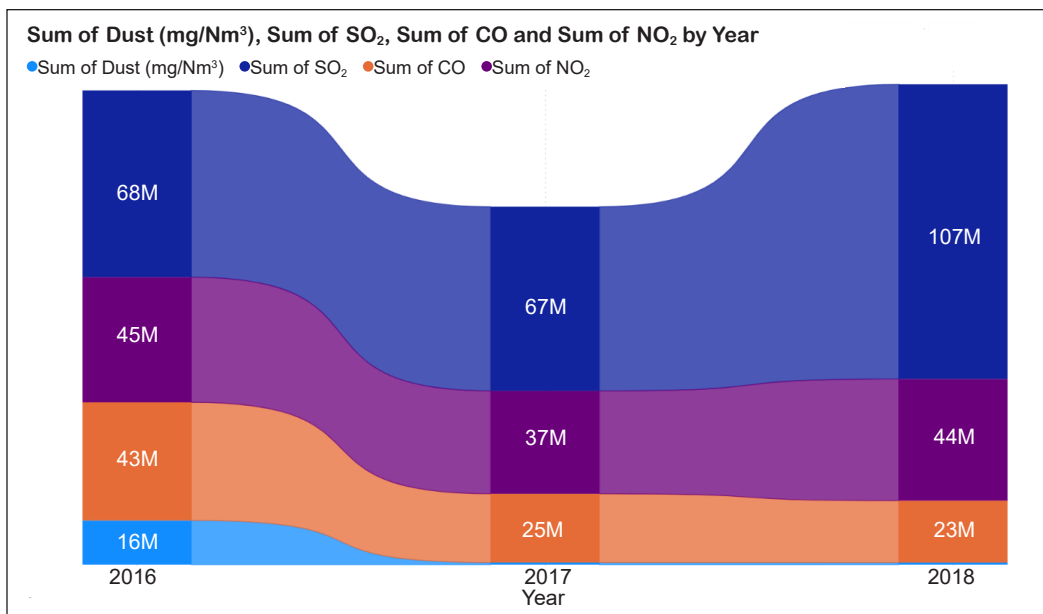


Figure 2. Evolution of emission levels in the period of 2016 to 2018

### Effect of Other Emissions on SO<sub>2</sub>

The findings from Figure 2 highlight that the period from 2016 to 2018 witnessed substantial emissions production from the CFPP, with SO<sub>2</sub> exhibiting the highest recorded emissions during this timeframe. A linear regression analysis was performed to gain a deeper understanding of the relationship between SO<sub>2</sub> emissions and other pollutants. The predictors used in the model were NO<sub>2</sub>, CO, and dust. The regression analysis results revealed significant associations between these pollutants and SO<sub>2</sub> emissions. Specifically, NO<sub>2</sub> emerged as a significant positive predictor, with a coefficient (B) of 0.534, a p-value of 0.000, and an odds ratio (OR) of 1.707. It indicates that an increase in NO<sub>2</sub> emissions is linked to a corresponding increase in SO<sub>2</sub> emissions.

On the other hand, CO and dust were identified as significant negative predictors of SO<sub>2</sub> emissions. The regression analysis showed that CO had a B value of -0.035, a p-value of 0.000, and an OR of 0.694, implying that higher CO emissions were associated with lower SO<sub>2</sub> emissions. Similarly, dust exhibited a B value of -0.017, a p-value of 0.000, and an OR of 0.983, indicating that increased dust emissions were correlated with decreased SO<sub>2</sub> emissions.

As presented in Table 4, these findings underscore the interplay between different pollutant emissions and their impact on SO<sub>2</sub> emissions from the CFPP. The positive relationship between NO<sub>2</sub> and SO<sub>2</sub> emissions suggests a potential co-occurrence or shared sources of these pollutants within the power plant. Conversely, the negative associations between CO and dust emissions with SO<sub>2</sub> emissions might indicate competing mechanisms or the effectiveness of emission control measures targeting these pollutants.

Table 4  
*Effect of other emissions on SO<sub>2</sub>*

Parameter	B	Sig.	Exp(B)	95% Interval for Exp(B)	
				Lower	Upper
(Intercept)	4.935E-14	1.000	1.000	0.999	1.001
NO <sub>2</sub>	0.534	0.000	1.707	1.705	1.709
CO	-0.035	0.000	0.965	0.964	0.967
Dust	-0.017	0.000	0.983	0.982	0.984
(Scale)	0.713 <sup>a</sup>				

### Comparison of Modelled Ground Level Concentration with Historical Data

This comparison aims to validate the simulation output's reliability and accuracy by benchmarking with the measured concentration. It is essential to validate the simulated output as atmospheric dispersion is a stochastic phenomenon. The simulation model is likely to have deviated from the measured concentration due to either a single factor or a combination of the model configuration, atmospheric chemistry and unpredictable human behaviour (Rao, 2005). Therefore, by analysing the simulated with the historical data, the source of errors can be pointed out, and corrective actions can be taken to improve the mathematical model.

The quarter-year interval historical data used for this study are an average of 24 hours of continuous data taken from February 20, 2018, to February 21, 2018, at 3 p.m. The data were taken at the AA4 sampling point in Kg. Chokoh Besar. The comparison between historical data and simulated data is summarised in Table 5, which shows the simulated concentrations of SO<sub>2</sub>, NO<sub>2</sub>, and CO are within the range of the historical data where the concentration of SO<sub>2</sub> is less than 30, while the concentration for both NO<sub>2</sub> and CO is less than 10. However, PM10 has a 63.67 square root mean error between the calculated and observed values.

One of the main factors contributing to this error is the meteorological data (Rao, 2005) because the meteorological data are obtained from Senai Station, where the distance between observation sites with the CFPP is about 30 km, and the reference height is 37.8 m. The second factor contributing to the high concentration reading is the location of the CFPP surrounded by other potential pollutant sources such as another power plant and port activities (Pelabuhan Tanjung Pelepas). Nevertheless, the simulated data are valid to be studied as possible adverse impacts on residents due to the pollutants emitted by the respective CFPP.

Table 5

*Comparison between simulated concentrations and historical data concentration*

Gases	Simulated data ( $\mu\text{g}/\text{m}^3$ )	Historical data ( $\mu\text{g}/\text{m}^3$ )
Sulphur dioxide ( $\text{SO}_2$ )	8.71	<30
Nitrogen dioxide ( $\text{NO}_2$ )	2.37	<10
Carbon monoxide (CO)	2.05	<10
Particulate matter ( $\text{PM}_{10}$ )	0.33	64

### Threat Zones for Worst-case Scenario by Using ALOHA Modelling

A graphical representation of the results is generated using ALOHA for  $\text{SO}_2$ ,  $\text{NO}_2$  and CO emissions. The modelled cases are then represented in Google Earth software to obtain an overview of the affected zone. Figure 3 shows that the concentration of  $\text{SO}_2$  dispersion does not exceed Acute Exposure Guideline Levels-3 (AEGL) and AEGL-2, where the threshold is  $78.6 \text{ mg}/\text{m}^3$  and  $1.965 \text{ mg}/\text{m}^3$ , respectively. However, the concentration of  $\text{SO}_2$  emission does exceed AEGL-1, where the concentration threshold is  $0.524 \text{ mg}/\text{m}^3$  for 60 min. A toxic zone whose concentration exceeds AEGL-1 reaches almost 3.5 linear km in the same wind direction and source point.  $\text{SO}_2$  AEGL-1 value is based on the No Observed Effect Level (NOEL) for bronchoconstriction in exercising asthmatics. It means that the affected population may have a disabling effect from  $\text{SO}_2$  exposure and only experience discomfort, irritation and certain asymptomatic (National Center for Biotechnology Information (NCBI), 2010).

In addition, the yellow line outside the threat zones in the Google Earth image is a wind direction confidence line. Wind direction confidence lines represent the range of uncertainty in the wind direction as wind rarely blows constantly from any one direction. In this simulated case, the wind direction confidence line is big due to the high potential of wind direction changes affected by low wind speed (EPA, 1999). The simulated case is also set at noon, where the stability class is B, which is moderately unstable.

Except for  $\text{SO}_2$  toxic threat zones, other gas pollutants such as  $\text{NO}_2$  and CO simulated outcomes show that the pollutant concentration does not exceed levels of concern (i.e., AEGL).  $\text{SO}_2$  gas pollutant emission rate is the highest compared to other gases. Gaussian



Figure 3. Toxic threat zone for the worst-case scenario of SO<sub>2</sub> emission

Plume Equation shows that the pollutant emission rate is directly proportional to the mean concentration of the diffusing substance at a point. Therefore, the higher the pollutant emission, the higher the ground level concentration and potentially exceed the levels of concern (LOC).

Furthermore, one of the factors for low gas concentration simulation outcomes is the stack height. In this study, the stack height is 200 m. According to Stack Height Vermont (2020), tall or adequate stack height is required to minimise the ground-level concentration of air pollutants. Hence, to obtain the ground-level concentration values for other gases, ALOHA's LOC is set to user-defined as below:

- (1) LOC-1 (yellow zone) = 0.001 mg/m<sup>3</sup>
- (2) LOC-2 (orange zone) = 0.01 mg/m<sup>3</sup>
- (3) LOC-3 (red zone) = 0.1 mg/m<sup>3</sup>

Figure 3 presents the results of the simulated NO<sub>2</sub> emissions. However, it is important to note a limitation of the ALOHA model, which does not make predictions for distances beyond 10 km from the release point (EPA, 1999). Therefore, the threat zone depicted in Figure 3 is truncated at the 10-km limit. Upon examination of Figure 4, it is evident that a zone exceeding LOC-3 extends up to approximately 5.3 km in the same wind direction and from the same source point. It indicates a higher level of concern within this zone. Moreover, it is worth mentioning that a zone surpassing both LOC-2 and LOC-1 extends beyond the 10-km mark. The simulation results also show that both threat points (Kampung Sungai Chengkeh and Kampung Sungai Boh) are in the LOC-3 threat zone.

Moving on to the simulation results for CO emission, as illustrated in Figure 5, it can be observed that the dispersion of CO emissions is slightly larger compared to NO<sub>2</sub>. It is evident in the threat zone, which extends beyond the LOC-3 threshold by approximately





Figure 4. Toxic zone for worst-case scenario of NO<sub>2</sub> emission

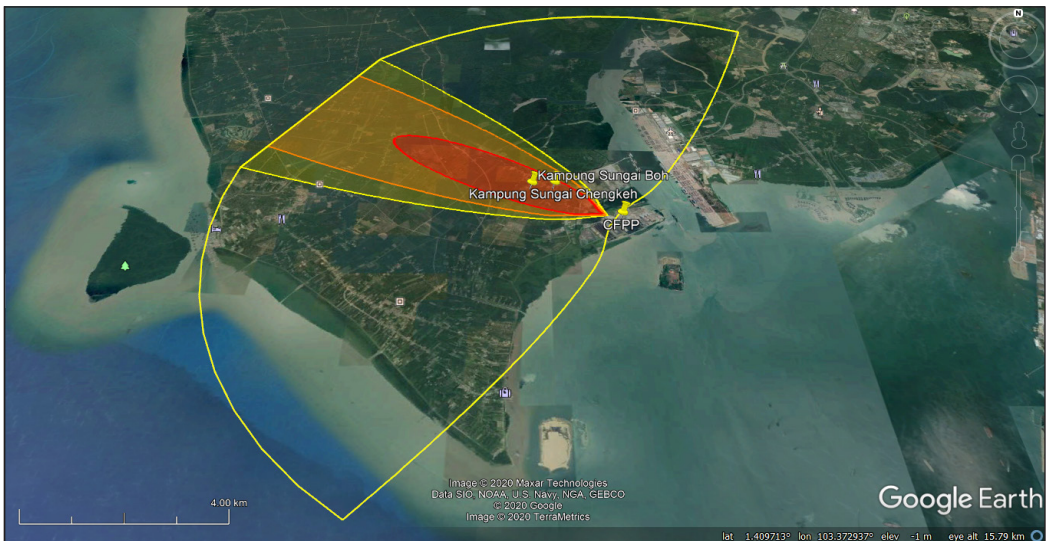


Figure 5. Threat zone for worst-case scenario of CO emission

1.8 km in the same wind direction and source point as NO<sub>2</sub>. Similar to the NO<sub>2</sub> dispersion modelling, the threat zones for LOC-2 and LOC-1 extend beyond the 10-km mark from the source point.

### Threat Zones for Best-case Scenario by Using ALOHA Modelling

In the best-case scenario, assessing the ground-level concentration of pollutants is essential. The ALOHA model's Level of Concern (LOC) parameter is set to user-defined values to



determine the ground-level concentration values. In this study, the following values were used for the LOC:

- (1) LOC-1 (yellow zone) =  $1\text{E-}12\text{ mg/m}^3$
- (2) LOC-2 (orange zone) =  $1\text{E-}10\text{ mg/m}^3$
- (3) LOC-3 (red zone) =  $1\text{E-}8\text{ mg/m}^3$

Figure 6 depicts the threat zones simulated for the best-case scenario of  $\text{SO}_2$  emissions. The results illustrate that Kampung Sungai Chengkeh falls within the threat zone LOC-1, indicating a relatively lower level of concern. On the other hand, Kampung Sungai Boh is observed to be within the threat zone LOC-2, indicating a slightly higher level of concern. Furthermore, it is worth noting that Kampung Sungai Boh is located closer to the downwind direction, which may affect the dispersion and concentration of  $\text{SO}_2$ .

The best-case scenario for  $\text{NO}_2$  emissions is illustrated in Figure 7, allowing the examination of the threat zones resulting from this emission. Upon closer inspection of Figure 7, it becomes apparent that Kampung Sungai Chengkeh does not fall within any identified threat zones. It suggests the concern for  $\text{NO}_2$  emissions in Kampung Sungai Chengkeh is relatively low. In contrast, Kampung Sungai Boh is within a threat zone that exceeds LOC-1, indicating a deeper concern for this area.

Figure 8 depicts the threat zones simulated for the best-case scenario of CO emissions. A similar result with the  $\text{SO}_2$  emission threat zones was obtained. Once again, Kampung Sungai Chengkeh is located within the threat zone designated as LOC-1. It implies a relatively lower concern for CO emissions in Kampung Sungai Chengkeh. Conversely, Kampung Sungai Boh is observed to be within the threat zone corresponding to LOC-2, indicating a moderately higher level of concern for this area in terms of CO emissions.



Figure 6. Toxic threat zone for the best-case scenario of  $\text{SO}_2$  emission



Figure 7. Toxic threat zone for the best-case scenario of NO<sub>2</sub> emission



Figure 8. Toxic threat zone for the best-case scenario of CO emission

## Health Risk Assessment

### Hazard Quotient (HQ) for Acute and Chronic Health Effects of Worst-case Scenario.

Figure 9 displays the estimated HQ values for acute health effects at Kampung Sungai Boh, while Figure 10 presents the HQ values for chronic health effects. Similarly, Figure 11

illustrates the estimated HQ values for acute health effects at Kampung Sungai Chengkeh, and Figure 12 showcases the HQ values for chronic health effects. It is important to note that when HQ values are less than 1, the likelihood of adverse health effects is considered low. Conversely, increased HQ indicates a higher probability of developing potential health risks. Furthermore, it is crucial to highlight that if HQ values exceed 10, the risk of chronic health effects is considered high (Chalvatzaki et al., 2019). Analysing the results for Kampung Sungai Boh (Figures 9 and 10), HQ values for adults and children exposed to SO<sub>2</sub> and NO<sub>2</sub> were observed to be above 1, suggesting a higher probability of potential health risks associated with these pollutants. On the other hand, HQ values for CO and

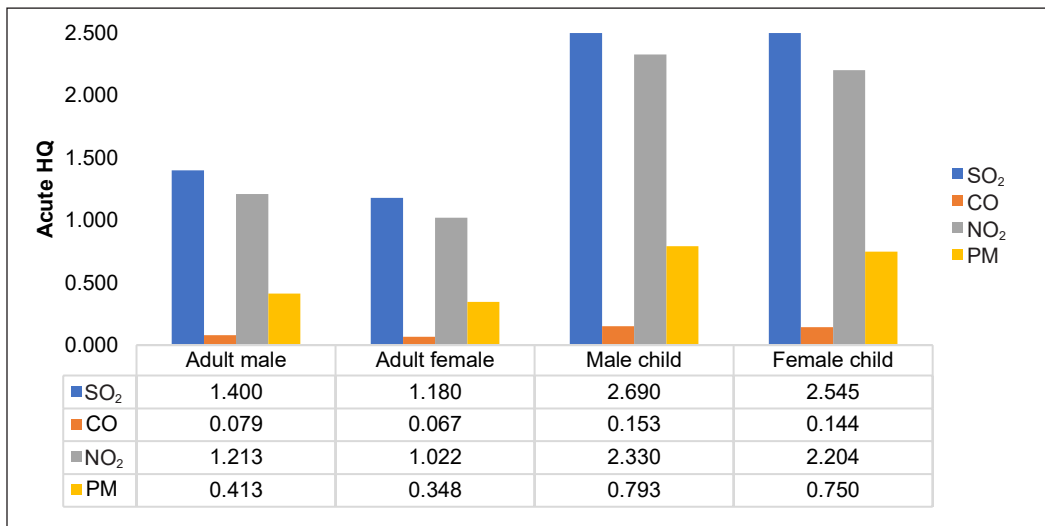


Figure 9. Worst-case scenario of acute HQ for different studied groups in Kampung Sungai Boh

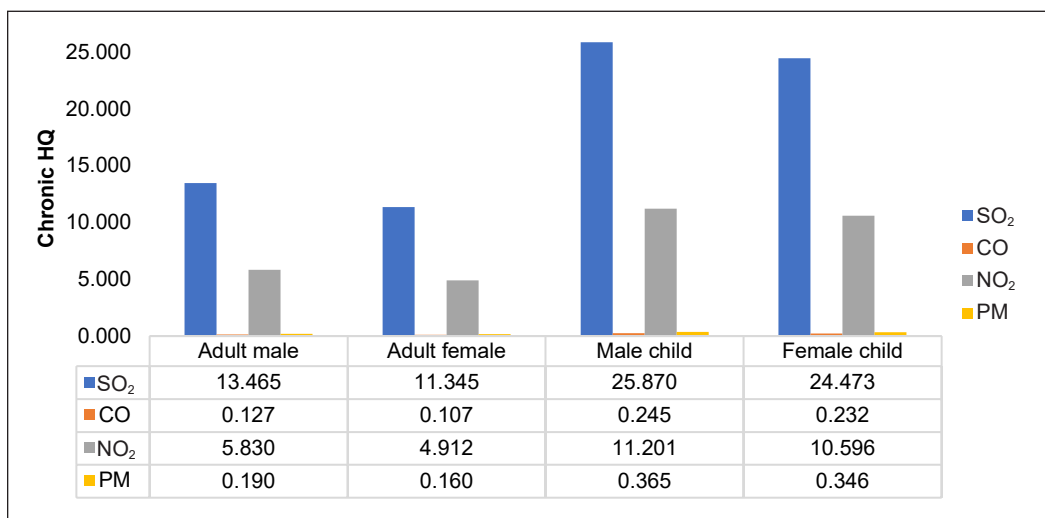


Figure 10. Worst-case scenario of chronic HQ for different studied groups in Kampung Sungai Boh

PM10 exposure in adults and children are below 1, indicating a relatively lower likelihood of adverse health effects from exposure to these gases.

Similar patterns emerge in the results for Kampung Sungai Chengkeh (Figures 11 and 12), where HQ values for NO<sub>2</sub> and SO<sub>2</sub> exposure exceed 1 for adults and children, while HQ values for CO and PM10 exposure remain below 1 for both acute and chronic health effects. Notably, HQ values for all age groups exposed to SO<sub>2</sub> in both villages and NO<sub>2</sub> exposure for children in both villages exceed 10, indicating a high chronic health risk.

Regarding the comparison across age groups, male children are at a higher risk than female children, similar to adults, where males are at a higher risk than females. These

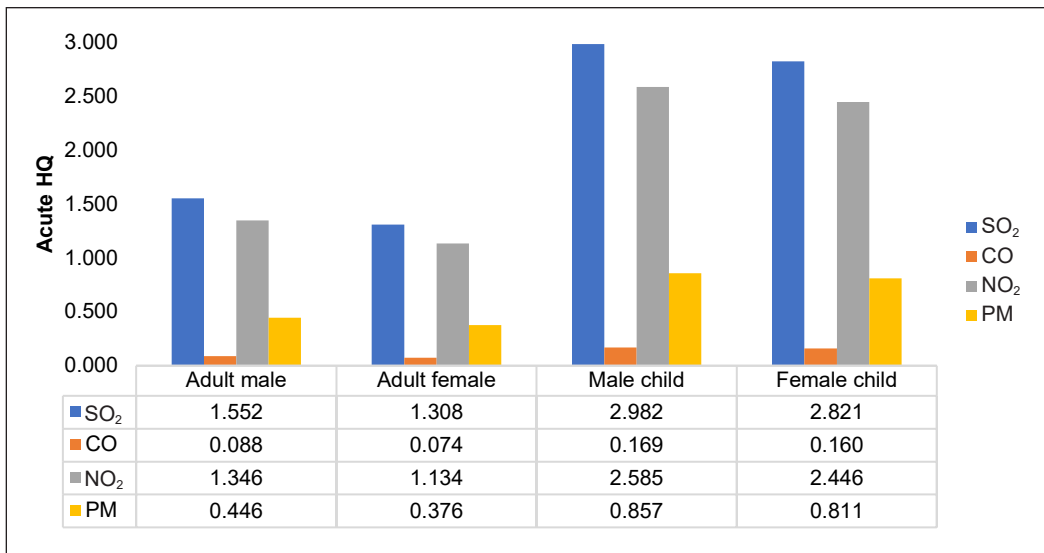


Figure 11. Worst-case scenario of acute HQ for different receptor groups in Kampung Sungai Chengkeh

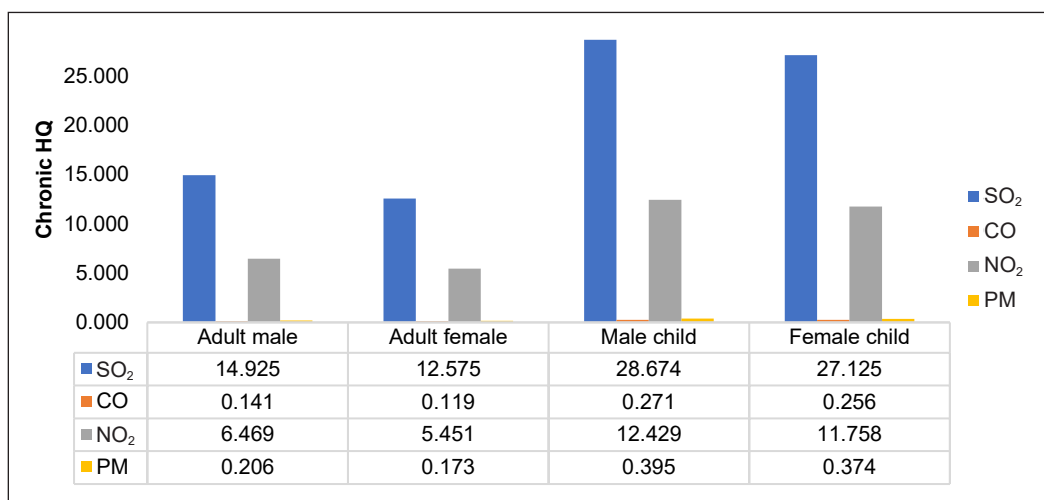


Figure 12. Worst-case scenario of chronic HQ for different receptor groups in Kampung Sungai Chengkeh



results can be attributed to the higher inhalation rate in males compared to females, resulting in a higher susceptibility to health risks from air pollution exposure. The findings illustrated in the four diagrams highlight that both short-term and long-term exposure to SO<sub>2</sub> pose the highest threat for both adults and children. It can be attributed to the high ground-level concentration of SO<sub>2</sub> and the low Reference Concentration (RfC) values associated with SO<sub>2</sub>. Consequently, the HQ values for SO<sub>2</sub> are significantly higher compared to the other gases. Conversely, the RfC values of CO for both acute and chronic exposures are relatively higher compared to the other three gases. As a result, the HQ values due to CO exposure are the lowest among the gases studied. These results suggest that neither adults nor children face a potential health risk when exposed solely to CO and PM10.

Compared to a study conducted in an urban-industrial area in South Africa, where the pollutants originated from a CFPP, metallurgical industries, and a manganese smelter, the acute HQ findings from our study show notable differences. Morakinyo et al. (2017) reported acute HQ values for SO<sub>2</sub> exposure of 0.1 for children and 0.07 for adults. Additionally, for acute HQ values of PM10 exposure, they reported 0.11 for children and 0.03 for adults. These results indicate that the acute HQ values from our study are considerably lower when compared to the study conducted in South Africa.

However, contrasting findings are observed for chronic HQ exposures, as the chronic HQ values from this study are substantially higher than those reported in the South African study. For instance, this study found a chronic HQ 449 for SO<sub>2</sub> exposure in children and 682 for adults. Furthermore, for chronic HQ values of PM10 exposure, this study reported 362 for children and 281 for adults. These significant differences can be attributed to various factors, such as the differing conditions of pollutant emission sources and the utilisation of different RfC values in the South African study. It is crucial to acknowledge that the South African study encompassed multiple pollutant sources, which may have contributed to the variations in the findings. Furthermore, using different RfC values, representing the safe exposure limits for chronic effects could also account for the differences in the chronic HQ values between the two studies.

### **Hazard Quotient (HQ) for Acute and Chronic Health Effects of Best-case Scenario.**

In the analysis of the best-case scenario, the assessment of acute and chronic health effects involved using specific values tabulated in Figures 13, 14, 15 and 16, which provide essential data for calculating the HQ for the mentioned effects. Upon careful examination of Figures 13, 14, 15 and 16, it becomes apparent that all the HQ values obtained are relatively low, suggesting that the potential health risks associated with the best-case scenario can be considered negligible. Interestingly, although the inhaled Average Daily Doses (ADDs) of CO are higher than those of ADDs of SO<sub>2</sub>, the acute and chronic HQs for SO<sub>2</sub> are significantly higher for both villages. This discrepancy arises from the fact that the

RfC values for SO<sub>2</sub> are 90% lower than the RfC values for CO. It is worth noting that the HQ values are inversely proportional to the RfC values. Therefore, the lower RfC values for SO<sub>2</sub> result in higher HQ values, indicating a potentially greater health risk associated with exposure to SO<sub>2</sub> in comparison to CO, despite the higher inhaled ADDs of CO in the best-case scenario.

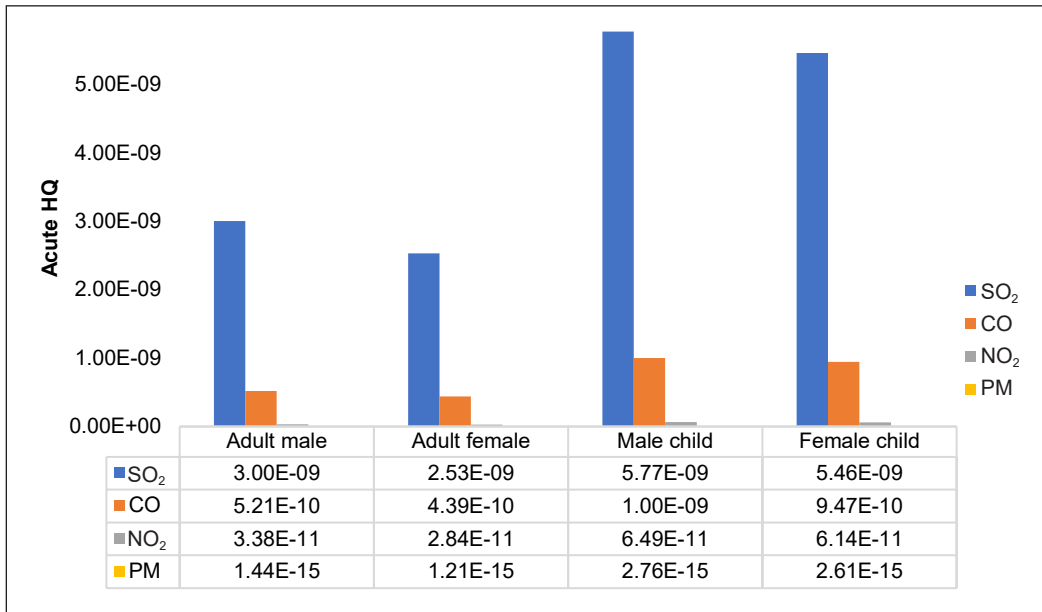


Figure 13. Best-case scenario of acute HQ for different studied groups in Kampung Sungai Boh

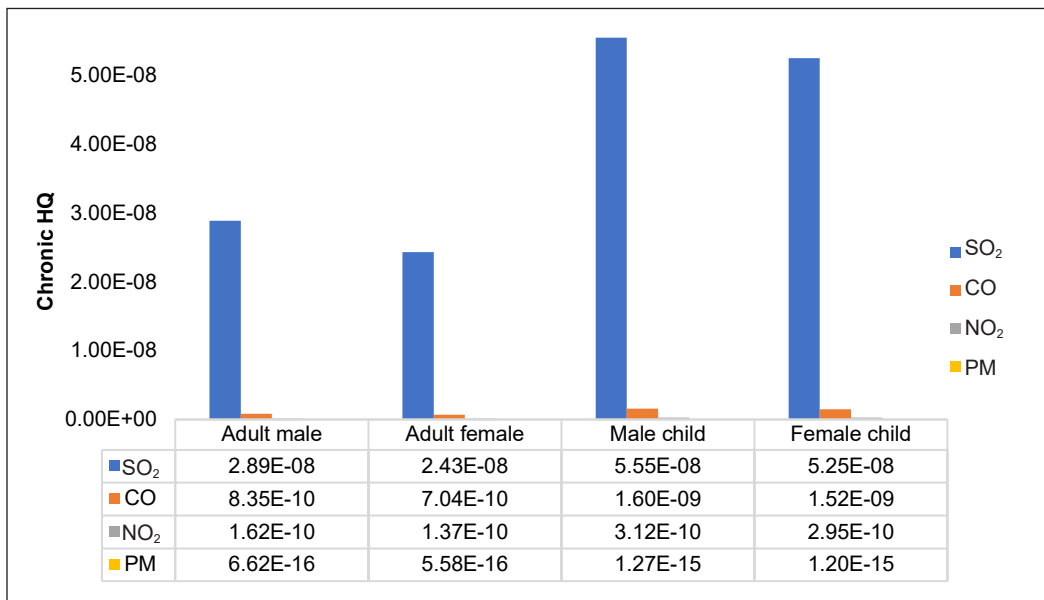


Figure 14. Best-case scenario of chronic HQ for different studied groups in Kampung Sungai Boh

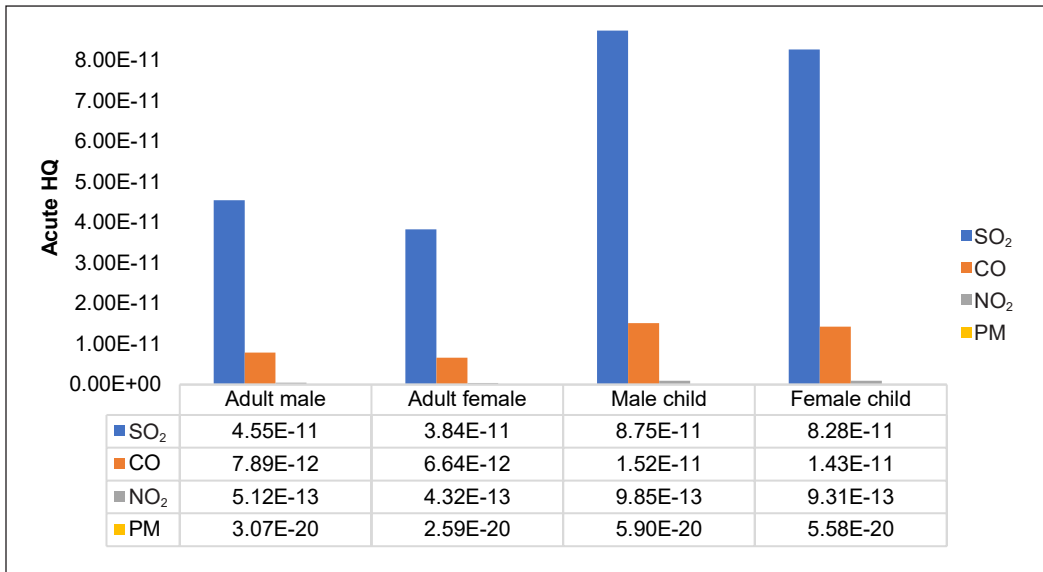


Figure 15. Best-case scenario of acute HQ for different studied groups in Kampung Sungai Chengkeh

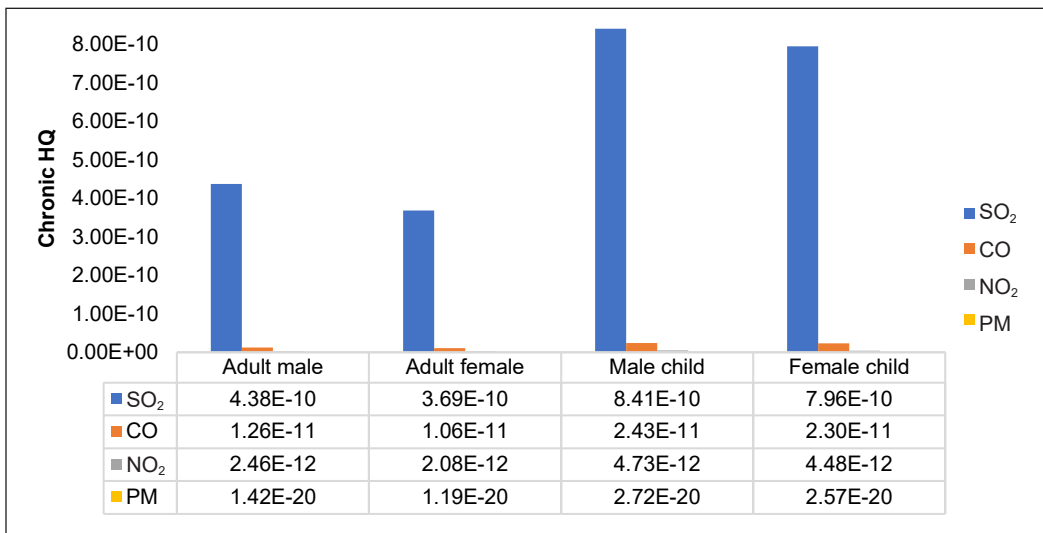


Figure 16. Best-case scenario of chronic HQ for different studied groups in Kampung Sungai Chengkeh

**Hazard Index (HI) for Acute and Chronic Health Effects.** Hazard Indexes (HI) were calculated using the previously determined HQ values to estimate the potential health risks associated with air pollutants. These HI represent the deposited dose resulting from the studied groups' exposure to the air pollutants. Figure 17 compares the acute exposure HI for different groups in Kampung Sungai Boh and the population residing in Kampung Sungai Chengkeh. On the other hand, Figure 18 displays the estimated HI for chronic exposure for all studied groups in both villages.



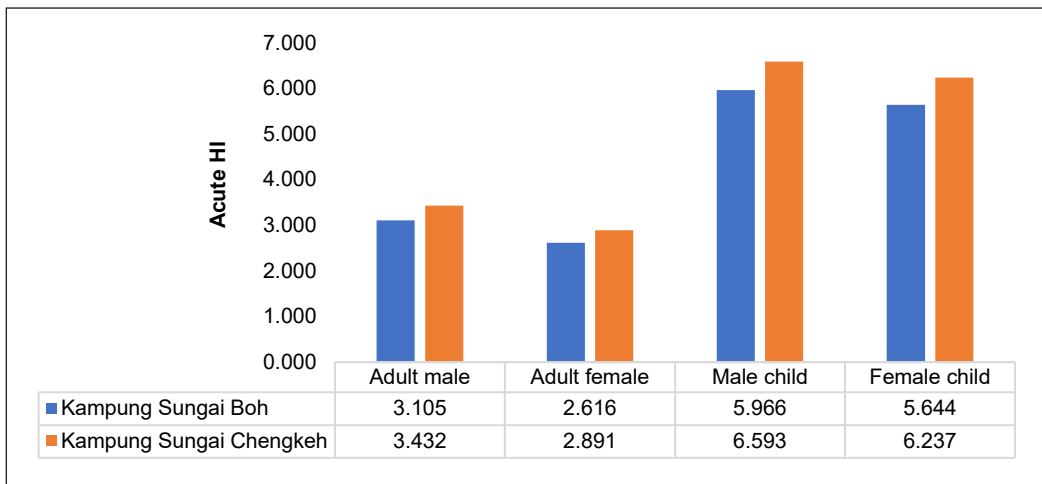


Figure 17. Comparison of acute exposure HI in Kampung Sungai Boh with Kampung Sungai Chengkeh

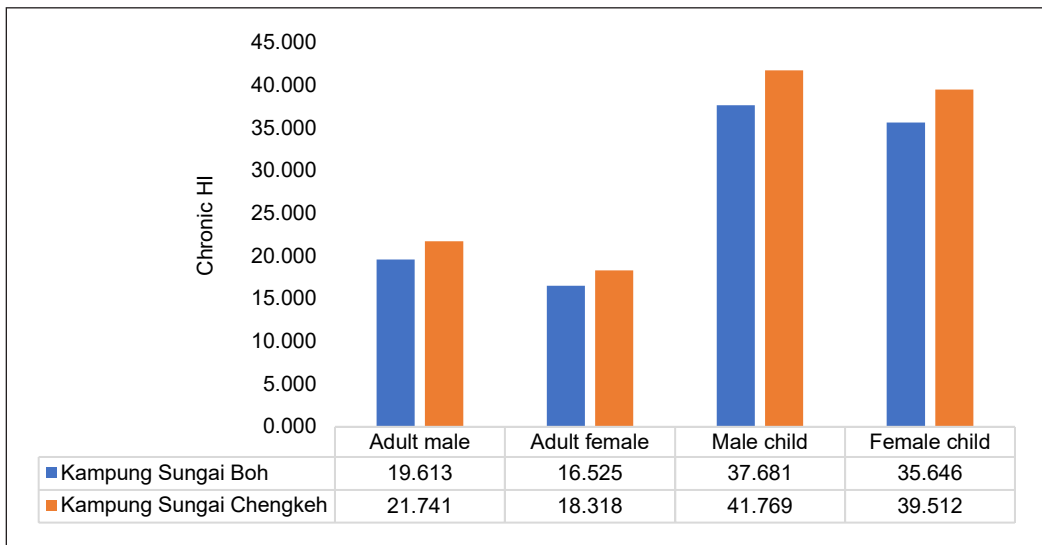


Figure 18. Comparison of chronic exposure HI in Kampung Sungai Boh with Kampung Sungai Chengkeh

According to the HQ recorded, it can be observed that SO<sub>2</sub>, NO<sub>2</sub>, PM<sub>10</sub> and CO accounted for approximately 45%, 39%, 13%, and 3%, respectively, of the HI values for adult males, adult females, male children, and female children in both threat areas. Regarding chronic exposure HI, the HQ values for SO<sub>2</sub> accounted for more than 50% of the total, followed by NO<sub>2</sub> (30%), CO (0.65%), and PM<sub>10</sub> (1%) for all studied groups. It was found that the values estimated for Kampung Sungai Chengkeh were 9.5% higher than those calculated for Kampung Sungai Boh when comparing the acute HI values. Similarly, for chronic health effects, the HI at Kampung Sungai Chengkeh were 9.8% higher than those at Kampung Sungai Boh for all studied groups. This observation aligns

with the consistency of ground-level concentration data, as indicated in Figure 18, where Kampung Sungai Chengkeh exhibited approximately 10% higher estimated concentration values compared to Kampung Sungai Boh.

Based on the recorded findings, it is concerning that the estimated HI values for all studied groups are significantly above one, indicating a likelihood of adverse health effects. The SO<sub>2</sub>, NO<sub>2</sub>, and CO exposures are particularly concerning as they target the cardiovascular and respiratory systems, suggesting a high risk of cardiovascular and respiratory diseases for all individuals residing in Kampung Sungai Boh and Kampung Sungai Chengkeh.

However, it should be noted that the HI values for the best-case scenario are not included in this document due to their significantly low levels. The range of acute HI for both villages varies from 4.54E-11 (for an adult female residing in Kampung Sungai Chengkeh) to 6.84E-09 (for a male child residing in Kampung Sungai Boh). Similarly, for chronic HI values, adult females in Kampung Sungai Chengkeh have the lowest value of 3.81E-10, while male children in Kampung Sungai Boh have the highest value of 5.74E-08. These findings are influenced by the ADD values discussed in Figure 18, where male children residing in Kampung Sungai Boh have the highest values, while women in Kampung Sungai Chengkeh have the lowest.

## CONCLUSION AND RECOMMENDATIONS

Reducing people's exposure to air pollution outdoors would significantly reduce mortality rates (Robert & Kevin, 2022). The uncontrolled release of SO<sub>2</sub> from SO<sub>x</sub> has the same deadly effects on plants and animals as acid rain and causes various illnesses in humans, such as irregular heartbeat, skin cancer, asthma, cough, headache, and throat and nose irritations. Another vital pollutant from coal power plants that generate electricity is NO<sub>x</sub>, which leads to fatal hypoxia in the respiratory system (Munawer, 2018). Most countries' greenhouse gas emissions come from power plants.

There is an immediate need for government action to reduce these emissions. Miller and Sullivan (2007) suggest that utilities, society, and the environment would benefit significantly from transitioning from energy generation based on fossil fuels to renewable fuels like hydropower. For example, the government may implement emissions tax to fund the construction of renewable energy power plants or the reforestation of Malaysian rain forests. The findings of this study have provided a solid foundation for building a programme to reduce emissions in Malaysia and conduct a cost-benefit analysis of switching to renewable energy sources for power generation. Furthermore, the impacts of pollution and particle matter can be reduced through individual interventions such as using face masks and air purifiers, regular physical activity, and a nutritious diet (Combes & Franchineau, 2019).

Regulations for health and the environment (protocols) should be defined globally to address these issues and encourage the widespread use of coal for power generation. Developing sound health and environmental safety protocols and providing adequate health and safety education will go a long way toward mitigating the effects on both the private and public sectors (Munawar, 2018). In addition, improved inhalation models are required to comprehend the impact of particle size and chemistry of various minerals on lung deposition and health impacts. Learning as much as possible about the health risks and exposure to PM from mining operations is essential to create estimating equations and a model with varying heights from the source. The outcomes can improve working conditions and environmental health (Patra et al., 2016).

This research demonstrates how prolonged exposure to air pollution causes health problems across all socioeconomic strata. This research backs up prior findings and suggests that the extent of the health effects may be more than previously thought. The findings suggest that reducing people's chronic exposure to fine particle pollution is necessary. Given that we only have data from a single CFPP and the authors used the participants' primary ZIP codes to determine their exposure levels, our estimation still has room for error. Despite the authors' best efforts to eliminate any potentially biased sensors, the degree to which ambient pollution monitors reflect the exposures of the subjects is imperfect. This research lacked information that would have allowed them to evaluate changes in exposure based on factors such as the possibility that some participants may have relocated and the subjects' activity and location, such as the amount of time they spent in traffic and indoors. Though they play a role in measurement and misclassification errors, they are unlikely to have introduced a bias that would account for the study's results.

## ACKNOWLEDGEMENT

The authors would like to thank the selected coal-fired power plants in Malaysia for providing the data used in this study.

## REFERENCES

- Azmi, Y., M., Junidah, R., Mariam, S., Safiah, M. Y., Fatimah, S., Norimah, A. K., Poh, B. K., Kandiah, M., Zalilah, M. S., Manan, W. M. W. A., Siti, H. M. D. & Tahir, A. (2009). Body mass index (BMI) of adults: Findings of the Malaysian adult nutrition survey (MANS). *Malaysian Journal of Nutrition*, 15(2), 97-119.
- Babatunde, K. A., Said, F. F., Nor, N. G., & Begum, R. A. (2018). Reducing carbon dioxide emissions from Malaysian power sector: Current issues and future directions. *Jurnal Kejuruteraan*, 1(6), 59-69. [http://dx.doi.org/10.17576/jkukm-2018-si1\(6\)-08](http://dx.doi.org/10.17576/jkukm-2018-si1(6)-08)
- Badman, D. D. G., & Jaffé, D. E. R. (1996). Blood and air pollution: State of knowledge and research needs. *Otolaryngology-Head and Neck Surgery*, 114(2), 205-208. <https://doi.org/10.1016/S0194-59989670166-3>

- Chalvatzaki, E., Chatoutsidou, S., Lehtomäki, H., Almeida, S., Eleftheriadis, K., Hänninen, O., & Lazaridis, M. (2019). Characterization of human health risks from particulate air pollution in selected European cities. *Atmosphere*, 10(2), Article 96. <https://doi.org/10.3390/atmos10020096>
- Clancy, L., Goodman, P., Sinclair, H., & Dockery, D. W. (2002). Effect of air-pollution control on death rates in Dublin, Ireland: An intervention study. *The Lancet*, 360(9341), 1210-1214. [https://doi.org/10.1016/S0140-6736\(02\)11281-5](https://doi.org/10.1016/S0140-6736(02)11281-5)
- Combes, A., & Franchineau, G. (2019). Fine particle environmental pollution and cardiovascular diseases. *Metabolism*, 100, Article 153944. <https://doi.org/10.1016/j.metabol.2019.07.008>
- Das, A., Singh, G., Habib, G., & Kumar, A. (2018). Non-carcinogenic and carcinogenic risk assessment of trace elements of PM<sub>2.5</sub> during winter and pre-monsoon seasons in Delhi: A case study. *Exposure and Health*, 12, 63-77. <https://doi.org/10.1007/s12403-018-0285-y>
- EPA. (1999). *ALOHA User's Manual*. US Environmental Protection Agency. <https://nepis.epa.gov/Exec/zyPURL.cgi?Dockey=P1003UZB.TXT>
- Gething, P. W., Smith, D. L., Patil, A. P., Tatem, A. J., Snow, R. W., & Hay, S. I. (2010). Climate change and the global malaria recession. *Nature*, 465(7296), 342-345. <https://doi.org/10.1038/nature09098>
- Hashim, J. H., & Hashim, Z. (2010). *Guidance document on health impact assessment (HIA) in environmental impact assessment (EIA)*. Department of Environment Malaysia. <https://mycep.doc.gov.my/guidelines/68e07de05a228b807477affcfae6685c.pdf>
- Höglund, L., Räisänen, J., Hämäläinen, A. M., Warholm, M., Hagen, M. van der, Suleiman, A., Kristjánsson, V., Nielsen, E., & Kopp, T. I. (2012). *Existing Default Values and Recommendations for Exposure Assessment - A Nordic Exposure Group Project 2011*. Nordic Council of Ministers. <https://doi.org/10.6027/TN2012-505>
- Hurt, S., Ollinger, J., Arce, G., Bui, Q., Tobia, A. J., & van Ravenswaay, B. (2001). Chapter 81 - Dialkylidithiocarbamates (EBDCs). In R. Krieger, J. Douli, E. Hodgson & H. Maibach, (Eds.), *Handbook of Pesticide Toxicology (Second Edition)* (pp. 1702-1779). Elsevier. <https://doi.org/10.1016/B978-012426260-7.50084-7>
- Kelsall, J. E., Samet, J. M., Zeger, S. L., & Xu, J. (1997). Air pollution and mortality in Philadelphia, 1974-1988. *American Journal of Epidemiology*, 146(9), 750-762. <https://doi.org/10.1093/oxfordjournals.aje.a009351>
- Kenessary, D., Kenessary, A., Adilgireiuly, Z., Akzholova, N., Erzhanova, A., Dosmukhametov, A., Syzdykov, D., Masoud, A. R., & Saliev, T. (2019). Air pollution in Kazakhstan and its health risk assessment. *Annals of Global Health*, 85(1), Article 133. <https://doi.org/10.5334/aogh.2535>
- Khalid, N. A., Tohir, M. Z. M., & Said, M. S. M. (2019). Consequence simulation of jet fire due to leakage of pipelines in a natural gas power plant in Malaysia. *PERINTIS eJournal*, 9(1), 15-29.
- Levy, J. I., Hammitt, J. K., Yanagisawa, Y., & Spengler, J. D. (1999). Development of a new damage function model for power plants: Methodology and applications. *Environmental Science & Technology*, 33(24), 4364-4372. <https://doi.org/10.1021/es990634>
- Mahlia, T. M. I. (2002). Emissions from electricity generation in Malaysia. *Renewable Energy*, 27(2), 293-300. [https://doi.org/10.1016/S0960-1481\(01\)00177-X](https://doi.org/10.1016/S0960-1481(01)00177-X)

- Malik, A. A., Nasif, M. S., Mokhtar, A. A., & Tohir, M. Z. M. (2021). Numerical investigation of the effect of weather conditions on the escalation and propagation of fire-induced domino effect. *Process Safety Progress*, 40(4), 296-308.
- Malik, A. A., Nasif, M. S., Arshad, U., Mokhtar, A. A., Tohir, M. Z. M., & Al-Waked, R. (2023). Predictive modelling of wind-influenced dynamic fire spread probability in tank farm due to domino effect by integrating numerical simulation with ANN. *Fire*, 6(3), Article 85. <https://doi.org/10.3390/fire6030085>
- Miller, K. A., & Sullivan, J. H. (2007). Long-term exposure to air pollution and incidence of cardiovascular events in women. *The New England Journal of Medicine*, 356(5), 447-458. <https://doi.org/10.1056/NEJMoa054409>
- Ministry of Natural Resources and Environment. (2015). *Malaysia Biennial Update Report to the UNFCCC*. <https://unfccc.int/sites/default/files/resource/MALBUR1.pdf>
- Morakinyo, O. M., Mokgobu, M. I., Mukhola, M. S., & Engelbrecht, J. C. (2017). Health risk assessment of exposure to ambient concentrations of benzene, toluene and xylene in Pretoria West, South Africa. *African Journal of Science, Technology, Innovation and Development*, 9(4), 489-496. <https://doi.org/10.1080/20421338.2017.1352123>
- Munawar, M. E. (2018). Human health and environmental impacts of coal combustion and post-combustion wastes. *Journal of Sustainable Mining*, 17(2), 87-96. <https://doi.org/10.1016/j.jsm.2017.12.007>
- National Center for Biotechnology Information (NCBI). (2010). *National Research Council (US) Committee on Acute Exposure Guideline Levels. Acute Exposure Guideline Levels for Selected Airborne Chemicals: Volume 8*. National Academies Press. <https://www.ncbi.nlm.nih.gov/books/NBK219999/>
- Patra, A. K., Gautam, S., & Kumar, P. (2016). Emissions and human health impact of particulate matter from surface mining operation - A review. *Environmental Technology & Innovation*, 5, 233-249. <https://doi.org/10.1016/j.eti.2016.04.002>
- Peters, A., Perz, S., Doring, A., Stieber, J., Koenig, W., & Wichmann, H. E. (1999). Increases in heart rate during an air pollution episode. *American Journal of Epidemiology*, 150(10), 1094-1098. <https://doi.org/10.1093/oxfordjournals.aje.a009934>
- Ranjan, M. Z., Baharudin, B. T. H. T., Mahadi, M. R., & Baharudin, M. R. (2019). Developing a construction occupational safety and health risk assessment matrix (COSHRAM) with modifying risk factors. *International Journal of Recent Technology and Engineering*, 8, 301-307.
- Rao, K. S. (2005). Uncertainty analysis in atmospheric dispersion modelling. *Pure and Applied Geophysics*, 162, 1893-1917. <https://doi.org/10.1007/s00024-005-2697-4>
- Ren, T., & Patel, M. K. (2009). Basic petrochemicals from natural gas, coal and biomass: Energy use and CO<sub>2</sub> emissions. *Resources, Conservation and Recycling*, 53(9), 513-528. <https://doi.org/10.1016/j.resconrec.2009.04.005>
- Robert, J. L., & Kevin R. C (2022). Personal interventions to reduce exposure to outdoor air pollution. *Annual Review of Public Health*, 43(1), 293-309. <https://doi.org/10.1146/annurev-publhealth-052120-103607>

- Shekarchian, M., Moghavvemi, M., Mahlia, T. M. I., & Mazandarani, A. (2011). A review on the pattern of electricity generation and emission in Malaysia from 1976 to 2008. *Renewable and Sustainable Energy Reviews*, 15(6), 2629-2642. <https://doi.org/10.1016/j.rser.2011.03.024>
- Standard Australia Committee and Standard New Zealand Committee. (2015). *Joint Standards Australia/Standards New Zealand Committee EV-007: Methods for Examination of Air. Stationary Source Emissions Part 3: Determination of Odour Concentration by Dynamic Olfactometry*. <https://www.saiglobal.com/PDFTemp/Previews/OSH/as/as4000/4300/43233.pdf>
- The IBR Asia Group Sdn. Bhd. (2019). *Towards a World-Class Energy Sector* [ST(P)04/03/2019; p. 52]. Energy Commission. <https://fliphtml5.com/vhro/clqa>
- Vermont. (2020). *Stack height and rain guard guidance*. Department of Environmental Conservation. <https://dec.vermont.gov/air-quality/permits/construction/stack-height-and-rain-guards>
- Yakubu, O. (2017). Addressing environmental health problems in Ogoniland through implementation of united nations environment program recommendations: Environmental management strategies. *Environments*, 4(2), Article 28. <https://doi.org/10.3390/environments4020028>
- Yang, W., Burrows, T., MacDonald-Wicks, L., Williams, L., Collins, C., Chee, W., & Colyvas, K. (2017). Body weight status and dietary intakes of urban Malay primary school children: Evidence from the family diet study. *Children*, 4(1), Article 5. <https://doi.org/10.3390/children4010005>
- Zubir, A. A. M., Rusli, N. S., Alkarkhi, A. F. M., & Yusup, Y. (2017). Emission inventory for power plants and passenger cars in peninsular Malaysia for the years 2008-2014. In. *Proceedings of the 5th International Conference on Environmental Research and Technology* (pp. 404-410). School of Industrial Technology, USM.

## Proportions of Green Area and Tree Health on University Campus: The Impact of Pavement Presence

Nur Ain Aiman Abd Rahim<sup>1</sup>, Rahmad Zakaria<sup>1</sup>, Asyraf Mansor<sup>1</sup>, Mohd Ashraf Mohamad Ismail<sup>2</sup> and Nik Fadzly N Rosely<sup>1\*</sup>

<sup>1</sup>*School of Biological Sciences, Universiti Sains Malaysia, 11800 USM, Minden, Penang, Malaysia*

<sup>2</sup>*School of Civil Engineering, Universiti Sains Malaysia, 14300 USM, Nibong Tebal, Penang, Malaysia*

### ABSTRACT

Sustainable urban ecosystems require healthy green spaces that provide ecological services to meet social and environmental needs. This study evaluates the proportion of green areas in the urban campus and assesses the effects of pavement on the health of *Samanea saman*. The MAPIR Normalised Difference Vegetation Index (NDVI) camera mounted on a drone captured a proportion of green areas of approximately 27.80%, which is roughly a quarter of the total study sites. However, the significance of green areas depends on the health of the trees. Therefore, the pavement effects on the health status of trees in green areas were studied using Visual Tree Assessment (VTA) and aerial image analysis using the MAPIR (NDVI) camera on a drone. Although both methods evaluate health status, the final outputs differed. VTA produced categorical outputs, which assigned trees into health categories based on a visual assessment of factors. In contrast, NDVI produced linear outputs, which provided a numerical value to demonstrate tree health. Both methods indicate that trees in non-paved areas are healthier, particularly for the excellent trees identified by the VTA, which suggests that pavement negatively impacts tree health. However, the effects of

paved and non-paved areas on tree health status analysis using the aerial image are not significant ( $p > 0.05$ ), which may be due to the low quality and accuracy of the images. The study provides insights into the importance of green areas and tree health in creating sustainable urban ecosystems.

**Keywords:** Green areas, non-paved, normalised difference vegetation index, paved, tree health, visual tree analysis

### ARTICLE INFO

#### Article history:

Received: 22 January 2023

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.11>

#### E-mail addresses:

ainaiman@student.usm.my (Nur Ain Aiman Abd Rahim)

rahmadz@usm.my (Rahmad Zakaria)

asyrafm@usm.my (Asyraf Mansor)

ceashraf@usm.my (Mohd Ashraf Mohamad Ismail)

nroselnik@outlook.com (Nik Fadzly N Rosely)

\* Corresponding author



## INTRODUCTION

Universiti Sains Malaysia (USM) is an urban space incorporating green areas and infrastructure as part of the landscape. USM green areas are predominantly surrounded by mature trees characterised by various trees such as *Samanea saman*, *Andira inermis*, and *Millettia pinnata*, including *Peltophorum petrocarpum* and *Pterocarpus indicus*. Considering that the ornamental trees in USM were planted, typically at the ages of 1 to 3, and the institution was established in 1969, the ornamental trees planted in the area may range from 54 to 60 years old. This age may be considered old for some species, depending on various factors such as their general lifespan, environmental conditions, basic needs, and other factors. Additionally, it is worth noting that naturally occurring trees that have been preserved may be older.

The potential problem that arises from the tree is inevitable, especially during storms. For example, according to Hauer and Peterson (2016) and Vogt et al. (2015), trees can cause damage to grey infrastructure, and managing tree-infrastructure conflicts is a significant concern of arboriculture practitioners. Tree problems generally include conflicts with infrastructure, potential impacts on human health and safety, and cultural, aesthetic, and social concerns. Moreover, these problems are compounded by environmental and energy-related challenges, resulting in a complex set of problems (Rotherham, 2020; Vaz et al., 2017). Research on tree risk management is essential to evaluate and reduce these disadvantages (Klein et al., 2019). When trunks and tree limbs fall on people, vehicles, and structures, big trees cause fatalities and injuries (Schmidlin, 2009). On the other hand, in Toronto, Canada, residents express storm-related safety worries, although deaths are uncommon (Conway & Yip, 2016).

*Samanea saman*, *P. pterocarpum*, *Cinnamomum iners*, *Ficus benjamina*, and other trees have been recorded as Malaysian cities' most prevalent ornamental trees (Hasan et al., 2018). *S. saman*, commonly called Rain Tree, is not native to Malaysia, but this species' existence is not harmful to native species. On the contrary, this species provides beneficial effects for urban ecosystems. This vast tree provides shade with its wide canopy. Matured *S. saman* is usually large, with a dome-like crown. However, this tree species tended to cause problems such as split branches and exposed roots that cause cracks in the pavement or, even worse, exacerbate the buildings' foundation. Concerns were raised due to the age and large tree size. In Malaysia, trees with a large diameter at breast height (DBH) are often considered large (Seng et al., 2004). Therefore, health monitoring of the affected trees is required. Despite being a time-consuming and demanding process, Visual Tree Assessment (VTA) remains a reliable traditional method for determining the health status of trees.

VTA takes into account numerous parameters. For example, tree characteristics which are the shape and density of the crown structure, the colour, texture, and density of the foliage, the texture and density of the bark, the presence of visible wounds, decay, or cavities in the trunk, the presence of any visible roots, the size, structure, and signs of root damage. Furthermore, other parameters are structural defects such as in the tree, such as cracks, split or leaning, and

site conditions. In addition, pest and disease status is one parameter that impacts overall health. Considering all the above factors, the final parameter is the tree's overall health and vigour. Another method, the Normalised Difference Vegetation Index (NDVI), which describes the difference between visible and near-infrared reflectance of plant cover, only considers green foliage pigment. Unlike NDVI, VTA considers the health of root systems, trunks and branches, hazards, and surroundings. Roots' health is affected by ground conditions.

However, the observations depend on the arborist or researcher, increasing the possibility of human errors. Besides the VTA's credibility in forecasting tree health, the NDVI concept is the difference between visible and infrared reflectance of vegetation. It can also estimate the density of green foliage from the leaves on an area of land (Weier & Herring, 2000). Moreover, NDVI has been used to observe the impacts of climate on vegetation biomass and phenological trends (Wang et al., 2003). Unlike other remote sensing devices, drones operated with the NDVI concept produce reliable data for health monitoring with fast and easy handling. Furthermore, lightweight drones are low in cost and have high flexibility for vegetation evaluation and deployment in harsh environments (Assmann et al., 2019; Duffy et al., 2018).

Urban green space loss threatens environmental sustainability and human health. Therefore, university campuses significantly contribute as urban sites that might promote sustainability and green infrastructure. This study evaluates the proportions of green space on USM's main campus using a drone equipped with NDVI technology. It examines how pavement affects the health of raintrees (*S. saman*). Assessing green space proportions is essential for understanding urban areas' environmental sustainability and quality of life and can inform urban planning and management strategies. Pavements can significantly negatively impact tree health by restricting water and nutrient absorption, leading to soil compaction and root damage. The results of this study can also contribute to broader efforts to promote the health and sustainability of urban green spaces.

The project involved collecting health data on paved and non-paved trees by performing aerial imagery analysis and VTA on Rain trees as the tree model to have a different perspective and investigate potential correlation. Furthermore, by providing insights into the relationship between pavement coverage and tree health, this study aims to inform the development of more effective urban planning and management strategies for promoting the health and sustainability of urban green spaces. The results of this study can also contribute to the broader field of knowledge on the benefits of urban green areas for environmental sustainability and human health.

## **MATERIALS AND METHODS**

### **Study Site**

This study was conducted on the main campus areas of the USM, located in Penang, Malaysia. The main campus areas of USM comprise a variety of green spaces such

as gardens, parks, and natural reserves. Additionally, USM is in Penang, a rapidly urbanising area in Malaysia, making it a particularly relevant site for studying the effects of urbanisation on natural environments. Based on the USM website = the whole campus is 240.13 hectares, including the RST compound, which is not included in this study. The calculations based on Google Maps are about 145 hectares (Figure 1).

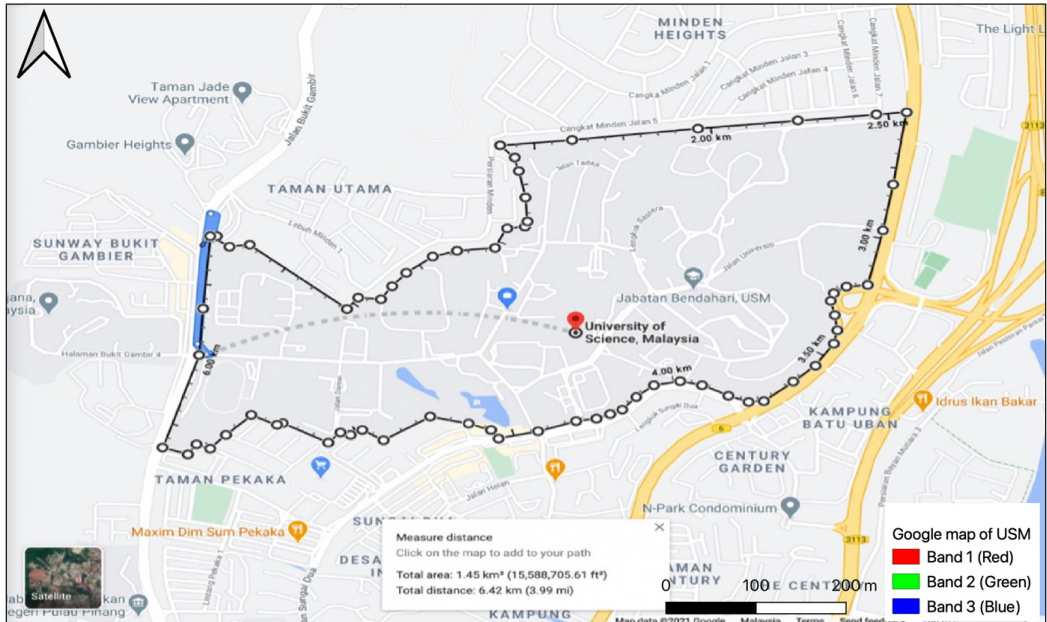


Figure 1. Google map of USM, Gelugor

### Drone Mapping

The data collection for Unmanned Aerial Vehicle (UAV) mapping was conducted from the 17<sup>th</sup> to 19<sup>th</sup> of November 2021 around the property of USM, Gelugor (5.3559° N, 100.3025° E), with an estimated area of 188.52 hectares. The methodology flowchart in Figure 2 shows the steps to acquire the proportions of green areas. We used a DJI Phantom 4 Drone. DJI Ground Stations (GS) Pro was used as the ground control app to create the waypoints for flight at 120 m height. OFO Sdn Bhd, a local drone survey company, provided proper flight permits and protocols.

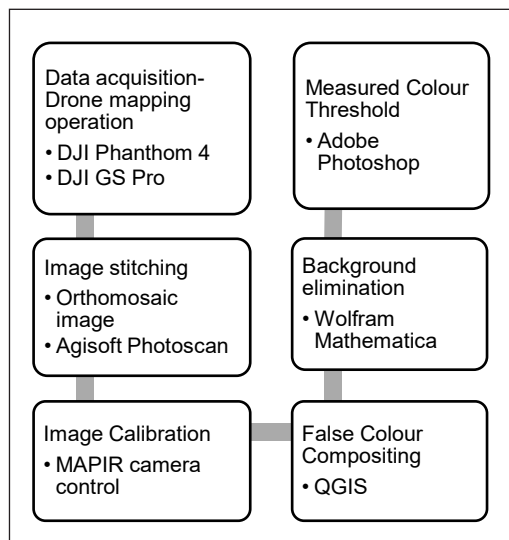


Figure 2. Flowchart of NVDI to measure Green Area Proportion

## Identification of Green Area Composition

The images were captured with a resolution of around 20 megapixels in multiple spectral bands. The mapped UAV data was stitched using Agisoft Photoscan for the first objective. Next, the overlapped images were collected in each flight mission to cover the whole area as ortho-mosaic images. Then, NVDI images were calibrated using MAPIR camera control software (<https://www.mapir.camera/>) and FIJI (<https://www.mapir.camera/pages/calibrating-survey-images>) and further analysed by QGIS. Later, the `SeparateColour` command in Wolfram Mathematica Image processing was applied to simplify the data to reduce background colour bias. The resulting images are in digital formats, such as JPEG or TIFF, with around 20 megapixels in multiple spectral bands. Finally, the Photoshop colour threshold command was used to calculate the percentage of a green area. The green space in this project refers to the regions occupied by vegetation from the aerial view.

## Tree Health Status on Paved and Non-paved Surface of *S. saman*

To assess the effects of pavement on tree health, we employed a two-pronged methodology. For the second objective, we comprehensively evaluated the health of *S. saman* trees in scattered locations on the USM campus, given that 30 trees were covered by pavement and the other 30 were not. It involved a VTA, directly observing and measuring tree health indicators. Additionally, we used the NDVI analysis of aerial imagery to extract pixel density information, which was used to calculate a quantitative measure of tree health. VTA provide qualitative data on tree health, and the NDVI analysis provides quantitative data on pixel density and overall tree health. VTA uses several parameters, such as the health of roots, trunk, leaves, crown and branches structures, tree characteristics and surrounding hazards. In addition, crown class, tree shapes, leaf density, and tree fertility were also assessed. Out of the 184 *S. saman* individuals, a random sample of 60 was chosen for the study. Thirty trees were selected in the areas covered with pavement, while the other thirty were not covered by pavement. The trees were recorded and analysed from 1/3/2021 to 15/3/2021 around six varying ecological landscape spots in the USM Gelugor Campus.

Two different surface conditions were assessed: paved and non-paved areas. Both conditions were studied in 3 different spots but shared the same characteristics. Paved spots are usually near the road, pavement, and footpath. Besides, they are commonly located in the central area where heavy traffic and vehicles are. In comparison, the non-paved area is located near the open space near the lake, field, and the edge of the university compound. The trees were diverse regarding species, but only the urban trees, *S. saman*, were studied. *S. saman* is abundant and easy to access throughout the property. It is an urban tree with a large umbrella canopy providing shade to pedestrians. This tree landscape was designed along the road on a paved area and in the lake's open spaces, including the fields. In paved areas, the trees were planted for pedestrians, while in non-paved areas, *S. saman* was planted

around the recreational place, especially the lakes and fields. These trees were planted in both areas, most likely for shade and to maintain thermal conditions.

The information from the landscape condition of the trees is valuable as an indication of a landscape condition influencing tree health. Thus, the paved and non-paved areas' trees health was compared. The NVDI in black and white bands was further evaluated using the ImageJ measure command with integrated density and then compared with the Visual Tree Assessment (VTA) to see the correlations. VTA depends on several chosen parameters that significantly influence the well-being status classified in tree characteristics, hazards and, most importantly, tree health. Intensity values are parametric; meanwhile, VTA produced non-parametric data. Finally, visual tree assessment was examined by ranking vegetation into excellent, good, fair and poor according to the conditions in Table 1.

ImageJ software was used to measure the integrated density of each tree. First, the sample area selected from tree-covered was retained by 22 × 22 square pixels of covered portions to maintain linear variables. Then, the covered portions were utilised to measure integrated density individually. Afterwards, the images were assessed using the false-colour features, black and white. Comparatively, the other method utilised stitched drone orthomosaic images.

Table 1  
*List of tree health status and the conditions*

Tree health status	Conditions
Excellent	<ul style="list-style-type: none"> <li>• Mature specimen</li> <li>• Well-balanced grand or outstanding appearance and structure</li> <li>• No evidence of insect or parasitic attack or disease, epicormic growth or dead wood and physical damage</li> </ul>
Good	<ul style="list-style-type: none"> <li>• Mature specimen</li> <li>• Tree structure, appearance form and balance are considered typical.</li> <li>• Little evidence of insect or parasitic attack, epicormic growth or dead wood</li> </ul>
Fair	<ul style="list-style-type: none"> <li>• Mature specimen</li> <li>• Sparse or pale-coloured foliage</li> <li>• Epicormic growth or deadwood throughout the crown</li> <li>• Evidence of some branch fall</li> </ul>
Poor	<ul style="list-style-type: none"> <li>• Limited life expectancy (less than 5–10 years)</li> <li>• Limited habitat value</li> <li>• Poor form, health and condition, significant dieback or sparse canopy.</li> <li>• Physical damage, disease, decay, and susceptibility to large limb drops, including bark forks.</li> </ul>

*Source.* Dunster et al. (2017)

## RESULTS AND DISCUSSION

### Proportions of Green Areas in an Urban Compound

The NDVI image, calibrated using MAPIR software and generated by QGIS, has been stitched together using Agisoft Photoscan (Figure 3). The colour has been colour-corrected



in QGIS, and the level of red intensity depicts the health status of the trees. However, this image also encompasses the Minden Sports field and the USM stadium's main football field. The most common method to calculate the area uses the threshold method, manually moving the saturation and brightness and then measuring the area using the Measure command in ImageJ. This method has drawbacks as it is subjective to the researcher's perception. Furthermore, QGIS generated a yellow background in the map to complicate things further, resulting in an extra colour channel in the image. We implemented an extra step to reduce the bias using the SeparateColour command in Wolfram Mathematica Image processing. Mathematica will extract different colour channels from the original image. The red colour extraction is done automatically using the software's predetermined method deemed best. Then, we used the Adobe Photoshop colour threshold method to calculate the percentage of the areas covered by tree grass in USM. We first outlined the exact map area using the smart selection tool and measured the number of pixels in that area. Only the red colour wavelength was selected, and the number of red pixels in that selection was measured. To illustrate this method, Figure 4 shows the red colour selection.

We then calculated the percentage of the red pixels from the total pixel number of the USM map. The method was repeated three times to increase the calculation accuracy by 95%. Finally, the red areas calculated are 27.80% of the studied urban areas. From the estimated 145 hectares, the green areas with vegetation are 39.15 hectares. This calculation considers the caveat of including the Minden Field and the USM main football stadium. In

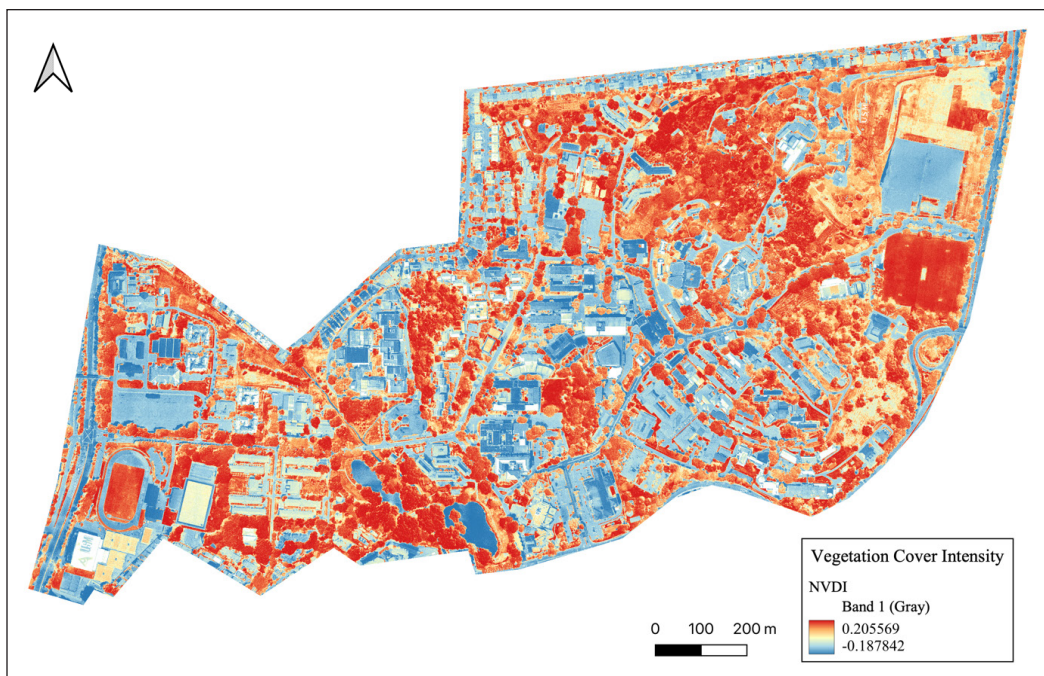


Figure 3. NVDI map of USM in RGB colour ramp



Figure 4. Map of USM in black and red

general, the proportion of green areas that are biologically active influences many aspects of the urban ecosystem. The university campus, specifically Universiti Sains Malaysia (USM), is an urban compound with a densely constructed infrastructure, roads, and others. According to the definition, urban areas have more than 50% built-up area in every 1 km by 1 km area, without considering the continuity of the built-up area (Schneider et al., 2010). Therefore, the compound is deemed urban, with only 27.80% green area. However, the urban environment here is surrounded by maintained trees.

“Green area” is partially or entirely covered with grass, trees, shrubs, or other vegetation. The proportion of green areas indicates the ability to provide green space efficacy services depends mainly on the ratio of green space to the city’s total area (Bihuňová et al., 2021). The polygon has 188.52 hectares here, including biologically active regions that can support or promote biological activity, such as non-biologically active regions that lack the environmental conditions necessary to support biological activity or are devoid of life. The USM campus’s green area covers 27.80% of the total area. According to previous studies, a city’s optimal proportion of green areas is 50% (Gupta et al., 2012). Therefore, compared to this optimal percentage, the percentage of green areas on the USM campus is not ideal. The policies of the urban area are usually established according to government green area goals. For instance, the optimal urban green areas targeted by the Indonesian government were around 30%, referring to Indonesia Law No. 26/2007.



However, the green area in Jakarta City in 2014 was only about 9.84% (Listyarini et al., 2014). Therefore, the government of Indonesia needed some improvement in the management to meet the goals. Researchers in Malaysia suggest that the urban green cover target should be 30% (Kanniah, 2017). In 2001, the green cover of Kuala Lumpur was around 27%, then reduced to 24% in 2013, followed by an increment to 30% in 2016. In comparing Kuala Lumpur's green area coverage in 2016 to USM, Kuala Lumpur's green coverage was slightly higher. However, the small difference in percentage indicates that USM green area coverage was decent. The increase in green area coverage in Malaysia in 2016 was due to the government initiative to plant trees along the streets and many other locations (Kanniah, 2017). However, Indonesia, for example, is limited in developing green areas as huge funds are required due to the land price skyrocketing, according to Listyarini et al. (2014), similar to cities in Malaysia such as Kuala Lumpur and Penang.

In another city, Changchun City was a green area in 2014 grounded on the National Economy and Society statistical bulletin was about 36.5% after undergoing rapid urbanisation (Yang et al., 2017). Compared to these cities, the proportion of green spaces in this urban campus is average. Therefore, there were more improvements in maintaining and increasing the green spaces on the campus. Large green areas provide ecological services to benefit humans and maintain a healthy ecosystem from ecological, social, and economic perspectives (Semeraro et al., 2021). According to respondents' perception on green spaces, most of them perceived good perception towards the benefits of these spaces, especially health (Gashu et al., 2019). The most substantial health benefit of green space in Hong Kong is obesity (Lachowycz & Jones, 2011). It is interconnected with the accessibility of green areas to provide spaces for recreation. Shades from trees provided comfort for society to have a healthy lifestyle. Urban green areas commonly refer to any natural or man-made areas covered with vegetation, such as parks, gardens, and street trees. While citizens tend to associate green spaces with parks, it is important to note that green spaces can take on various forms and have been shown to provide numerous benefits to cities regarding liveability and sustainability (Lyytimäki & Sipilä, 2009). USM is one of the types of urban green space that benefits the community without being labelled as a park. However, it has been established that the health advantages of green spaces are proportional to urbanisation (Milligan & Bingley, 2007; Mitchell & Popham, 2007).

### **Tree Health Status Between Paved and Unpaved Ground Surfaces**

The effects of pavement on the overall health of 60 *S. saman* trees were evaluated using both NVDI image analysis and VTA. A comprehensive evaluation of the impact of pavement on tree health was carried out using these two methods. Six different patches were assessed, which were Math School (MS), Aman Damai (AD), Minden (MD), Rumah Staff (RS), Fajar Harapan (FH) and Tapak Convex (TC). The initial approach involved evaluating the

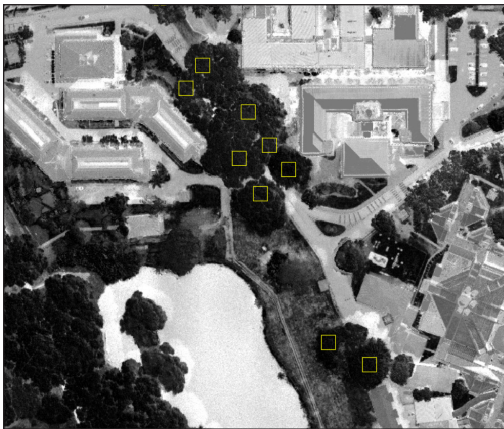
health of trees on an individual level using NVDI aerial imagery for both paved and non-paved soil surfaces. In Figures 5(a), 5(b), and 5(c), as well as Figures 6(a), 6(b), and 6(c), images were generated using ImageJ in black and white. The yellow squares within each image indicate the covered portions extracted to measure integrated density. Each square within the images represents a single tree.

Table 2 below shows the value of max, min, and mean integrated density based on surface conditions (Paved and Unpaved). Statistically, the pixel’s values were insignificant at the t-value of 1.58686, and the p-value was more than 0.5 by 0.58991. Thus, the result is insignificant at  $p < 0.05$  between paved and non-paved areas. The health of the trees is determined by the pixels retrieved from the photos. The foliage colour has been altered from the tones of green to grayscale’s false colour. The darker green becomes dark grey, whereas the lighter green becomes light grey. The greyish tones were observable. However, the integrated density generated from ImageJ visualises the colour into numbers. Integrated density measures luminescence or commonly describes the emission of light by a substance. Therefore, integrated density area is the product of area and means intensity. Besides, the results may be influenced by limiting factors affecting the quality of images. The higher resolution and quality contribute to more accurate analysis and vice versa. One of the limiting factors is flight altitude. Studies show that flight at low altitudes produces more reconstruction details and better precision than at high altitudes, especially with high image overlaps (Seifert et al., 2019). When drones fly at a relatively high altitude, they can take low-quality photos while observing many target nodes and edges. On the other hand, high-quality photos and narrow observation areas can be captured from a relatively low altitude.

Nonetheless, the images used in this work were orthomosaic imagery assembled from overlapping the images into the complete USM map. It may lack accuracy and quality as orthomosaic images require a slightly higher altitude to yield the whole area due to high-rise infrastructure requiring UAVs to fly higher than the trees and infrastructure for a comprehensive aerial view. Regarding altitude, rather than using an orthomosaic map for the integrated density extraction, a single aerial image for an individual tree is suggested. It yields better images and resolutions as the flight altitude will be much shorter and higher in quality. However, it is much more applicable for small sample sizes. Therefore, this method would be inefficient in sampling tree health for a vast area with many trees. However, an orthomosaic image is still applicable by increasing the overlap

Table 2  
*Integrated density based on surface conditions (paved and unpaved)*

Locations	Max (pixels)	Min (pixels)	Mean (pixels)
Paved	22706.0	4753.0	12234.3
Non-Paved	22291.0	3225.0	10312.0



(a)



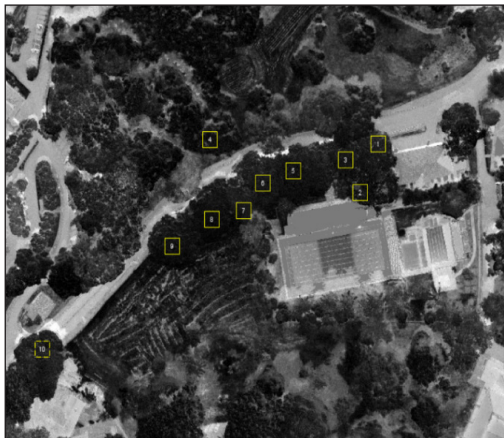
(a)



(b)



(b)



(c)



(c)

*Figure 5.* ImageJ tree selection of paved area: (a) Math School; (b) Aman Damai; and (c) Minden

*Figure 6.* ImageJ tree selection of unpaved area: (a) Rumah Staff; (b) Fajar Harapan; and (c) Tapak Convex

percentages, which can offer more details of the vegetation geometry (Seifert et al., 2019). For example, a quality photogrammetric model requires 80% on-lap and 60-80% side-lap overlap. On-lapping means overlapping images in the flight direction while side-lapping overlaps perpendicularly (Colomina & Molina, 2014). Therefore, increasing overlapping percentages are suggested to yield significant results.

Figures 7(a) and 7(b) show the health status of trees planted in paved and non-paved areas using field assessment (VTA) according to four classes: excellent, good, fair and poor. In this project, both VTA and aerial images were used to evaluate the health status. However, the final health status rating was determined solely by the results of the VTA survey, which classified the area as ranging from excellent to poor.

The health status between paved and non-paved soil surfaces in VTA (Visual Tree Assessment) shows different compositions between all four statuses divided into excellent, good, fair, and poor. VTA status for paved soil surface health is visibly lower than in non-paved land. Thus, the pavement surrounding the trees affects the health and growth of the trees. The trees grown on paved surfaces tend to lose their optimal needs of water, oxygen and nutrients from the impermeable surface that restricts ecological processes (Mullaney, 2015). Therefore, excellent status was difficult to achieve. However, due to other factors, the good status is achievable by 33%, making it fair at 54%, leaving the lowest health status, poor, with 13%. It is clearly shown that paved areas lead to poorer health than non-paved land. Non-paved land attains better health for 27% of the population, which is excellent. Besides, good status is a majority of 43%. Therefore, it leaves fair and poor at 23% and 7%, respectively. The methodologies used in this study generate several health statuses from qualitative and quantitative evaluations. Even though the methods applied were different, the result was intended for the same purpose, which is to evaluate health.

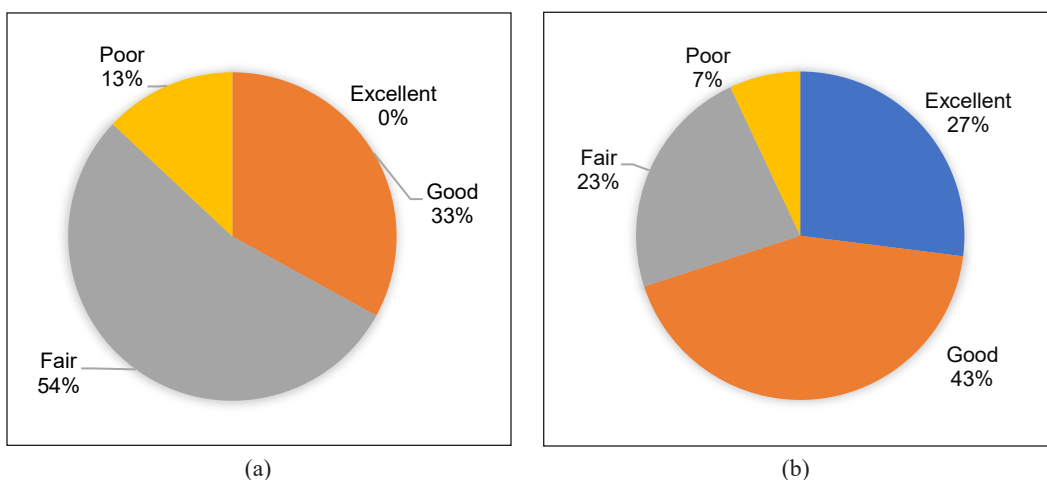


Figure 7. Health status percentages: (a) paved areas; and (b) non-paved areas

Overall, the results from both methodologies suggest that trees growing in non-paved or open areas are far healthier than trees growing in close-knit paved patches. The pavement is one of the parameters used to evaluate the health of individual trees in VTA based on observable evidence.

Despite the prominent importance of trees providing oxygen and reducing heat, the trees were grown along the pavement to provide shade. The paved surface, such as the black surface of asphalt pavement, absorbs more heat from the atmosphere and provides pedestrians for urban citizens. Pavements serve as a pedestrian-friendly surface for urban residents. However, pavements with dark asphalt surfaces absorb more heat from the sun, leading to higher temperatures on the pavement surface. As a result, the urban heat island effect worsens. Moreover, the impervious nature of these pavements reduces water infiltration, resulting in dry soil conditions, lower evapotranspiration rates, and ultimately affecting urban forestry (Cheela et al., 2021). Dryer soil reduces the water availability, leading to dehydration among the trees. Hence, deteriorating the optimal environment provides a better tree health status. The pavement here was primarily made of impervious materials such as asphalt. There were two materials widely used, classified as impervious and pervious pavement. Pervious pavements have the same strength as impermeable pavements but are more permeable to air and water. Because of these functional properties, it has been suggested that pervious paving, rather than impervious asphalt, could benefit urban trees (Mullaney et al., 2015). Asphalt is the primary pavement type based on previous research (Guo et al., 2018). The pavement primarily affected the soil condition.

Trees growing under pavement show lower health than those in non-paved areas. In the paved area, the lack of healthy trees in the excellent category suggests that the trees could not achieve optimal health overall. Additionally, the proportion of trees in the lowest health category is higher (about 13%) than in non-paved areas, suggesting a greater risk of health deterioration in paved areas. These results demonstrate a decline in tree health on paved surfaces. Conversely, the prominent percentage of health status, excellent and suitable for non-paved land, was about 70%. VTA results leaned towards the non-paved land surface rather than the land covered by pavement. The matured trees in this urban campus grew much better on an open surface as they were free from disturbance in an asphalt bed. However, the health of non-paved still trees was listed in the poor classes. These conditions may occur owing to other parameters of the trees.

Researchers assessed every part of VTA, such as roots, trunks, branches, leaves, and hazards. To conclude the overall health status, all the factors were considered. Guided from raw VTA data for the trees with poor health, the main factor affecting the trees' condition was the condition of the roots. Exposing roots was explicitly common, but the concern parts of roots was the infestation of ants and termites from the nest close to the root systems.



The infestation can be seen on the surface of exposed roots. Besides, the exposed roots among matured *S. saman* were common as they have shallow root systems (Staples & Elevitch, 2006). Therefore, it is most likely to be visible on the soil surface. A root was declared infested if it bores characteristic dark scarred spots on the root's surface, which are a common indication of weevil penetration and feeding. Roots that were undamaged on the surface were considered uninfested (Stathers et al., 2003).

Although the effects of pavement on tree health have been demonstrated through the outcome of health assessments, other factors may still influence overall health. Evaluating the overall health of trees solely based on pavement factors may not be sufficient. However, it shows that the pavement seems to have been influenced by modifying micro-environmental conditions (e.g., soil nutrients, temperature, and aeration). Other studies also agreed that enclosed surfaces could significantly alter microclimate conditions (Rahman et al., 2020; Rahman et al., 2017). Thus, paved surfaces were known to influence the growth potential and affect the biomass allowances. Likewise, the water usage of vegetation in city areas was also affected (Wang et al., 2020). For example, the rate of water availability of plant life was modified by the permeability rate of the pavement (de la Mota Daniel et al., 2018; Rahman et al., 2013). In addition, some claim that the soil structure was also shifted exponentially, accountable to changes in the growth and transpiration rate (Rahman et al., 2011).

On the other hand, as this pavement study was done in an urban setting, some patches were not entirely covered by pavement. Allow the ecological process to occur optimally at the open edges. Compared to experiment settings, this study tends to be more lenient. It permits the lack of precision. Still, the study should also be done in actual conditions for researchers to learn about the tree's health in its natural state. It makes it possible to estimate the mitigation for better management. *Samanea saman* is a deciduous tree on this campus. However, impervious pavement has been proven to limit deciduous tree species' development, restrict plant photosynthesis, and increase leaf budburst (Chen et al. 2018). Nevertheless, it has been shown that the strength of pavement's effects on plant physiological and growth features varies by tree species (Song et al., 2015; Wang et al., 2019).

Rain trees, on the other hand, threatened the pavement's durability as it aged, as the tree's shallow roots tended to heave up and damage the pavement. Although the pavement break improves water flow and aeration, it also greatly reduces the utility of the pavement and detracts from urban areas' aesthetics. Overall, the integrated density values from modified NVDI images were reciprocal to VTA, suggesting that non-paved trees were healthier than paved ones. The diversified values were expected as the intensity of chlorophyll pigment from the leaf differed from each. Good health status is anticipated to establish a low Integrated density value and vice versa. Even so, the maximum value

point indicating the worst health was found on the non-paved land. The differences in maximum values between each land were minor. The data observation demonstrates that the cause was from the root's conditions. Thus, using aerial images to assess tree health was not entirely accurate.

Based on the data observations, it appears that the cause of the issue lies with the root conditions of the trees. Therefore, relying solely on aerial images to assess tree health may not be entirely accurate as the aerial image may not be able to capture the details of the root system due to perspective and may not provide a complete understanding of the root conditions. The same goes for other parts, such as trunks (Xiao & McPherson, 2005). Even with less data at the individual level, it would still be useful. On the other hand, large, difficult-to-reach areas can be monitored more effectively and more cheaply. It facilitates traditional monitoring (Booth et al., 2008). However, visual assessment methods are considered adequate for evaluating the likelihood of urban tree breakage (de La Barra et al., 2018). It is also very time-consuming and labour-intensive.

According to the study, physical and aerial sampling results may contradict because remote sensing techniques' ability to deliver relevant data on vegetation structure is limited (Hestir et al., 2008). For example, the intensity calculated by ImageJ of colour relies on the intensity of green chlorophyll pigment, indicating the capacity of a tree to undergo photosynthesis and the rate of tree growth. However, does green leaf pigment enough to assess tree health? Leaf colour and density are fundamental parameters in assessing tree health. However, it does not represent the whole tree rank. For example, Kaewkhaw and Srivinit (2020) discovered 35 parameters to assess the tree. However, through the studies, 11 parameters scored 1.00 IOC values from damage to trees, mechanical damage, insects, or disease. Discoloured leaves are one of the parameters. Hence, indicating leaves provides health information but only as a portion.

Depending solely on leaf colour intensity to assess tree health is may not enough to indicate the whole tree's health rank since remote sensing is most sensitive to detecting changes in the health of the foliage. Depending on the species, the changes in leaves occur depending on the adaptation of the tree itself. For example, after a dry period, trees such as the Rain Tree or *S. saman* are semi-deciduous, shedding their leaves for a short period. During a specific dry season, they may remain leafless for weeks. However, if there is adequate moisture, they will quickly resprout. In moister climates, it appears that rain tree is "evergreen." Therefore, the density and intensity of the leaves of raintrees do not reflect the tree's health. Besides, the changes on the leaves might appear as an adaption to reduce moisture loss and maintain their well-being. However, assessing tree health based on foliage is still helpful. Most diseases or abiotic stresses to the roots or bole are ultimately expressed as changes in the foliage's density, chemistry, colour, and moisture content.



## CONCLUSION

The findings of this study highlight the importance of maintaining an optimal proportion of green areas in urban settings to support a sustainable urban ecosystem. Using drone-mounted NVDI technology provided insight into the green area proportion of USM's main campus, approximately 27.80%. However, to maintain the proportion of green space, the health of the trees in these green spaces is also important. The pavement was found to negatively impact the health of *S. saman* trees, with non-paved areas exhibiting better tree health than paved areas. A combination of VTA and the NVDI aerial image analysis provided a more comprehensive understanding of tree health, emphasising the importance of using multiple methods to ensure accuracy. Further research on additional urban tree species in real-world contexts is needed to better understand the degree of pavement effects between species. Overall, this study provides insights into the importance of green areas and tree health in creating sustainable urban ecosystems, which can inform urban planning and management strategies for promoting the health and sustainability of urban green spaces.

## ACKNOWLEDGEMENT

This work was supported by USM RUI grant PBIOLOGI/1001/ 8011124 and Universiti Sains Malaysia Industry grant 304 /PBIOLOGI /6501144 /I145.

## REFERENCES

- Assmann, J. J., Kerby, J. T., Cunliffe, A. M., & Myers-Smith, I. H. (2019). Vegetation monitoring using multispectral sensors - best practices and lessons learned from high latitudes. *Journal of Unmanned Vehicle Systems*, 7(1), 54-75. <https://doi.org/10.1139/juvs-2018-0018>
- Bihuňová, M., Supuka, J., Tóth, A., Šinka, K., & Kuczman, G. (2021). Urban green areas and woody plant composition: Dwelling space quality factor in the Klokočina Housing Estate. *Ekológia(Bratislava)*, 40(1), 80-90. <https://doi.org/10.2478/eko-2021-0010>
- Booth, D. T., Cox, S. E., Meikle, T., & Zuuring, H. R. (2008). Ground-cover measurements: Assessing correlation among aerial and ground-based methods. *Environmental Management*, 42(6), 1091-1100. <https://doi.org/10.1007/S00267-008-9110-X/TABLES/5>
- Cheela, V. R. S., John, M., Biswas, W., & Sarker, P. (2021). Combating urban heat island effect - A review of reflective pavements and tree shading strategies. *Buildings*, 11(3), 1-21. <https://doi.org/10.3390/BUILDINGS11030093>
- Chen, Y., Wang, X., Jiang, B., & Li, L. (2018). The leaf phenophase of deciduous species altered by land pavements. *International Journal of Biometeorology*, 62(6), 949-959. <https://doi.org/10.1007/s00484-018-1497-3>
- Colomina, I., & Molina, P. (2014). Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 92, 79-97. <https://doi.org/10.1016/j.isprsjprs.2014.02.013>

- Conway, T. M., & Yip, V. (2016). Assessing residents' reactions to urban forest disservices: A case study of a major storm event. *Landscape and Urban Planning*, *153*, 1-10. <https://doi.org/10.1016/j.landurbplan.2016.04.016>
- de La Barra, J. R., Ponce-Donoso, M., Vallejos-Barra, O., Daniluk-Mosquera, G., & Coelho Duarte, A. P. (2018). Comparison of four methods of visual risk tree assessment in urban areas. *Colombia Forestal*, *21*(2), 161-173. <https://doi.org/10.14483/2256201X.12604>
- de la Mota Daniel, F. J., Day, S. D., Owen, J. S., Stewart, R. D., Steele, M. K., & Sridhar, V. (2018). Porous-permeable pavements promote growth and establishment and modify root depth distribution of *Platanus × acerifolia* (Aiton) willd. in simulated urban tree pits. *Urban Forestry and Urban Greening*, *33*, 27-36. <https://doi.org/10.1016/j.ufug.2018.05.003>
- Duffy, J. P., Cunliffe, A. M., DeBell, L., Sandbrook, C., Wich, S. A., Shutler, J. D., Myers-Smith, I. H., Varela, M. R., & Anderson, K. (2018). Location, location, location: considerations when using lightweight drones in challenging environments. *Remote Sensing in Ecology and Conservation*, *4*(1), 7-19. <https://doi.org/10.1002/rse2.58>
- Dunster, J. A., Smiley, E. T., Matheny, N., & Lilly, S. (2017). *Tree Risk Assessment Manual* (2nd ed.). International Society of Arboriculture.
- Gashu, K., Gebre-Egziabher, T., & Wubneh, M. (2019). Local communities' perceptions and use of urban green infrastructure in two Ethiopian cities: Bahir Dar and Hawassa. *Journal of Environmental Planning and Management*, *63*(2), 287-316. <https://doi.org/10.1080/09640568.2019.1578643>
- Guo, M., Tan, Y., Wang, L., & Hou, Y. (2018). A state-of-the-art review on interfacial behavior between asphalt binder and mineral aggregate. In *Frontiers of Structural and Civil Engineering* (Vol. 12, Issue 2, pp. 248-259). Springer. <https://doi.org/10.1007/s11709-017-0422-x>
- Gupta, K., Kumar, P., Pathan, S. K., & Sharma, K. P. (2012). Urban neighborhood green index – A measure of green spaces in urban areas. *Landscape and Urban Planning*, *105*(3), 325-335. <https://doi.org/10.1016/j.landurbplan.2012.01.003>
- Hasan, R., Othman, N., & Ismail, F. (2018). Developing Malaysian roadside tree species selection model in urban areas. *Planning Malaysia Journal*, *16*(3), 248-260. <https://doi.org/10.21837/pm.v16i7.515>
- Hauer, R. J., & Peterson, W. D. (2016). Municipal tree care and management in the United States: A 2014 urban & community forestry census of tree activities. *Special Publication*, *16*(1), 1-71.
- Hestir, E. L., Khanna, S., Andrew, M. E., Santos, M. J., Viers, J. H., Greenberg, J. A., Rajapakse, S. S., & Ustin, S. L. (2008). Identification of invasive vegetation using hyperspectral remote sensing in the California Delta ecosystem. *Remote Sensing of Environment*, *112*(11), 4034-4047. <https://doi.org/10.1016/J.RSE.2008.01.022>
- Kaewkhaw, S., & Srivanit, M. (2020). Aggregation of Thai arborist judgments on urban tree hazard inventories used to determine tree health at single-tree level. *IOP Conference Series: Materials Science and Engineering*, *910*, Article 012023. <https://doi.org/10.1088/1757-899X/910/1/012023>
- Kanniah, K. D. (2017). Quantifying green cover change for sustainable urban planning: A case of Kuala Lumpur, Malaysia. *Urban Forestry & Urban Greening*, *27*, 287-304. <https://doi.org/10.1016/J.UFUG.2017.08.016>

- Klein, R. W., Koester, A. K., Hauer, R. J., Hansen, G., & Escobedo, F. J. (2019). Risk assessment and risk perception of trees: A review of literature relating to arboriculture and urban forestry. *Arboriculture and Urban Forestry*, 45(1), 26-38. <https://doi.org/10.48044/JAUF.2019.003>
- Lachowycz, K., & Jones, A. P. (2011). Greenspace and obesity: A systematic review of the evidence. *Obesity Reviews*, 12(5), e183-e189. <https://doi.org/10.1111/J.1467-789X.2010.00827.X>
- Listyarini, S., Warlina, L., Indrawati, E., & Pardede, T. (2014). Optimization model for predicting green areas in Jakarta to minimize impacts of climate change. *WIT Transactions on Ecology and the Environment*, 191, 305-315. <https://doi.org/10.2495/SC140261>
- Lyytimäki, J., & Sipilä, M. (2009). Hopping on one leg – The challenge of ecosystem disservices for urban green management. *Urban Forestry & Urban Greening*, 8(4), 309-315. <https://doi.org/10.1016/J.UFUG.2009.09.003>
- Milligan, C., & Bingley, A. (2007). Restorative places or scary spaces? The impact of woodland on the mental well-being of young adults. *Health & Place*, 13(4), 799-811. <https://doi.org/10.1016/J.HEALTHPLACE.2007.01.005>
- Mitchell, R., & Popham, F. (2007). Greenspace, urbanity and health: Relationships in England. *Journal of Epidemiology & Community Health*, 61(8), 681-683. <https://doi.org/10.1136/JECH.2006.053553>
- Mullaney, J. (2015). *Using Permeable Pavements to Promote Street Tree Growth*. [Doctoral dissertation]. University of the Sunshine Coast, Queensland. <https://doi.org/10.25907/00454>
- Mullaney, J., Lucke, T., & Trueman, S. J. (2015). The effect of permeable pavements with an underlying base layer on the growth and nutrient status of urban trees. *Urban Forestry & Urban Greening*, 14(1), 19-29. <https://doi.org/10.1016/j.ufug.2014.11.007>
- Rahman, M. A., Moser, A., Rötzer, T., & Pauleit, S. (2017). Microclimatic differences and their influence on transpirational cooling of *Tilia cordata* in two contrasting street canyons in Munich, Germany. *Agricultural and Forest Meteorology*, 232, 443-456. <https://doi.org/10.1016/j.agrformet.2016.10.006>
- Rahman, M. A., Smith, J. G., Stringer, P., & Ennos, A. R. (2011). Effect of rooting conditions on the growth and cooling ability of *Pyrus calleryana*. *Urban Forestry and Urban Greening*, 10(3), 185-192. <https://doi.org/10.1016/j.ufug.2011.05.003>
- Rahman, M. A., Stringer, P., & Ennos, A. R. (2013). Effect of pit design and soil composition on performance of *Pyrus calleryana* street trees in the establishment period. *Arboriculture and Urban Forestry*, 39(6), 256-266.
- Rahman, M., Haque, S., & Rahman, Z. (2020). Identifying and categorizing opinions expressed in Bangla sentences using deep learning technique. *International Journal of Computer Applications*, 176(17), 13-17. <https://doi.org/10.5120/ijca2020920119>
- Rotherham, I. D. (2020). Routledge handbook of urban forestry. *Arboricultural Journal*, 42(3), 180-183. <https://doi.org/10.1080/03071375.2020.1738787>
- Schmidlin, T. W. (2009). Human fatalities from wind-related tree failures in the United States, 1995-2007. *Natural Hazards*, 50(1), 13-25. <https://doi.org/10.1007/s11069-008-9314-7>

- Schneider, A., Friedl, M. A., & Potere, D. (2010). Mapping global urban areas using MODIS 500-m data: New methods and datasets based on 'urban ecoregions.' *Remote Sensing of Environment*, *114*(8), 1733-1746. <https://doi.org/10.1016/J.RSE.2010.03.003>
- Seifert, E., Seifert, S., Vogt, H., Drew, D., van Aardt, J., Kunneke, A., & Seifert, T. (2019). Influence of drone altitude, image overlap, and optical sensor resolution on multi-view reconstruction of forest images. *Remote Sensing*, *11*(10), Article 1252. <https://doi.org/10.3390/rs11101252>
- Semeraro, T., Scarano, A., Buccolieri, R., Santino, A., & Aarvevaara, E. (2021). Planning of urban green spaces: An ecological perspective on human benefits. *Land*, *10*(2), Article 105. <https://doi.org/10.3390/land10020105>
- Seng, H. W., Ratnam, W., Noor, S. M., & Clyde, M. M. (2004). The effects of the timing and method of logging on forest structure in Peninsular Malaysia. *Forest Ecology and Management*, *203*(1-3), 209-228. <https://doi.org/10.1016/J.FORECO.2004.07.050>
- Song, Y., Li, F., Wang, X., Xu, C., Zhang, J., Liu, X., & Zhang, H. (2015). The effects of urban impervious surfaces on eco-physiological characteristics of Ginkgo biloba: A case study from Beijing, China. *Urban Forestry and Urban Greening*, *14*(4), 1102-1109. <https://doi.org/10.1016/j.ufug.2015.10.008>
- Staples, G. W., & Elevitch, C. R. (2006). *Samanea saman* (rain tree). In C. R. Elevitch (Ed.), *Species Profiles for Pacific Island Agroforestry* (pp. 1-14). Permanent Agriculture Resources (PAR).
- Stathers, T. E., Rees, D., Kabi, S., Mbilinyi, L., Smit, N., Kiozya, H., Jeremiah, S., Nyango, A., & Jeffries, D. (2003). Sweetpotato infestation by *Cylas* spp. in East Africa: I. Cultivar differences in field infestation and the role of plant factors. *International Journal of Pest Management*, *49*(2), 131-140. <https://doi.org/10.1080/0967087021000043085>
- Vaz, A. S., Kueffer, C., Kull, C. A., Richardson, D. M., Vicente, J. R., Kühn, I., Schröter, M., Hauck, J., Bonn, A., & Honrado, J. P. (2017). Integrating ecosystem services and disservices: Insights from plant invasions. *Ecosystem Services*, *23*, 94-107. <https://doi.org/10.1016/j.ecoser.2016.11.017>
- Vogt, J., Hauer, R. J., & Fischer, B. C. (2015). The costs of maintaining and not maintaining the urban forest: A review of the urban forestry and arboriculture literature. *Arboriculture and Urban Forestry*, *41*(6), 293-323.
- Wang, J., Rich, P. M., & Price, K. P. (2003). Temporal responses of NDVI to precipitation and temperature in the central Great Plains, USA. *International Journal of Remote Sensing*, *24*(11), 2345-2364. <https://doi.org/10.1080/01431160210154812>
- Wang, X. M., Wang, X. K., Su, Y. B., & Zhang, H. X. (2019). Land pavement depresses photosynthesis in urban trees especially under drought stress. *Science of the Total Environment*, *653*, 120-130. <https://doi.org/10.1016/j.scitotenv.2018.10.281>
- Wang, X., Wang, X., Sun, X., Berlyn, G. P., & Rehim, A. (2020). Effect of pavement and water deficit on biomass allocation and whole-tree transpiration in two contrasting urban tree species. *Urban Ecosystems*, *23*(4), 893-904. <https://doi.org/10.1007/s11252-020-00953-z>
- Weier, J., & Herring, D. (2000). *Measuring Vegetation (NDVI & EVI)*. Earth Observatory. National Aeronautics and Space Administration.

- Xiao, Q., & McPherson, E. G. (2005). Tree health mapping with multispectral remote sensing data at UC Davis, California. *Urban Ecosystems*, 8(3-4), 349-361. <https://doi.org/10.1007/s11252-005-4867-7>
- Yang, C., He, X., Wang, R., Yan, F., Yu, L., Bu, K., Yang, J., Chang, L., & Zhang, S. (2017). The effect of urban green spaces on the urban thermal environment and its seasonal variations. *Forests*, 8(5), Article 153. <https://doi.org/10.3390/F8050153>

## The Effect of Nutrients in Anodic Chamber to the Performance of Microbial Fuel Cell (MFC)

Nashley Ursula Mundi Ujai<sup>1</sup>, Siti Kudnie Sahari<sup>1\*</sup>, Marini Sawawi<sup>1</sup>, Kuryati Kipli<sup>1</sup>, Asmahani Awang<sup>2</sup>, Mohamad Rusop Mahmood<sup>3</sup>, Lilik Hasanah<sup>4</sup>, Abdul Rahman Kram<sup>1</sup> and Zainab Ngaini<sup>5</sup>

<sup>1</sup>Faculty of Engineering, Universiti Malaysia Sarawak, 94300 UNIMAS, Kota Samarahan, Sarawak, Malaysia

<sup>2</sup>Faculty of Science and Natural Resources, Universiti Malaysia Sabah, 88400 UMS, Kota Kinabalu, Sabah, Malaysia

<sup>3</sup>Faculty of Electrical Engineering Universiti Teknologi Mara, 40450 UiTM, Shah Alam, Selangor, Malaysia

<sup>4</sup>Faculty of Mathematics and Sciences Education, Indonesia University of Education, Bandung, Jawa Barat 40154, Indonesia

<sup>5</sup>Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, 94300 UNIMAS, Kota Samarahan, Sarawak, Malaysia

### ABSTRACT

This paper describes a device known as a Single-chamber Microbial Fuel Cell (SMFC) that was used to generate bioelectricity from plant waste containing lignocellulosic components, such as bamboo leaves, rice husk and coconut waste, with various anodic chamber substrate compositions. The maximum power density among all assembled SMFCs was determined to be 231.18  $\mu\text{W}/\text{m}^2$ , generated by coconut waste. This model's bioelectricity production was enhanced by adding organic compost to the anodic chamber, which acts as a catalyst in the system. The maximum power density of 788.58  $\mu\text{W}/\text{m}^2$  was attained using a high proportion of coconut waste (CW) and organic compost. These results show that the higher percentage of lignin in CW improved the bioelectricity of SMFC.

### ARTICLE INFO

#### Article history:

Received: 29 January 2023

Accepted: 20 July 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.12>

#### E-mail addresses:

nashleyujai@gmail.com (Nashley Ursula Mundi Ujai)

sskudnie@unimas.my (Siti Kudnie Sahari)

smarini@unimas.my (Marini Sawawi)

kkuryati@unimas.my (Kuryati Kipli)

asmahani\_awang@yahoo.com (Asmahani Awang)

rusop@salam.uitm.edu.my (Mohamad Rusop Mahmood)

lilikhasanah@upi.edu (Lilik Hasanah)

karahman@unimas.my (Abdul Rahman Kram)

nzainab@unimas.my (Zainab Ngaini)

\* Corresponding author

*Keywords:* Bamboo leaves, coconut waste, lignocellulosic, rice husk, single-chamber microbial fuel cell

### INTRODUCTION

In 2050, the global population is projected to increase by 2 billion, from 7.7 billion to 9.7 billion (United Nations, 2019). Population growth will increase the energy demand, leading to the depletion of energy resources

such as fossil fuels. Fossil fuels, made up of decomposing plants and animals buried by layers of rocks from over a million years ago, are a non-renewable energy resource consisting of coal, oil and natural gas (Energy.Gov, 2017; National Geographic Society, 2019). Eventually, these non-renewable energy sources will be depleted due to the continued use of fossil fuels to generate electricity. In addition to being non-renewable, fossil fuels contribute to global warming by emitting large amounts of CO<sub>2</sub> into the atmosphere due to the carbon emission's ability to retain heat (Denchak, 2022). It was estimated that over 8 million people worldwide die every year due to inhaling polluted air from burning fossil fuels containing particles such as greenhouse gases (Kottasová & Dewan, 2021). That is why finding a renewable energy source to produce electricity for all consumers is essential. Numerous renewable energy sources have been developed, including discovering Microbial Fuel Cell (MFC) to meet the ever-increasing energy demand. MFC is an environmentally friendly device with minimal carbon emissions. MFC can generate electricity from organic residues, such as food or water waste, as it converts biodegradable substances into simpler substances to generate bioelectricity (Koch et al., 2016; Ucar et al., 2017).

According to a study done in 2020, MFC, however, cannot support bioelectricity for a large population due to its low power density (Khoo et al., 2020), and a few factors influence the operation of MFC. The type of substrate and its concentration in the anodic chamber were identified as the factors that influenced the voltage production of MFC (Aghababaie et al., 2015; Kumar et al., 2017). In addition, solid organic waste has gained significant interest as MFC substrates due to its high organic matter content, a crucial component for these systems (Kumar et al., 2022). Previous research found that solid fruit waste generates more electricity than solid food waste (Moqsud, 2021). This finding suggests that plant waste has the potential to function as an MFC substrate for bioelectricity production.

Due to the higher concentration of lignocellulosic, which functions as bacteria's food source, plant wastes have the potential to be utilised as MFC substrates. Certain microorganisms, when fed lignocellulosic materials, can convert the complex polysaccharides (cellulose and hemicellulose) present in lignocellulosic materials into simpler carbohydrates that can be used for energy production (Chandra & Madakka, 2019). To the best of our knowledge, researchers are still investigating the optimal composition of plant waste for producing MFCs with a higher voltage. Therefore, this research employed lignocellulosic materials with various compositions as a substrate for a single-chamber MFC (SMFC).

## **METHODOLOGY**

### **Construction of a Single-chamber Microbial Fuel Cell (SMFC) Without Adding Organic Compost**

This study used a single-chamber MFC to generate electrical energy from electrons derived by bacteria in the anodic chamber.



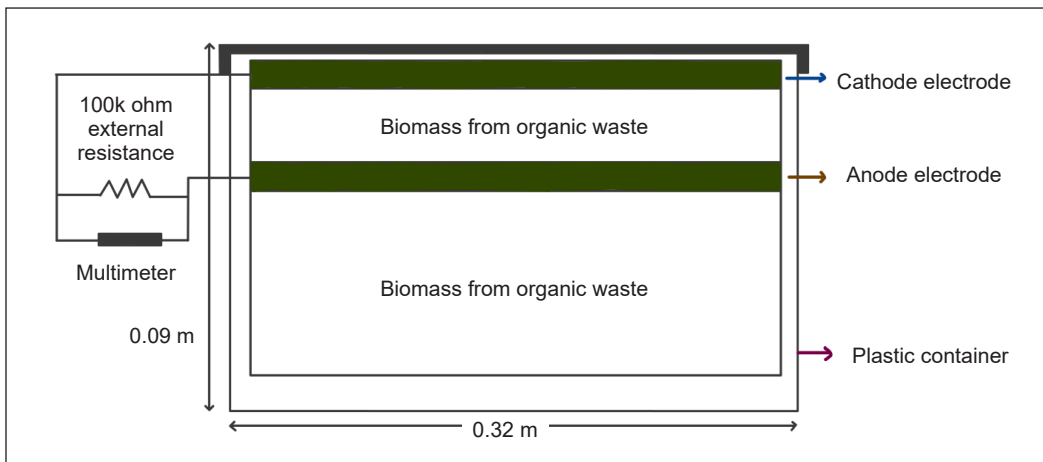


Figure 1. The schematic design of SMFC

This study uses an SMFC because of its affordability and effectiveness. Compared to a double-chamber MFC, an SMFC generates less internal resistance due to the shorter distance between its two electrodes (Flimban et al., 2019; Nawaz et al., 2022). In terms of design, it is easy to construct without any unnecessary hassle, which is excellent for research (Kumar et al., 2017).

The  $5.03 \times 10^{-3} \text{ m}^2$  of anode and cathode electrodes used in this study were constructed from granulated active carbon from coconut shells that were attached using epoxy to a circular shape stainless-steel mesh (0.24 mm stainless-steel grade 304) with a 4 cm radius. A previous study stated that an activated carbon electrode is biocompatible, less toxic and has a rough surface that can enhance the attachment of bacteria as well as generate greater bioelectricity (Sahari et al., 2022; Zhang et al., 2014).

Figure 1 shows the experimental setup of the SMFC, which is constructed by layering a mixture sample of 50 g plant waste (bamboo leaves, rice husk or coconut waste) that has been finely ground with 300 g water. The wires from both terminals of the electrodes are then connected to an external 100 k $\Omega$  resistor to ensure that the SMFC operates in a closed circuit. The details of the anodic chamber substrates are presented in Table 1. Figure 2 demonstrates the measuring of generated voltage and current by the single-chamber MFC prototype using a multimeter while connecting to an external 100 k $\Omega$  resistor, which indicates that this is conducted in a closed-circuited manner.



Figure 2. Measuring voltage and current generated by the SMFC using a multimeter

## Construction of a Single-chamber Microbial Fuel Cell (SMFC) by Adding Organic Compost

To gain a more thorough understanding of the effect of nutrients on SMFC efficiency, plant nursery-purchased organic compost was added to the SMFC substrate containing plant waste. Organic compost and plant waste weights varied from 20 to 30 g, while the water weight was fixed at 300 g per SMFC. In total, nine distinct SMFC prototypes were assembled.

Table 1  
Summary of organic substrates used for the SMFCs

SMFC No.	Organic substrates used in anodic chamber
1	50 g BL + 300 g water
2	50 g RH + 300 g water
3	50 g CW + 300 g water
4	20 g BL + 30 g OC + 300 g water
5	20 g RH + 30 g OC + 300 g water
6	20 g CW + 30 g OC + 300 g water
7	30 g BL + 20 g OC + 300 g water
8	30 g RH + 20 g OC + 300 g water
9	30 g CW + 20 g OC + 300 g water

Note. BL: bamboo leaves, RH: rice husk, CW: coconut waste, OC: organic compost

## Electricity Generating Parameter

Each SMFC prototype's power density was compared to assess the SMFC's efficiency. The magnitude of the power density was determined by normalising the system's current to the electrodes' surface area. The model system's current flow was directly measured using a digital multimeter. The formula used to calculate the electrical parameters of the SMFC is depicted in Equation 1 (Ullah & Zeshan, 2020).

$$\text{Power Density } (\mu\text{W}/\text{m}^2) = V^2/(\text{R} \times \text{A}) \quad [1]$$

where, A is the surface area of the electrodes ( $\text{m}^2$ ), V is the voltage generated (V), and R is the external resistance ( $\Omega$ ) connected to the system.

The SMFCs were observed four times per day for seven consecutive days to provide further explanation. At precisely 0600, 1200, 1800 and 2400 each day, the voltage and current produced by each SMFC were measured with a multimeter and logged. As previously mentioned, the power density of each SMFC is calculated using Equation 1.

## RESULTS AND DISCUSSION

### Comparison of Different Contents of Lignocellulosic Component of Plant Waste in Anodic Chamber

A study evaluated the power density generation of an SMFC from different plant wastes containing different lignocellulosic component compositions. Table 2 shows the percentage of each component of lignocellulosic, which consists of lignin, cellulose and hemicellulose, in each plant waste.

Table 2 shows that coconut waste (CW) contained the greatest percentage of lignin, whereas bamboo leaves contained the highest percentage of cellulose. These observations are related to the power density output of the SMFC with various anodic substrates, as

Table 2  
Lignocellulosic component of plant waste

Plant waste	Lignocellulosic component			Reference
	% Cellulose	% Hemicellulose	% Lignin	
Rice husk	35–40	15–20	20–25	Gao et al., 2018
Coconut waste	20–30	15–30	50	Anuchi et al., 2022
Bamboo leaves	47.2	23.9	23.8	Bai et al., 2013

depicted in Figure 3. Figure 3 demonstrates that an SMFC containing coconut waste as the anodic chamber substrate produced the highest maximum power density, while an anodic chamber containing bamboo leaves produced the lowest. The power density of 50 g of coconut waste was 50% higher than that of an SMFC-containing rice husk. As discussed in a previous study, the increased power density may be attributable to lignin’s ability to decrease internal resistance (Sakdaronnarong et al., 2015). This argument was founded on a previous study in which the presence of humic acid in the MFC chamber decreased internal resistance and increased power density. Due to their similar structures, it is believed that lignin has the same effect on the internal resistance of MFC as humic acid (Huang & Angelidaki, 2008). The fact that the highest power density was associated with the highest proportion of lignin proves that lignin is responsible for the highest power density. The lowest power density from an SMFC contained the highest percentage of cellulose, possibly due to its complex structure that is difficult for microorganisms to degrade, whereby special enzymes are required to convert it into simpler compounds that can be used as a fuel source (Boisset et al., 2000).

As stated previously in the methodology, SMFC data was collected four times daily at 0600, 1200, 1800 and 2400 (Figure 4). The highest power density was observed at

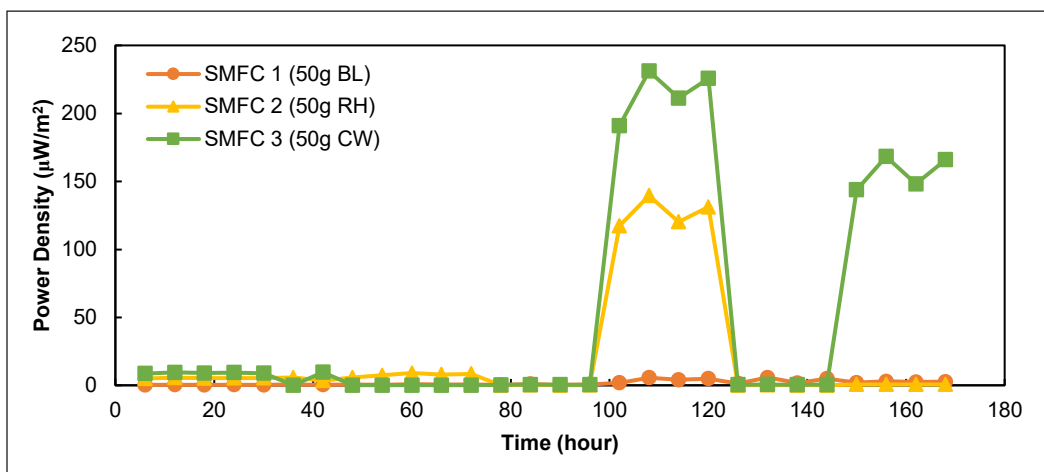


Figure 3. Power density of SMFC with different types of anodic substrate

1200 hours for SMFCs 1, 2 and 3 because the microorganisms or bacteria are actively operating currently, causing the biochemical reactions they promote to produce more electrons (Koch et al., 2016). To fully understand this phenomenon, it is necessary to investigate the effect of the surrounding temperature on the power density of SMFCs, which was not measured in this study.

Figure 5 illustrates the power density produced by various anodic substrates—clearly, all SMFCs containing CW as a substrate generated a higher power density than the others. When the volume of plant

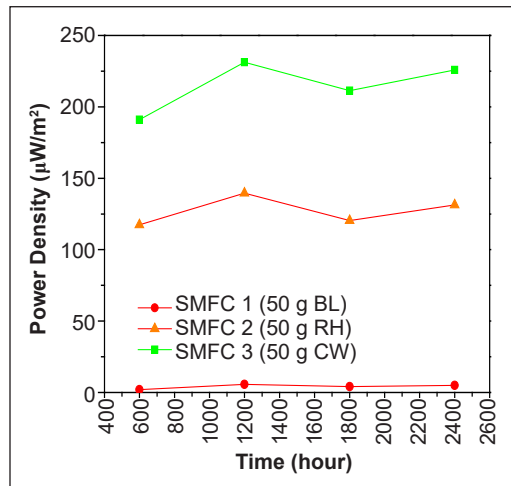


Figure 4. The generation of power density of SMFC with different time

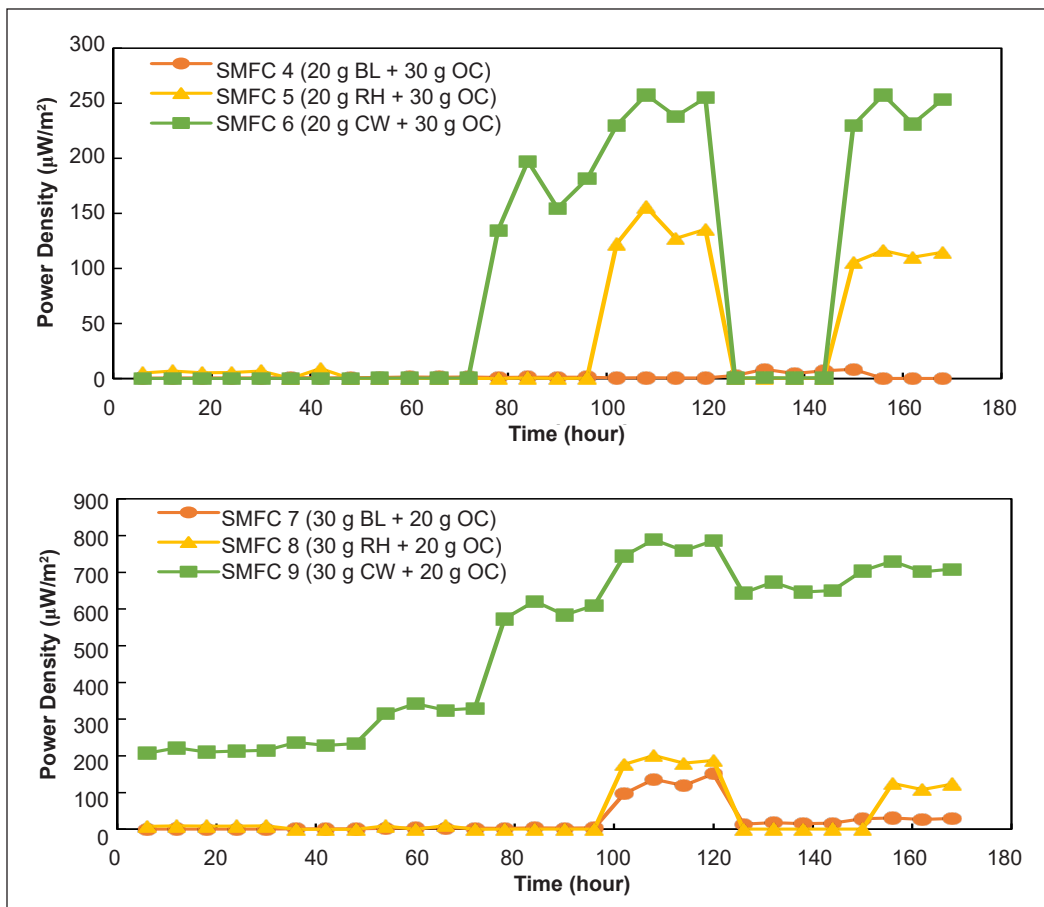


Figure 5. Different compositions of anodic substrates

waste and organic compost was increased, we discovered that the power density of all SMFCs increased by more than 100%. These results indicate that the substrate volume substantially affects the power generation of SMFCs, which is in good agreement with a previous report (Ullah & Zeshan, 2020).

We have plotted graphs summarising our findings to investigate the effect of plant waste and organic compost on the efficacy of SMFC. Figure 6 illustrates only the power generation of SMFCs containing CW, which produced the highest power density compared to the others. The SMFC with the highest power density was the one that contained more CW and less organic compost. In the SMFC, organic compost acted as a catalyst to encourage bacterial growth. The higher power density observed in the SMFC containing more CW and less organic compost may be attributable to sufficient food (CW) for the microbes in the organic compost. In contrast, the SMFC containing less CW and more organic compost revealed that the microbes had little food to sustain their lives and produce bioelectricity (Rahman et al., 2021). Table 3 summarises the findings for all SMFCs in this study.

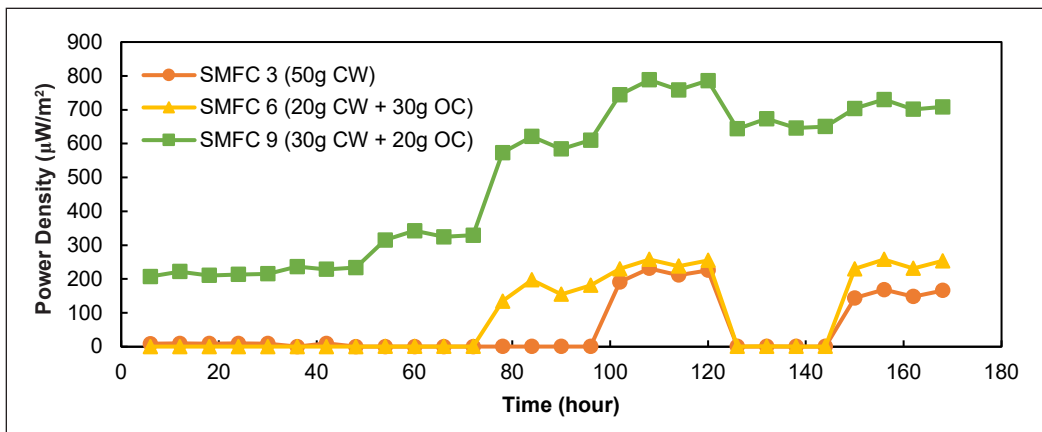


Figure 6. Different concentrations of catalyst (organic compost) to the performance of SMFC

Table 3  
Summary of the power density generation for all SMFCs

SMFC No.	Types of Anodic Chamber Substrates	Maximum Power Density (µW/m <sup>2</sup> )
1	50 g BL + 300 g water	5.84
2	50 g RH + 300 g water	139.61
3	50 g CW + 300 g water	231.18
4	20 g BL + 30 g OC + 300 g water	8.45
5	20 g RH + 30 g OC + 300 g water	155.87
6	20 g CW + 30 g OC + 300 g water	257.65
7	30 g BL + 20 g OC + 300 g water	151.44
8	30 g RH + 20 g OC + 300 g water	201.04
9	30 g CW + 20 g OC + 300 g water	788.58

## COMPARISON OF PREVIOUS STUDIES

Table 4 shows a summary of recent studies on MFCs with regard to anodic chamber substrates. Our findings imply that coconut waste has the potential to be used as a substrate in MFC systems that generate bioelectricity. However, the amount of power density was very small, and optimisation of parameters such as the size of electrodes, size of chamber, and configuration of chambers for the MFC system should be reconsidered.

Table 4  
*Recent literature on MFC with respect to various anodic chamber substrates*

Types of MFC	Substrates in		Power Density (W/m <sup>2</sup> )	Reference
	Anode Chamber	Cathode Chamber		
Single-chamber	Molasses solution	-	0.545 m	Flores et al., 2022
Dual-chamber	Wastewater with glucose	Wastewater	253.84 m	Ye et al., 2019
Dual-chamber	Rice straw with CDB	Nutrient mineral buffer	145.2 m	Hassan et al., 2014
Single-chamber	Kitchen garbage Bamboo waste	-	60 m 40 m	Moqsud et al., 2014
Terrestrial	Soil	Rocks	3 m	Zhang et al., 2013
Single-chamber	Agricultural soil Forest soil	-	42.49 m 2.44 m	Dunaj et al., 2012
Single-chamber	Coconut waste Bamboo waste Rice husk	-	788.58 $\mu$ 151.44 $\mu$ 201.04 $\mu$	This study

## CONCLUSION

We examined the impact of various plant waste types and their quantities on the efficacy of MFCs. Due to its reduced internal resistance, coconut waste that contained the highest proportion of lignin had a greater power density. In contrast, the lower power density observed in SMFCs containing bamboo waste was attributed to its higher percentage of cellulose, whose complex structure requires special enzymes for degradation, making it inappropriate for microorganism digestion. Sufficient food is essential for microorganisms to survive and produce bioelectricity. Careful selection of substrate types, catalysts, and their composition is necessary to enhance the performance of MFCs.

## ACKNOWLEDGEMENTS

A part of this work was supported by the SDG@ Borneo grant (GL/F02/MCUN/18/2020) and Universiti Malaysia Sarawak.

## REFERENCES

- Aghababae, M., Farhadian, M., Jeihanipour, A., & Biria, D. (2015). Effective factors on the performance of microbial fuel cells in wastewater treatment - A review. *Environmental Technology Reviews*, 4(1), 71-89. <https://doi.org/10.1080/09593330.2015.1077896>
- Anuchi, S. O., Campbell, K. L. S., & Hallett, J. P. (2022). Effective pretreatment of lignin-rich coconut wastes using a low-cost ionic liquid. *Scientific Reports*, 12, Article 6108. <https://doi.org/10.1038/s41598-022-09629-4>
- Bai, Y. Y., Xiao, L. P., Shi, Z. J., & Sun, R. C. (2013). Structural variation of bamboo lignin before and after ethanol Organosolv pretreatment. *International Journal of Molecular Sciences*, 14(11), 21394-21413. <https://doi.org/10.3390/ijms141121394>
- Boisset, C., Frascini, C., Schüle Martin, Henrissat, B., & Chanzy, H. (2000). Imaging the enzymatic digestion of bacterial cellulose ribbons reveals the endo character of the cellobiohydrolase Cel6A from *Humicola insolens* and its mode of synergy with cellobiohydrolase CEL7A. *Applied and Environmental Microbiology*, 66(4), 1444-1452. <https://doi.org/10.1128/aem.66.4.1444-1452.2000>
- Chandra, M. R., & Madacka, M. (2019). Comparative biochemistry and kinetics of microbial lignocellulolytic enzymes. In V. Buddolla (Ed.), *Recent Developments in Applied Microbiology and Biochemistry* (pp. 147-159). Academic Press. <https://doi.org/10.1016/b978-0-12-816328-3.00011-8>
- Denchak, M. (2022, June 1). *Fossil Fuels: The Dirty Facts*. NRDC. <https://www.nrdc.org/stories/fossil-fuels-dirty-facts>
- Dunaj, S. J., Vallino, J. J., Hines, M. E., Gay, M., Kobyljanec, C., & Rooney-Varga, J. N. (2012). Relationships between soil organic matter, nutrients, bacterial community structure, and the performance of microbial fuel cells. *Environmental Science & Technology*, 46(3), 1914-1922. <https://doi.org/10.1021/es2032532>
- Energy.Gov. (2017). *Fossil*. U. S. Department of Energy. <https://www.energy.gov/science-innovation/energy-sources/fossil>
- Flimban, S. G., Ismail, I. M., Kim, T., & Oh, S. E. (2019). Overview of recent advancements in the microbial fuel cell from fundamentals to applications: Design, major elements, and scalability. *Energies*, 12(17), Article 3390. <https://doi.org/10.3390/en12173390>
- Flores, S. R., Pérez-Delgado, O., Naveda-Renny, N., Benites, S. M., De La Cruz-Noriega, M., & Narciso, D. A. D. (2022). Generation of bioelectricity using molasses as fuel in microbial fuel cells. *Environmental Research, Engineering and Management*, 78(2), 19-27. <https://doi.org/10.5755/j01.arem.78.2.30668>
- Gao, Y., Guo, X., Liu, Y., Fang, Z., Zhang, M., Zhang, R., You, L., Li, T., & Liu, R. H. (2018). A full utilization of rice husk to evaluate phytochemical bioactivities and prepare cellulose nanocrystals. *Scientific Reports*, 8, Article 10482. <https://doi.org/10.1038/s41598-018-27635-3>
- Hassan, S. H. A., Gad El-Rab, S. M. F., Rahimnejad, M., Ghasemi, M., Joo, J. H., Yong, S. O., Kim, I. S., & Oh, S. E. (2014). Electricity generation from rice straw using a microbial fuel cell. *International Journal of Hydrogen Energy*, 39(17), 9490-9496. <https://doi.org/10.1016/j.ijhydene.2014.03.259>
- Huang, L., & Angelidaki, I. (2008). Effect of humic acids on electricity generation integrated with xylose degradation in microbial fuel cells. *Biotechnology and Bioengineering*, 100(3), 413-422. <https://doi.org/10.1002/bit.21786>



- Khoo, K. S., Chia, W. Y., Tang, D. Y. Y., Show, P. L., Chew, K. W., & Chen, W. H. (2020). Nanomaterials utilization in biomass for biofuel and bioenergy production. *Energies*, *13*(4), Article 892. <https://doi.org/10.3390/en13040892>
- Koch, C., Aulenta, F., Schröder, U., & Harnisch, F. (2016). Microbial electrochemical technologies: Industrial and environmental biotechnologies based on interactions of microorganisms with electrodes. In M. Y. Murray (Ed.), *Comprehensive Biotechnology* (pp. 545-563). Elsevier. <https://doi.org/10.1016/b978-0-12-409548-9.09699-8>
- Kottasová, I., & Dewan, A. (2021, February 9). Fossil fuel air pollution causes almost 1 in 5 deaths globally each year. *Cable News Network*. <https://edition.cnn.com/2021/02/09/world/climate-fossil-fuels-pollution-intl-scn/index.html>
- Kumar, R., Singh, L., & Zularisam, A. W. (2017). Microbial fuel cells: Types and applications. In L. Singh & V. C. Kalia (Eds.), *Waste Biomass Management - A Holistic Approach* (pp. 367-384). Springer. [https://doi.org/10.1007/978-3-319-49595-8\\_16](https://doi.org/10.1007/978-3-319-49595-8_16)
- Kumar, S. D., Yasasve, M., Karthigadevi, G., Aashabharathi, M., Subbaiya, R., Karmegam, N., & Govarthanam, M. (2022). Efficiency of microbial fuel cells in the treatment and energy recovery from Food Wastes: Trends and applications - A review. *Chemosphere*, *287*, Article 132439. <https://doi.org/10.1016/j.chemosphere.2021.132439>
- Moqsud, M. A. (2021). Bioelectricity from organic solid waste. In H. M. Saleh (Ed.), *Strategies of Sustainable Solid Waste Management* (pp. 1-10). IntechOpen. <https://doi.org/10.5772/intechopen.95297>
- Moqsud, M. A., Omine, K., Yasufuku, N., Bushra, Q. S., Hyodo, M., & Nakata, Y. (2014). Bioelectricity from kitchen and bamboo waste in a microbial fuel cell. *Waste Management & Research: The Journal for a Sustainable Circular Economy*, *32*(2), 124-130. <https://doi.org/10.1177/0734242x13517160>
- National Geographic Society. (2019, May 31). *Fossil Fuels*. National Geographic Society. <https://www.nationalgeographic.org/encyclopedia/fossil-fuels/>
- Nawaz, A., ul Haq, I., Qaisar, K., Gunes, B., Raja, S. I., Mohyuddin, K., & Amin, H. (2022). Microbial fuel cells: Insight into simultaneous wastewater treatment and bioelectricity generation. *Process Safety and Environmental Protection*, *161*, 357-373. <https://doi.org/10.1016/j.psep.2022.03.039>
- Rahman, W., Yusup, S., & Mohammad, S. N. (2021). Screening of fruit waste as substrate for microbial fuel cell (MFC). *AIP Conference Proceedings*, *2332*(1), Article 020003. <https://doi.org/10.1063/5.0043556>
- Sahari, S. K., Rosli, M. Z. F., Butit, A. M., Kipli, K., Anyi, M., Awang, A., Sawawi, M., Mahmood, M. R., Hasanah, L., Kram, A. R., Embong, Z., & Nahrawi, H. (2022). Fabrication of single chamber microbial fuel cell (SMFC) using soil as a substrate. *Pertanika Journal of Science and Technology*, *30*(2), 1103-1114. <https://doi.org/10.47836/pjst.30.2.14>
- Sakdaronnarong, C., Ittitanakam, A., Tanubumrungsuk, W., Chaithong, S., Thanosawan, S., Sinbuathong, N., & Jeraputra, C. (2015). Potential of lignin as a mediator in combined systems for biomethane and electricity production from ethanol stillage wastewater. *Renewable Energy*, *76*, 242-248. <https://doi.org/10.1016/j.renene.2014.11.009>
- Ucar, D., Zhang, Y., & Angelidaki, I. (2017). An overview of electron acceptors in microbial fuel cells. *Frontiers in Microbiology*, *8*, Article 643. <https://doi.org/10.3389/fmicb.2017.00643>

- Ullah, Z., & Zeshan, S. (2020). Effect of substrate type and concentration on the performance of a double chamber microbial fuel cell. *Water Science & Technology*, 81(7), 1336-1344. <https://doi.org/10.2166/wst.2019.387>
- United Nations. (2019, June 17). Growing at a slower pace, world population is expected to reach 9.7 billion in 2050 and could peak at nearly 11 billion around 2100. *United Nations News*. <https://www.un.org/development/desa/en/news/population/world-population-prospects-2019.html>
- Ye, Y., Ngo, H. H., Guo, W., Chang, S. W., Nguyen, D. D., Liu, Y., Nghiem, L. D., Zhang, X., & Wang, J. (2019). Effect of organic loading rate on the recovery of nutrients and energy in a dual-chamber microbial fuel cell. *Bioresource Technology*, 281, 367-373. <https://doi.org/10.1016/j.biortech.2019.02.108>
- Zhang, D., Ge, Y., & Wang, W. (2013). Study of a terrestrial microbial fuel cell and the effects of its power generation performance by environmental factors. In *Proceedings of the 2013 International Conference on Advanced Mechatronic Systems* (pp. 445-448). IEEE Publishing. <https://doi.org/10.1109/icamechs.2013.6681825>
- Zhang, J., Li, J., Ye, D., Zhu, X., Liao, Q., & Zhang, B. (2014). Tubular bamboo charcoal for anode in microbial fuel cells. *Journal of Power Sources*, 272, 277-282. <https://doi.org/10.1016/j.jpowsour.2014.08.115>



## Investigation of Blended Seaweed Waste Recycling Using Black Soldier Fly Larvae

Iva Yenis Septiariva<sup>1</sup>, I Wayan Koko Suryawan<sup>2\*</sup>, Wisnu Prayogo<sup>3</sup>, Sapta Suhardono<sup>4</sup> and Ariyanti Sarwono<sup>2</sup>

<sup>1</sup>Civil Engineering Study Program, Faculty of Engineering, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

<sup>2</sup>Department of Environmental Engineering, Faculty of Infrastructure Planning, Universitas Pertamina, Jakarta, 12220, Indonesia

<sup>3</sup>Department of Civil Engineering, Faculty of Engineering, Universitas Negeri Medan, Medan, Indonesia

<sup>4</sup>Department of Environmental Science, Faculty of Engineering, Universitas Sebelas Maret, Surakarta, 57126, Indonesia

### ABSTRACT

Seaweed waste is often found in the coastal areas of Indonesia. Along with other waste, large quantities of seaweed will disrupt the ecosystem and harm the surrounding environment. Therefore, treating seaweed waste effectively and efficiently is necessary to mitigate such adverse effects. This study processed seaweed waste using Black Soldier Fly (BSF) larvae to accelerate waste degradation, reducing the waste buildup. The characteristics of waste, the feeding treatment, chemical content, and parameters of waste processing were assessed. Based on the results, seaweed waste contains highly organic compounds, potentially used as fertilizer or animal feed. The feeding method demonstrated that the mass of BSF larvae increased according to the intensity of the feeding treatment. The total waste degradation obtained from the BSF larvae was 72.61% with a waste reduction index of 1.87 g/day, and

the efficiency of digested feed was 7.61% with feeding treatment every three days. The chemical composition of BSF pre-pupae, pupae, and residues was also evaluated to determine their feasibility as animal feed and biomass. The results suggested that pre-pupae and pupae comprise protein, fat, calcium, and phosphorus. At the same time, the waste residue has a C-organic and N-organic ratio of 13%–14% and is an alternative for future waste-to-energy

### ARTICLE INFO

#### Article history:

Received: 01 February 2023

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.13>

#### E-mail addresses:

[ivayenis@gmail.com](mailto:ivayenis@gmail.com) (Iva Yenis Septiariva)

[i.suryawan@universitaspertamina.ac.id](mailto:i.suryawan@universitaspertamina.ac.id) (I Wayan Koko Suryawan)

[wisnuprayogo@unimed.ac.id](mailto:wisnuprayogo@unimed.ac.id) (Wisnu Prayogo)

[sapta.suhardono@staff.uns.ac.id](mailto:sapta.suhardono@staff.uns.ac.id) (Sapta Suhardono)

[ariyanti.sarwono@universitaspertamina.ac.id](mailto:ariyanti.sarwono@universitaspertamina.ac.id) (Ariyanti Sarwono)

\* Corresponding author

applications. Overall, BSF offers a great prospect in generating high-value products and simultaneously serves as a strategy for efficient waste management.

*Keywords:* BSF larvae, coastal area, seaweed waste, waste to energy

---

## INTRODUCTION

Municipal solid waste has long been one of the main environmental issues worldwide, both on land and in the ocean. The higher the population on earth, the more it is produced. The increased amount of waste can profoundly impact the lives of living things on Earth. Waste can be the main trigger for various human health problems, poor life quality of animals and plants, and environmental degradation. Today, different types of waste generated on land, in a river, and an ocean can damage the surrounding environment (Cordova & Nurhati, 2019; Nurhati & Cordova, 2020). Seaweed waste is one type of marine debris commonly found in coastal areas, particularly in Indonesia (Bijang et al., 2018). Seaweed waste typically occurs during east wind season; naturally, tides and sea currents carry the seaweed waste to the coastline. When present in large quantities, seaweed waste can disrupt the ecosystem around the coast and interfere with the activities of humans, animals, and plants. Areas with vast seaweed waste are Tanjung Benoa Beach and Nusa Dua Beach, Bali. At least two to three tons of seaweed waste are observed daily. The accumulated seaweed waste can disturb the lives of other living things (NusaBali.com, 2021); therefore, an effective and efficient treatment of seaweed waste plays a key role in minimizing various environmental threats and avoiding multiple health issues for humans in the coastal environment.

This seaweed can be processed into high-value products. Seaweed, a plant-derived organic waste, can be easily degraded or decomposed along with other organic waste such as vegetables, fruit, and straw. This type of waste can be easily decomposed with microorganisms and animals. Seaweed waste, in general, contains a high amount of complex carbohydrates (Salgado et al., 2021), protein, water, and cellulose (Hidayati et al., 2021). Several studies have been reported to treat seaweed waste to minimize its negative impact on the surrounding environment. Seaweed waste can be converted into feed and organic fertilizer in the form of granules or liquids and into inorganic fertilizers by utilizing pickling waste from the metal coating industry as a mixture (Sulistiyowati & Prayitno, 2021). Certain microbes or Black Soldier Fly (BSF) can assist in the degradation of seaweed and accelerate the decomposition process of seaweed (Liland et al., 2017; Swinscoe et al., 2020).

The main objective of this study is to explore the effectiveness and efficiency of BSF larvae in the degradation and conversion of seaweed waste, thereby providing insights into sustainable waste management practices and potential applications for the resulting

products. Piles of seaweed waste could be easily and rapidly reduced while concurrently curtailing environmental risks in coastal areas. Investigating blended seaweed waste recycling using BSF larvae addresses significant challenges and unsolved problems in waste management and resource utilization. Coastal areas in Indonesia often face the accumulation of seaweed waste, which poses environmental and economic burdens. Current waste management practices may not effectively handle the increasing volumes of seaweed waste, necessitating sustainable and efficient methods for waste degradation. Additionally, the underutilization of seaweed waste as a valuable resource for fertilizer, animal feed, and waste-to-energy applications remains a pressing issue.

Investigating blended seaweed waste recycling using BSF larvae is of great importance. Firstly, it offers a sustainable approach to managing seaweed waste by utilizing the degrading capabilities of BSF larvae, thereby reducing waste accumulation and minimizing environmental impact. Secondly, this study contributes to resource recovery by converting seaweed waste into valuable products, including fertilizer, animal feed, and potentially even energy sources. By embracing the principles of the circular economy, the study promotes the efficient use of resources and reduces waste generation. The novelty of this investigation lies in its specific focus on blended seaweed waste recycling using BSF larvae. While previous studies have explored waste management and resource utilization through the application of BSF larvae, this study extends the knowledge base by targeting the specific challenge of seaweed waste. The study introduces a novel and innovative approach to waste management and resource utilization by employing BSF larvae to process and convert seaweed waste into valuable products.

## **MATERIALS AND METHODS**

### **Feeding**

Seaweed waste was dried and homogenized by a blender for the feeding substrate. This seaweed was dried to maintain the salinity of seawater. The research was conducted with two repetitions of treatment. The waste feeding rate into the reactor can be calculated using an average of 40 mg/larva.day, according to Diener (2010). Various feeding treatment times of once a day (F24), every two days (F48), and every three days (F72) were performed in this study.

The water content of the waste profoundly affects the larval consumption time of the given waste, and 60%–90% wt/wt of water content is considered optimal (Diener, 2010). The higher the water content in the given waste, the more BSF larvae escape from the breeding reactor. On the contrary, the lack of water content might inhibit the digestive process of BSF larvae. Therefore, in this study, water content was adjusted to 70% wt/wt by adding water to the waste.

## BSF Larvae

In this study, the reactor was seeded with 200 larvae of one-week-old. When sufficient food reserves are available, BSF larvae have a better tolerance to lower temperatures and grow faster at a temperature of 30°C–36°C. At the same time, young larvae show low survival at less than 20°C and higher than 45°C. Young larvae are typically very susceptible to the influence of external factors, including temperature, low oxygen pressure, mold, moisture content, and toxic materials. Its resistance to these factors will increase after 1 week (5–10 mg in size). After 10 days of age, these larvae will compete with the older ones in the breeding incubator. After hatching, BSF can reduce waste from the larval stage to the prepupa stage to approximately 55% wt/wt of waste (Diener, 2010).

Data on larval wet weight gain is measured using analytical balance to determine the effect of the waste type and feeding frequency on larval growth. During the experiment, 10% of the larvae were weighed in each reactor, and 20 batches were taken. Once weighted, the larvae were immediately returned to the reactor (Diener et al., 2011).

## Data Analysis

Feed residues were recorded to determine the overall degradation (OD) and waste reduction index (WRI) as presented by Equations 1 and 2. The larval digestibility of the feed consumed is represented in the digested feed conversion efficiency (ECD), as shown in Equation 3. The metabolism of feed consumed (g) was calculated based on the mass balance.

$$\text{Overall degradation (OD)} = \frac{\text{Total feed offered (g)} - \text{Dry residue remained (g)}}{\text{Total feed offered (g)}} \quad (1)$$

$$\text{Waste reduction index (WRI)} = \frac{\text{Total feed offered (g)} - \text{Dry residue remained (g)}}{\text{Rearing duration (day)}} \quad (2)$$

$$\text{Efficiency of conversion of digested feed (ECD)} = \frac{\text{BSF Larvae biomass (g)}}{\text{Total feed offered (g)} - \text{Dry residue remained (g)}} \quad (3)$$

The protein, fat, calcium, phosphorus, and ash content of the BSF prepupa and pupae were measured to explore further the opportunities for cultivating BSF prepupa and pupae as animal feed. The Standar Nasional Indonesia (SNI) methods can be utilized to analyze the parameters of protein, fat, calcium, phosphorus, and ash content in solid fertilizer. These methods are widely recognized and accepted as standard practices in Indonesia. The SNI 01-2891-1992 method can be followed for protein content analysis, which employs the Kjeldahl or Dumas combustion methods. Fat content can be determined using the Soxhlet extraction method according to the procedures outlined in SNI 01-3555-1998. The appropriate technique specified in SNI 01-2894-1992 can be employed to analyze the calcium content, which may involve atomic absorption spectroscopy or other suitable



methods described in the standard. Similarly, the phosphorus content can be determined using the methods outlined in SNI 01-2894-1992, which may include colorimetric methods or other applicable techniques. Lastly, the ash content can be analyzed using the gravimetric method described in SNI 01-3554-1998, following the specific steps and procedures mentioned in the standard. Adhering to these SNI methods ensures compliance with the Indonesian standards and facilitates accurate and reliable analysis of the mentioned parameters in solid fertilizer samples.

Analysis of residues resulting from the processing by BSF larvae is adjusted to the Indonesian government's solid fertilizer standards. This study tested solid fertilizer parameters such as pH, C-organic, N-organic, C/N ratio, and water content. The SNI 06-6989-2004 method can be used for pH measurement, which involves using a pH meter or indicator to measure the pH after creating a suspension of the sample in water. The SNI 19-7030-2004 method provides guidelines for determining C-organic (Organic Carbon) and N-organic (Organic Nitrogen) contents. The C-organic content can be measured using the Walkley-Black method or other suitable methods, while the N-organic content can be determined through the Kjeldahl method or other appropriate techniques. The SNI 19-7030-2004 method also outlines the procedure for calculating the C/N ratio based on the determined organic carbon and organic nitrogen contents. Lastly, the water content can be analyzed following the SNI 19-7030-2004 method, which involves drying the sample at a specific temperature and calculating the weight difference before and after drying. Adhering to these standard methods ensures accurate and consistent analysis of the specified parameters in accordance with SNI guidelines.

In addition to the mentioned solid fertilizer parameters, further assessments were conducted to evaluate the quality and composition of the residues. Two additional tests were performed: Thermogravimetric Analysis (TGA) and Gas Chromatography-Mass Spectrometry (GC-MS). TGA analysis was employed to investigate the thermal properties and stability of the residues. This test allows for the determination of weight changes as a function of temperature, providing insights into the decomposition and volatilization behavior of the samples. On the other hand, GC-MS analysis was utilized to examine the chemical composition of the residues and identify the presence of organic compounds. By separating and analyzing the individual components present in the sample, GC-MS provides valuable information about the organic constituents and their relative abundance. These additional tests contribute to a comprehensive understanding of the residues' characteristics and aid in assessing their suitability as solid fertilizers.

## RESULTS

In this study, the seaweed waste was characterized prior to being fed to maggot to obtain the nutritional content of maggot feed before being used as a medium for maggot growth.

The characteristics of seaweed waste are summarized in Table 1. Based on these data, seaweed waste has a neutral pH between 7.6–7.8, a water content of 74.2%, and an ash content of 10.1%. Maintaining a near-neutral pH is crucial as a low pH excessively produces nitrogen and might kill eggs from insects (Setyaningsih et al., 2017). pH can affect the amount of nutrients absorbed by BSF larvae by affecting the performance of enzymes in the larvae intestines (Kim et

Table 1

*The characteristics of seaweed waste*

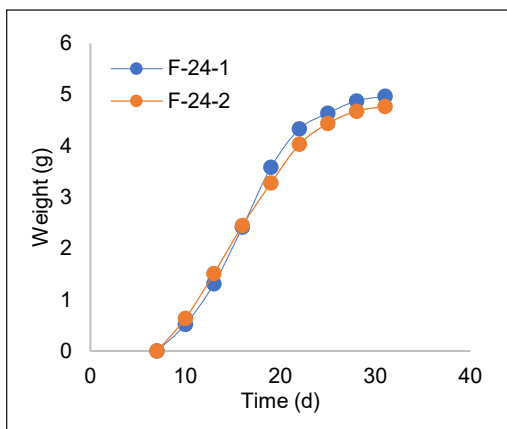
Parameters	Results
Temperature	31°C–33°C
Moisture Content	74.2%±0.8
Ash Content	10.1%±0.4
pH	7.6–7.8
C-organic	39.12%
N-organic	3.85%
C/N	10.16

al., 2011). The pH treatments impacted BSF larval weight largely on the first, third, and fifth day instead of the final stage (Meneguz et al., 2018). Data from Table 1 suggest that the breeding site conditions are particularly suitable for breeding BSF larvae as it is in accordance with the optimum temperature for larval growth of 30°C–36°C (Green & Popa, 2012) and an optimum air humidity of 60%–70% (Holmes et al., 2012).

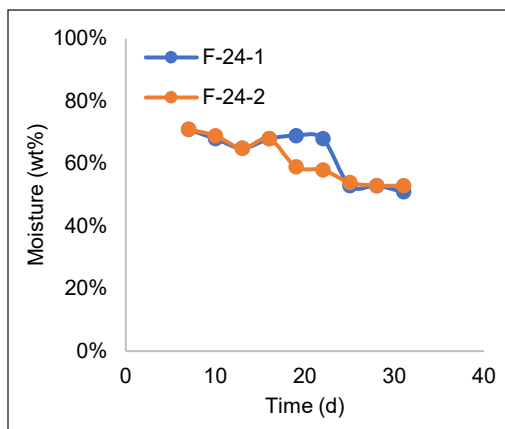
On day 20, when prepupa transforms into a fly, this adult BSF does not require food but merely utilizes fat reserves in the body as an energy source (Diener et al., 2011). In the prepupa phase, BSF lends itself to self-harvesting (Richard, 2017), and based on this, once turning into a prepupa, the wet and moist prepupa is removed from the reactor. For the feeding treatment every two days (Figure 1b), the BSF larvae weight gain was 4.63 g. Meanwhile, feeding every three days (Figure 1c) showed an increase in the average mass of BSF larvae from 0 to 4.42 g. Figure 1 implies that daily feeding treatment had the highest gain weight of BSF larvae, followed by feeding every two and finally every three days. With more intensity of feeding, the BSF larvae will grow rapidly and have nutrients-rich for breeding.

The BSF larvae breeding was carried out using the feeding method at a specified time. In the 7<sup>th</sup> to the 31<sup>st</sup>-day range, the feeding time was performed in three ways: every one day, every two days, and every three days. Batch feeding allowed the shortest development time, while daily feeding revealed the heaviest final weight of larvae (Meneguz et al., 2018). Different feeding time aims to identify the optimum conditions for BSF larvae to reduce seaweed waste in large quantities and a short period, as illustrated in Figures 1 and 2. Figure 1 presents the mass changes in BSF larvae after feeding every day, every two days, and every 3 days. All feeding conditions show an increasing trend every day from the 7<sup>th</sup> to the 31<sup>st</sup>. In daily feeding treatment (Figure 1a), the mean larval weight reached 3.59 g and 4.87 on the 19<sup>th</sup> and 31<sup>st</sup> days, respectively.

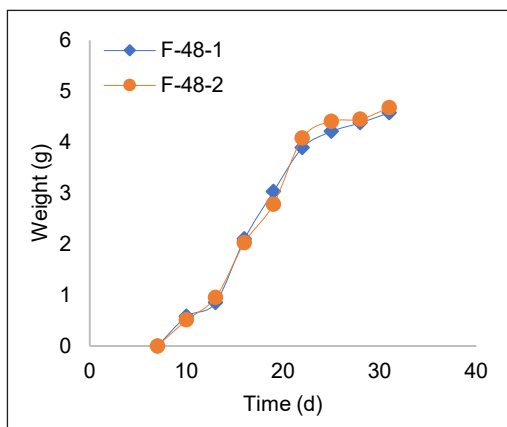
To evaluate the water content in the seaweed waste changes in moisture content of seaweed waste against feeding time were measured and presented in Figure 2. This feeding



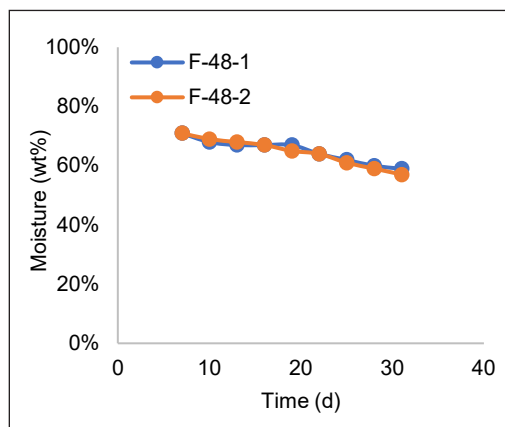
(a)



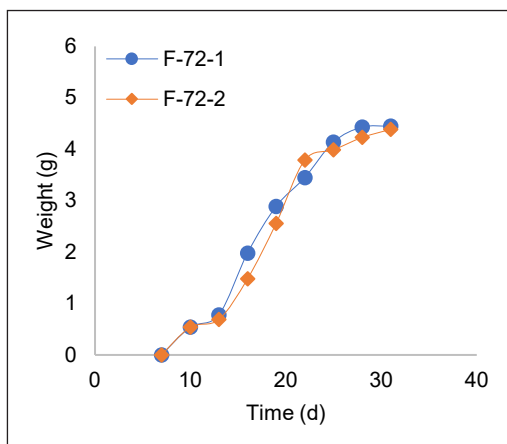
(a)



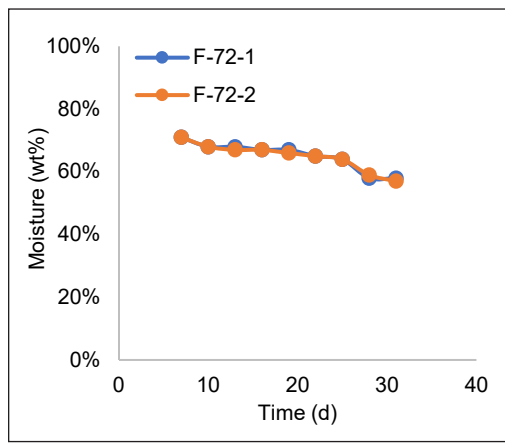
(b)



(b)



(c)



(c)

Figure 1. Changes in the mass of BSF larvae from day 7 to day 31: (a) feeding every day; (b) feeding every 2 days; and (c) feeding every 3 days

Figure 2. Changes in the waste moisture of the seaweed waste from day 7 to day 31: (a) feeding every day; (b) feeding every 2 days; and (c) feeding every 3 days

time parameter was performed similarly to in Figure 1. The results reveal that the overall water content of seaweed waste slightly decreased over time. In the first condition of daily feeding treatment (Figure 2a), the water content in the seaweed waste declines from 71% to 52%. It indicates a noticeable reduction in the water content of the seaweed waste throughout the experimental duration. During the feeding process, the larvae consume the organic material and extract nutrients from it. As the larvae consume the seaweed waste, they break it into smaller particles through mechanical action and enzymatic digestion. During this process, the larvae extract moisture from the waste to support their metabolic activities and growth. This moisture extraction from the seaweed waste contributes to reducing its water content. Moreover, BSF larvae have efficient digestive systems that can efficiently process organic matter, including cellulose-rich materials like seaweed. In addition, they have specialized gut enzymes that help break down complex carbohydrates and release water as a byproduct of digestion. It further contributes to the decrease in water content in the seaweed waste.

In addition to OD and WRI analysis, the efficiency of feed conversion by BSF larvae efficiency of conversion of digested food (ECD) was also estimated to determine the conversion of digested food to body substance. Based on the calculations in Table 2, the average ECD value is 7.44% for feeding every day, 7.23% for feeding every two days, and 7.61% for feeding every three days. The ECD illustrates the efficiency level of BSF larvae in converting the consumed feed into their biomass (Abduh et al., 2017). The higher the ECD value, the higher the level of efficiency. These results show that feeding every three days increases the conversion efficiency of seaweed waste by BSF larvae. The difference in the percentage efficiency of the various feeding treatments was possibly influenced by several factors, such as the amount of food given to the larvae, environmental factors like temperature, humidity, and pH, as well as internal factors such as disease, presence of parasites, and the sex of the BSF larvae.

In addition to waste processing parameters, in this study, the composition of mass in seaweed processing using BSF larvae was also depicted in Figure 3. From Figure 3, the percentage of dry feed and biomass larvae every three-day feeding treatment had the highest

Table 2  
*The results of seaweed waste processing parameters with BSF larvae*

Parameters	Unit	F-24		F-48		F-72	
		1	2	1	2	1	2
Total dry feed offered	g	248	248	120	120	80	80
Dry residue remained	g	185.58	178.98	51.33	59.81	22.39	21.43
OD	%	25.17	27.83	57.23	50.16	72.01	73.21
WRI	g/day	2.01	2.23	2.22	1.94	1.86	1.89
ECD	%	7.96	6.91	6.67	7.78	7.72	7.50

percentage, followed by feeding treatment every two days and daily feeding treatment. In contrast, daily feeding treatment residue holds the highest percentage.

For further analysis, measurements of chemical content in pre-pupa and pupae from waste processing with BSF larvae were performed, as summarized in Table 3. Based on the measurement results, s and both comprise the BSF's prepupa and pupae and comprise a high portion of protein and fat. Table 3 shows the average protein, fat, and ash contents of 40%, 34%, and 19%, respectively, for both pre-pupa and pupae. As for calcium and phosphorus, the average value is 4% and 1%. The feeding treatment for BSF larvae did not significantly influence the percentage of chemical content present in pre-pupa and BSF pupae. The high protein content of BSF larvae might be affected by dry weight conversion as well as the nutrition of larval food ingredients (Fadhillah & Bagastyo, 2020).

Although BSF larvae are able to decompose seaweed waste efficiently, they still generate residue. This residue can be further utilized as plant nutrients or composting. Therefore, this study assessed the chemical composition of BSF larvae processing residue.

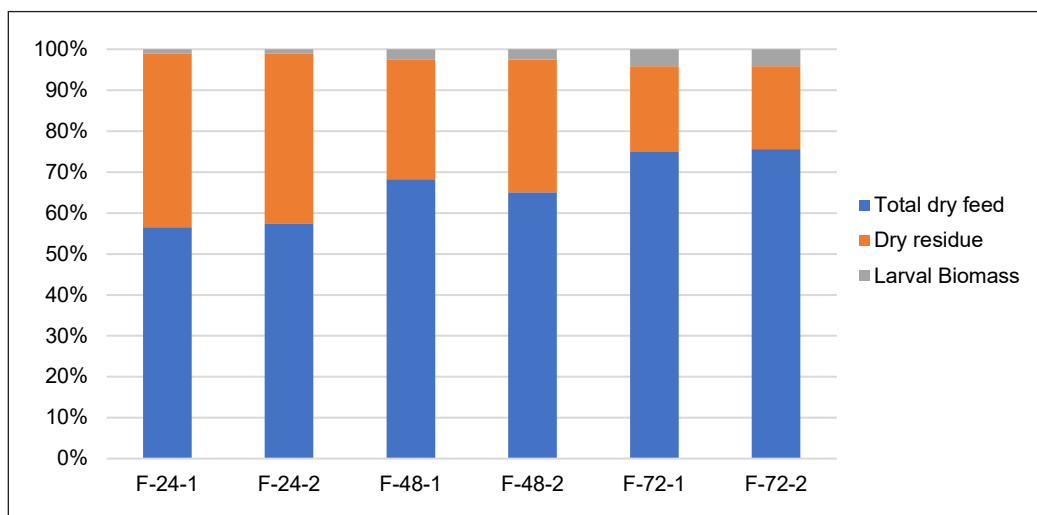


Figure 3. Mass composition in seaweed waste treatment with BSF larvae

Table 3

Chemicals content in pre-pupa and pupae from waste processing with BSF larvae

Chemical content	Unit	Pre-pupa			Pupa		
		F-24	F-48	F-72	F-24	F-48	F-72
Protein	%	42.2	42.8	40.6	41.3	39.4	39.2
Fat	%	34.3	33.3	34.3	35.4	37.8	36.7
Calcium	%	4.32	3.43	4.84	3.84	4.35	4.78
Phosphor	%	0.98	0.93	0.89	1.05	1.06	1.01
Ash	%	18.2	19.54	19.37	18.41	17.39	18.31

The results in Table 4 suggest that the feeding treatment does not significantly impact the chemical content of the residue. These measurements show that waste residue has an acidic pH ranging from 4.6–4.8, C-organic in the 36%–37% range, N-organic from 2.6%–2.8% with a C/N ratio of 13–14, and water content from 14%–19%. These parameters still fall in the standard range, indicating a safe percentage of the chemical content of the residue. Overall, the residue gives a lower percentage of chemical content than seaweed waste prior to processing (Table 1), implying that decomposing seaweed waste using BSF larvae has been successfully carried out.

GC-MS test was conducted to identify residue compounds under the thermal process, as shown in Figure 4. Based on the GC-MS chromatogram, the seaweed residue decomposes into simpler compounds, as many as 50 compound fragments. Some of the compounds produced from the pyrolysis process of seaweed waste residues include

Table 4  
Chemical content in seaweed waste residue with BSF larvae

Chemical content	Unit	F-24		F-48		F-72		Standard
		F-24-1	F-24-2	F-48-1	F-48-2	F-72-1	F-72-2	
pH		4.8	4.7	4.89	4.8	4.6	4.8	4-8
C-organic	%	36.2	37.6	36.8	36.5	37.3	36.4	≥12
N-organic	%	2.7	2.8	2.6	2.7	2.6	2.7	-
C/N		13.41	13.43	14.15	13.52	14.35	13.48	10-25
Water Content	%	18.3	19.2	15.2	14.5	14.3	14.8	13-20

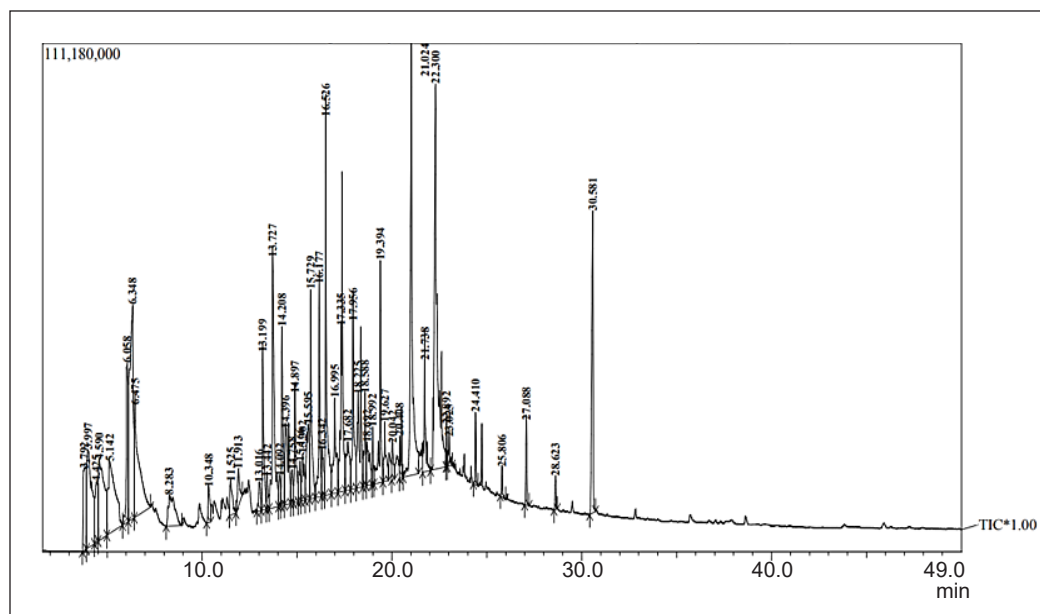


Figure 4. GC-MS chromatogram of seaweed waste residue after processing with BSF larvae

aromatic compounds, alcohol, and ketones such as butyrolactone at 6.30%, hexadecanoic acid at 6.89%, benzene at 4.18%, 2,3-butanedione of 4.14%, phenol of 4.36%, and other compounds (Figure 4). These above compounds can be utilized as biomass precursors if processed with the right method to increase the value of benefits. From GC-MS analysis of seaweed waste residue after processing with BSF larvae, it can be concluded that the residue can be treated using a thermal process to generate biomass fuel.

## DISCUSSION

The organic N levels are generally derived from proteins and chlorophyll in the seaweed. Safia et al. (2020) reported 4.16% of protein, 0.36% of fat, and 25.50% of carbohydrate content in seaweed. In addition to the above nutrients, seaweed possesses beneficial bioactive compounds like alkaloids, flavonoids, phenol hydroquinone, and tannins (Safia et al., 2020). The used seaweed waste has a fairly high nutrient content from organic C-organic and N-organic levels. Microorganisms utilize the C and N organic content as an energy source and the formation of body cells (Widyastuti et al., 2021). The above high nutritional content indicates that seaweed waste has great potential as a BSF diet. Due to relatively high water content, seaweed waste needs to be preliminary treated to reduce its water content and steadily maintain larvae consumption. The seaweed waste used in the study had C-organic and N-organic levels of 39.12% and 3.85%, respectively, with a C/N ratio of 10.16. The organic C is derived from carbohydrates such as cellulose, glucose, and lignin. The C/N ratio 15 can reportedly enhance BSF larvae's nutrient consumption, shortening the maturation feed time to five weeks (Beesigamukama et al., 2021). The nutritional composition of the larvae strongly depends on the consumed feeding medium (Sheppard et al., 2002; Tschirner & Simon, 2015).

Daily feeding time showed the highest water content decrement among the three feeding times. The decrease in the water content in the sample indicated that the feeding of BSF larvae through seaweed waste was successfully performed. The amount of feeding for BSF larvae will affect the water condition in waste. The more intense feeding will stimulate the BSF larvae to decompose the seaweed waste, thereby reducing the water content of the seaweed waste.

Previous research stated that low water content of the media is recommended during cultivating BSF. High water content above 70% reportedly hinders the development and growth of the larvae (Diener, 2010; Li et al., 2011). The high water content in the growing media/feed for larvae results in anaerobic conditions (Elvita & Arseto, 2015). The decomposition of organic matter under anaerobic conditions produces ammonia and methane, which can inhibit the feed consumption process by larvae and disturb their growth (Elvita & Arseto, 2015).

Various analyses related to the optimum conditions for feeding BSF larvae, water content in the sample, and parameters associated with seaweed waste processing using



BSF larvae are carried out in this study. Overall, the average seaweed waste degradation (OD) by BSF larvae is 26.50% w/w for feeding every day, 53.69% w/w for feeding every two days, and 72.61% w/w for feeding every three days. The OD value implies that the feeding treatment has the most substantial degradation every three days. It might be due to less food in the same period, stimulating BSF larvae to devour more. The waste reduction index (WRI), a calculation to determine the ability of BSF larvae to consume seaweed waste within a predetermined period (Rofi et al., 2021), is also carried out to indicate the high preference of BSF larvae for the food given. Table 2 shows that after the 31st day, the number of WRIs in the daily feeding treatment (average) was 2.12 g/day, 2.08 g/day for feeding every two days, and 1.87 g/day for a three-day feeding treatment. The daily feeding treatment had the highest number of WRIs because of more available food for BSF.

The higher the feed consumption value, the greater the potential for using larvae to break down feed or waste. A feed consumption value of 26.2%–39.7% was reported by Diener et al. (2009) using chicken feed as bait for BSF larvae. Another feed consumption value of 9.29%–36.82% was reported using cassava waste (Supriyatna et al., 2016). The value gap was likely due to the different quality of the bait, which subsequently affected the nutrient contribution for the development of BSF larvae. Fish waste provides a greater source of protein than cassava waste and chicken feed. While the nutrient composition affected the growth and development of BSF larvae, the bait with high protein and fat content accelerated the increase in larval weight (Sheppard et al., 2002).

Direct and indirect utilization of BSF larvae as animal feed has its advantages. BSF larvae and pre-pupae can be reared in organic waste, used as poultry feed, and simultaneously decompose unused waste. However, further studies on economic utilization are urgently needed, including increasing the ability of maggots to recycle organic waste (Handayani et al., 2021; Pendyurin et al., 2021; Wangko, 2014). BSF larvae have been recommended as an alternative protein source for animal feed, as in corn and soybeans (Kawasaki et al., 2019). The reported protein content of BSF larvae was 12.71% (young jackfruit as a feed substrate) and 11.30% (banana peel as a feed substrate) (Pangestu et al., 2017). In addition, the proximate test of 7 days old pupa stage is crude protein content of 35.40%–42.31%, crude fat of 3.33%–36.41%, crude fiber of 18.68%–37.60%, and beta-N of 0.03%–10.33% (Mujahid et al., 2017).

Among all the above methods, using black soldier fly (BSF) to degrade organic waste, especially seaweed waste, has recently gained attention and is considered effective. BSF is a species of *Hermetia illucens*, originating from the genus *Hermetia*, and is a family group of *Stratiomyidae* with the order Diptera having a physical shape like a wasp. The BSF larvae of these flies are widely used to accelerate the decomposition of organic waste by breaking down waste through heat generated from the larvae (Raksasat et al., 2020). Using BSF larvae, organic waste reduction can be fastened with the reduction acceleration reaching

50%–60% (Ojha et al., 2020), which is an effective way to decompose organic waste. BSF larvae possess high nutritional value, facilitating their growth and metamorphosis. The body weight of the larvae serves as an indirect indicator of not only the quantity of nitrogen successfully absorbed by the larvae but also the stored energy utilized in the development of organs and tissues during metamorphosis. The significant nutrient content and widespread availability of BSF larvae, coupled with their utilization that does not compete with human resources and the ease of cultivating them using readily available growth media, indicate the promising potential of BSF larvae as a valuable source of animal feed or for aquaculture purposes, particularly in coastal regions.

Using BSF larvae has proven highly effective and efficient in degrading and converting seaweed waste. The study results demonstrate the rapid decomposition of the waste material, with significant waste reduction achieved through the feeding activities of the larvae. It indicates the effectiveness of BSF larvae in breaking down the organic components of seaweed waste. Moreover, the waste degradation process by BSF larvae offers a sustainable waste management practice. It provides an environmentally friendly solution for handling and reducing the accumulation of seaweed waste in coastal areas.

Furthermore, the larvae contribute to developing sustainable waste-to-resource strategies by effectively converting the waste into valuable products, such as protein-rich pre-pupae and pupae. BSF larvae's waste conversion efficiency is evident in the high digested feed efficiency achieved during the study. In addition, the larvae efficiently convert the organic matter in the seaweed waste into biomass, demonstrating their ability to utilize and transform waste materials into valuable resources. Furthermore, conducting an environmental impact assessment is crucial to understanding the sustainability of implementing BSF larvae-based waste management systems. This assessment should include evaluating the carbon footprint, energy efficiency, and potential emissions associated with large-scale implementation and assessing the effects on local ecosystems, soil health, and water quality. Lastly, exploring the economic viability and market potential of utilizing BSF larvae for seaweed waste management is essential. Assessing the cost-effectiveness of large-scale production, examining potential business models, and understanding market demand for the resulting products will provide valuable information for this waste management approach's practical implementation and commercialization (Suryawan et al., 2023).

## CONCLUSION

From the study results, the decomposition of seaweed waste can be carried out rapidly using BSF larvae. Different feeding treatments of BSF larvae induce changes in larval mass and water content in seaweed waste—the more intense the feeding, the heavier the larval mass, and the less water content in the seaweed waste. With the optimum feeding time of

once every three days, BSF larvae could degrade waste by 72.61% with a waste reduction index of 1.87 g/day and digested feed efficiency of 7.61%. Further, chemical composition was also determined for pre-pupae, pupae, and waste residues to determine their potential as animal feed. Pre-pupa and pupae contain large amounts of protein (40%) and fat (34%), followed by calcium, phosphorus, and dust.

In contrast, all parameters for residue are lower than waste before processing, with the highest C-organic content of 36%–37% and water content of as much as 14%–19%. BSF larvae-treated seaweed waste has advantages such as easy processing, high decomposition speed, and vast applications such as waste-to-biofuel and animal feed. Therefore, BSF's processing of seaweed waste offers a solution to reduce environmental pollution in coastal areas.

Future research recommendations can be proposed to build upon these results and explore additional aspects of sustainable waste management practices and potential applications for the resulting products. Firstly, further research can focus on optimizing the feeding strategies for BSF larvae to enhance waste degradation efficiency. It could involve investigating different feeding frequencies, quantities, and nutrient compositions to determine the optimal conditions for maximizing waste decomposition rates and the quality of the resulting products. In addition, there is a need to explore innovative approaches for valorizing the waste residues left after the BSF larvae processing. Finally, future studies could evaluate the potential applications of these residues as biofertilizers, biofuels, or other value-added products, thereby creating a more comprehensive and sustainable waste management process.

## ACKNOWLEDGMENT

We extend our profound gratitude to the coastal communities of Indonesia for granting us access and invaluable insights into the challenge posed by seaweed waste. Their cooperation and firsthand accounts were instrumental in shaping the direction of this research. We also acknowledge the team's dedication responsible for the meticulous collection, treatment, and analysis of seaweed samples. Our sincere thanks go to the experts in the Black Soldier Fly cultivation field, whose guidance ensured the success of the waste degradation process. Furthermore, we appreciate the financial and infrastructural support provided by Universitas Sebelas Maret, Surakarta, Indonesia, and Universitas Pertamina, Jakarta, Indonesia, without which this research would not have been possible.

## REFERENCES

- Abduh, M. Y., Manurung, R., Faustina, A., Affanda, E., Siregar, I. R. H., & Muhammad, C. (2017). Bioconversion of *Pandanus tectorius* using black soldier fly larvae for the production of edible oil and protein-rich biomass. *Journal Of Entomology and Zoology Studies*, 5(1), 803-809.

- Beesigamukama, D., Mochoge, B., Korir, N. K., K.M. Fiaboe, K., Nakimbugwe, D., Khamis, F. M., Subramanian, S., Wangu, M. M., Dubois, T., Ekesi, S., & Tanga, C. M. (2021). Low-cost technology for recycling agro-industrial waste into nutrient-rich organic fertilizer using black soldier fly. *Waste Management*, *119*, 183-194. <https://doi.org/https://doi.org/10.1016/j.wasman.2020.09.043>
- Bijang, C. M., Latupeirissa, J., & Ratuhanrasa, M. (2018). Biosorption of copper metal ions (Cu<sup>2+</sup>) on brown seaweed (*Padina australis*) Biosorbent. *Indonesian Journal of Chemical Research*, *6*(1), 538-549. <https://doi.org/10.30598/ijcr.2018.6-cat>
- Cordova, M. R., & Nurhati, I. S. (2019). Major sources and monthly variations in the release of land-derived marine debris from the Greater Jakarta area, Indonesia. *Scientific Reports*, *9*(1), Article 18730. <https://doi.org/10.1038/s41598-019-55065-2>
- Diener, S. (2010). *Valorisation of Organic Solid Waste using the Black Soldier Fly, Hermetia illucens*, in *Low and Middle-Income Countries* (Doctoral dissertation). Eidgenössische Technische Hochschule Zürich. <https://doi.org/10.3929/ethz-a-6559779>
- Diener, S., Solano, N. M. S., Gutiérrez, F. R., Zurbrügg, C., & Tockner, K. (2011). Biological treatment of municipal organic waste using black soldier fly larvae. *Waste and Biomass Valorization*, *2*(4), 357-363. <https://doi.org/10.1007/s12649-011-9079-1>
- Diener, S., Zurbrügg, C., & Tockner, K. (2009). Conversion of organic material by black soldier fly larvae: Establishing optimal feeding rates. *Waste Management & Research*, *27*(6), 603-610. <https://doi.org/10.1177/0734242X09103838>
- Elvita, S. S., & Arseto, Y. B. (2015, November 23-24). *Reduction of organic solid waste by black soldier fly (Hermetia illucens) larvae*. [Paper Presentation]. 5th Environmental Technology and Management Conference “Green Technology towards Sustainable Environment”, Bandung, Indonesia.
- Fadhillah, N., & Bagastyo, A. Y. (2020). Utilization of *Hermetia illucens* larvae as a bioconversion agent to reduce organic waste. In *IOP Conference Series: Earth and Environmental Science* (Vol. 506, No. 1, p. 012005). IOP Publishing. <https://doi.org/10.1088/1755-1315/506/1/012005>
- Green, T. R., & Popa, R. (2012). Enhanced ammonia content in compost leachate processed by black soldier fly larvae. *Applied Biochemistry and Biotechnology*, *166*(6), 1381-1387. <https://doi.org/10.1007/s12010-011-9530-6>
- Handayani, D., Naldi, A., Larasati, R. R. N. P., Khaerunnisa, N., & Budiarmaka, D. D. (2021). Management of increasing economic value of organic waste with Maggot cultivation. In *IOP Conference Series: Earth and Environmental Science* (Vol. 716, No. 1, p. 012026). IOP Publishing. <https://doi.org/10.1088/1755-1315/716/1/012026>
- Hidayati, S., Zulferiyenni, Maulidia, U., Satyajaya, W., & Hadi, S. (2021). Effect of glycerol concentration and carboxy methyl cellulose on biodegradable film characteristics of seaweed waste. *Heliyon*, *7*(8), Article e07799. <https://doi.org/https://doi.org/10.1016/j.heliyon.2021.e07799>
- Holmes, L. A., Vanlaerhoven, S. L., & Tomberlin, J. K. (2012). Relative humidity effects on the life history of *Hermetia illucens* (Diptera: Stratiomyidae). *Environmental Entomology*, *41*(4), 971-978. <https://doi.org/10.1603/EN12054>

- Kawasaki, K., Hashimoto, Y., Hori, A., Kawasaki, T., Hirayasu, H., Iwase, S., Hashizume, A., Ido, A., Miura, C., Miura, T., Nakamura, S., Seyama, T., Matsumoto, Y., Kasai, K., & Fujitani, Y. (2019). Evaluation of black soldier fly (*Hermetia illucens*) larvae and pre-pupae raised on household organic waste, as potential ingredients for poultry feed. *Animals*, 9(3), Article 98. <https://doi.org/10.3390/ani9030098>
- Kim, W., Bae, S., Park, K., Lee, S., Choi, Y., Han, S., & Koh, Y. (2011). Biochemical characterization of digestive enzymes in the black soldier fly, *Hermetia illucens* (Diptera: Stratiomyidae). *Journal of Asia-Pacific Entomology*, 14(1), 11-14. <https://doi.org/https://doi.org/10.1016/j.aspen.2010.11.003>
- Li, Q., Zheng, L., Cai, H., Garza, E., Yu, Z., & Zhou, S. (2011). From organic waste to biodiesel: Black soldier fly, *Hermetia illucens*, makes it feasible. *Fuel*, 90(4), 1545-1548. <https://doi.org/https://doi.org/10.1016/j.fuel.2010.11.016>
- Liland, N. S., Biancarosa, I., Araujo, P., Biemans, D., Bruckner, C. G., Waagbø, R., Torstensen, B. E., & Lock, E.-J. (2017). Modulation of nutrient composition of black soldier fly (*Hermetia illucens*) larvae by feeding seaweed-enriched media. *Plos One*, 12(8), Article e0183188. <https://doi.org/10.1371/journal.pone.0183188>
- Meneguz, M., Gasco, L., & Tomberlin, J. K. (2018). Impact of pH and feeding system on black soldier fly (*Hermetia illucens*, L; Diptera: Stratiomyidae) larval development. *Plos One*, 13(8), Article e0202591. <https://doi.org/10.1371/journal.pone.0202591>
- Mujahid, M., Amin, A. A., Hariyadi, H., & Fahmi, M. R. (2017). Oil palm empty bunches bioconversion using *Trichoderma* sp. and black soldier fly larvae as poultry feed composition. *Jurnal Ilmu Produksi Dan Teknologi Hasil Peternakan*, 5(1), 5-10. <https://doi.org/10.29244/jipthp.5.1.5-10>
- National Standardization Agency of Indonesia. (1992). *Cara uji makanan dan minuman* [Methods for analysing food and beverages], SNI 01-2891-1992. Badan Standardisasi Nasional, Jakarta.
- National Standardization Agency of Indonesia. (1992). *Cara uji bahan pengawet makanan dan bahan tambahan yang dilarang untuk makanan* [Analytical methods for food additives/preservatives], SNI 01-2894-1992. Badan Standardisasi Nasional, Jakarta.
- National Standardization Agency of Indonesia. (1998). *Cara uji minyak dan lemak* [Methods for analysing oil and fat], SNI 01-3555-1998. Badan Standardisasi Nasional, Jakarta.
- National Standardization Agency of Indonesia. (1998). *Cara uji air minum dalam kemasan* [Analytical methods for bottled drinking water], SNI 01-3554-1998. Badan Standardisasi Nasional, Jakarta.
- National Standardization Agency of Indonesia. (1998). *Spesifikasi kompos dari sampah organik domestik* [Compost specifications from domestic organic waste], SNI 19-7030-2004. Badan Standardisasi Nasional, Jakarta.
- Nurhati, I. S., & Cordova, M. R. (2020). Marine plastic debris in Indonesia: Baseline estimates (2010-2019) and monitoring strategy (2021-2025). *Marine Research in Indonesia*, 45(2), 1-6. <https://doi.org/10.14203/mri.v45i2.581>
- NusaBali.com. (2021). *Sampah Rumput Laut Menumpuk di Pesisir Timur* [Seaweed Trash Piles Up on the East Coast]. <https://www.nusabali.com/berita/96000/sampah-rumput-laut-menumpuk-di-pesisir-timur>

- Ojha, S., Bußler, S., & Schlüter, O. K. (2020). Food waste valorisation and circular economy concepts in insect production and processing. *Waste Management*, *118*, 600-609. <https://doi.org/10.1016/j.wasman.2020.09.010>
- Pangestu, W., Prasetya, A., & Cahyono, R. B. (2017). Pengolahan limbah kulit pisang dan nangka muda menggunakan larva black soldier fly (*Hermetia illucens*) [Banana peel waste processing and young jackfruit using black soldier fly larvae (*Hermetia illucens*)]. *Simposium Nasional Rapi XVI*, *2*, 97-101.
- Pendyurin, E. A., Rybina, S. Y., & Smolenskaya, L. M. (2021). Research of black soldier fly (*Hermetia illucens*) maggots zoocompost's influence on soil fertility. In *International Conference Industrial and Civil Construction* (pp. 42-49). Springer International Publishing. [https://doi.org/10.1007/978-3-030-68984-1\\_7](https://doi.org/10.1007/978-3-030-68984-1_7)
- Raksasat, R., Lim, J. W., Kiatkittipong, W., Kiatkittipong, K., Ho, Y. C., Lam, M. K., Font-Palma, C., Zaid, H. F. M., & Cheng, C. K. (2020). A review of organic waste enrichment for inducing palatability of black soldier fly larvae: Wastes to valuable resources. *Environmental Pollution*, *267*, Article 115488. <https://doi.org/https://doi.org/10.1016/j.envpol.2020.115488>
- Richard, N. (2017). Process manual for the establishment of a black soldier fly (*Hermetia illucens*) production system. *Haettu*, *20*(43), 1-8.
- Rofi, D. Y., Auvaria, S. W., Nengse, S., Oktorina, S., & Yusrianti, Y. (2021). Modification of black soldier fly (*Hermetia illucens*) larvae Feed as an effort to accelerate the reduction of fruit and vegetable waste. *Jurnal Teknologi Lingkungan*, *22*(1), 130-137. <https://doi.org/10.29122/jtl.v22i1.4297>
- Safia, W., Budiyanti, & Musrif. (2020). Kandungan nutrisi dan senyawa bioaktif rumput laut (*Euchema cottonii*) yang dibudidayakan dengan teknik rakit gantung pada kedalaman berbeda [The nutritional content and bioactive compounds of grass sea (*Euchema cottonii*) which is cultivated using the hanging raft technique different depth]. *Jurnal Pengolahan Hasil Perikanan Indonesia*, *23*(2), 261-271.
- Salgado, C. L., Muñoz, R., Blanco, A., & Lienqueo, M. E. (2021). Valorization and upgrading of the nutritional value of seaweed and seaweed waste using the marine fungi *Paradendryphiella salina* to produce mycoprotein. *Algal Research*, *53*, Article 102135. <https://doi.org/https://doi.org/10.1016/j.algal.2020.102135>
- Setyaningsih, E., Astuti, D. S., & Astuti, R. (2017). Kompos daun solusi kreatif pengendali limbah [Leaf compost is a creative solution waste control]. *Bioeksperimen: Jurnal Penelitian Biologi*, *3*(2), 45-51. <https://doi.org/10.23917/bioeksperimen.v3i2.5181>
- Sheppard, D. C., Tomberlin, J. K., Joyce, J. A., Kiser, B. C., & Sumner, S. M. (2002). Rearing methods for the black soldier fly (Diptera: Stratiomyidae). *Journal of Medical Entomology*, *39*(4), 695-698. <https://doi.org/10.1603/0022-2585-39.4.695>
- Sulistiyowati, H. I., & Prayitno, P. (2021). Pengolahan air limbah industri pengolahan rumput laut menggunakan nano adsorben tersuspensi [Processing industrial wastewater treatment seaweed using suspended nano adsorbents]. *DISTILAT: Jurnal Teknologi Separasi*, *7*(9), 514-521. <https://doi.org/10.33795/distilat.v7i2.272>
- Supriyatna, A., Manurung, R., Esyanti, R. R., & Putra, R. E. (2016). Growth of black soldier larvae fed on cassava peel wastes, An agriculture waste. *Journal of Entomology and Zoology Studies*, *4*(6), 161-165.

- Suryawan, I. W. K., Septiariva, I. Y., Sari, M. M., Ramadan, B. S., Suhardono, S., Sianipar, I. M. J., Tehupeior, A., Prayogo, W., & Lim, J. W. (2023). Acceptance of waste to energy (WtE) technology by local residents of Jakarta City, Indonesia to achieve sustainable clean and environmentally friendly energy. *Journal of Sustainable Development of Energy, Water and Environment Systems*, 11(2), 1-17.
- Swinscoe, I., Oliver, D. M., Ørnsrud, R., & Quilliam, R. S. (2020). The microbial safety of seaweed as a feed component for black soldier fly (*Hermetia illucens*) larvae. *Food Microbiology*, 91, Article 103535. <https://doi.org/https://doi.org/10.1016/j.fm.2020.103535>
- Tschirner, M., & Simon, A. (2015). Influence of different growing substrates and processing on the nutrient composition of black soldier fly larvae destined for animal feed. *Journal of Insects as Food and Feed*, 1(4), 249-259. <https://doi.org/10.3920/JIFF2014.0008>
- Wangko, S. (2014). *Hermetia illucens* aspek forensik, kesehatan, dan ekonomi [*Hermetia illucens* forensic, health and economic aspects.]. *Jurnal Biomedik: JBM*, 6(1), 23-29. <https://doi.org/10.35790/jbm.6.1.2014.4159>
- Widyastuti, R. A. D., Rahmat, A., Warganegara, H. A., Ramadhani, W. S., Prasetyo, B., & Riantini, M. (2021). Chemical content of waste composting by black soldier fly (*Hermetia illucens*). In *IOP Conference Series: Earth and Environmental Science* (Vol. 739, No. 1, p. 012003). IOP Publishing. <https://doi.org/10.1088/1755-1315/739/1/012003>



# An Attribute-based Data Privacy Classification Through the Bayesian Theorem to Raise Awareness in Public Data Sharing Activity

Nur Aziana Azwani Abdul Aziz<sup>1</sup>, Masnida Hussin<sup>1,2\*</sup> and Nur Raidah Salim<sup>2</sup>

<sup>1</sup>Department of Communication Technology and Networking, Faculty of Computer Science and Information Technology, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

## ABSTRACT

The growth of the digital era with diverse existing electronic platforms offers information sharing and leads to the realization of a culture of knowledge. Vast amounts of data and information can be reached anywhere at any time, fingertips away. These data are public because people are willing to share them on digital platforms like social media. It should be noted that not all information is supposed to be made public; some is supposed to be kept private or confidential. However, people always misunderstand and are misled about which data needs to be secured and which can be shared. We proposed an attribute-based data privacy classification model using a Naïve Bayesian classifier in this work. It aims to identify and classify metadata (attributes) commonly accessible on digital platforms. We classified the attributes that had been collected into three privacy classes. Each class represents a level of data privacy in terms of its risk of breach. The public (respondent) is determined according to different ages to gather their perspective on the unclassified attribute data. The input from the survey is then used in the Naïve Bayesian classifier to formulate data weights. Then, the sorted privacy data in the class is sent back to the respondent to get their agreement on the class of attributes. We compare our approach

with another classifier approach. The result shows fewer conflicting reactions from the respondents to our approach. This study could make the public aware of the importance of disclosing their information on open digital platforms.

## ARTICLE INFO

### Article history:

Received: 03 February 2023

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.14>

### E-mail addresses:

azianaazwani00@gmail.com (Nur Aziana Azwani Abdul Aziz)

masnida@upm.edu.my (Masnida Hussin)

nurraidah@upm.edu.my (Nur Raidah Salim)

\* Corresponding author

**Keywords:** Naïve Bayes, privacy classification, public data attribute

## INTRODUCTION

The digital platforms transformed people's lives by making daily tasks much easier. In this era, data and records have been widely used to improve storing, accessibility, and sharing, such as the collaboration of information among various entities, i.e., organizations and businesses. Moreover, the fact that the data is stored in the Cloud computing environment and can reside anywhere beyond the geographical boundary might cause the user to lose control over their data. Data communication via digital platforms makes news, rumors, feedback, comments, reports, and other information available for retrieval and response by anyone. Such information might consist of confidential data that has been implicitly shared for several purposes. Some information might not be important to users but might be vital to others. The publicly accessible information includes personal details (e.g., name, phone number, electronic mail, home, or office address), job and affiliation, and health information (e.g., medication lists, diagnostic tests, physical assessments, and history observations). The high demand for such data input collected on the digital platform has confused in determining the sensitivity levels of the data. This unnoticed data-sharing activity results in data leaks and exploitation by irresponsible parties.

An open digital platform is an online space where people can interact with it because technological advancements have drastically altered people's daily lives. Public data is mostly shared in an open digital space. There are lots of public data attributes that can be found on open digital platforms. Some examples of public data attributes are gender, age, education status, hometown, nationality, and email address. In our project, the information refers to "public data," which anyone can access from digital platforms such as social media and company websites. Noticeably, the open digital platform creates great potential for the economy and society; however, it may invade someone's privacy as public data is shared freely. The Australian Cyber Security Centre advises people to remain cautious when disclosing too much personal information online (Analysis & Policy Observatory, 2020). Someone who discloses their personal information risks becoming a victim of identity theft, stalking, and harassment. For instance, protecting personal information can help avoid phishing scams. The term "phishing" refers to a scam when fraudsters send emails or pop-up messages that seem to have been sent by a bank, a government agency, or an online retailer. The message directs the victims to a website or phone number that they can contact to update their account information or to receive a reward. It can imply that something negative will occur if the victims ignore it. It implies that someone's privacy has been violated.

Individuals' perceptions of privacy vary, which may result in a lack of knowledge. For instance, some individuals could believe that their workplace is unrelated to privacy, while others would hold the opposite opinion. In our work, we focus on developing a data privacy classification using the Bayesian theorem to identify and classify the public data attributes into different levels of privacy (i.e., low, medium, and high). Then, a data

catalog containing our data attributes and their privacy classes is created and meant to be shared with the public to increase their data privacy awareness. The remainder of this paper follows the study background, public data classification process, results, discussion, and concluding remarks.

## LITERATURE REVIEW

Public data refers to available data and information accessed through an electronic or digital platform. Organizations or individuals gather and make this information available for various purposes. According to the MyGovernment (2019), each piece of data must have a precise definition and set of characteristics. This document guideline aims to support the implementation of open data, where the public should cultivate the basic skills to evaluate information and determine its importance. It assists in justifying data and information privacy and measuring before the information is publicly shared on any digital platform. The data can be explained from various perspectives, including data acquisition, opinion target recognition, feature identification, sentiment analysis, opinion summarization, and sampling (Sanderson et al., 2019; Reza et al., 2020).

The ability of an individual to disclose just certain personal information about themselves is known as privacy. In the social media era, people became concerned about their data being visible online. However, the privacy policy on social networking sites is unclear and not well-defined. Salim et al. (2022) and Cain and Imre (2022) identified the different privacy options the social media site offers users to protect their information. These security options strengthen users' ability to divulge information while allowing the information privacy settings to those requiring it. Another researcher stated that sharing information on social media demands a lot of security and privacy settings (Rehman et al., 2022). They also stated that it is difficult to enforce data privacy on social media when people are willing to reveal their personal information publicly. Ravn et al. (2019) specified that users disclose much personal information on Instagram without realizing it is putting them in danger. From a Facebook security perspective, Rashid and Zaaba (2020) mentioned that Facebook users frequently store and exchange different types of personal data, which raises the risk of privacy breaches.

Researchers and practitioners conduct surveys to understand how users' opinions can be processed, analyzed, and used to raise public data privacy awareness. Their survey findings are classified into several key levels for classifying the users' responses in various variables. Bibhu et al. (2021) conducted a survey analysis that revealed that around 56% of people are very concerned about their privacy being publicly shared. Algarni (2019) highlighted that those four sensitivity classifications—high, medium, low, and unclassified are mostly used by researchers to leverage the data's sensitivity. They also mentioned that information that the public can view falls under the "unclassified" category. The information

on the medium level is intended for a specific group of people, and the information on the high level is the information that cannot be revealed, is extremely sensitive, and can be accessed by certain people only. Wu et al. (2021) highlighted that the privacy concept is based on 4 types, person, preference, event, and trait, for protecting privacy leakage. They employed a deep learning model and an ontology-based classification approach for grouping those privacy features.

The Bayesian network as a classifier has a transparent probability function and considers the prior information of samples, making Naive Bayes a simple and powerful classification model (Liu et al., 2013; Vu et al., 2022). The Bayes theorem is used by Abraham et al. (2019) to find highly confidential and less confidential data. They proposed a method of classifying emails according to their security levels, which are categorized as highly confidential and lowly confidential emails, using a Naïve Bayesian classifier. Vu (2022) proposed a privacy-preserving Naive Bayes classification based on secure multi-party computation. While achieving a high level of privacy, their model has higher accuracy compared to another classifier. Those works, however, do not investigate public data attributes. Meanwhile, Wibawa et al. (2019) showed the ability of the Naive Bayes classifier to classify the quality of a journal that ranks in Q1, Q2, Q3, Q4, and NQ. Their classifier approach classified the quality of journals and achieved 71.60% accuracy. Zanella-Béguelin et al. (2022) employed the Bayesian method for interpreting differential privacy to obtain a posterior for the confidence interval of the false positive and false negative rates of membership inference attacks. The result is promising but has not been analyzed from the public usability point of view. Besides utilizing the Naïve Bayes classifier, Shallal et al. (2020) proposed a *k-nearest neighbors* algorithm (*k-NN*) to divide Internet of Things (IoT) data into three security levels. However, the computation process of *k-NN* is very difficult due to the procedure of classification, which will utilize whole training samples. However, their three security levels significantly make the classification process more effective and less complex. Our work employs the Naïve Bayes classifier by classifying data privacy into three levels. The ranking of data attribute privacy is initially established from the viewpoint of the public. The Bayes theorem will then be used to quantify each data attribute. We solicit the public's opinion once more to confirm that they agree with the classified data attribute.

## **METHOD OF CLASSIFICATION PROCESS**

This discussion details how the Naive Bayes method retrieves each attribute's privacy class. Additionally, we described each stage of the privacy classification process.

### **Naïve Bayes Classifier for Privacy Classification**

A classifier is a machine learning model using certain attributes to distinguish between objects. The Bayes theorem can be used to calculate the likelihood of A occurring given

the presence of B. A is the hypothesis in this case, while B is the proof or evidence (Vu, 2022). The features are thought to be independent. That is, one feature’s existence does not change another’s behavior. As a result, it is known as naïve. The fundamental idea of Bayes’ theorem is shown in Equation 1.  $P(A|B)$  is the posterior probability, where the conditional probability distribution represents what parameters are likely after observing the data object (Liu et al., 2013).  $P(B|A)$  is the likelihood function, expressing the probability of falling under a specific category or class.  $P(A)$  is the prior probability distribution representing prior knowledge or uncertainty about a data object.

$$P(A|B) = \frac{P(B|A) \times P(A)}{P(B)} \tag{1}$$

By adding new variables that matched our public data attributes privacy setting, we modified Equation 1—given that Equation 2 modifies the Bayes theorem used in this work.

$$P(\text{weightage} | \text{privacy level}) = \frac{P(\text{privacy level} | \text{weightage}) \times P(\text{weightage})}{P(\text{privacy level})} \tag{2}$$

It is described using Equation 2, where the posterior probability is  $P(\text{weightage} | \text{privacy level})$  or  $P(W|PL)$ . It is the probability that the weight is True, given the evidence from the chosen privacy level based on the public count. The public count is the number of votes from survey respondents on the chosen privacy level.  $P(\text{privacy level} | \text{weightage})$  or  $P(PL|W)$  defines the probability of the chosen privacy level based on the public count if the weight is true for the likelihood function. In other words, it uses the sum of the two preceding values. The prior involved in this study is  $P(\text{weightage})$ , which is the probability of assumed weightage, and  $P(\text{privacy level})$ , which is the evidence and is the probability of the chosen privacy level based on the public count.

**Methodology for Privacy Classification**

This work takes several stages for the classifying process (Figure 1). The stages are inter-operable with each other to ensure the consistency of input and output between stages.

**Data Observation and Collection.** We observed the public data attributes and collected 15 from various sources based on Table 1. We believe there are 15 data attributes that the public normally shares

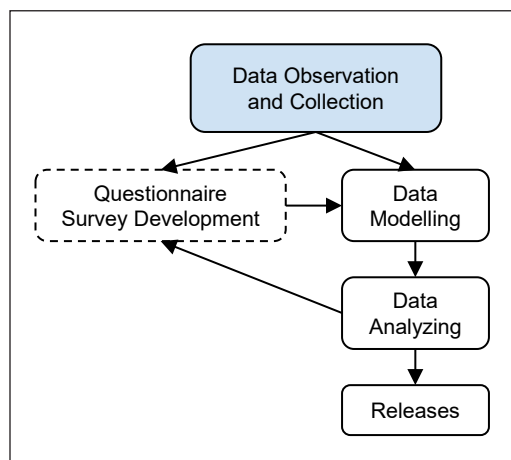


Figure 1. Privacy classification process

on an open digital platform. We then designed the respondent to represent the public. We identified 40 people as our respondents (Budiu & Moran, 2021), representing a range of ages. The focus of 40 respondents as hard-core Internet users is sufficient for obtaining comprehensive insights on data attribute privacy levels. The respondents come from a variety of social backgrounds, including friends, relatives, and individuals selected at random, and are not part of any specific group. The respondent is engaged in our study through emails based on their willingness to be part of this data collection. The age factor chosen is significant in analyzing their digital data-sharing behaviors. University students and young adults between 20 and 29 are the first age group. Another age group is between 30 and 50 since they have more experience dealing with privacy concerns in real life. Respondents over 50, who are older individuals or may be retirees and are potentially exposed to data privacy exposure, make up the last age group. As a result, there are 3 age groups: 13 respondents between the ages of 20 and 29, 14 respondents between the ages of 30 and 50, and 13 respondents above 50. Figure 2 displays the pie chart of the respondents' age group.

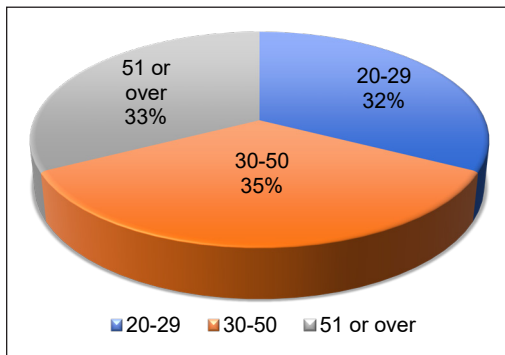


Figure 2. Respondents' age group

Table 1  
Public data attributes and their sources

Public data attributes	Sources
Address	Dokuchaev et al. (2020)
Age	Reza et al. (2020)
Date of Birth	Reza et al. (2020)
Education Level	Wu et al. (2021)
Email	Reza et al. (2020)
Gender	Reza et al. (2020)
Hometown	Reza et al. (2020)
Identity Card Number	Wu et al. (2021)
Visited Location	Dokuchaev et al. (2020)
Name	Dokuchaev et al. (2020)
Nationality	Dokuchaev et al. (2020)
Phone Number	Indeed (2021)
Profile Picture	Wu et al. (2021)
Relationship Status	Salim et al. (2022)
Workplace	Salim et al. (2022)

**Data Modeling.** We created a survey through Google Forms to determine the respondents' view on the privacy level of the selected public data attributes. Figure 3 shows the survey question that requires the respondents to select the privacy level of 15 public data attributes based on their point of view. The public data attributes are classified into three privacy classes: low, medium, and high. It aims to observe initial respondents' awareness of public data privacy. Their response will then become the foundation element in our Bayes Theorem model. The privacy classes—high, medium, or low are represented as follows:

- **High privacy class:** The attribute that significantly affects individuals' privacy and may cause security risks. For example, when data classified in this privacy class

gets into the wrong hands, it may result in people having their privacy compromised, identities stolen, or fraud committed in their names.

- **Medium privacy class:** The attribute that has little effect on the privacy of individuals and may slightly cause security risks. The real damage may lessen in this privacy class compared to the high privacy class. However, it could occur to any individual as the data in this privacy class might be a turning point for the security risks.
- **Low privacy class:** The attribute that does not significantly affect individuals' privacy and may not cause security risks. For example, if the individual decides to expose data that belongs in this privacy class, the invasion of privacy for this person may not occur. It somehow will not impact the individual's daily life as it is not really involved with real damage.

From your point of view, select the privacy level of common public data attributes below. \*  
 Dari sudut pandangan anda, pilih tahap privasi atribut data awam di bawah.

*Mark only one oval per row.*

	Low / Rendah	Medium / Sederhana	High / Tinggi
Address / Alamat	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Age / Umur	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Date of Birth / Tarikh Lahir	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Education Level / Tahap Pendidikan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Email / Emel	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Gender / Jantina	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Hometown / Tempat Asal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Identity Card (IC) Number / Nombor Kad Pengenalan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Visited Location / Lokasi yang dilawati	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Name / Nama	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Nationality / Kewarganegaraan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Phone Number / Nombor Telefon	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Profile Picture / Gambar profil	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Relationship Status / Status Hubungan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Workplace / Tempat Kerja	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Figure 3. Survey questions in the first-time verification process

The Bayesian value is calculated for each public data attribute using the designated Bayes theorem from Equation 2. It includes the probability of assuming weight and the probability of choosing a privacy level based on a public count. The public count is the number of votes from the respondents' feedback on the chosen privacy level. The weighting is designed to prioritize the privacy of public data attributes. It is determined by 15 public data attributes, divided into 3 classification classes. Hence, every rank increment of a public data attribute is assumed to be 0.33. Each attribute's weight is considered a multiple of 0.33, with the highest privacy having the most weight. The highest weight is 4.95, given to the highest privacy public data attribute, while the lowest weight is 0.33, given to the lowest privacy public data attribute. This data modeling is repeated for the second time after the respondents' responses are analyzed and formulated through the Bayesian theorem. The two-time public verification process ensures that our data privacy model is reliable and accurately reflects our users' perspectives.

**Data Analysis.** This stage analyzes the result obtained from two ranking steps: (1) the respondents' responses and (2) the Bayesian value. Public data attributes are initially ranked



based on the survey analysis according to the public count. For instance, one attribute will be classified into the “high privacy” class if most respondents select that privacy level for that attribute. Then, the respondents’ perspectives were further analyzed using the Bayesian data model. The public data attributes are further ranked and compared between the survey’s public count and the Bayesian value. The later discussion details the comparison of public data attributes’ rankings and their privacy classes.

**Releases and Feedback.** The public data attributes catalog is developed based on the latest privacy ranking generated from a two-step classifying process. This catalog aims to raise public awareness about the privacy of their information and their right to keep it private.

## RESULT AND DISCUSSION

We presented the results of the classification process according to mixed-method research, which employs quantitative research (by utilizing the Bayesian theorem) and qualitative research (by conducting surveys to get feedback from the respondents). Then, the final ranking of public data attributes will be given.

### The Ranking of the Privacy Level of each Public Data Attribute Based on the Number of Public Counts

According to the first survey output, the public data attributes classified into a high privacy class are the identity card number, phone number, visited location, and address. Table 2 shows the number of counts from the public for each privacy class of the public data attribute.

Based on Table 2, we can see that the first row of public data attributes has the highest privacy as the public count is the highest, which is 25 votes, and the subsequent row follows in descending order of public count, as the address has 20 votes, the phone number has 18 votes, and visited location has 16 votes. Most of the common public data attributes fall into the medium privacy class that is ranked accordingly, which includes workplace, education level,

Table 2  
*Public count for each attribute based on chosen privacy class*

Public data attributes	Privacy class		
	Low	Medium	High
Identity Card Number	0	15	25
Address	7	13	20
Phone Number	9	13	18
Visited Location	10	14	16
Workplace	8	16	16
Education Level	9	27	4
Age	12	20	8
Date of Birth	10	19	11
Email	11	19	10
Relationship Status	10	18	12
Profile Picture	11	16	13
Hometown	11	16	13
Name	14	14	12
Nationality	21	12	7
Gender	21	13	6

age, date of birth, email, relationship status, profile picture, hometown, and name because the workplace has 16 votes, education level has 27 votes, age has 20 votes, date of birth has 19 votes, email has 19 votes, relationship status has 18 votes, profile picture has 16 votes, hometown has 16 votes, and name has 14 votes. While nationality has 21 votes and gender has 21 votes, both are classified as having low privacy. However, public data attributes with similar public counts on certain privacy classes are now classified into medium privacy classes. It is evident when the workplace falls within the medium to high privacy class, and the name falls within the low to medium privacy class since both have the same public count.

### The Ranking of the Privacy Level of each Public Data Attribute Based on the Number of Public Counts Versus Bayes Value

The Bayes value is assessed using Equation 2 for each public data attribute. Table 3 is an example of calculation in producing Bayes value. The value is calculated and assigned to each public data attribute initially ranked based on the public count (Table 4).

We can see that the Bayes value is disorganized due to the different weights of each data attribute and that it contradicts its public count. We then sorted the ranking of public data attributes according to Bayes value and produced the new ranking list shown in Table 5. The difference only affects the high-priority class, where the address has replaced the identity card number rank, the phone number has replaced the address rank, and the identity card number has replaced the phone number rank. Meanwhile, other public data attributes do not require re-ranking or exchanging the privacy classes.

The output of classifying the public data attributes in Table 5 was then distributed again to the former respondents. Here, we assert a two-step verification process that ensures the dependability of our public data attributes catalog. This second survey required the respondents to state whether they “agree” or “disagree” with the current privacy level for each attribute, as depicted in Figure 4. Table 6 shows the result of the second survey analysis, which is the number of respondents that agree and disagree with the stated privacy class of each public data attribute.

Table 3  
Example data calculation (High privacy class)

#	Public data attributes	P (PL  W)	P (W)	P (PL)	Bayes value
1	Identity Card Number	5.575	4.95	$25/40 = 0.625$	44.154
2	Address	5.12	4.62	$20/40 = 0.5$	47.308
3	Phone Number	4.74	4.29	$18/40 = 0.45$	45.188
4	Visited Location	4.36	3.96	$16/40 = 0.4$	43.164

Table 4  
Initial ranking of public data attributes based on the public count

Privacy class	Public data attribute	Bayes value
High	Identity Card Number	44.154
	Address	47.308
	Phone Number	45.188
	Visited Location	43.164
Medium	Workplace	36.572
	Education Level	19.433
	Age	20.611
	Date of Birth	17.312
	Email	13.543
	Relationship Status	10.692
	Profile Picture	8.456
	Hometown	5.676
Low	Name	3.790
	Nationality	1.489
	Age	0.537

Table 5  
Final ranking of public data attributes after arranging Bayes value

Privacy class	Public data attribute	Bayes value
High	Address	47.308
	Phone Number	45.188
	Identity Card Number	44.154
	Visited Location	43.164
	Workplace	36.572
Medium	Age	20.611
	Education Level	19.433
	Date of Birth	17.312
	Email	13.543
	Relationship Status	10.692
	Profile Picture	8.456
	Hometown	5.676
	Name	3.790
Low	Nationality	1.489
	Gender	0.537

	Yes / Ya	No / Tidak
Address (Privacy Class : High)	<input type="radio"/>	<input type="radio"/>
Phone Number (Privacy Class : High)	<input type="radio"/>	<input type="radio"/>
Identity Card Number (Privacy Class : High)	<input type="radio"/>	<input type="radio"/>
Visited Location (Privacy Class : High)	<input type="radio"/>	<input type="radio"/>
Workplace (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Education Level (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Age (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Date of Birth (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Email (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Relationship Status (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Profile Picture (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Hometown (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Name (Privacy Class : Medium)	<input type="radio"/>	<input type="radio"/>
Nationality (Privacy Class : Low)	<input type="radio"/>	<input type="radio"/>
Gender (Privacy Class : Low)	<input type="radio"/>	<input type="radio"/>

Table 6  
Number of respondents that agree and disagree

Public data attributes	Privacy class	Number of respondents	
		Agree	Disagree
Address	High	39	1
Phone Number		37	3
Identity Card Number		38	2
Visited Location		32	8
Workplace		34	6
Education Level	Medium	33	7
Age		32	8
Date of Birth		34	6
Email		32	8
Relationship Status		30	10
Profile Picture		30	10
Hometown		33	7
Name		36	4
Nationality	Low	34	6
Gender		33	7

Figure 4. Survey questions in the second-time verification process

### Comparing with the *k*-NN Learning Machine Method

We also compare the Naïve Bayes classification with the *k*-nearest neighbors algorithm (*k*-NN) for an unsorted data set given in Table 7. In *k*-NN learning, the number of trained data will be kept as a sample to identify different classes = (high, medium, low). The training samples are taken from (Shallal et al., 2020), and the unsorted data set is given as  $UnS = \{da_1, da_2, \dots, da_n\}$  where  $(da_1, da_2, da_3, \dots, da_n)$  is independent of each other. For the *k* neighbor, we set *k* = the ratio of  $da_n$  and 15, where 15 is the total of data sampling. The comparison result in Table 8 shows slight differences in the data sorted within the group. We also request that our respondents (the same group of 40 participants) give their opinion on the classification using the *k*-NN algorithm. It is shown in Table 9 that the five data received disagreeing responses from respondents regarding the group. It might be due to the *k*-NN algorithm’s required training sample, where a few training cycles lead to misclassification. By comparing our Naïve Bayes approach, which received agreement for all data attributes, we conclude that our approach is suitable for identifying the respective privacy levels for the data attributes.

Table 7  
*Unsorted data set*

Public data attributes
Address, Age, Date of Birth, Education Level, Email, Gender, Hometown, Identity Card Number, Visited Location, Name, Nationality, Phone Number, Profile Picture, Relationship Status, Workplace

Table 8  
*Comparison group of data classification*

Privacy class	Public data attribute through Bayes value	Privacy class	Public data attribute through k-NN value
High	Address	High	Address
	Phone Number		Email
	Identity Card Number		Visited Location
	Visited Location		Age
Medium	Workplace	Medium	Gender
	Age		Name
	Education Level		Workplace
	Date of Birth		Education Level
	Email		Identity Card Number
	Relationship Status		Date of Birth
	Profile Picture		Hometown
Low	Hometown	Low	Profile Picture
	Name		Phone Number
	Nationality		Nationality
	Gender		Relationship Status

Finally, we formed a catalog of public data attributes for easier notifying the users on privacy levels of their common information, given in Figure 5.

### CONCLUSION

The development of digital data sharing and analysis procedures is important in the digital era, especially with the advancement of mobile technologies and social media platforms. The data privacy policy should be revealed to ensure public rights and the usefulness of available electronic and digital media. Aside from a thorough understanding of public data privacy, awareness of information sharing via digital media improves socio-economic conditions. It implicitly attracts many investors to make

Table 9  
Number of respondents that agree and disagree with k-NN approach

Public data attributes	Privacy class	Number of respondents	
		Agree	Disagree
Address	High	39	1
Email		10	30
Visited Location		32	8
Age	Medium	32	8
Gender		7	33
Name		36	4
Workplace		34	6
Education Level		33	7
Identity Card Number		3	37
Date of Birth		34	6
Hometown		33	7
Profile Picture		30	10
Phone Number		2	38
Nationality	Low	34	6
Relationship Status		14	26

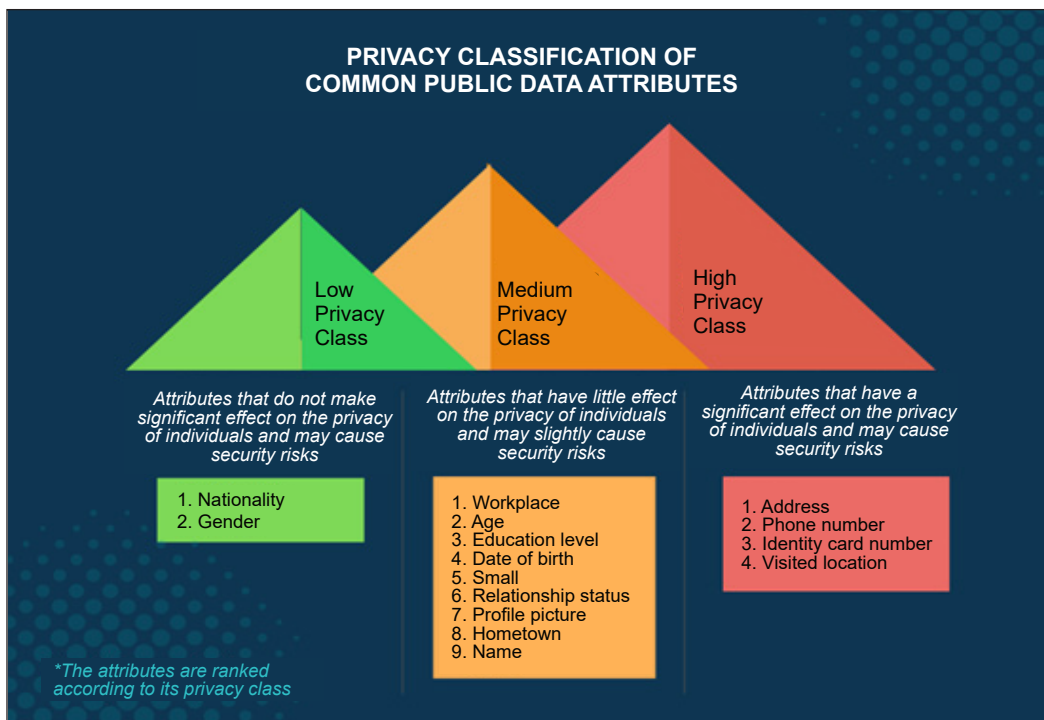


Figure 5. The catalogue

investments in the country. This work may lead one to perceive that public data also has privacy. We used the Naïve Bayesian classifier to classify public data attributes into low, medium, and high privacy classes. Mixed-method research in verifying the classification process further makes our public data attributes privacy ranking list reliable. Furthermore, this acknowledges that public data privacy has its level that should be highlighted. It may encourage individuals to be more cautious when exposing their information, especially on open digital platforms.

## ACKNOWLEDGEMENTS

This work has been supported by Cyber Security Malaysia (CSM) through a collaboration project of CSM-INSPEM, Universiti Putra Malaysia under Public Data Classification Program-Privacy and User Right parked under Tabung Amanah INSPPEM (vot.no.: 6300410-10065).

## REFERENCES

- Abraham, A., Kanjamala, E. R., Thomas, E. M., & Akhila, G. P. (2019). Email security classification of imbalanced data using naive Bayes classifier. *International Journal of Wireless Communications and Network Technologies*, 8(3), 16-20. <https://doi.org/10.30534/ijwcnt/2019/04832019>
- Algarni, A. (2019). A survey and classification of security and privacy research in smart healthcare systems. *IEEE Access*, 7, 101879-101894. <https://doi.org/10.1109/ACCESS.2019.2930962>
- Analysis & Policy Observatory. (2020). *ACSC Annual Cyber Threat Report: July 2019 to June 2020*. Australian Cyber Security Centre. <https://apo.org.au/node/308071> <https://www.cyber.gov.au/acsc/view-all-content/advice/personal-information-and-privacy>
- Bibhu, V., Salagrama, S., Lohani, B. P., & Kushwaha, P. K. (2021). An analytical survey of user privacy on social media platform. In *2021 International Conference on Technological Advancements and Innovations (ICTAI)* (pp. 173-176). IEEE Publishing. <https://doi.org/10.1109/ICTAI53825.2021.9673402>
- Budiu, R., & Moran, K. (2021). *How many participants for quantitative usability studies: A summary of sample-size recommendations*. Nielsen Normal Group. <https://www.nngroup.com/articles/summary-quant-sample-sizes/>
- Cain, J. A., & Imre, I. (2022). Everybody wants some: Collection and control of personal information, privacy concerns, and social media use. *New Media & Society*, 24(12), 2705-2724. <https://doi.org/10.1177/14614448211000327>
- Dokuchaev, V. A., Maklachkova, V. V., & Statev, V. Y. (2020). Classification of personal data security threats in information systems. *T-Comm*, 14(1), 56-60. <https://doi.org/10.36724/2072-8735-2020-14-1-56-60>
- Indeed. (2021). *A guide to data classification (with types and examples)*. Indeed. <https://www.indeed.com/career-advice/career-development/data-classification>
- Liu, S., Zhu, M., & Yang, Y. (2013). A Bayesian classifier learning algorithm based on optimization model. *Mathematical Problems in Engineering*, 2013, Article 975953. <https://doi.org/10.1155/2013/975953>

- MyGovernment. (2019). *Mygov - The government of Malaysia's Official Portal*. MyGovernment. <https://www.malaysia.gov.my/portal/content/30588>
- Rashid, A. F. A., & Zaaba, Z. F. (2020). Facebook, Twitter, and Instagram: The privacy challenges. In *2020 International Conference on Promising Electronic Technologies (ICPET)* (pp. 122-127). IEEE Publishing. <https://doi.org/10.1109/ICPET51420.2020.00032>
- Ravn, S., Barnwell, A., & Neves, B. B. (2019). What is “publicly available data”? exploring blurred public-private boundaries and ethical practices through a case study on Instagram. *Journal of Empirical Research on Human Research Ethics*, *15*(1-2), 40-45. <https://doi.org/10.1177/1556264619850736>
- Rehman, S. U., Manickam, S., & Al-Charchafchi, A. (2022). Privacy calculus model for online social networks: A study of Facebook users in a Malaysian University. *Education and Information Technologies*, *28*, 7205-7223. <https://doi.org/10.1007/s10639-022-11459-w>
- Reza, K. J., Islam, M. Z., & Estivill-Castro, V. (2020). Protection of user-defined sensitive attributes on online social networks against attribute inference attack via adversarial data mining. In *Information Systems Security and Privacy: 5th International Conference, ICISSP 2019* (pp. 230-249). Springer International Publishing. [https://doi.org/10.1007/978-3-030-49443-8\\_11](https://doi.org/10.1007/978-3-030-49443-8_11)
- Salim, S., Turnbull, B., & Moustafa, N. (2022). Data analytics of social media 3.0: Privacy protection perspectives for integrating social media and Internet of Things (SM-IoT) systems. *Ad Hoc Networks*, *128*, Article 102786. <https://doi.org/10.1016/j.adhoc.2022.102786>
- Sanderson, T., Reeson, A., & Box, P. (2019). Optimizing open government: An economic perspective on data sharing. In *Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance* (pp. 140-143). ACM Publishing. <https://doi.org/10.1145/3326365.3326383>
- Shallal, Q. M., Hussien, Z. A., & Abbood, A. A. (2020). Method to implement K-NN machine learning to classify data privacy in IOT environment. *Indonesian Journal of Electrical Engineering and Computer Science*, *20*(2), 985-990. <https://doi.org/10.11591/ijeecs.v20.i2.pp985-990>
- Vu, D. H. (2022). Privacy-preserving Naive Bayes classification in semi-fully distributed data model. *Computers & Security*, *115*, Article 102630. <https://doi.org/10.1016/j.cose.2022.102630>
- Vu, D. H., Vu, T. S., & Luong, T. D. (2022). An efficient and practical approach for privacy-preserving Naive Bayes classification. *Journal of Information Security and Applications*, *68*, Article 103215. <https://doi.org/10.1016/j.jisa.2022.103215>
- Wibawa, A. P., Kurniawan, A. C., Murti, D. M., Adiperkasa, R. P., Putra, S. M., Kurniawan, S. A., & Nugraha, Y. R. (2019). Naïve Bayes classifier for journal quartile classification. *International Journal of Recent Contributions from Engineering, Science & IT (IJES)*, *7*(2), 91-99. <https://doi.org/10.3991/ijes.v7i2.10659>
- Wu, J., Li, W., Bai, Q., Iko, T., & Moustafa, A. (2021). Privacy information classification: A hybrid approach. *ArXiv Preprint*. <https://doi.org/10.48550/arXiv.2101.11574>
- Zanella-Béguelin, S., Wutschitz, L., & Tople, S. (2022). Bayesian estimation of differential privacy. *ArXiv Preprint*. <https://doi.org/10.48550/arXiv.2206.05199>



# High-performance THz Metallic Axial Mode Helix Antenna with Optimised Truncated Hollow Cone Ground Plane for 6G Wireless Communication System

Zahraa Raad Mayoof Hajiyat<sup>1</sup>, Alyani Ismail<sup>1\*</sup>, Aduwati Sali<sup>1</sup> and Mohd. Nizar Hamidon<sup>2,3</sup>

<sup>1</sup>Wireless and Photonic Network Research Centre (WiPNET), Department of Computer and Communication Systems Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>3</sup>Institute of Nanoscience and Nanotechnology (ION2), Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

## ABSTRACT

The Terahertz (THz) band antenna configuration operates in the 0.1–10 THz frequency range and offers a stable performance for future 6th Generation (6G) wireless communication systems. However, the available metallic axial mode helix antenna designs exhibit a peak directivity of lower than 18 dBi within 0.5–1 THz, making it inappropriate to be applied in wireless communication systems. Therefore, this study proposed a high-performance THz metallic five-turn axial mode helix antenna with an optimised truncated hollow cone ground plane for 6G wireless communication systems. Following the creation of the proposed antenna design using cost-effective copper (annealed), the truncated hollow cone ground plane of the THz axial mode helix antenna was optimised via simulation in a Computer Simulation Technology Microwave Studio (CST MWS) software and a verification of the proposed THz antenna design in Analysis System High-Frequency

Structure Simulator (Ansys HFSS) software for a fair comparison. Based on the results, the proposed THz metallic axial mode helix antenna with optimised truncated hollow cone ground plane recorded an impedance bandwidth of 0.46 THz, Fractional Bandwidth (FBW) of 61.33% for  $|S_{11}| \leq -10$  dB, and a maximum directivity and realised gain of 21.8 dBi and 21.5 dBi at 0.85 THz,

## ARTICLE INFO

### Article history:

Received: 14 February 2023

Accepted: 14 June 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.15>

### E-mail addresses:

glancing@ymail.com (Zahraa Raad Mayoof Hajiyat)

alyani@upm.edu.my (Alyani Ismail)

aduwati@upm.edu.my (Aduwati Sali)

mnh@upm.edu.my (Mohd. Nizar Hamidon)

\* Corresponding author

respectively. Within the 0.5–1 THz, the proposed optimised THz antenna design achieved an outstanding performance, including an FBW of more than 50%, excellent directivity of higher than 15.8 dBi, radiation efficiency of greater than 87%, circular polarisation, and low-profile helix turns. In short, the proposed THz metallic axial mode helix antenna with optimised truncated hollow cone ground plane design is appropriate for various THz 6G wireless applications.

*Keywords:* 6G wireless communication system, axial mode, helix antenna, high-performance, metallic, terahertz

---

## INTRODUCTION

The Terahertz (THz) band antenna configuration has received overwhelming interest over its compatible application in 6<sup>th</sup> Generation (6G) wireless communication systems. The 0.1–10 THz range offers numerous advantages, such as non-ionic high-frequency range, excellent resolution, broad bandwidth (Singhal, 2019), more robust information security, and stable communication, making it a more favourable alternative for future wireless (6G) communication (P. K. Singh et al., 2021). Nevertheless, one of the main drawbacks of the THz frequency band is the high attenuation path loss and molecular absorption loss (P. K. Singh et al., 2021; Ullah et al., 2019; Xia & Jornet, 2019). In particular, frequencies over 1 THz are inappropriate for wireless communication due to the absorption wavelengths of humidity and other gases in the THz band (P. K. Singh et al., 2021). Thus, only the low THz frequency range (0.1–1 THz) is suitable for future wireless communication, given its fewer signal losses due to the effect of small fog and dust particles (P. K. Singh et al., 2021; Ullah et al., 2019). Additionally, several low attenuation windows appear below 1 THz, specifically around 0.3, 0.35, 0.41, 0.65, and 0.85 THz (P. K. Singh et al., 2021),  $w_1=0.38\text{--}0.44$  THz,  $w_2=0.45\text{--}0.52$  THz,  $w_3=0.62\text{--}0.72$  THz, and  $w_4=0.77\text{--}0.92$  THz (M. Singh et al., 2021). In view of this, broad bandwidth THz band antennas could be utilised for various THz band applications in future 6G wireless communication, such as medical diagnosis (0.42, 0.50, and 0.80 THz), homeland defence (0.6–0.8 THz), video-rate imaging (0.6 THz), explosive and weapon detection (0.41 THz), Terabits/sec links (0.2375, 0.350, and 0.840 THz), space communication, and energy conservation devices (Ghalamakri & Mokhtari, 2022).

The propagation path loss can be compensated by employing two key components, namely, the high-gain and directive THz band antenna (Aqlan et al., 2020; Fan et al., 2016; Faridani & Khatir, 2018; Kürner, 2018; Pillai et al., 2022; Wu et al., 2019; Xia & Jornet, 2019). However, using metallic materials (Hajiyat et al., 2021b) to design Ultra-Wideband (UWB) THz band antennas with high directional gain (Chen et al., 2020) is extremely difficult due to the impact of antenna size on gain and bandwidth (Harrington, 1960; Huang & Boyle, 2008). Besides, the decreased skin depth and conductivity of the copper metal at

the THz band minimises the radiation efficiency of the antenna elements in the THz band (Dash & Patnaik, 2018; Jamshed et al., 2020).

Antennas on-chip based on meta-materials and meta-surfaces are interesting techniques to achieve high-performance THz antennas without altering their physical size (Alibakhshikenari et al., 2022). Various studies have demonstrated the advantages of these techniques in terms of smaller size, low profile, wider bandwidth, and better radiation properties (Alibakhshikenari et al., 2022; Alibakhshikenari et al., 2021). However, these studies only reported fractional bandwidth of lower than 50%.

Interestingly, Chetioui et al. (2022) and most research by Acharya et al. (2015), Boudkhal et al. (2018), Guo et al. (2014) and Guo et al. (2016) operated THz band helix antenna above 1 THz frequency. Although they provide wideband or UWB bandwidth for THz 6G communication, such approaches are unsuitable for wireless communication systems due to their frequency of over 1 THz. On the contrary, very few studies on THz band axial mode helix antennas have reported such applications below 1 THz (Boudkhal et al., 2019; Hajiyat et al., 2021a; Hajiyat et al., 2021c). Moreover, the study by Hajiyat et al. (2021c) achieved a peak directivity of 13.4 dBi. Thus, the result implies that further investigation is required to design a THz axial mode helix antenna with greater peak directivity than 18 dBi to compensate for the propagation path loss in the THz range, ultimately achieving effective UWB 6G wireless communication systems.

To the best of the author's knowledge, a high-performance metallic axial mode helix antenna with a peak directivity of higher than 18 dBi has not been reported in 6G wireless communication systems at a frequency range of 0.5–1 THz. Therefore, this study proposed a high-performance THz metallic axial mode helix antenna with an optimised truncated hollow cone ground plane for the UWB 6G wireless communication system. The truncated hollow cone ground plane parameters of the proposed metallic axial mode helix antenna were optimised. Subsequently, the simulation results were computed using the Computer Simulation Technology Microwave Studio (CST MWS) software. The performance of the proposed antenna design was then verified via an Analysis System High-Frequency Structure Simulator (Ansys HFSS). Finally, both findings were compared and revealed a strong association. The computed results notably outperform similar research in terms of directivity for frequencies between 0.5 and 1 THz.

## METHODS

### Design of the Proposed Axial Mode Helix Antenna with Optimised Truncated Hollow Cone Ground Plane

This study investigated the performance of an optimised three-dimensional (3D) copper ground structure added to the THz axial mode helix antenna. Figure 1(a) shows the proposed optimised THz antenna design schematic diagram. Copper (annealed) was applied in the

design due to its widespread use in antenna manufacturing and past literature studies on copper antennas at THz frequencies. The ground plane of the proposed THz metallic axial mode helix antenna was reshaped, designed, and optimised to reach exceptionally high directional gain performance. Notably, this study introduced a truncated hollow cone ground plane (3D-shaped), which differs from that of the circular ground plane (2D-shaped) in Hajiyat et al. (2021c). Regardless, the same number of helix turns was used in this study to maintain a low profile of helix turns.

The proposed optimised THz antenna design was fed with a 50-Ω copper (annealed) coaxial cable (vacuum as an insulator), where the inner conductor diameter of the coaxial feed is equivalent to the diameter of the helix wire to simplify the antenna geometry for enhanced and reproducible antenna design at THz frequency. The values of the optimised truncated hollow cone ground plane parameters for the proposed axial mode helix antenna are shown in Table 1.

Figure 1(b) presents the proposed five-turn axial mode helix antenna geometry with an optimised truncated hollow cone ground plane fed with the 50-Ω coaxial waveguide port in the Computer Simulation Technology Microwave Studio (CST MWS) software based on the Finite Integration Technique (FIT) (SIMULIA, 2019). A waveguide port was used in the feeding line due to extremely low reflection levels at high frequencies. For analysis purposes, a time-domain solver with a hexahedral mesh type of -50 dB accuracy and open (add space) boundary conditions for all directions were utilised for the simulation setting.

Experimental measurements of THz band antennas are hampered by unresolved issues in THz band antenna manufacturing technology, equipment availability, and expensive

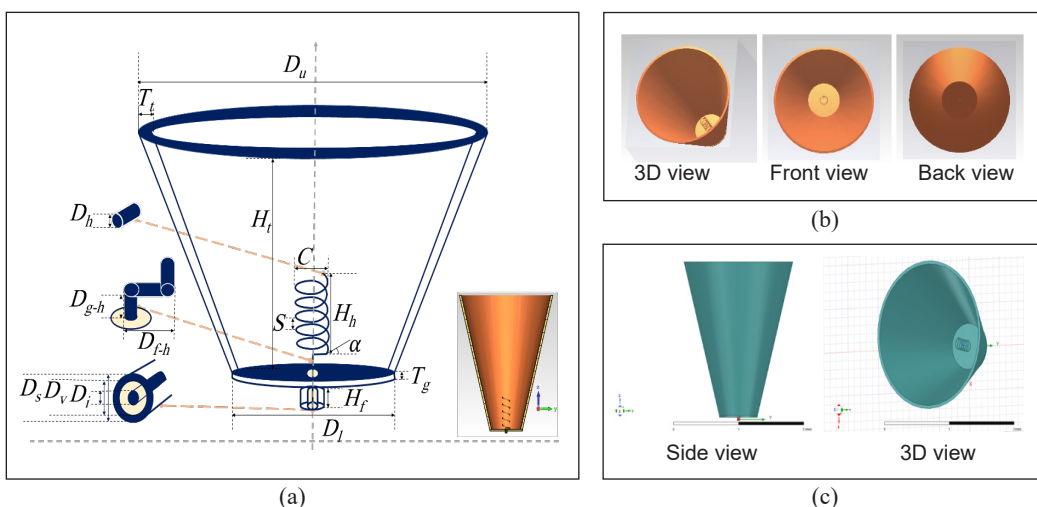


Figure 1. The proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane: (a) Schematic diagram; (b) 3D view, front view, and back view in the CST MWS software with air surrounding, respectively; and (c) Side view and 3D view in the Ansys HFSS Software plane  
 Note. Design drawn not to scale

facilities (Akyildiz et al., 2020; Hajiyat et al., 2021b). Therefore, further software comparisons were used in this work to validate the proposed optimised THz antenna design. The proposed antenna design and optimisation were performed using the CST MWS software, while Analysis System High-Frequency Structure Simulator (Ansys HFSS) software version 2021 R1, which is based on the Finite Element Method (FEM) (Ansys Inc, 2021), was used for antenna design validation and comparison. Figure 1(c) shows the proposed optimised THz antenna design in the Ansys HFSS software.

Table 1

*Design parameters for the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane*

Parameter	Value (mm)
<b>Helix antenna with truncated hollow cone ground plane</b>	
Circumference of a turn ( $C$ )	0.40
Spacing between turns ( $S$ )	0.10
Height of the helix antenna ( $H_h$ )	0.50
Diameter of the helix wire ( $D_h$ )	0.0150
Distance between the feed wire and the helix antenna ( $D_{f-h}$ )	0.0790
Distance between the ground plane and the helix antenna ( $D_{g-h}$ )	0.0085
Upper diameter of the truncated hollow cone ( $D_u$ )	1.70
Ground lower diameter of the truncated hollow cone ( $D_l$ )	0.60
Height of the truncated hollow cone ( $H_t$ )	1.950
Ground thickness of the truncated hollow cone ( $T_g$ )	0.020
Thickness of the truncated hollow cone (shell thickness of loft) ( $T_l$ )	0.035
<b>Coaxial waveguide port</b>	
Inner conductor diameter of the coaxial feed ( $D_i$ )	0.015
Vacuum diameter of the coaxial feed (insulator) ( $D_v$ )	0.034
Outer conductor diameter of the coaxial feed (outer sheath) ( $D_s$ )	0.036
Height of the coaxial feed ( $H_f$ )	0.030
Parameter	Value
Helix turn number ( $N$ )	5
Helix angle ( $\alpha$ )	14.03°

### Optimisation of Truncated Hollow Cone Ground Plane Parameters

The geometrical dimensions of the truncated hollow cone ground plane were determined to maximise the antenna directivity while maintaining a high Fractional Bandwidth (FBW) in the selected frequency range. The upper diameter of the truncated hollow cone, the ground lower diameter of the truncated hollow cone, the height of the truncated hollow cone, and the thickness of the truncated hollow cone were assessed as independent variables. Other helix antenna dimensions were kept constant as control variables. Table 2 shows the data input in the CST MWS software.

**Factors Affecting the Performance of the Metallic Axial Mode Helix Antenna**

The following sections describe the effects of the truncated hollow cone ground plane parameters (height, upper diameter, ground lower diameter, and thickness) on the directivity and FBW performance of the proposed optimised THz antenna design.

**Effect of Varying Height**

The effect of the height of the truncated hollow cone ground plane ( $H_t$ ) of the proposed optimised THz antenna design was assessed at fixed  $D_u$ ,  $D_l$ , and  $T_t$  values. The  $H_t$  was varied to determine the height

effect on the directivity and FBW performance of the metallic axial mode helix antenna in the THz band. Figures 2 (a) and 2(b) show that the increased  $H_t$  led to an increase in the directivity performance of the axial mode helix antenna but decreased the FBW performance. Additionally, the directivity performance of the proposed optimised THz antenna design has achieved over 18 dBi when the  $H_t$  value was more than 1.95 mm. In short, the analysis suggests that the directivity and FBW performance of the proposed optimised THz antenna design are positively and negatively correlated with the  $H_t$  respectively.

**Effect of Varying Upper Diameter**

The effect of the upper diameter of the truncated hollow cone ground plane ( $D_u$ ) of the proposed optimised THz antenna design was analysed at fixed  $H_t$ ,  $D_l$ , and  $T_t$  values. As shown in Figures 3(a) and 3(b), the directivity performance of the metallic axial mode helix antenna enhanced as the  $D_u$  increased. The FBW performance also increased until the  $D_u$  value reached 0.8 mm, where the FBW performance decreased and maintained within a  $D_u$  range of 1.3–1.8 mm. Based on the analysis, the directivity performance of the metallic axial mode helix antenna is more positively correlated to the  $D_u$  compared to the FBW performance, with higher FBW recorded at specific  $D_u$  values.

**Effect of Varying Ground Lower Diameter**

The effect of the ground lower diameter of the truncated hollow cone ground plane ( $D_l$ ) of the proposed optimised THz antenna design was determined at fixed  $H_t$ ,  $D_u$ , and  $T_t$  values. According to Figures 4(a) and 4(b), increasing the  $D_l$  decreased the directivity performance

Table 2  
*Investigated parameter range of truncated hollow cone ground plane optimisation*

Variable	Parameter	Investigation range
Independent	$H_t$ (unit)	0.218–2.5 (mm)
	$D_u$ (unit)	0.5–1.8 (mm)
	$D_l$ (unit)	0.5–0.9 (mm)
	$T_t$ (unit)	0.01–0.07 (mm)
Control	C (unit)	0.40 (mm)
	S (unit)	0.10 (mm)
	$H_h$ (unit)	0.50 (mm)
	$D_h$ (unit)	0.0150 (mm)
	$D_{r-h}$ (unit)	0.0790 (mm)
	$D_{g-h}$ (unit)	0.0085 (mm)
	N	5
$\alpha$	14.03°	

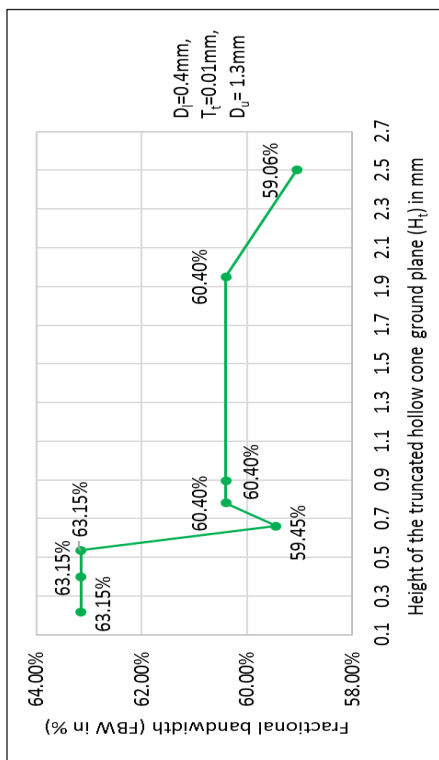
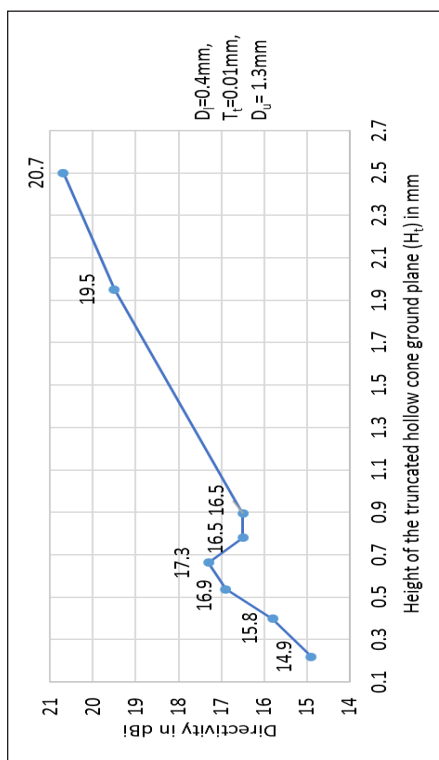


Figure 2. The effect of varying  $H_t$  values on the performance of the metallic axial mode helix antenna in terms of: (a) directivity; and (b) FBW

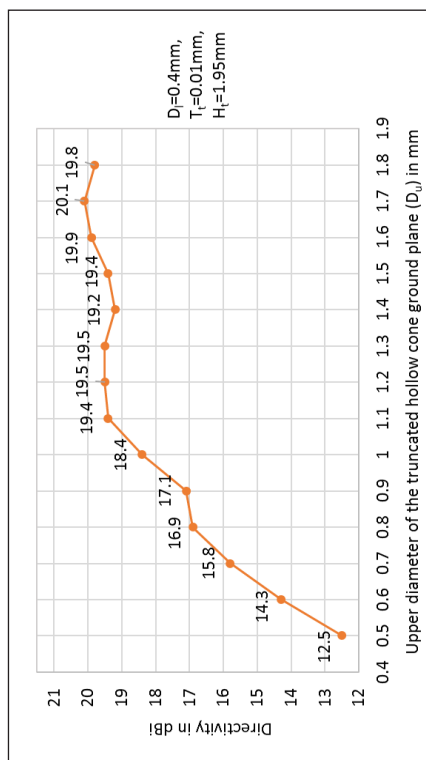
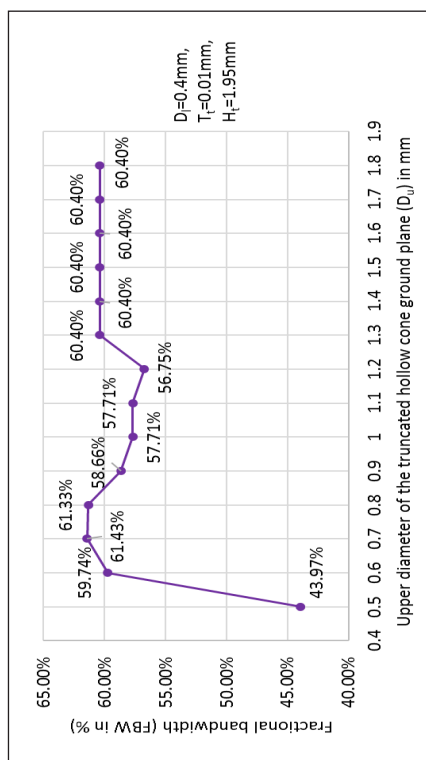


Figure 3. The effect of varying  $D_u$  values on the performance of the metallic axial mode helix antenna in terms of: (a) directivity; and (b) FBW



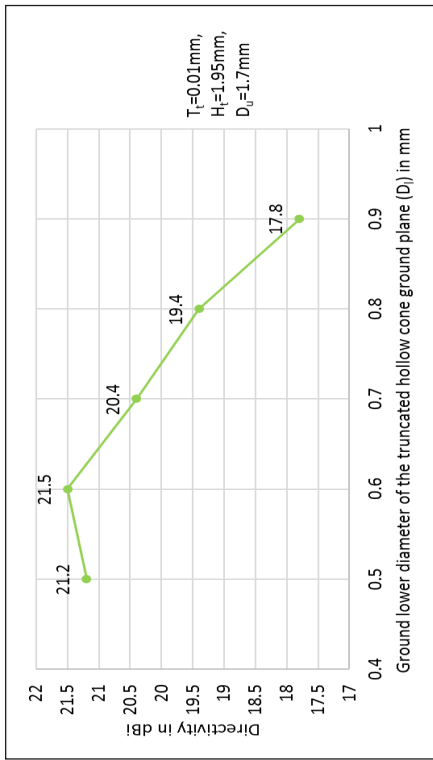
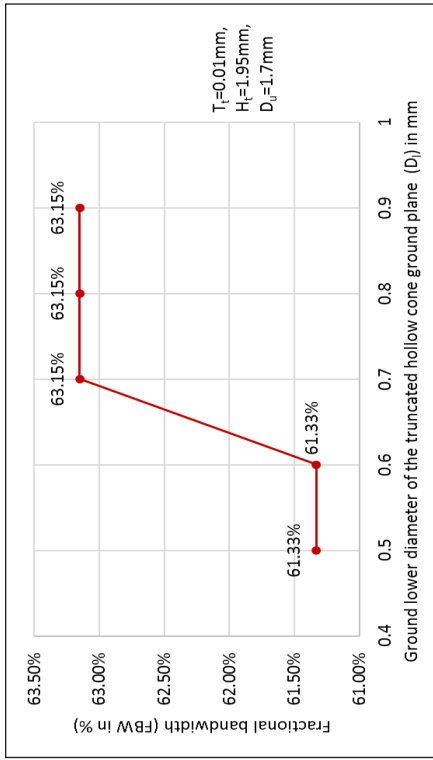


Figure 4. The effect of varying  $D_1$  values on the performance of the metallic axial mode helix antenna in terms of the (a) directivity and (b) FBW

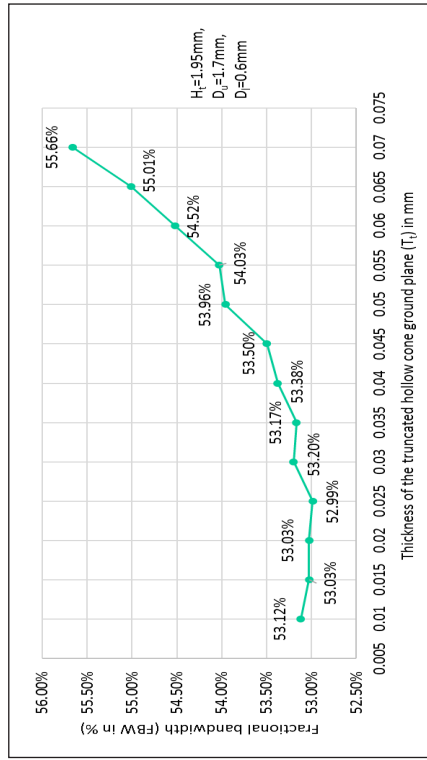
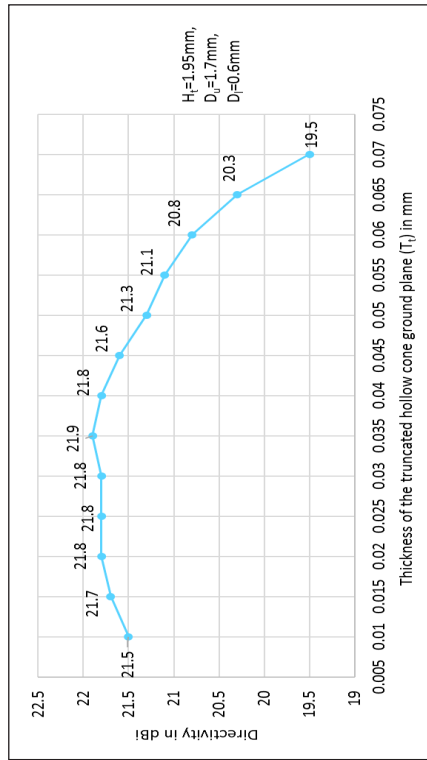


Figure 5. The effect of varying  $T_1$  values on the performance of the metallic axial mode helix antenna in terms of: (a) directivity; and (b) FBW

of the proposed optimised THz antenna design. However, the FBW performance was enhanced as the  $D_1$  increased and stabilised at a  $D_1$  range of 0.7–0.9 mm.

### Effect of Varying Thickness

The effect of the loft ( $T_1$ ) shell thickness of the truncated hollow cone ground plane of the proposed optimised THz antenna design was evaluated at fixed  $H_t$ ,  $D_u$ , and  $D_1$  values. Figures 5(a) and 5(b) show that the increase in  $T_1$  reduced the directivity performance but increased the FBW performance of the metallic axial mode helix antenna. Thus, thickening the ground plane of the truncated hollow cone significantly improved the FBW performance of the proposed optimised THz antenna design compared to the directivity performance.

## RESULTS AND DISCUSSION

### Design Results of the CST MWS Software

Figures 6(a) and 6(b) depict the performance of the proposed optimised THz antenna design using the CST MWS software in terms of the  $S_{11}$  parameter and total efficiency and radiation efficiency, respectively. Based on the results, the UWB bandwidth ( $|S_{11}| \leq -10$  dB) operates between 0.52 to 0.98 THz with an FBW of 61.33%, impedance bandwidth (BW) of 0.46 THz, and a deeper return loss of -29.73 dB at 0.709 THz. In addition, the optimised metallic axial mode helix antenna exhibits two resonant frequencies at 0.710 THz and 0.924 THz, with a total efficiency of 0.53–0.95 and radiation efficiency of 0.87–0.95 within a frequency range of 0.5–1 THz.

Furthermore, the directivity and realised gain of the proposed optimised THz antenna design at 0.5–1 THz were 15.8–21.8 dBi and 15.1–21.5 dBi, respectively. Figure 7 shows that the proposed optimised THz antenna design recorded a maximum directivity of 21.8 dBi with a realised gain of 21.5 dBi at 0.85 THz.

### Validation Results of the Ansys HFSS Software

Figures 8(a) and 8(b) illustrate the performance of the proposed optimised THz antenna design using the Ansys HFSS Software in terms of the  $S_{11}$  parameter and radiation efficiency, respectively. The UWB bandwidth was attained within the 0.52–0.98 THz range, which also recorded an FBW ( $|S_{11}| \leq -10$  dB) of 61.33%, impedance BW of 0.46 THz, and a minimum return loss of -31.45 dB at 0.71 THz. The optimised proposed THz antenna design also recorded three resonant frequencies at 0.57 THz, 0.71 THz, and 0.91 THz, with a radiation efficiency of 0.87–0.99 at a frequency range of 0.5–1 THz.

Moreover, the proposed optimised THz antenna design recorded a directivity of 15.87–21.8 dBi and a realised gain of 14.1–21.6 dBi within the 0.5–1 THz frequency range. Figure 9 shows a maximum directivity of 21.8 dBi with a realised gain of 21.6 dBi at 0.85 THz by the proposed optimised THz antenna design.

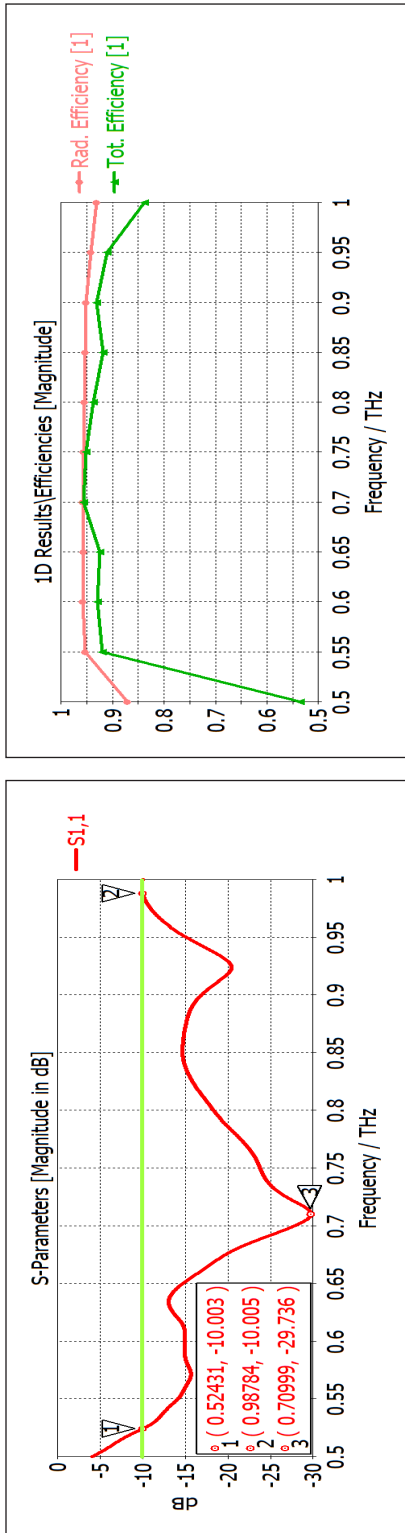


Figure 6. The CST MWS performance analysis of the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane in terms of: (a)  $S_{11}$  parameter; and (b) total efficiency and radiation efficiency

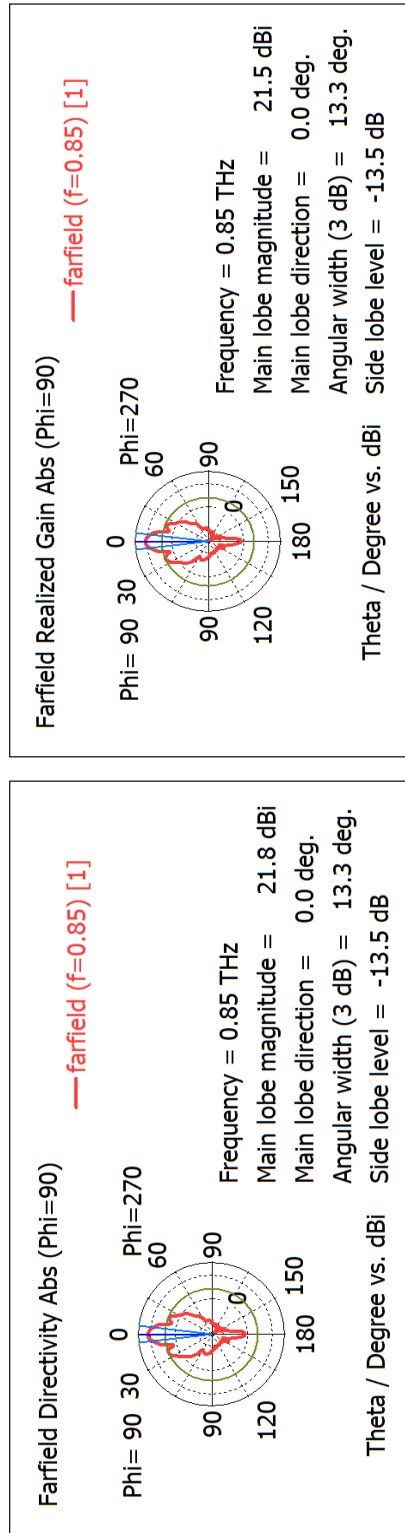


Figure 7. The CST MWS farfield polar view of: (a) directivity; and (b) realised gain of the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane at 0.85 THz

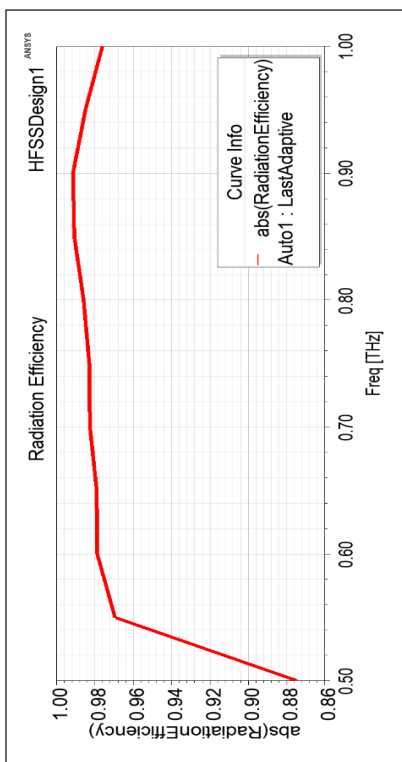


Figure 8. The Ansys HFSS performance analysis of the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane in terms of: (a)  $S_{11}$  parameter; and (b) radiation efficiency

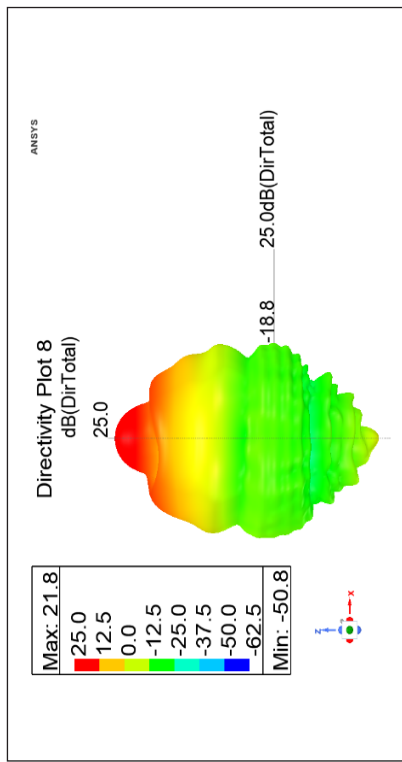
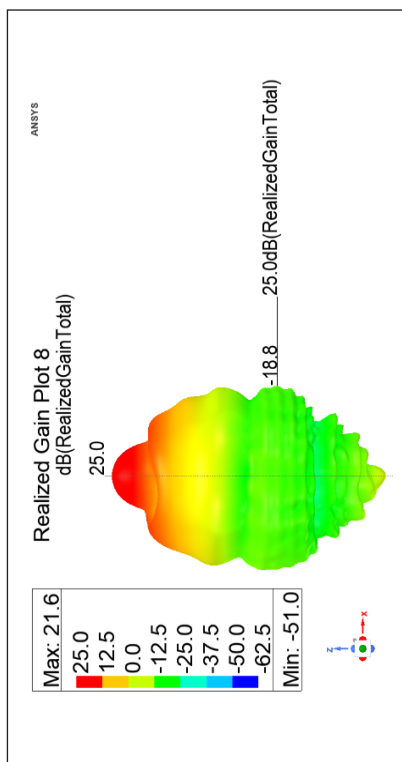
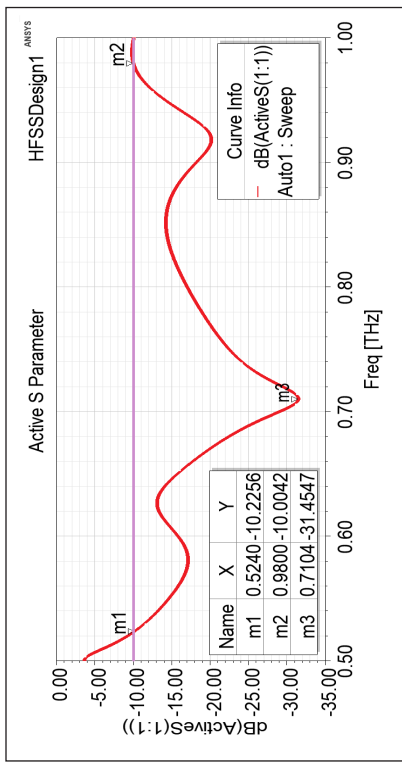


Figure 9. The Ansys HFSS farfield 3D polar view of: (a) directivity; and (b) realised gain of the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane at 0.85 THz

### Comparison Between CST MWS and Ansys HFSS Results

Figure 10 compares the directivity findings from the CST MWS and the  $S_{11}$  parameter (BW) from the Ansys HFSS software of the proposed optimised THz antenna design. With respect to the CST MWS, the average relative difference<sup>1</sup> of the  $S_{11}$  parameter and directivity between CST MWS and Ansys HFSS within the 0.5–1 THz range is 5.5% and 0.65%, respectively, signifying a good agreement between the two simulation results. The two simulators’ different numerical approaches and mesh structure contributed to the varying outcomes. These variations also match those reported in past literature. Remarkably, both simulation results recorded an FBW of 61.33% within the 0.52–0.98 THz range and a peak directivity of 21.8 dBi at 0.85 THz, which is effective for UWB 6G communication systems.

<sup>1</sup> Relative difference (%) = (Ansys HFSS value - CST MWS value / CST MWS value) × 100%

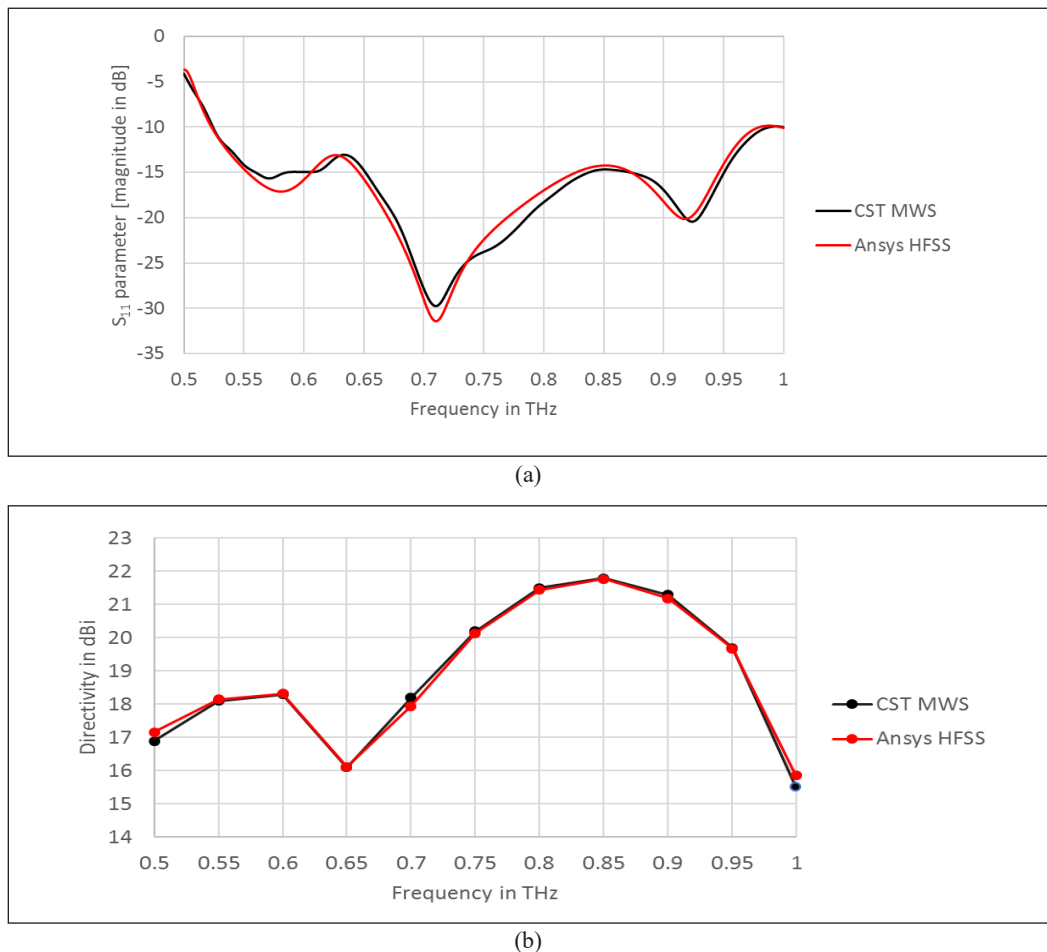


Figure 10. Comparison between the Ansys HFSS and CST MWS results of the proposed metallic axial mode helix antenna with optimised truncated hollow cone ground plane in terms of: (a)  $S_{11}$  parameter; and (b) directivity

## Performance Comparison with Past Literature Studies

Table 3 displays the performance comparison of the proposed optimised THz antenna design in this study to that of currently available THz helix antennas operating at similar frequency ranges in terms of the FBW, directivity, radiation efficiency, physical dimensions, applied antenna design method, and antenna design complexity. It is worth mentioning that there are only a few axial mode helix antenna studies in the 0.5–1 THz frequency band. Comparatively, most existing THz axial mode helix antennas achieved a lower directivity with over 50% FBW (or less) compared to the proposed optimised THz antenna design in this study.

Additionally, the physical antenna dimensions applied in past studies are smaller than that of this study's proposed optimised THz antenna design due to their simple ground plane structure. In contrast, the proposed optimised THz antenna design recorded lower profile helix turns without the ground plane compared to the smallest THz helix antenna dimension (Boudkhil et al., 2019). Therefore, the results suggest that the optimised truncated hollow cone ground plane structure (3D copper) enhanced the helix antenna directivity with a slightly lower FBW in the THz band and large antenna size. Overall, the comparative analysis verified the excellent performance of the newly proposed metallic axial mode helix antenna in the 0.5–1 THz frequency range for potential application in 6G wireless communication systems.

Table 3

*Performance comparison of the proposed THz axial mode helix antenna design with other THz axial mode helix antennas*

References	Boudkhil et al., 2019	Hajiyat et al., 2021a	Hajiyat et al., 2021c	This work
Frequency range (THz)	0.8 - 1.05	0.52 - 1.03	0.52 - 1.057	0.52 - 0.98
BW % (FBW %)	25	51 (65.40)	53.70 (68.10)	46 (61.33)
Peak directivity, realised gain in dBi (frequency)	-	12.1, 11.8 (1 THz)	13.4, 13 (0.95 THz)	21.8, 21.5 (0.85 THz)
Radiation efficiency (%)	-	90 - 96.92	89.69 - 96.6	87.09 - 95.77
Physical dimensions in mm (width, length, height)	$0.07981 \times 0.080 \times 0.18173$	$0.40 \times 0.40 \times 0.329$	$0.60 \times 0.60 \times 0.52$	$1.7 \times 1.7 \times 1.97$ ( $D_u, D_s, H_t + T_g$ )
Applied antenna design method	10.74 turns of tapered helix antenna with coplanar waveguide	Three-turns of uniform helix antenna with circular ground plane	Five-turns of uniform helix antenna with large circular ground plane	Five-turns of uniform helix antenna with truncated hollow cone ground plane
Antenna design complexity	Higher number of helix turns with a simple ground plane	Low number of helix turns with a simple ground plane	High number of helix turns with a simple ground plane	High number of helix turns with a complex ground plane

## CONCLUSION

A high-performance THz metallic axial mode helix antenna with an optimised truncated hollow cone ground plane for 6G wireless communication systems was proposed in this paper. The simulation results show that the optimised copper (annealed) axial mode helix antenna performed well in the 0.52–0.98 THz frequency band with an impedance BW of 0.46 THz and FBW of 61.33%. Additionally, the highest directivity and realised gain recorded were 21.8 dBi and 21.5 dBi at 0.85 THz, respectively. The comparative analysis between the CST MWS and Ansys HFSS showed good agreement, further validating the proposed antenna design. Moreover, the performance comparison of this study shows that the proposed optimised THz antenna design offered an outstanding directivity performance compared to other available THz axial mode helix antennas. In conclusion, the proposed axial mode helix antenna design offers a promising approach for THz 6G wireless applications, including directional UWB communication links, data centre networks, satellite communication, environmental pollution monitoring, augmented reality, and entertainment technology.

## ACKNOWLEDGEMENT

The authors express their gratitude to the Wireless and Photonics Network Research Centre (WiPNET), Faculty of Engineering and Institute of Nanoscience and Nanotechnology (ION2), Universiti Putra Malaysia (UPM) for their support and facilities used in the research.

## REFERENCES

- Acharya, I., Chauhan, A. S., & Sengupta, S. (2015). Gain enhancement of MEMS helix antenna using double substrate and fractal structures. In *2015 2nd International Conference on Electronics and Communication Systems (ICECS)* (pp. 757-761). IEEE Publishing. <https://doi.org/10.1109/ECS.2015.7125013>
- Akyildiz, I. F., Kak, A., & Nie, S. (2020). 6G and beyond: The future of wireless communications systems. *IEEE Access*, *8*, 133995-134030.
- Alibakhshikenari, M., Ali, E. M., Soruri, M., Dalarsson, M., Naser-Moghadasi, M., Virdee, B. S., Stefanovic, C., Pietrenko-dabrowska, A., Koziel, S., Szczepanski, S., & Limiti, E. (2022). A comprehensive survey on antennas on-chip based on metamaterial, metasurface, and substrate integrated waveguide principles for millimeter-waves and terahertz integrated circuits and systems. *IEEE Access*, *10*, 3668-3692. <https://doi.org/10.1109/ACCESS.2021.3140156>
- Alibakhshikenari, M., Virdee, B. S., Salekzamankhani, S., Aïssa, S., See, C. H., Soin, N., Fishlock, S. J., Althwayb, A. A., Abd-Alhameed, R., Huynen, I. McLaughlin, J. A., Falcone, F., & Limiti, E. (2021). High-isolation antenna array using SIW and realized with a graphene layer for sub-terahertz wireless applications. *Scientific Reports*, *11*, Article 10218. <https://doi.org/10.1038/s41598-021-87712-y>
- Ansys Inc. (2021). *Ansys High Frequency Structure Simulator (HFSS) Software Version 21*. Ansys Inc. <https://www.ansys.com/products/electronics/ansys-hfss>



- Aqlan, B., Himdi, M., Le Coq, L., & Vettikalladi, H. (2020). Sub-THz circularly polarized horn antenna using wire electrical discharge machining for 6G wireless communications. *IEEE Access*, 8, 117245-117252. <https://doi.org/10.1109/ACCESS.2020.3003853>
- Boudkhil, A., Chetioui, M., Benabdallah, N., & Benahmed, N. (2018). Development and performance enhancement of MEMS helix antenna for THz applications using 3D HFSS-based efficient electromagnetic optimization. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 16(2), 210-216. <http://doi.org/10.12928/telkomnika.v16i1.8000>
- Boudkhil, A., Chetioui, M., Damou, M., & Benahmed, N. (2019). A new design of a THz  $\Lambda$ -helical antenna based on MEMS technology using SNLPs. In *2019 6th International Conference on Image and Signal Processing and their Applications (ISPA)* (pp. 1-5). IEEE Publishing. <http://doi.org/10.1109/ISPA48434.2019.8966892>
- Chen, S., Liang, Y. C., Sun, S., Kang, S., Cheng, W., & Peng, M. (2020). Vision, requirements, and technology trend of 6G: How to tackle the challenges of system coverage, capacity, user data-rate and movement speed. *IEEE wireless Communications*, 27(2), 218-228. <http://doi.org/10.1109/MWC.001.1900333>
- Chetioui, M., Boudkhil, A., Damou, M., Bouallem, M., Bouasria, F., & Benahmed, N. (2022). A MEMS cone-shaped helix antenna for THz applications using ANN. In *2022 7th International Conference on Image and Signal Processing and their Applications (ISPA)* (pp. 1-5). IEEE Publishing. <http://doi.org/10.1109/ISPA54004.2022.9786345>
- Dash, S., & Patnaik, A. (2018). Material selection for THz antennas. *Microwave and Optical Technology Letters*, 60(5), 1183-1187. <https://doi.org/10.1002/mop.31127>
- Fan, K., Hao, Z. C., & Hong, W. (2016). A 325–500 GHz high gain antenna for terahertz applications. In *2016 International Symposium on Antennas and Propagation (ISAP)* (pp. 780-781). IEEE Publishing.
- Faridani, M., & Khatir, M. (2018). Wideband hemispherical dielectric lens antenna with stabile radiation pattern for advanced wideband terahertz communications. *Optik*, 168, 355-359. <https://doi.org/10.1016/j.ijleo.2018.04.028>
- Ghalamakri, B., & Mokhtari, N. (2022). Wide-band octagon-star fractal microstrip patch antenna for terahertz applications. *Optik*, 259, Article 168990. <https://doi.org/10.1016/j.ijleo.2022.168990>
- Guo, L., Huang, F., & Tang, X. (2014). A novel integrated MEMS helix antenna for terahertz applications. *Optik*, 125(1), 101-103. <https://doi.org/10.1016/j.ijleo.2013.06.016>
- Guo, L., Meng, H., Zhang, L., & Ge, J. (2016). Design of MEMS on-chip helical antenna for THz application. In *2016 IEEE MTT-S International Microwave Workshop Series on Advanced Materials and Processes for RF and THz Applications (IMWS-AMP)* (pp. 1-4). IEEE Publishing. <https://doi.org/10.1109/IMWS-AMP.2016.7588385>
- Hajiyat, Z. R., Ismail, A., Sali, A., & Hamidon, M. N. (2021a). Design and analysis of helical antenna for short-range ultra-high-speed THz wireless applications. *Optik*, 243, Article 167232. <https://doi.org/10.1016/j.ijleo.2021.167232>
- Hajiyat, Z. R. M., Ismail, A., Sali, A., & Hamidon, M. N. (2021b). Antenna in 6G wireless communication system: Specifications, challenges, and research directions. *Optik*, 231, Article 166415. <https://doi.org/10.1016/j.ijleo.2021.166415>

- Hajiyat, Z. R., Ismail, A., Sali, A., & Hamidon, M. N. (2021c). Three dimensional all-metal high aspect ratio directive helix antenna for UWB THz 6G communications. (2021). In *2021 IEEE Asia-Pacific Conference on Applied Electromagnetics (APACE)* (pp. 1-4). IEEE Publishing. <https://doi.org/10.1109/APACE53143.2021.9760540>
- Harrington, R. F. (1960). Effect of antenna size on gain, bandwidth, and efficiency. *Journal of Research of the National Bureau of Standards*, *64*(1), 1-12.
- Huang, Y., & Boyle, K. (2008). *Antennas: From Theory to Practice*. John Wiley & Sons.
- Jamshed, M. A., Nauman, A., Abbasi, M. A. B., & Kim, S. W. (2020). Antenna selection and designing for THz applications: Suitability and performance evaluation: A survey. *IEEE Access*, *8*, 113246-113261. <https://doi.org/10.1109/ACCESS.2020.3002989>
- Kürner, T. (2018). THz communications: Challenges and applications beyond 100 Gbit/s. In *2018 International Topical Meeting on Microwave Photonics (MWP)* (pp. 1-4). IEEE Publishing. <https://doi.org/10.1109/MWP.2018.8552889>
- Pillai, J., Yadav, D., Prajapati, J., Upadhayay, M. D., & Babu, N. (2022). Compact wideband H-shaped slot antenna for sub-THz applications. In *2022 IEEE Delhi Section Conference (DELCON)* (pp. 1-6). IEEE Publishing.
- SIMULIA. (2019). *Computer Simulation Technology Microwave Studio version 19, Finite Integration in Technique Package*. SIMULIA <https://www.3ds.com/products-services/simulia/products/electromagnetic-simulation/>
- Singh, M., Singh, S., & Islam, M. T. (2021). Highly efficient ultra-wide band MIMO patch antenna array for short range THz applications. In A. Biswas, A. Banerjee, A. Acharyya & H. Inokawa (Eds.), *Emerging Trends in Terahertz Engineering and System Technologies* (pp. 193-207). Springer. [https://doi.org/10.1007/978-981-15-9766-4\\_10](https://doi.org/10.1007/978-981-15-9766-4_10)
- Singh, P. K., Mallik, S., Das, P., Krishna, H., & Tiwary, A. K. (2021). Terahertz antennas for future communications. In A. Acharyya & P. Das (Eds.), *Advanced Materials for Future Terahertz Devices, Circuits and Systems* (pp. 315-330). Springer. [https://doi.org/10.1007/978-981-33-4489-1\\_18](https://doi.org/10.1007/978-981-33-4489-1_18)
- Singhal, S. (2019). Ultrawideband elliptical microstrip antenna for terahertz applications. *Microwave and Optical Technology Letters*, *61*(10), 2366-2373. <https://doi.org/10.1002/mop.31910>
- Ullah, S., Ruan, C., Haq, T. U., & Sadiq, M. S. (2019). Circular polarized microstrip planar THz antenna using Z-shape slots. In *2019 International Symposium on Antennas and Propagation (ISAP)* (pp. 1-3). IEEE Publishing.
- Wu, G. B., Zeng, Y. S., Chan, K. F., Qu, S. W., & Chan, C. H. (2019). High-gain circularly polarized lens antenna for terahertz applications. *IEEE Antennas and Wireless Propagation Letters*, *18*(5), 921-925. <https://doi.org/10.1109/LAWP.2019.2905872>
- Xia, Q., & Jornet, J. M. (2019). Expedited neighbor discovery in directional terahertz communication networks enhanced by antenna side-lobe information. *IEEE Transactions on Vehicular Technology*, *68*(8), 7804-7814. <https://doi.org/10.1109/TVT.2019.2924820>

# Soft Set-based Parameter Reduction Algorithm Through a Discernibility Matrix and the Hybrid Approach for the Risk-Factor Prediction of Cardiovascular Diseases by Various Machine Learning Techniques

Menaga Anbumani<sup>1</sup> and Kannan Kaniyaiah<sup>2\*</sup>

<sup>1</sup>Department of Computer Science and Engineering, Srinivasa Ramanujan Centre, SASTRA Deemed to be University, Kumbakonam-612001 Tamil Nadu, India

<sup>2</sup>Department of Mathematics, Srinivasa Ramanujan Centre, SASTRA Deemed to be University, Kumbakonam-612001 Tamil Nadu, India

## ABSTRACT

Parameter reduction without performance degradation is a promising task in decision-making problems. For instance, a great challenge exists in constructing cost functions in gaming theory. Nevertheless, soft set theory handles all its drawbacks conveniently through a new tool for the choice function mathematically. In this paper, we propose an algorithm (SSPRDM) for parameter reduction of soft sets through discernibility matrices, and it is implemented in detecting the risk factor of heart disease problems by using six types of machine learning techniques. The parameters are extracted from the heart disease patient data by the SSPRDM algorithm, and then six machine learning techniques (LDA, KNN, SVM, CART, NB, RF) are performed in the prediction of risk factors for heart disease. The experimental results showed that the present hybrid approach provides an accuracy of 88.46% in the Random Forest technique, whereas the same Random Forest classifier provides an accuracy of 69.23% in the prediction of risk factors of cardiovascular disease (CVD) diagnosis in the earlier work which is a drastic improvement. Moreover, out of 18 parameter reductions, the core component is identified as Total Cholesterol, which is to be considered in all types of

CVD diagnosis, whereas Sugar-Fasting (C), Total-Cholesterol (G), and HDL-Cholesterol (I) are the core components identified in three parameter reductions ABCEGHI, ACFGIJ, and BCFGJK.

## ARTICLE INFO

### Article history:

Received: 04 April 2023

Accepted: 17 August 2023

Published: 24 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.16>

### E-mail addresses:

[anbumenakaa@gmail.com](mailto:anbumenakaa@gmail.com) (Menaga Anbumani)

[anbukannan@rediffmail.com](mailto:anbukannan@rediffmail.com) (Kannan Kaniyaiah)

\* Corresponding author

**Keywords:** Discernibility function, discernibility matrix, parameter reduction, risk factor, soft sets

## INTRODUCTION

The difficulties that existed in the dissimilarities of set theory were overcome by the concept of a soft set with suitable parameterizations (Moldostov, 1999). According to Moldostov (1999), the definition of a soft set is a couple  $(S, A)$  over  $(X)$  iff there exists a mapping from  $S$  on  $A$  to the collection of all subsets of the set  $X$ . Soft sets sidestep the limitations on suitable interpretation in such a way it gets an apt and easy tactics in exercise. There is a challenge in constructing cost function as far as gaming theory is concerned. Nevertheless, soft set theory handles all its drawbacks conveniently through a new tool for the choice function mathematically. The parameter reduction (Xie, 2016) on soft sets has been widely exposed in the history of soft sets in recent days.

The usage of rough sets (Pawlak, 1982; Pawlak, 1991) in decision-making applications is given by Maji et al. (2002). In the interim, some unsuitable and illogical reports in minimizing the parameter are projected (Maji et al. 2002). These irrational reports were pointed out and enhanced to reduce the complexity of collecting information and making better conclusions through soft sets (Chen et al. 2005). Kovkov et al. (2007) used optimization problems to reflect the idea of soft sets and also concentrated on validating the idea of the estimated description of objects. The clue for a standard decrease of parameters in soft sets is taken into account to find solutions to the problems of suboptimal choice and supplemented parameter sets of soft sets (Kong et al., 2008).

The AND operation is used to reduce multi-valued information systems based on soft-set. This strategy is recommended by Herawan et al. (2009). Rose et al. (2010) have developed a novel approach for finding maximal supported sets using parameterizations and diminution. These supported sets have been determined from the Boolean value information system. This proposed methodology is considered to be a reliable decision-making system. An uni-int decision-making method was provided (Cagman & Enginoglu, 2010). This method links the decision function to choose the optimal choice fruitfully in day-to-day problems that have no certainties naturally by descending as wide-ranging replacements.

A novel standard parameter-lessening procedure using soft sets was recommended (Ma et al., 2011). It is grounded on oriented-parameter sum ignoring parameter significant grade and choice partition. A method was proven to diminish the number of parameters for a soft set by not varying its unique classification ability (Ali, 2012). Considering the age factors, prostate volume (PV), and prostate-specific antigen (PSA) of patients, a Soft Expert System (SES) was developed to predict prostate cancer (Yuksel et al., 2013).

Normal parameter reductions based on soft sets are used to design the harmony search algorithm and the intelligent optimization algorithm and are applied to solve the same problems in data mining. By removing the unessential core, Kong et al. (2014) used the particle swarm optimization method to minimize the attributes in the soft set. A technologically progressive HPC algorithm is used to advance the choice of search area,

which is also used to reduce choice costs by streamlining the real categorizations by choice partition order (Mohammed et al., 2017). The restriction in the current prevailing algorithm is broken by Khan and Zhu (2019), who proposed a procedure for standard parameter reduction of soft set that finds application in the medical field. Sadiq et al. (2020) principally focused on lessening implementation expenses by adapting the innovative organizations by choice barrier order and refining the likelihood of probing the domain's realm through an advanced Markov chain model.

In the medical domain, there are many works related to the application of machine learning in forecasting diseases (Mitchell, 1997). The major impact on the heart is due to critical conditions caused by cardiovascular disease (Boukhatem et al., 2022). It requires early disease prediction by finding reliable, accurate, and sensible approaches to identifying the numerous risk factors related to heart (Shah et al., 2020). Jindal et al. (2021) classified the patients with heart disease using logistic regression and KNN algorithms. The prediction of sudden bursts of heart-related diseases among patients on dialysis grounded on machine learning is given accurately by Mezzatesta et al. (2019). Bhat et al. (2022) created a strategy for predicting the risk of developing diabetes in North Kashmir using machine learning algorithms. Six MLA, including Random Forest (RF), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), Gradient Boost (GB), Decision Tree (DT), and Logistic Regression (LR), have been used successfully in the experimental study. For the balanced data set, RF has the highest accuracy rate of any classifier.

A method to predict cardiac disease was proposed by Ansari et al. using a variety of machine learning algorithms (MLA), including logistic regression (LR), k-nearest neighbor (KNN), support vector machine (SVM), Naive Bayes (NB), random forest (RF), and decision tree (DT) (Ansari et al., 2023). They assessed the model's performance in predicting heart disease using the testing data set, and their findings indicated that, compared to other algorithms, the k-nearest neighbor method and random forest best fit the data.

In the ground field of medical sciences, these methods were beneficial in examining and estimating a diversity of medicinal syndromes (Alotaibi, 2022). Nonalcoholic fatty liver disease (NAFLD) is common among patients and results in cardiovascular disease (CVD), which acts as the key source of death (Sharma et al., 2022). Palaniappan et al. (2021) developed a machine learning model on modified SEIR, which is used to forecast the epidemic disease dynamics on the contamination risk. The model used Multivariate Logistic Regression. In early 2020, AlArjani et al. (2022) witnessed an outburst of COVID-19, and the entire globe was affected by the same. Machine learning is an important tool in screening, analyzing, forecasting, tracking, and predicting the features and tendencies of COVID-19 (Rahman et al., 2021). Even though many articles on COVID-19 were published in 2020, no effective estimation means still exist to identify the disease with a hundred percent efficiency (Abirami & Kumar, 2022). The most common heart disease is coronary

artery disease (CAD). Over 80% of the deaths are due to this disease, common among developing countries, including Nigeria. This disease's victims are under 70 (Muhammad et al., 2021).

According to research, Alzheimer's disease, a frequent form of dementia, has no known cure. However, by raising the patient's quality of life and offering a solution to strengthen the patient's cognitive capacities with Smartphones, the disease's progression can be slowed down. It highlights the necessity for relatively straightforward cross-platform mobile application development with interactive GUIs to improve users' cognitive capacities (Gupta et al., 2018).

For more effective and accurate recognition of Tetralogy of Fallot (TOF), Wang et al. (2021) suggest a new artificial intelligence model based on cardiovascular computed tomography. Their model combines stochastic pooling, structural optimization, and convolutional neural networks. It is known as SOSPCNN (structurally optimized stochastic pooling convolutional neural network). Multiple-way data augmentation is also performed to avoid overfitting. The proposed SOSPCNN model is given explainability using Grad-CAM. Meanwhile, this SOSPCNN model is used to create desktop and web apps.

Zhang and colleagues used two-level stationary wavelet entropy (SWE) to extract features from brain images. Next, they examined three machine learning-based classifiers: the decision tree, k-nearest neighbors (kNN), and support vector machine (Zhang et al., 2016). According to their experimental findings, the kNN performed the best among the three classifiers. Furthermore, the proposed SWE+kNN technique outperforms four cutting-edge approaches.

It motivates that more work on hybrid approach is essential in the routine life problems. The current article aims to highlight effective procedures for parameter lessening using soft sets with a discernibility matrix and to identify the core parameters involved in diagnosing heart disease. Further, it is utilized to estimate the influence of risk associated with heart ailment using machine learning techniques.

### Soft Sets and Information Systems

Consider that  $X$  is a determinate collection of objects, and  $E$  is a finite set of attributes. The pair  $(X, E, V, f)$  is known as an information system if  $f$  is an information function from  $X \times E$  to  $V = \bigcup_{\alpha \in E} V_\alpha$  where every  $V_\alpha = \{f(x, \alpha), \alpha \in E, x \in X\}$  is the value of the attribute  $\alpha$ . An information system  $(X, E, V, f)$  is called a bi-value if  $V = \{0, 1\}$ . Suppose that  $S = f_E$  is a soft set over  $X$ . Then  $I_S = (X, E, V, f_S)$  is called the bi-value information system persuaded by  $S$  where  $f_S: V \times E \rightarrow V$  is well-defined by  $f_S(x, \alpha) = \begin{cases} 1, & x \in F(\alpha) \\ 0, & x \notin F(\alpha) \end{cases}$ , for any  $x \in X, \alpha \in E$ . Let  $I = (X, E, V, f)$  be a bi-value information system. Then,  $S_1 =$



$(f_I, E)$  is called the soft set over  $U$  persuaded by  $I$ , where  $F_I: A \rightarrow 2^X$  is well-defined by  $F_I(\alpha) = \{x \in X / f(x, \alpha) = 1, \alpha \in E\}$ . Suppose that  $\Sigma = \{S / S = f_A \text{ is a soft set over } U\}$  and  $\Gamma = \{I / I = (X, E, V, f) \text{ is a bi-value information system}\}$ . Then, there occurs a one-to-one relation between  $\Sigma$  and  $\Gamma$ . Let  $f_E$  be a soft set over  $X$  and let  $(X, E, V, f)$  be the bi-value information system persuaded by  $f_E$  over  $X$ . For any  $B \subseteq A$ , the association  $R_B$  is stated as follows:  $R_B = \{(x, y) \in X \times X / f(x, \alpha) = f(y, \alpha), \forall \alpha \in B\}$ . One can straightforwardly witness that  $R_{\{\alpha\}} = R_\alpha$  and  $R_B = \bigcap_{\alpha \in B} R_\alpha$ . Let  $f_E$  be a soft set over  $X$ . (1) Any  $B \subseteq A$  is a parameter lessening of  $f_E$  if  $R_A = R_B$  and  $R_A \supseteq R_{B-\{\alpha\}}$  for any  $\alpha \in B$ . (2) The common collection of all parameters lessening  $f_E$  is called the softcore and is designated by  $\text{core}(f_E)$ . Suppose  $f_E$  is a soft set over  $|X| = n$  and let  $(X, L, V, f)$  be the bi-value info system persuaded by  $f_E$  over  $X$ . For  $x, y \in X$ ,  $d(x, y)$  is defined as follows:  $d(x, y) = \{e \in E / g(x, e) \neq f(y, e)\}$ . (1)  $d(x, y)$  is the collection of parameters that can distinguish the objects  $x$  and  $y$ . (2)  $D(f_E) = (d_{ij})_{n \times n}$  is called the discernibility matrix of  $f_E$ , where  $X = \{x_1, x_2, \dots, x_n\}$  and  $d_{ij} = d(x_i, x_j)$ .

**MATERIALS AND METHODS**

Vignesh Clinical Laboratory, Karthik Laboratory and Keerthi X-Ray, and Vaishnavi Clinical Laboratory are renowned laboratories in Kumbakonam, Tamilnadu, India. Data is collected from the laboratories mentioned above. The dataset contains details: Blood-Glucose, Blood-Urea, Serum-Creatinine, Total-Cholesterol, HDL-Cholesterol, LDL-Cholesterol, Triglycerides, Risk-Factor-I and Risk-Factor-II. The total number of records is 110. First, some efficient procedures to minimize the parameters using soft sets are proposed. It is then used to predict the estimation of risk factors associated with heart ailment using machine learning techniques. Finally, the results are discussed and compared with those obtained by the authors with standard parameter lessening (without soft sets) and are visualized in Figure 1.

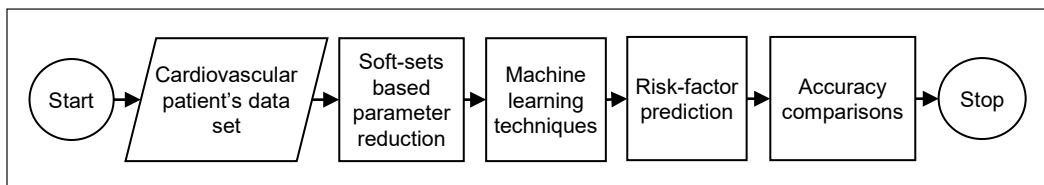


Figure 1. Methodology

**Proposed SSPRDM Algorithm for Minimizing the Parameters Using Soft Sets**

The succeeding discussion studies the procedure for effective parameter lessening through a discernibility matrix, which contains the resulting steps and applies to decision-making problems.



### Step 1: Finding Discernibility\_Matrix

The algorithm Discernibility\_Matrix takes `map_matrix[][]` as input (Kannan & Menaga, 2022). It constructs the `dmatrix` of size  $n \times n$ . Here, 'n' defines the total number of objects. The `dmatrix[i][j] = null` for every  $\text{rowvalue}(i) = \text{columnvalue}(j)$ . That is, the diagonal entries in the `dmatrix` are set to null. Each entry in the `dmatrix` is found by concatenating the features[k] where  $k = 1$  to  $f$  provided the condition `map_matrix[k][j]` not equal to `map_matrix[k][i]` is satisfied. The process is repeated for  $n$  times. In the succeeding algorithm,  $n$  is denoted as `num`.

#### Algorithm 1 Discernibility\_Matrix(`map_matrix[][]`)

```

1: value = empty
2: for x ← 1 to num do
3:   for y ← 1 to num do
4:     for k ← 1 to f do
5:       if matrix[k, y] <> matrix[k, x] then
6:         value = value + features[k]
7:       end if
8:     Next k
9:   end for
10:  dmatrix[x, y] = value
11:  value = empty
12: Next y
13: end for
14: Next x
15: end for
16: Display the dmatrix

```

The number of comparisons done gives the time complexity to generate the `discernibility_matrix`. The inner loop consists of one comparison repeated for  $n \times n \times f$ . Hence, the algorithm's efficiency is given by  $O(n^3)$  as  $f$  approaches  $n$ .

### Step 2: Finding Discernibility\_Function

The algorithm takes `Discernability_matrix[][]` as input and finds unique entries such that no two entries are exactly equal. The first step removes the spaces in the leading and trailing ends of the entries using `trim()`. Then, it takes the entries one by one, compares them with the remaining entries, and sets null for the exactly equal entries. After  $n \times n$  iterations, the matrix consists of unique parameters and null entries. Excluding null entries, the remaining parameters are stored in the `Discernibilityfn_matrix` for further processing. Regarding time complexity, each entry is compared with all the remaining entries, and  $n$  is the count of the objects. Hence,  $O(n^2)$  gives the overall complexity.

**Algorithm 2 Discernibility\_Fucntion (Discernibility\_matrix[[[]],num)**

```

1: for x ← 1 to num do
2:   for y ← 1 to num do
3:     Discernibility_matrix[x][y] = Discernibility_matrix[x][y].Trim
4:     Next y
5:   end for
6:   Next x
7: end for
8: for x ← 1 to n do
9:   for y ← 1 to n do
10:    S = Discernibility_matrix[x][y]
11:    for h ← 1 to num do
12:      for k ← 1 to num do
13:        if S.Equals(Discernibility_matrix[h][k]) then
14:          Discernibility_matrix[h][k]=""
15:        end if
16:      Next k
17:    end for
18:    Next h
19:  end for
20:  Next y
21: end for
22: Next x
23: end for
24: q = 1
25: for x ← 1 to num do
26: for y ← 1 to num do
27:   if (Discernibility_matrix[i][j] <> "") then
28:     Discernibilityfn_matrix[q] = Discernibility_matrix[h][k]
29:     q = q + 1
30:   end if
31:   Next y
32: end for
33: Next x
34: end for

```

**Step 3: Parameter Reduction\_Discernibility\_function**

The chief goal of finding discernibilityfn\_matrix is to find the parameter reduction from the discernibility matrix. It consists of unique entries. The next step is to take each parameter and to find the list of parameters containing it, and it can be concluded that all the parameters

containing it can be reduced to the parameter taken for the substring comparison. For this, the Instr() function returns the index of the first occurrence of the substring contained in the original string. Nevertheless, Instr() finds the pattern in another string bit, and it does not return true if the original string contains the substring but not in the exact pattern. Another function named check is used to find that.

The check function takes two parameter entries as strings as input. It finds the length of the two strings. The function also converts the two strings to a character array. It compares and finds the minimum length between the two strings. It then checks whether one string is present in another character by character using string.contains(). If so, the function returns true or false. It helps find the parameter reduction. It is used if Instr() fails to find the occurrence.

**Algorithm 3 Algorithm Reduction\_Discernibility\_function(Discernibilityfn\_matrix[],q)**

1: // q is the total count of distinct entries in the Discernibilityfn\_matrix[]

2: for i ← 1 to q do

3: S = Discernibilityfn\_matrix[i]

4: for k ← i + 1 to q do

5: if (k <> 1) then

6: If instr(Discernibilityfn\_matrix[ik], s) <> 0 or check(Discernibilityfn\_matrix[k], s) = true then

7: Discernibilityfn\_matrix[k] = ""

8: end if

9: end if

10: Next k

11: end for

12: Next i

13: end for

14: m = 1

15: for i ← 1 to q do

16: if Discernibilityfn\_matrix[i] <> "" then

17: Reduction[m] = Discernibilityfn\_matrix[i]

18: m = m + 1

19: end if

20: Next i

21: end for

22: for i ← 1 to m do

23: Display Reduction[i]

24: Next i

25: end for

**Algorithm 4 Algorithm Check(S1, S2)**

```

1: // S1 and S2 are two string entries form Discernibilityfn_matrix[][]
2: //Algorithm returns Yes if either S1 is in S2 or S2 is in S1
3: i = s1.length
4: j = s2.length
5: c1[] = s1.toCharArray()
6: c2[] = s2.toCharArray()
7: if i > j) then
8:   min = i
9: else
10:  min = j
11: end if
12: if i <= j then
13:   for k ← 0 to i - 1 do
14:     if s2.Contains(c1(k)) then
15:       count = count + 1
16:     end if
17:   Next k
18: end for
19: else
20:   for k ← 0 To j - 1 do
21:     if s1.Contains(c1(k)) then
22:       count = count + 1
23:     end if
24:   Next k
25: end for
26: end if
27: if count = min then
28:   return Yes
29: else
30:   return No
31: end if

```

The running time performance of the algorithm `Reduction_Discernibility_function` is  $O(q^2)$ , where  $q$  defines the order of the `discernibilityfn_matrix`. Also, it includes the time taken to complete the execution of the function call. The major part of the check function is the execution of `contains()`, which is repeated for the times, and it is equivalent to the total count of the characters in the string that is minimum among two strings passed. Therefore,  $O(q^2)$  gives the complete complexity of the procedure.



$$V_3, V_6, V_9\} \wedge \{V_1, V_3, V_4, V_5, V_6, V_7, V_9\} \wedge \{V_1, V_4, V_6, V_7, V_8, V_9\} \wedge \{V_6, V_7, V_9\} \wedge \{V_3, V_5, V_8\} \wedge \{V_1, V_4, V_6, V_8, V_9\} \wedge \{V_1, V_4, V_8\} \wedge \{V_1, V_3, V_4, V_5, V_6, V_9\} \wedge \{V_6, V_9\} \wedge \{V_1, V_3, V_4, V_5\} \wedge \{V_1, V_2, V_3, V_6, V_8, V_9\}.$$

Here, the various parameter reductions returned as output from the proposed soft-set-based parameter reduction algorithms are taken as the data set for which the standard machine learning classifiers are applied, the model accuracies are evaluated, and a comparison is made to identify the predominant parameters with better accuracy.

In a hybrid approach, all 11 independent data set parameters are considered for reductions (Table 2). For the convenience of implementation, the parameters are labeled using the alphabet indicated in Table 2.

The data set is represented as a soft-set, and a sample of the data set is passed as inputs to the proposed soft-set parameter reduction algorithms—the set of parameter reductions returned as output is indicated in Table 3.

### Risk Factor Prediction with Machine Learning Techniques

The cardiovascular disease data set is taken, and a machine learning technique is applied to identify the accuracy of model predictions. In our earlier work, out of 11 independent parameters namely Gender, Blood Hemoglobin, Sugar Fasting, Sugar PP, Urea, Creatinine, Total Cholesterol, Triglycerides, HDL Cholesterol, LDL Cholesterol and V LDL Cholesterol, there occurs positive correlation between Sugar Fasting and Sugar PP with a coefficient equal to 0.847; Creatinine and Urea has correlation coefficient equal to 0.625; LDL cholesterol and Total Cholesterol has

Table 2  
All 11 parameters with label

Label	Parameters
A	Gender
B	Blood–Hemoglobin
C	Sugar–Fasting
D	Sugar–PP
E	Urea
F	Creatinine
G	Total–Cholesterol
H	Triglycerides
I	HDL–Cholesterol
J	LDL–Cholesterol
K	VLDL–Cholesterol

Table 3  
List of parameter reductions

S. No	Parameter Reductions
1.	BCFGIJK
2.	BCFGHIJK
3.	CFGHIJK
4.	ACFGIK
5.	ACFGIJ
6.	ABEGHI
7.	ACEGIJ
8.	ACEGHI
9.	AFGHJK
10.	ACEFGI
11.	BEGHIK
12.	AEGHIJ
13.	AIEFGHI
14.	BCEGIK
15.	CEGHIK
16.	ABCEGI
17.	ACEGK
18.	AEGHK

correlation coefficient equal to 0.983; VLDL cholesterol and Triglycerides has correlation coefficient equal to 0.998 (Kannan & Menaga, 2022). Hence, to diagnose further Sugar PP, creatinine, LDL cholesterol, and VLDL cholesterol were not considered. Henceforth, 8 parameters were considered for Risk Forecasting.

Subsequently, various supervised machine learning classification algorithms, namely LDA, KNN, CART, Random Forest, SVM, and Nave Bayes were applied for the data set with the model accuracies as 61.54%, 65.38%, 65.38%, 69.23%, 61.08%, and 57.69% respectively. The result shows that the Random Forest method gives better accuracy than all the applied classifiers.

Now, a hybrid approach is attempted to improve the prediction accuracy. The machine learning classifiers KNN, LDA, SVM, CART, Random Forest, and Naive Bayes are implemented for the above 18 parameter reductions, and the accuracy of model predictions is evaluated.

True positive and negative are dealings of a test's capability that properly categorize an individual with or without illness. True positive mentions a test's capability to designate an individual with illness as positive. A highly sensitive test designates that there are few false negative consequences, and thus, fewer belongings of illness are missed. The true negative test is its ability to designate an individual who does not have an illness. A highly specific test indicates that there are few false positive results. Having specificity as low will not be appropriate for screening since many people without the illness will screen positive and hypothetically collect redundant analytical measures. It is observed that the highest values 0.8462, 0.9231 are the sensitivity and specificity, respectively, in the parameter reductions ACFGIJ and BCFGJK, whereas 0.8462 and 0.5485 are the values for sensitivity and specificity in ABCEGHI in Kannan and Menaga (2022). Table 4 recapitulates the true positive and specificity metrics of various classifiers.

Accuracy is one of the measurements for assessing the efficiency of the model's classification. The informal meaning of accuracy is defining the segment of estimates for the right model. Formally, accuracy is defined as 
$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}}$$

. For two-fold organization, accuracy can also be computed as positives and negatives as

follows: 
$$\text{Accuracy} = \frac{\text{T.P} + \text{T.N}}{\text{T.P} + \text{T.N} + \text{F.P} + \text{F.N}}$$
, where TP = True Positives, TN = True Negatives,

FP = False Positives, and FN = False Negatives.

From Table 5, it is known that improved accuracies are obtained by the Random Forest method. The two-parameter reductions, namely Blood Hemoglobin, Sugar Fasting, Creatinine, Total Cholesterol, HDL Cholesterol, LDL, VLDL (BCFGJK) and Gender, Sugar Fasting, Creatinine, Total Cholesterol, HDL Cholesterol, LDL (ACFGIJ), provide the highest accuracy 88.46% out of all remaining parameter reductions. However, the



Table 4  
Sensitivity and specificity table

Parameters	Sensitivity and specificity metric of various classifiers											
	LDA		KNN		SVM		CART		NB		RF	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
ABCEGHI in (Kamran, 2022)	0.6923	0.5385	0.8462	0.4615	0.7692	0.5385	0.6154	0.6923	1.0000	0.2308	0.8462	0.5385
ACFGIJ	0.5385	0.6923	0.7692	0.6154	0.8462	0.6923	0.6923	0.7692	0.4615	0.7692	0.8462	0.9231
BCFGHIJK	0.6923	0.6154	0.6154	0.5385	0.7692	0.7692	0.7692	0.4615	0.5385	0.8462	0.8462	0.9231
ACEGIJ	0.5385	0.4615	0.6923	0.6154	0.9231	0.7692	0.6923	0.7692	0.7692	0.5385	0.7692	0.9231
ABCEGI	0.5385	0.5385	0.6154	0.6923	0.6154	0.6154	0.5385	0.6154	0.6923	0.5385	0.7692	0.8462
BCFGHIJK	0.9231	0.6154	0.7692	0.5385	0.7692	0.5385	0.8462	0.4615	0.7692	0.2308	0.7692	0.7692
AEFGHI	0.6154	0.6154	0.8462	0.3846	0.9231	0.7692	0.3846	0.4615	0.7692	0.5385	0.6923	0.8462
CFGHIJK	0.4615	0.6923	0.6154	0.5385	0.7692	0.6923	0.6923	0.7692	0.6923	0.5385	0.6923	0.7692
ACEFGI	0.4615	0.6154	0.6154	0.6923	0.6923	0.6154	0.76923	0.07692	0.15385	0.92308	0.6154	0.7692
BEGHIK	0.5385	0.6154	0.6923	0.3077	0.6923	0.3846	0.3846	0.3077	0.5385	0.3077	0.6154	0.6923
ACFGJK	0.5385	0.6923	0.6923	0.6154	0.6923	0.4615	0.6923	0.7692	0.6923	0.6154	0.6154	0.6154
ABEGHI	0.6154	0.6923	0.6923	0.3077	0.8462	0.5385	0.3846	0.3077	0.9231	0.5385	0.6154	0.6154
AEFGHI	0.6154	0.6923	0.6923	0.3077	0.9231	0.5385	0.3846	0.3077	0.5385	0.3077	0.6154	0.6154
ACEGHI	0.6154	0.6154	0.7692	0.5385	0.6923	0.4615	0.6923	0.5385	0.6154	0.3846	0.5385	0.6154
AFGHJK	0.4615	0.6923	0.7692	0.3846	0.7692	0.3846	0.3846	0.4615	0.6154	0.5385	0.5385	0.6154
BCEGIK	0.5385	0.5385	0.6154	0.5385	0.6923	0.5385	0.5385	0.6154	0.5385	0.3846	0.6154	0.5385
CEGHIK	0.6154	0.6154	0.7692	0.5385	0.7692	0.5385	0.6923	0.5385	0.6154	0.3846	0.5385	0.6154
ACEGK	0.5385	0.4615	0.6923	0.5385	0.5385	0.4615	0.6923	0.5385	0.6923	0.3846	0.4615	0.4615
AEGHK	0.7692	0.4615	0.6923	0.3077	0.6923	0.3846	0.3846	0.3077	0.5385	0.3846	0.3846	0.4615

Table 5  
Accuracy

Accuracy Parameters	Metric of various classifiers					
	LDA in %	KNN in %	SVM in %	CART in %	NB in %	RF in %
ABCEGHI in (Kannan, 2022)	61.54	65.38	65.38	65.38	61.34	69.23
ACFGIJ	61.54	69.23	76.92	73.08	61.54	88.46
BCFGIJK	61.16	57.69	76.92	61.54	69.23	88.46
ACEGIJ	50	63.58	84.62	73.08	63.58	84.62
ABCEGI	53.85	63.58	61.54	57.69	61.54	80.77
BCFGHIJK	76.92	61.54	73.08	53.08	69.23	76.92
AEGHIJ	61.54	61.54	84.62	42.31	65.38	76.92
CFGHIJK	57.69	57.69	73.08	73.08	61.54	73.08
ACEFGI	53.85	65.38	65.38	42.31	53.85	69.23
BEGHIK	57.69	50	53.85	34.62	42.31	65.38
ACFGJK	61.54	65.38	57.69	73.08	65.38	61.54
ABEGHI	65.38	50	69.23	34.62	73.08	61.54
AIEFGHI	65.38	50	73.08	34.62	42.31	61.54
ACEGHI	61.54	65.38	57.69	61.54	50	57.69
AFGHJK	57.69	57.69	57.69	42.31	57.69	57.69
BCEGIK	53.85	57.69	61.54	57.69	46.15	57.69
CEGHIK	61.54	65.38	65.38	61.54	50	57.69
ACEGK	50	61.54	50	61.54	53.85	46.15
AEGHK	61.54	50	53.85	34.62	46.15	42.31

accuracies of LDA, KNN, SVM, CART, and RF in ACFGIJ are improved than that of LDA, KNN, SVM, CART, NB, and RF in BCFGIJK respectively, whereas in NB, the accuracy is decreased. At the same time, the accuracies of LDA, KNN, SVM, CART, NB, and RF in ACFGIJ are either improved or remain the same as that of LDA, KNN, SVM, CART, NB, and RF in ABCEGHI in Kannan and Menaga (2022), respectively. Hence, the proposed hybrid approach drastically improves accuracy over the conventional ACFGIJ and BCFGIJK methods (Figures 2, 3, and 4).

Moreover, Sugar Fasting (C), Total Cholesterol (G), and HDL Cholesterol (I) are the core components identified in all these three parameter reductions. Then, the importance of the parameters is identified (Random Forest) in BCFGIJK and ACFGIJ, shown in Tables 6 and 7, respectively.

Initially, the order of importance of the parameters in BCFGIJK [Figure 5(a)] is HDL cholesterol, LDL cholesterol, Total cholesterol, VLDL cholesterol, Creatinine, Blood Hemoglobin, and Sugar fasting, whereas, after tuning, it is LDL cholesterol, HDL

cholesterol, Total cholesterol, VLDL cholesterol, Creatinine, Blood Hemoglobin and Sugar fasting. Similarly, the order of importance of the parameters in ACFGIJ (Figure 5b) is HDL cholesterol, LDL cholesterol, Total cholesterol, Creatinine, Gender, and Sugar fasting, whereas, after tuning, it is LDL cholesterol, HDL cholesterol, Total cholesterol, Sugar fasting, Creatinine and Gender.

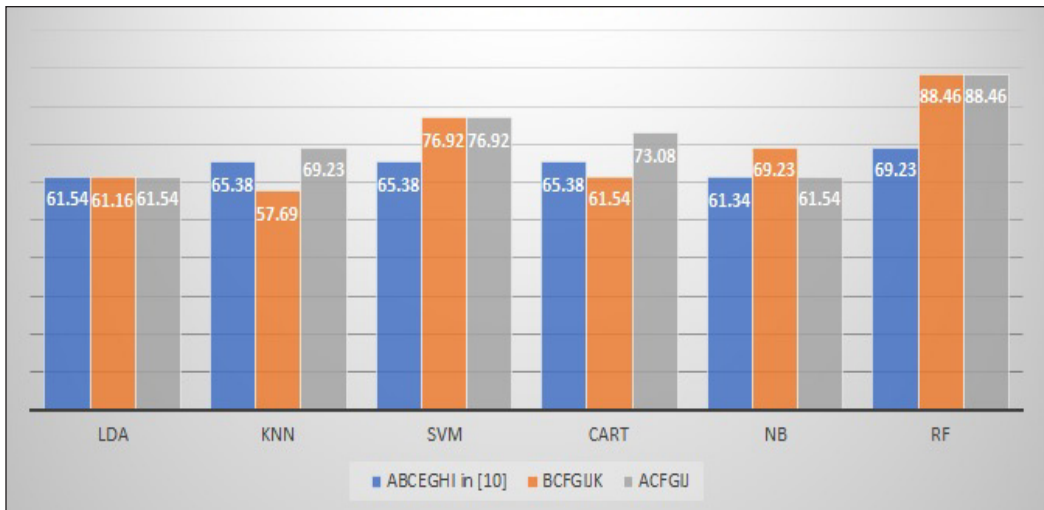


Figure 2. Comparison of accuracy

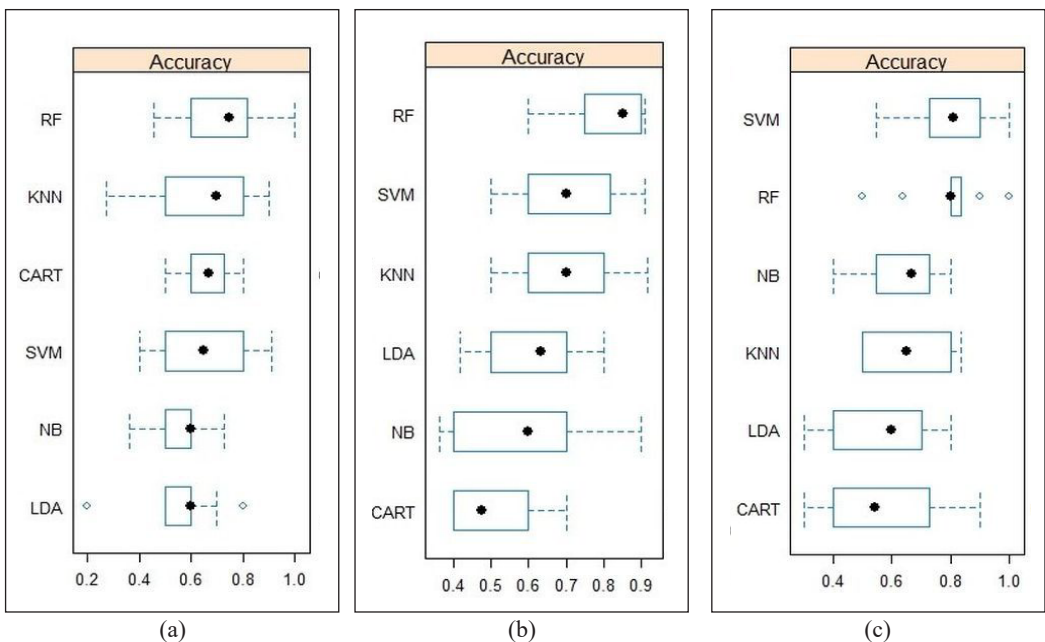


Figure 3. Comparison of accuracy: (a) ABCEGHI in (Kannan, 2022); (b) BCFGUJ; and (c) ACFGU

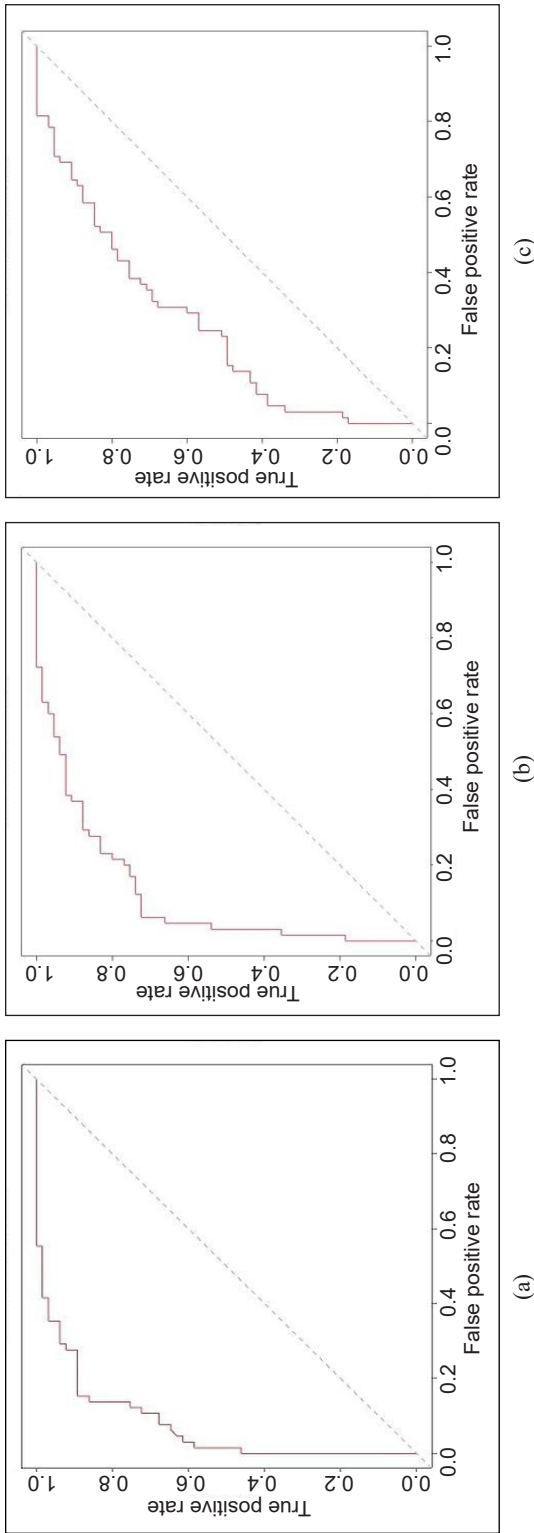


Figure 4. ROC Curves: (a) ABCEGHI in (Kannan, 2022); (b) BCFGJJK; and (c) ACFGIJ

Table 6  
Importance of parameters in random forest in BCFGJJK

Parameters	There is no risk	There is risk	Mean decrease accuracy	Mean decrease Gini
Blood Hemoglobin	4.7493817	1.337873	4.352876	7.332752
Sugar Fasting	0.8752779	4.387068	3.477273	7.237084
Creatinine	4.0794305	2.948084	4.899098	5.083962
Total Cholesterol	15.8845728	16.667775	20.519157	11.189501
H.D.L Cholesterol	22.1786327	16.667775	25.162436	12.299851
L.D.L Cholesterol	17.9061889	17.977023	23.582813	13.653628
V.L.D.L Cholesterol	1.6163369	7.961650	7.252641	7.665264

Table 7  
Importance of parameters in random forest in ACFGIJ

Parameters	There is no risk	There is risk	Mean decrease accuracy	Mean decrease Gini
Gender	0.7070882	6.234695	5.186608	1.745934
Sugar Fasting	1.6658202	4.722036	4.421667	9.171790
Creatinine	3.1746004	5.079762	5.521531	6.397909
Total Cholesterol	17.3298524	17.734249	24.069675	14.070509
H.D.L Cholesterol	26.4972019	26.996690	33.388424	16.315529
L.D.L Cholesterol	19.8903453	23.907039	29.075341	16.724294

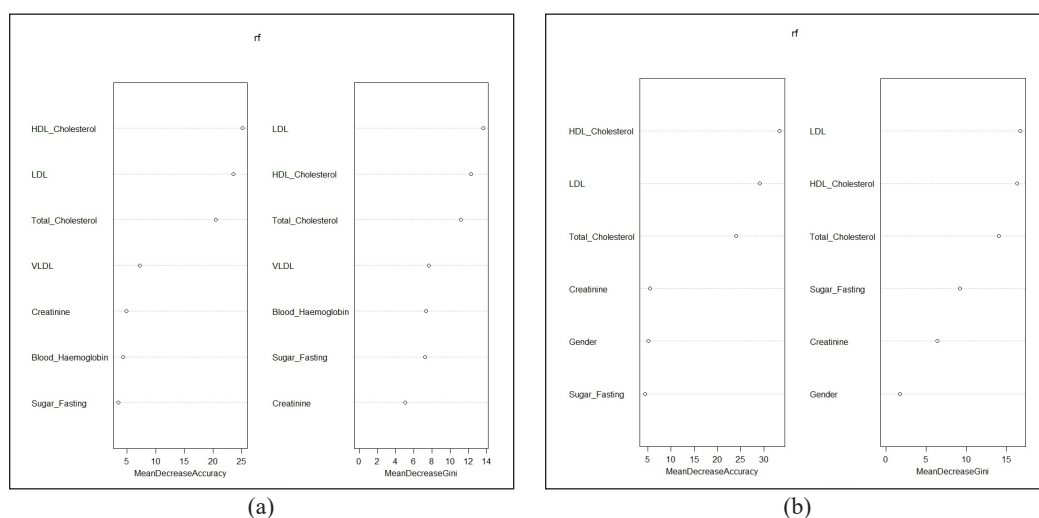


Figure 5. Importance of parameters in: (a) BCFGIIJK; and (b) ACFGIJ

## CONCLUSION

The present paper proposes an algorithm to minimize the parameters based on a soft-set discernibility matrix. The execution time analysis of the suggested procedure is calculated. Further, parameter reduction algorithms are implemented in a real-time decision-making application, i.e., for predicting cardiovascular disease risk factors. Also, a hybrid approach combining soft-set and machine learning techniques is proposed for efficient parameter reduction, and accuracy comparisons were made. In the existing approach, the Random Forest provided better accuracy (69.23%), and the proposed hybrid approach provided improved accuracy (88.46%) for the same classifier, which is a drastic improvement. A more comparative analysis between multiple datasets and their properties might be undertaken to find all the essential elements for forecasting the risk factor. Furthermore, rather than conventional machine learning approaches, deep learning techniques like artificial neural networks and structural equation modeling can be used in the future.

## ACKNOWLEDGEMENT

The authors thank Dr. D. Rajalakshmi, Department of Mathematics, Srinivasa Ramanujan Centre, SASTRA Deemed to be University, Kumbakonam, Tamil Nadu, INDIA, for providing the data set to enhance the quality of the paper.

## REFERENCES

- Abirami, R. S., & Kumar, G. S. (2022). Comparative study based on analysis of coronavirus disease (COVID-19) detection and prediction using machine learning models. *SN Computer Science*, 3(1), Article 79. <https://doi.org/10.1007/s42979-021-00965-2>
- AlArjani, A., Nasseef, M. T., Kamal, S. M., Rao, B. V. S., Mahmud, M., & Uddin, M. S. (2022). Application of mathematical modeling in prediction of COVID-19 transmission dynamics. *Arabian Journal for Science and Engineering*, 47(8), 10163-10186. <https://doi.org/10.1007/s13369-021-06419-4>
- Ali, M. I. (2012). Another view on reduction of parameters in soft sets. *Applied Soft Computing*, 12(6), 1814-1821. <https://doi.org/10.1016/j.asoc.2012.01.002>
- Alotaibi, A. S. (2022). Hybrid model based on relief algorithm and k-nearest neighbor for erythematous-squamous diseases forecasting. *Arabian Journal for Science and Engineering*, 47(2), 1299-1307. URL <https://doi.org/10.1007/s13369-021-05921-z>
- Ansari, G. A., Bhat, S. S., Ansari, M. D., Ahmad, S., Nazeer, J., & Eljaly, A. E. M. (2023). Performance evaluation of machine learning techniques (MLT) for heart disease prediction. *Computational and Mathematical Methods in Medicine*, 2023, Article 8191261. <https://doi.org/10.1155/2023/8191261>
- Bhat, S. S., Selvam, V., Ansari, G. A., Ansari, M. D., & Rahman, M. H. (2022). Prevalence and early prediction of diabetes using machine learning in North Kashmir: A case study of District Bandipora. *Computational Intelligence and Neuroscience*, 2022, Article 2789760. <https://doi.org/10.1155/2022/2789760>
- Boukhatem, C., Youssef, H. Y., & Nassif, A. B. (2022). Heart disease prediction using machine learning. In *2022 Advances in Science and Engineering Technology International Conferences (ASET)* pp. 1-6. IEEE Publishing. <https://doi.org/10.1109/ASET53988.2022.9734880>
- Cagman, N., & Enginoglu, S. (2010). Soft set theory and uni-int decision making. *European Journal of Operational Research*, 207(2), 848-855. <https://doi.org/10.1016/j.ejor.2010.05.004>
- Chen, D., Tsang, E. C. C., Yeung, D. S., & Wang, X. (2005). The parameterization reduction of soft sets and its applications. *Computers & Mathematics with Applications*, 49(5), 757-763. <https://doi.org/10.1016/j.camwa.2004.10.036>
- Gupta, G., Gupta, A., Jaiswal, V., & Ansari, M. D. (2018). A review and analysis of mobile health applications for alzheimer patients and caregivers. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 171-175). IEEE Publishing. <https://doi.org/10.1109/PDGC.2018.8745995>
- Herawan, T., Rose, A. N. M., & Deris, M. M. (2009). Soft set theoretic approach for dimensionality reduction. In D. Ślęzak, T. Kim, Y. Zhang, J. Ma & K. Chung (Eds.), *Database Theory and Application* (pp. 171-178). Springer. [https://doi.org/10.1007/978-3-642-10583-8\\_20](https://doi.org/10.1007/978-3-642-10583-8_20)

- Jindal, H., Agrawal, S., Khera, R., Jain, R., & Nagrath, P. (2021). Heart disease prediction using machine learning algorithms. In *IOP conference series: materials science and engineering* (Vol. 1022, No. 1, p. 012072). IOP Publishing. <https://doi.org/10.1088/1757-899X/1022/1/012072>
- Kannan, K., & Menaga, A. (2022). Some efficient algorithms on the parameter reduction of soft sets for decision making problems. *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, 92(3), 319-329. <https://doi.org/10.1007/s40010-021-00730-3>
- Khan, A., & Zhu, Y. (2019). An improved algorithm for normal parameter reduction of soft set. *Journal of Intelligence and Fuzzy Systems*, 37(2), 2953-2968. <https://doi.org/10.3233/JIFS-190071>
- Kong, Z., Gao, L., Wang, L., & Li, S. (2008). The normal parameter reduction of soft sets and its algorithm. *Computers and Mathematics with Applications*, 56(12), 3029-3037. <https://doi.org/10.1016/j.camwa.2008.07.013>
- Kong, Z., Zhang, G. D., & Wang, L. F. (2014). Normal parameter reduction in soft set based on harmony search algorithm. *Applied Mechanics and Materials*, 644-650, 2173-2176. <https://doi.org/10.4028/www.scientific.net/AMM.644-650.2173>
- Kovkov, D. V., Kolbanov, V. M., & Molodtsov, D. A. (2007). Soft sets theory-based optimization. *Journal of Computer and Systems Sciences International*, 46(6), 872-880. <https://doi.org/10.1134/S1064230707060032>
- Ma, X., Sulaiman, N., Qin, H., Herawan, T., & Zain, J. M. (2011). A new efficient normal parameter reduction algorithm of soft sets. *Computer & Mathematics with Applications*, 62(2), 588-598. <https://doi.org/10.1016/j.camwa.2011.05.038>
- Maji, P. K., Roy, A. R., & Biswas, R. (2002). An application of soft sets in a decision making problem. *Computers & Mathematics with Applications*, 44(8), 1077-1083.
- Mezzatesta, S., Torino, C., Meo, P. D., Fiumara, G., & Vilasi, A. (2019). A machine learning-based approach for predicting the outbreak of cardiovascular diseases in patients on dialysis. *Computer Methods and Programs in Biomedicine*, 177, 9-15. <https://doi.org/10.1016/j.cmpb.2019.05.005>
- Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill.
- Mohammed, M. A. T., Sadiq, A. S., Arshah, R. A., Ernawan, F., & Mirjalili, S. M. (2017). Soft set decision/forecasting system based on hybrid parameter reduction algorithm. *Journal of Telecommunication, Electronic and Computer Engineering*, 9(2-7), 143-148.
- Molodtsov, D. (1999). Soft set theory - First results. *Computers & Mathematics with Applications*, 37(4-5), 19-31. [https://doi.org/10.1016/S0898-1221\(99\)00056-5](https://doi.org/10.1016/S0898-1221(99)00056-5)
- Muhammad, L. J., Al-Shourbaji, I., Haruna, A. A., Mohammed, I. A., Ahmad, A., & Jibrin, M. B. (2021). Machine learning predictive models for coronary artery disease. *SN Computer Science*, 2(5), Article 350. <https://doi.org/10.1007/s42979-021-00731-4>
- Palaniappan, S., Raghavi, V., David, B., & Nisha, S. P. (2021). Prediction of epidemic disease dynamics on the infection risk using machine learning algorithms. *SN Computer Science*, 3(1), Article 47. <https://doi.org/10.1007/s42979-021-00902-3>



- Pawlak, Z. (1982). Rough sets. *International Journal of Computer & Information Sciences*, 11(5), 341-356. <https://doi.org/10.1007/BF01001956>
- Pawlak, Z. (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data (Vol. 9)*. Springer Science & Business Media.
- Rahman, M. M., Islam, M. M., Manik, M. M. H., Islam, M. R., & Al-Rakhami, M. S. (2021). Machine learning approaches for tackling novel Coronavirus (COVID-19) pandemic. *SN Computer Science*, 2(5), Article 384. <https://doi.org/10.1007/s42979-021-00774-7>
- Rose, A. N. M., Herawan, T., & Deris, M. M. (2010). A framework of decision making based on maximal supported sets. In L. Zhang, B. L. Lu & J. Kwok (Eds.), *Advances in Neural Networks - ISNN 2010* (pp. 473-482). Springer. [https://doi.org/10.1007/978-3-642-13278-0\\_61](https://doi.org/10.1007/978-3-642-13278-0_61)
- Sadiq, A. S., Tahir, M. A., Ahmed, A. A., & Alghushami, A. (2020). Normal parameter reduction algorithm in soft set based on hybrid binary particle swarm and biogeography optimizer. *Neural Computing and Applications*, 32(16), 12221-12239. <https://doi.org/10.1007/s00521-019-04423-2>
- Shah, D., Patel, S., & Bharti, S. K. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science*, 1(6), Article 345. <https://doi.org/10.1007/s42979-020-00365-y>
- Sharma, D., Gotlieb, N., Farkouh, M. E., Patel, K., Xu, W., & Bhat, M. (2022). Machine learning approach to classify cardiovascular disease in patients with nonalcoholic fatty liver disease in the UK Biobank Cohort. *Journal of the American Heart Association*, 11(1), Article e022576. <https://doi.org/10.1161/JAHA.121.022576>
- Wang, S. H., Wu, K., Chu, T., Fernandes, S. L., Zhou, Q., Zhang, Y. D., & Sun, J. (2021). SOSPCNN: Structurally optimized stochastic pooling convolutional neural network for tetralogy of fallot recognition. *Wireless Communications and Mobile Computing*, 2021, Article 5792975. <https://doi.org/10.1155/2021/5792975>
- Xie, N. X. (2016). An algorithm on the parameter reduction of soft sets. *Fuzzy Information and Engineering*, 8(2), 127-145. <https://doi.org/10.1016/j.fiae.2016.06.001>
- Yuksel, S., Dizman, T., Yildizdan, G., & Sert, U. (2013). Application of soft sets to diagnose the prostate cancer risk. *Journal of Inequalities and Applications*, 2013(1), Article 229. <https://doi.org/10.1186/1029-242X-2013-229>
- Zhang, Y., Lu, S., Zhou, X., Yang, M., Wu, L., Liu, B., Phillips, P., & Wang, S. (2016). Comparison of machine learning methods for stationary wavelet entropy-based multiple sclerosis detection: Decision tree, k-nearest neighbors, and support vector machine. *Simulation*, 92(9), 861-871. <https://doi.org/10.1177/0037549716666962>

*Review Article*

## Understanding the Degradation of Carbofuran in Agricultural Area: A Review of Fate, Metabolites, and Toxicity

Nurul Syuhada Baharudin<sup>1</sup>, Harlina Ahmad<sup>1\*</sup> and Md Sohrab Hossain<sup>2</sup>

<sup>1</sup>*Environmental Technology Division, School of Industrial Technology, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*

<sup>2</sup>*HICoE-Centre for Biofuel and Biochemical Research, Institute of Self-Sustainable Building, Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS, Seri Iskandar, 32610 UTP, Perak Darul Ridzuan, Malaysia*

### ABSTRACT

Pesticides are commonly applied in agriculture to manage pests, parasites, weeds, and other crop diseases to maximise product output value and minimise or reduce yield losses. Pesticides are widely and indiscriminately used in crop production in most countries. Carbofuran is a highly toxic insecticide commonly used to protect crops in agricultural areas. Exposure to carbofuran can cause harmful effects on both the ecological environment and human health, particularly on non-target species such as birds and aquatic organisms. Carbofuran continues to be used, although it has been banned in some countries. This review paper highlights carbofuran usage, its residue, toxicity, and mechanisms of degradation pathways in water, soil, and food, especially in agriculture. It has been shown that hydrolysis, photolysis, and microbial degradation are the main pathways for carbofuran degradation. Carbofuran phenol is significantly less toxic than all photo-metabolites, while 3-ketocarbofuran and 3-hydroxycarbofuran are practically as harmful as the parent compound. Hence, more study is needed on degradation pathways that may thoroughly minimise the toxicity of this deadly pesticide.

#### ARTICLE INFO

*Article history:*

Received: 04 January 2023

Accepted: 14 June 2023

Published: 21 December 2023

DOI: <https://doi.org/10.47836/pjst.32.1.17>

*E-mail addresses:*

[nurulsyuhadahjbaharudin@gmail.com](mailto:nurulsyuhadahjbaharudin@gmail.com) (Nurul Syuhada Baharudin)

[harlinaa@usm.my](mailto:harlinaa@usm.my) (Harlina Ahmad)

[sohrab.hossain@utp.edu.my](mailto:sohrab.hossain@utp.edu.my) (Md Sohrab Hossain)

\* Corresponding author

*Keywords:* Carbofuran, degradation, fate, metabolic pathways, metabolites, toxicity, usage

### INTRODUCTION

Pesticides are compounds or a mix of chemicals designed to kill, prevent, or control pests (Cossi et al., 2020). According to the terminology used in the agricultural

industry, “synthetic pesticide” refers to any pesticide manufactured and not found naturally. Synthetic pesticides are generally produced through chemical synthesis and are designed to be more potent and persistent than their natural counterparts (Ahmadi, 2019). Organochlorines, organophosphates, carbamates, and pyrethroids are classes of synthetic pesticides. In contrast to synthetic pesticides, pesticides derived from natural sources, such as plants or microbes, are often called “biopesticides” or “organic pesticides”. Biopesticides are derived from living organisms and their by-products, and they typically have low toxicity levels and little environmental impact (Ayilara et al., 2023).

Pesticides are used to control a wide range of pests, such as insects, plant pathogens, weeds, microbes, nematodes (roundworms), molluscs, mammals, birds, and fish, which cause massive losses in plantations and properties, spread infectious diseases, and act as disease vectors. Moreover, applying pesticides against pests and air-borne diseases improves the agricultural sector’s economic output and boosts crop yields (Idayu et al., 2014; Okoli et al., 2017). Despite having significant benefits, there are serious concerns regarding the effects of pesticides, such as their potential toxicity toward humans and other animals (Daliya et al., 2020).

Malaysia is an agriculture-based country with vast areas reserved for plantations and farming. 70% of the cropland is devoted to oil palm and rubber plantations, which occupy more than 3.8 million hectares (Fatah et al., 2017). Given the massive demand for plantation produce in recent years, pesticides have been extensively used. It has been reported that applying pesticides leads to ground and surface water contamination and causes many undesirable effects and problems in tropical countries with heavy rainfall, high humidity, and high temperatures (Farahani et al., 2007; Ismail et al., 2004). In Japan, China, and the Republic of Korea, more than 10 kg/ha of pesticide are applied, while most Asian countries are below 1 kg/ha (Figure 1).

It represents an average annual growth rate of 1.6% in the net value of agricultural pesticides used during the past 15 years. Recently, about two million metric tonnes of pesticides have been used annually: herbicides (50%), insecticides (30%), fungicides (18%), rodenticides (9%), and nematicides (Sharma et al., 2019). Their transport, transfer, and transformation mechanisms determine pesticides’ environmental fate. Herbicides in paddy fields, vegetable plantations in the Cameron Highlands, Pahang, Johor and tobacco-growing areas were studied (Boh & Yeang, 2002; Nashriyah & Azimahtol, 2002; Halimah et al., 2005; Ismail et al., 2015). It was also reported that surface and groundwater pollution due to pesticides had become a growing concern in tropical countries, including Malaysia (Sapari & Ismail, 2012).

Carbofuran is a highly hazardous pesticide that harms people, animals, and the environment. Due to its adverse effects, many countries, including Malaysia, have prohibited its use. However, as an alternative to chemical pesticides, there has been a recent

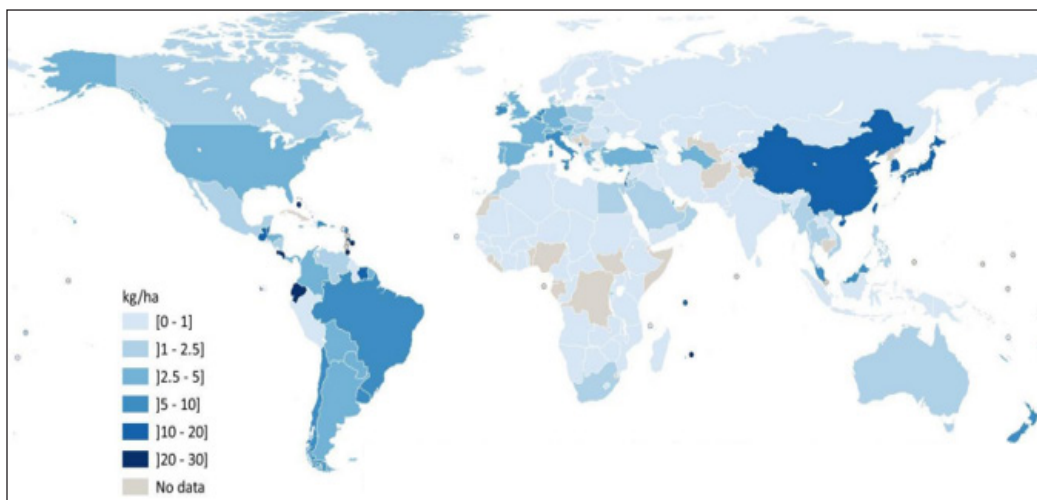


Figure 1. Pesticides use per cropland area, 2018 (FOA, 2020)

interest in using biological agents and natural pesticides. Natural pesticides are derived from natural sources such as plants, minerals, or microorganisms, whereas biological control agents are living organisms that feed on or parasitise pests (Yubak-Dhoj et al., 2021).

Biological control and natural pesticides are effective in pest control while posing the least risk to humans and the environment. They are frequently more specific and targeted than chemical pesticides, which can have unintended consequences for non-target organisms. Biological agents and natural pesticides are becoming more popular as safer and more sustainable alternatives to chemical pesticides such as carbofuran (Torres & Bueno, 2018).

In a study by Mustapha et al. (2019), 3-hydroxycarbofuran and 3-ketocarbofuran as carbofuran by-products are more polar but equally toxic to target and non-target organisms. Other known by-products of carbofuran include carbofuran phenol, 3-hydroxy-7-carbofuran phenol, and 3-keto-7-carbofuran phenol (through phenyl-ring oxidation/reduction and hydroxylation reactions), N-hydroxymethyl carbofuran, and 3-hydroxy-N-hydroxymethylcarbofuran (through methyl hydroxylation reactions). In plant tissue samples, 3-hydroxycarbofuran was higher and lasted longer than the parent compound (Ramasubramanian & Paramasivam, 2018; Ramesh et al., 2015).

The reaction pathways and breakdown of pesticides are crucial to determining the process involved, especially the nature of the reactive intermediates and their final products. In this way, proper management and control of the discharge of pesticides into the environment can be implemented (Fenner et al., 2013; Bose et al., 2021). For instance, water management is vital in rice production since the growth of the paddy field requires a large amount of water supply. Water from the main canal is channelled into each plot and retained for a certain period before being drained from the paddy field. Therefore,

the pesticides used in surface water would significantly affect the water quality and could threaten the lives of aquatic animals in the surrounding plantation.

This review aims to assess the reported concentrations of carbofuran residues and their effects on the environment and human health, as well as to discuss their insights on carbofuran usage, toxicity, and degradation, particularly in agriculture. Thus, we aim to emphasise the comprehensive mechanism of the carbofuran degradation pathway via hydrolysis, photolysis, and microbial degradation into its significant metabolites.

## CARBOFURAN

Carbofuran (2,3-dihydro-2,2-dimethyl benzofuran-7-yl-N-methylcarbamate) is a broad-spectrum insecticide, acaricide, and nematicide. Carbofuran is classified under the general group of carbamate derivatives called furadan, curaterr, carbodan, carbosip, and chinufur, with its chemical structure shown in Figure 2 (Pyne, 2015; Mukaj et al., 2017).

It was during World War II that a lot of new methods were developed for dealing with pests. Many manufactured chemicals were tested for their insecticidal efficacy because of the pressing need to reduce human disease transmission by pests in tropical areas. Figure 3 displays a timeline of carbofuran usage throughout history. Carbamates, a third class of chemically synthesised insecticides, were identified by Swiss scientists in the 1940s. In the mid-1940s, the Geigy Chemical Company aimed to create an insect repellent and began working on carbofuran. A set of tests on carbamate compounds found that these compounds were ineffective insect repellents and poisonous to flies, mosquitoes, aphids, and other smaller pests. Geigy was prompted by this situation to explore the development of dimethyl carbamate insecticides instead of carbamate repellents (Mahindru, 2009).

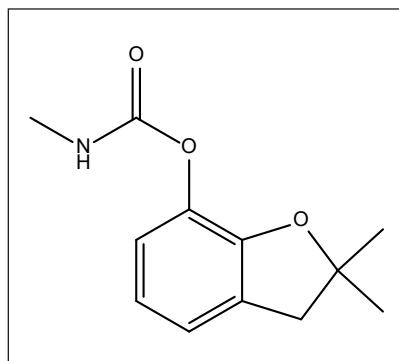


Figure 2. Carbofuran chemical structure

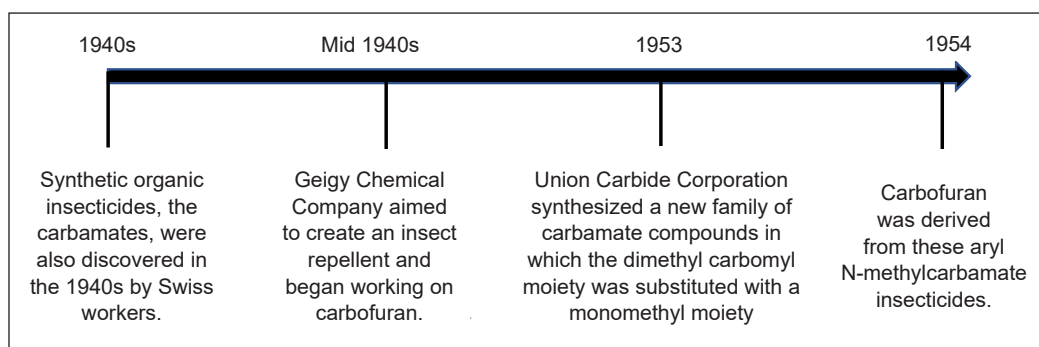


Figure 3. Timeline showing the history of the production of carbofuran

In 1953, Union Carbide Corporation synthesised a new family of carbamate compounds in which the dimethyl carbamoyl moiety was substituted with a monomethyl moiety. Compared to dimethyl carbamic acids, the insecticidal properties of aryl N-methylcarbamate were superior (Park et al., 2013). Carbofuran was derived from these aryl N-methylcarbamate insecticides.

Carbofuran (2,3-dihydro-2,2-dimethyl-7-benzofuranyl-methylcarbamate) is a nematicide generally applied to soil or grassland. It is highly mobile and soluble in water (351 g/mL), posing a danger of environmental water pollution (Table 1) (Ćwieląg-Piasecka et al., 2021). The pesticide has moderate sorption to soil, based on organic matter and clay mineral composition, though it is unclear which soil element has the strongest affinity (Cáceres et al., 2019; Khairatul et al., 2013). Carbofuran degrades faster in an alkaline medium but is relatively stable in neutral or acidic environments. It is hydrophilic at concentrations up to 700 mg/L, while its solubility decreases in organic solvents at less than 30 mg/L (Brasel et al., 2007).

As with carbofuran, it has a low solubility in water but a high solubility in acetone, acetonitrile, benzene, and cyclohexane. It degrades at temperatures above 130°C and stimulates combustion when ignited. The substance is available as a powder, granules, and flowable formulation in a solvent (USEPA, 2006). It has a shorter environmental half-life than chlorinated hydrocarbon insecticides, degrades faster, and is excreted by nontarget organisms. Carbofuran may reach the aquatic environment by runoff from treatment plantation fields, direct application or broadcasting of granular formulations, or drift from sprayable mixtures (Valente-Campos et al., 2019).

## CARBOFURAN USAGE/RESIDUE IN AGRICULTURAL PRODUCTS

Carbofuran is a pesticide that is commonly used to repel pests and nematodes on a wide range of crops, including rice, potatoes, corn, and soybeans, because of its broad-spectrum biological action and comparatively low persistence compared to organochlorine pesticides (Kuswandi et al., 2017; Bedair et al., 2022).

In Asia, Australia, and South America, its use is widespread. Meanwhile, Canada, Kenya, Brazil, the United States, and the European Union are among the countries that

Table 1  
*Carbofuran's physical and chemical characteristics (Otieno et al., 2010a)*

Chemical formula	C <sub>12</sub> H <sub>15</sub> N O <sub>3</sub>
Molecular weight	221.6 g/mol
Physical state	White crystalline solid
Flammability	Not flammable
Melting point	15 to 154°C
Density	1.180 (20 °C)
Vapour pressure	2 × 10 <sup>-5</sup> mmHg (33°C) 1.1 × 10 <sup>-4</sup> mmHg (50°C)
Octanol/water partition coefficient	42.5
Water solubility	351 µg/ml (250°C)
K <sub>oc</sub>	9–36 ml/g



have banned or restricted the use of carbofuran. Canada banned carbofuran-containing pesticides in 2009 when Health Canada's Pest Management Regulatory Agency (PMRA) cancelled all carbofuran registrations (Health Canada, 2020). In 2010, Kenya prohibited carbofuran import, distribution, and sale. Carbofuran's registration in Brazil was suspended in 2011, effectively banning its use there. In 2008, the European Union banned the use of carbofuran in all its member countries. The United States Environmental Protection Agency (EPA) banned all carbofuran uses in the country in 2009, claiming risks to human health and the environment (U.S. EPA, 2009).

Aubergines and other vegetables in Malaysia are often treated with carbofuran (Sim et al., 2019). It is an approved insecticide until May 2023, when it will be banned in agriculture (Pesticides Act 1974, 2021). The Pesticides Board of Malaysia, which oversees controlling pesticides in Malaysia, has said that carbofuran will be banned on April 20, 2021. The ban was put in place because the chemical was harmful to people and the environment, as well as to wildlife and organisms that were not its intended target.

About 1% of the earth's surface is used to grow paddy rice (*Oryza sativa*) because it is grown on every continent except Antarctica. Over half of the world's population uses rice as its primary food, which ranks second to wheat in global cultivation (Kumar et al., 2020). Given that rice is the only food crop cultivated during the rainy season in the tropical areas of Asia, the region has contributed over 90% of the world's rice production. Malaysia's land area used for rice cultivation has remained consistent since the 1980s at roughly 0.7 million ha. Nevertheless, rice productivity increased yearly, from 2.1 t/ha in 1961 to 3.6 t/ha in 2008. Similarly, total rice production has increased yearly since 1985, with an average annual increase of 28,000 metric tons (Ibitoye et al., 2014).

Table 2 summarises the amounts of carbofuran residues in water from the different agricultural industries. Carbofuran is a pesticide that can leave residues in treated products and is subject to maximum residue limits (MRL). However, it is important to note that MRLs for carbofuran may vary by country or region and are subject to periodic updates. Carbofuran levels in Canadian freshwater ranged from 0.03 to 158.5 mg/L, which is high. Carbofuran was discovered once (at 3.0 µg/L) in 678 samples from 1971 to 1986 (Health Canada, 2020). As for the USA, 1 µg/L of carbofuran was detected in streams (Papadakis et al., 2015), and 1 to 30 µg/L was discovered in groundwater (Conway & Pretty, 2013). For example, carbofuran has been heavily used in Kenya, increasing environmental contamination. (Otieno et al., 2011). Several samples of surface water from paddy fields were found to contain carbofuran at concentrations ranging from 0–3.395 µg/L, and those from lakes had carbofuran concentrations ranging from 0.949–1.671 µg/L, respectively (Chowdhury et al., 2012). It has been discovered that concentrations of carbamate insecticides in the environment that are higher than the limit set by the European Community (0.5 g/L) are unusual (Plese et al., 2005; Zhang et al., 2016).



Table 2

*Concentration of carbofuran residues in water from the various matrixes*

Concentration	Matrix	Reference
0.03–158.5 mg/L	Freshwater	Health Canada (2020)
1 µg/L	Streams	Papadakis et al. (2015)
1–30 µg/L	Groundwater	Conway and Pretty (2013)
0.18–14 µg/L	Drinking water	Howard (2017)
1–5 µg/L	Water table aquifers	Howard (2017)
0.1–0.5 µg/L	Rainwater	Howard (2017)
0.005–0.495 mg/L	Rivers	Otieno et al. (2011)
2.301 mg/L	Ponds and dams	Otieno et al. (2011)
0.949–1.671 µg/L	Lake's water	Chowdhury et al. (2012)
0–3.395 µg/L	Paddy water	Chowdhury et al. (2012)
233 µg/L	Paddy water	Clasen et al. (2014)
16–28 µg/L	Runoff water	Larson et al. (2019)
4.3 µg/L	Ground water	Moreira et al. (2015)

The highest values found in the paddy water exceed the values approved by the American Environmental Protection Agency (EPA) (40.0 g/L) and the European Community (<0.10 µg/L). The concentration of paddy water, 0–3.395 µg/L reported by Chowdhury et al. (2012), was below both MRL, but Clasen et al. (2014) showed carbofuran concentration exceeded the MRL with 233 µg/L. The United States Environmental Protection Agency (EPA) has set a Maximum Residual Disinfectant Level (MRDL) of carbofuran in drinking water at 0.04 parts per million (ppm), which is the highest level of a disinfectant allowed in drinking water, while the World Health Organization (WHO) recommends a detection limit of 0.9 µg/litre for carbofuran in water and provides analytical methods to measure its concentration (Edzwald, 2011; World Health Organization, 2020).

Also, Farahani et al. (2007) and Khuntong et al. (2010) measured the amount of carbofuran residue in agricultural fields, rivers, and estuaries worldwide. However, understanding the breakdown of carbofuran and its transformation products is limited solely to the parent compounds. To date, researchers have not considered the potential hazards of its derivative products. Carbofuran pollution in rice-field waterways, mainly contributed through runoff, is plausible due to its widespread use in paddy cultivation and relatively good solubility in water (320 mg/L at 20°C) (Clasen et al., 2014).

Table 3 shows carbofuran concentrations in surface soil. Even though soil residue levels were low, water contamination by runoff and secondary transfer to smaller birds were still feasible. Carbofuran concentrations were relatively higher in the rainy season than in the dry season, indicating that it dissolves quickly and can be detected in the soil soon after application. Moreover, Otieno et al. (2012) reported that pesticide residues, including

Table 3

*Carbofuran concentration residues in soil from various agricultural areas*

Concentration	Soil type	Region	Reference
0.146 mg/kg	Farm soil	Isiolo	Otieno et al. (2011)
0.176 mg/kg	Farm soil	Laikipia	Otieno et al. (2011)
4.9 mg/dm	Sandy soil	Netherland	Leistra and Boesten (2008)
0.01–0.54 µg/kg	Polish soil	Poland	Ukalska-Jaruga et al. (2020)
0.2479 mg/L	Paddy soil	Brazil	de Melo Plese et al. (2005)
2.35 µg/g	Sandy clay soil	N. Sembilan	Farahani et al. (2008)
5.80 µg/g	Clay soil	Perak	Farahani et al. (2008)
148 µg/g	Alluvial soil	Orissa	Ripley and Chau (2020)
202 µg/g	Laterite soil	Kerala	Ripley and Chau (2020)
0.178 ± 0.01 mg/kg	Paddy soil	Bangkok	Khuntong et al. (2010)
0.01–1.08 mg/g	Clay soil	Kenya	Jemutai-Kimosop et al. (2014)

carbofuran were elevated during the monsoon season due to increased agricultural runoff that discharged large amounts of pesticide residues into waterways.

In addition, carbofuran is known to be more resistant to insecticides than other organophosphorus (OP) or carbamate insecticides. Since 0.15–0.33 mg/kg of carbofuran was found in milk products, hens' tissues, and eggs (World Health Organization, 2020), carbofuran has been banned in several countries (Mishra et al., 2020). Plants are more tolerant to carbofuran than animals. Migratory waterflow risks can be decreased by restricting granular formulations with more than 3% active components.

Trevisan et al. (2004) found more carbofuran traces in the bagasse than in the whole fruit, but they decreased quickly during the sampling time. This insecticide was found to stay in the bagasse for at least 28 days, and there was evidence that it tended to build up when treated with multiple sprays and went down more slowly during this time. It shows that carbofurans are degraded quickly into their metabolites. Also, 3-hydroxy-carbofuran was found in a few bagasse samples from the treatment with the double rate (20 g a.i. 100 L<sup>-1</sup> water), but only in very small amounts (0.05–0.09 mg/kg). The summary of concentrations of carbofuran residues in vegetables, fruits, and food from the various agricultural industries is shown in Table 4.

Food safety is a major public concern worldwide, and most governments, private companies, and international organisations have established MRLs for food commodities to control an unacceptable risk of human toxicity. To avoid unacceptable health risks, pesticide residues in food commodities such as fruits and vegetables must be below MRLs. In developing countries, more cases of pesticide poisoning were reported than in developed countries (Azam et al., 2020; Latif et al., 2011). Carbamate concentrations significantly exceeded MRLs in many samples; carbofuran concentrations exceeded MRLs in 80 samples, respectively. Carbamates, like OPs, are acetylcholinesterase enzyme system

Table 4

*Concentration of carbofuran residues in vegetables, fruits, and food from various agricultural products*

Concentration	Vegetable, fruit, and food	Reference
0.15–0.33 mg/kg	Milk product & eggs	Health Canada (2020)
48.3±5.3 µg/g	Long bean	Lan et al. (2019)
73.7±4.9 µg/g	Chives	Lan et al. (2019)
22.6±3.2 µg/g	Mango	Lan et al. (2019)
51.9±6.1 µg/g	Watermelon	Lan et al. (2019)
0.001 mg/kg	Apples	EFSA (2014)
0.001 mg/kg	Potato	EFSA (2014)
0.002 mg/kg	Lettuce	EFSA (2014)
0.01 mg/kg	Citrus fruit	EFSA (2014)
8–17 µg/kg	Tobacco	USEPA (2006)
19–36 µg/kg	Corn	USEPA (2006)
14–18 µg/kg	Cotton	USEPA (2006)
0.075 mg/kg	Sugarcane	Kabir et al. (2007)
0.86 ± 0.10 mg/kg	Orange (bagasse)	Trevisan et al. (2004)
0.07 ± 0.02 mg/kg	Orange (Whole fruit)	Trevisan et al. (2004)

inhibitors, but because of their high mammalian toxicity and the even greater toxicity of their intermediates, carbamates are a potential environmental concern (Liu et al., 2019; Syed et al., 2014).

## TOXICITY OF CARBOFURAN

The acute toxicity level of carbofuran differs between organisms. The toxicity of carbofuran to mammals is higher through the oral and inhalation routes than the dermal route. Furthermore, carbofuran effects in humans are associated with reproductive disorders, endocrine-disrupting activity, and cytotoxic and genotoxic abnormalities (Mishra et al., 2020). In past years, carbofuran has emerged as a global contaminant that can harm health and the environment (Figure 4). Massive carbofuran use can significantly affect non-target species via many pathways.

The World Health Organization (2020) classified carbofuran as a highly hazardous class 1b pesticide (rat oral [lethal dose 50%] LD<sub>50</sub> ranging from 5 to 50 mg/kg). Despite the decision by WHO, the classification does not take into consideration the fact that the technical product provided a safer profile through the dermal route (LD<sub>50</sub> > 1,000 mg/kg and LD<sub>50</sub> > 2,000 mg/kg in rats and rabbits, respectively), which should have classified carbofuran as class II or III (moderately or slightly hazardous) under WHO guidelines (Numan et al., 2018). Although such classifications help standardise the comparison of technical products, the information becomes less relevant under practical field applications as pesticide users are in direct contact with the formulations and not the technical product.

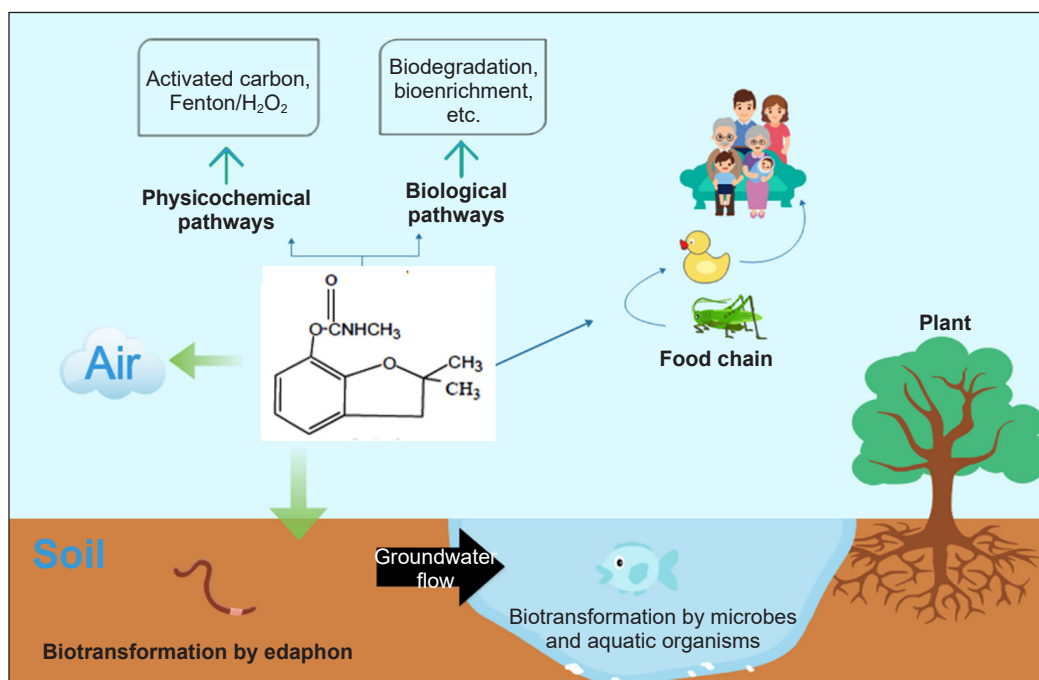


Figure 4. Contamination and degradation of carbofuran from the natural environment (Zhang et al., 2021)

Hence, the information regarding the toxicity of the end-use formulation only serves as a guideline to establish relevant warnings for users.

Carbofuran has caused the deaths of irregular fish, animals, and invertebrates recommended application rates and compositions. In short-term laboratory studies, sensitive aquatic biota died at 200 ppb carbofuran (in water), 238 ppb (acute oral), 190,000 ppb for birds, 2,000 ppb (acute oral) and 100,000 ppb (dietary) for mammals. Carbofuran is toxic and often life-threatening to fish at > 15 ppb and aquatic invertebrates at > 2.5 ppb. A level of 10 to 50 ppb in food and 1,000 ppb in drinking water is toxic to birds and mammals (Eisler, 1985; Kanedi, 2017). “Safe” carbofuran levels in beef are 50 ppb. The current allowable carbofuran aerosol levels of 0.05 ppb (50.0 g/m<sup>3</sup>) appear sufficient to preserve animals, but data suggests they should never exceed 2 ppb (Eisler, 1985; Otieno et al., 2010b).

According to the National Centre for Biotechnology Information (2023), carbofuran has one of the highest acute toxicities to humans of any insecticide widely used on field crops, and it is stated that 1 ml (1/4 teaspoon) of carbofuran can be fatal to humans. This statement is supported by the National Oceanic and Atmospheric Administration’s Chemical Emergency Response Tool, which indicates that a lethal oral dose for humans is 5 to 50 mg/kg or 7 drops to 1 teaspoon for a 150 lb. person (Alexeeff, 2000). It is important to note that carbofuran can cause serious effects on the maternal-placental foetal unit, and its toxicity can be potentiated by simultaneous exposure to other cholinesterase inhibitors.

The acceptable daily intake (ADI) estimates the amount of a substance that can be ingested daily over a lifetime without posing an appreciable health risk. The Food and Agriculture Organization (FAO) and the World Health Organization (WHO) have derived the ADI of carbofuran as 1.0 mg/kg bw per day (Ostergaard & Knudsen, 1998). The National Institute for Occupational Safety and Health (NIOSH) and American Conference of Governmental Industrial Hygienists (ACGIH) stated that the recommended airborne exposure limit for carbofuran is 0.1 mg/m<sup>3</sup> averaged over an 8 or 10-hour work shift (Satar et al., 2005).

In rats, mice, and lactating cows, carbofuran metabolism and excretion have been the subject of significant research. Absorption, degradation, and elimination from the oral dosage occur very quickly. Mixed-function oxidases are mammals' main enzymes responsible for metabolism, as 3-hydroxycarbofuran and 3-ketocarbofuran are the most prevalent carbamate metabolites. The toxicological effects of the carbofuran range on various matrixes are shown in Table 5.

A recent study (Mishra et al., 2020) found that grain farmers exposed to a low dose of carbofuran through the skin and the air for four days had carbofuran in their urine. Even though none of the farmers showed outward signs of being poisoned, the blood cholinesterase was significantly slowed down. Moreover, another study thoroughly discussed the metabolic

Table 5  
*Toxicological effects of carbofuran range on various matrixes*

Study Sample	Concentration/Median Lethal Concentration (LC50)/Median Lethal Dose (LD50)	Reference
Cattle	3–19 mg/kg	Choez and Evaristo (2018)
Chickens	2.5 mg/kg	Lehel et al. (2010)
Hamster	5–100 mg/mL	Soloneski et al. (2008)
Rat	75 mg/kg	Gammon et al. (2012)
Rat	1 mg/kg mg/kg	Islam et al. (2008)
Rat	$1.2 \times 10^{-8}$ mg/kg	Ferguson et al. (1984)
Human	$4.6 \times 10^{-8}$ mg/kg	Soreq and Zakut (1989)
Fly	$2.5 \times 10^{-8}$ µg/kg	Metcalf et al. (1968)
Birds	238 µg/kg	Eisler (1985)
Mammals	2000 µg/kg	Eisler (1985)
Catfish	23 µg/kg	Liong (1988)
Fish fillets	10.50 µg/kg	Tejada (1995)
Cat	2.5–3.5 mg/kg	Daul et al. (2012)
Guinea pig	9.2 mg/kg	Daul et al. (2012)
Dog	15.38 mg/k	Daul et al. (2012)
Rabbit	7.5 mg/kg	Daul et al. (2012)
Honeybees	0.16 µg/bee	USEPA (2006)
Japanese quail	1.7–1.9 mg/kg	Daul et al. (2012)
Pheasant	4.15 mg/kg	Smith (1992)

activity of carbofuran since the metabolites pose a substantial effect on the general toxicity of carbofuran (Ruíz-Hidalgo et al., 2016). Due to frequent reports of bird poisoning, carbofuran has been prohibited in several countries, particularly the United States, Canada, and the European Union (Sparling, 2016). The Royal Society for the Protection of Birds reported 316 incidents of bird poisoning due to carbofuran in the United Kingdom between 2002 and 2011 (Vincent, 2014). In addition, 3-ketocarbofuran and 3-hydroxycarbofuran, secondary metabolites of carbofuran, were also identified as poisonous. It was recorded that 187 vultures and hyenas died in 2004 after they scavenged bird carcasses containing carbofuran residue and its by-products (Plaza et al., 2019). Regardless of its severe effects, carbofuran is still broadly applied in Asia, South America, Australia, and Malaysia, which registered carbofuran as a legal pesticide (Sim et al., 2019).

### CARBOFURAN DEGRADATION PATHWAYS IN THE ENVIRONMENT

Pesticide degradation breaks down pesticides into less toxic derivatives or their original elements (Sidhu et al., 2019). The degradation occurs through various processes or dissipation mechanisms once the pesticide enters the water and soil. Pesticides from the soil usually migrate into surface water, such as rivers, lakes, streams, and estuaries. Once in the water, it degrades according to the pseudo-first-order rate equation (Villaverde et al., 2018). The half-life of carbofuran in the environment is 30–120 days, based on where it is, the temperature, the pH of the soil or water, and medium humidity. The most common ways carbofuran and its by-products degrade are through hydrolysis, sunlight exposure, and biodegradation (Morales et al., 2012).

Various factors have been identified to influence the dispersion and degradation of pesticides, including the country's climate and the properties of the water and soil. In addition, the physical-chemical properties of the pesticides also determine how they disperse and degrade in the environment (Kumar et al., 2015). Because of its widespread use in agriculture, Carbofuran can react with water to form C–N bonds and break C–X bonds (X representing –OH group) (Jaiswal et al., 2017). Biodegradation regulates the fate and transformation of carbofuran use in nature. Research into the degradation of pesticides is widespread. Generally, carbofuran is degraded by hydrolysis, photolysis, microbial degradation and oxidation, as described in Figure 5 (Prosen, 2012; Elbashir & Aboul-Enein, 2015).

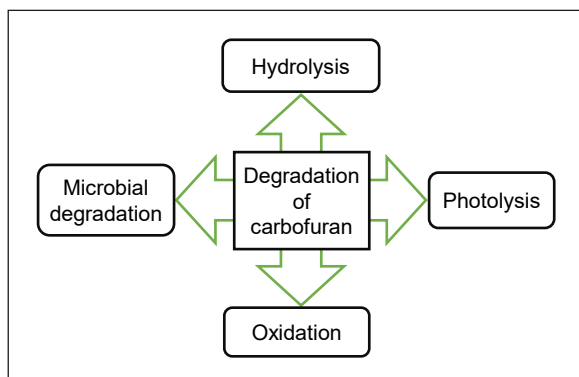


Figure 5. Degradation pathways for carbofuran

## Hydrolysis Degradation

Most studies have concluded that hydrolysis and photolysis are the two most important degradation pathways for pollutants. Understanding the fate of a pesticide in water requires knowledge of the transformation mechanisms. Pesticide hydrolysis is a secondary and nucleophilic substitution reaction and fits first-order reaction kinetics (Singh et al., 2017). Morrica et al. (2001) state that temperature and pH are common hydrolysis factors. The degradation pathway takes place in both water and sediment through base-catalysed hydrolysis, which generates carbofuran phenol as well as other degradation products, such as 3-hydroxy-7-phenolcarbofuran and N-methylcarbamic acid via hydroxylation of the benzofuranyl moiety (Evert, 2002; Howard, 2017; Nafeesa et al., 2017; Sandín-España & Sevilla-Morán, 2012). The carbofuran (2,3-dihydro-2,2-dimethyl-7-benzofuranol methylcarbamate) molecule degrades when it reacts with water to form a new C–N bond and break a C–X bond in the parent molecule. The net reaction is a direct removal of X by OH (Plese et al., 2005; Ripley & Chau, 2020).

Carbofuran is generally stable in water at acidic and neutral pH, but the hydrolysis rate increases rapidly with increasing pH. The chemical degradation controls the environmental persistence of carbofuran, as shown in Figure 6. Hydrolysis is an effective alternative degradation route, with average lifetimes in waters ranging from 2 days at pH = 9.5 to 1,700 days at pH = 5.2 (Marine, 1999). Besides the pH, the temperature directly influences the degradation rate in hydrolysis.

It was reported that the hydrolysis half-life,  $t^{1/2}$ , of carbofuran was 35 days at pH 7.0 and 350 days at pH 6.0 (Gammon et al., 2012). Acidic hydrolysis occurs under acidic conditions, which degrades slower than in alkaline conditions. For carbofuran, hydrolysis involves the cleavage of the carbamate bond in the compound, resulting in the formation of 3-hydroxy-carbofuran, which is less toxic than the original compound. Another previous study reported that carbofuran spikes recovered between 80% and 95% after 1, 3, and 6 hours at 25 °C and pH 8. On the contrary, the original amount of carbofuran recovered at pH 10 after 1, 3, and 6 hours was only 65, 35, and 10%, respectively (Bailey et al., 1996;

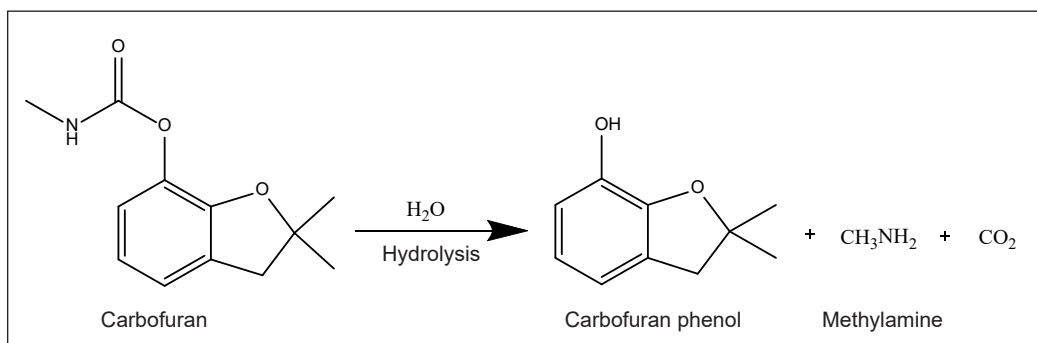


Figure 6. Hydrolysis mechanism of carbofuran (Filik and Çekiç, 2011)



Cycoń et al., 2017). The measured half-lives of carbofuran under acidic and alkaline conditions were 320 days and 150 days, respectively (Pinakini & Kumar, 2006).

Seiber and colleagues demonstrated that at pH 10, carbofuran hydrolysed 700 times faster than at pH 7, with half-lives of 1.2 and 864 hours, respectively. These results were collected from granule-treated paddy water. In addition, paddy water was hydrolysed faster than deionised water. At pH 7, paddy water half-lives were 240 and DI water 864 hours. At pH 8.7, paddy water and deionised water each had half-lives of 13.9 hours. Based on the results, the carbofuran half-life in paddy water (laboratory) at pH 8 was 40 hours, which was similar to the 57-hour half-life recorded from field samples at the same pH level, given slight variations in several factors, such as the source of light, pH level, the surrounding temperature, and the rate of microbial degradation (Howard, 2017; Mackay et al., 2006; Seiber et al., 1978).

Moreover, a higher rate of carbofuran hydrolysis was reported in anaerobic conditions. The hydrolysis rate of carbofuran under aerobic environments was compared to that in stable and undisturbed flooded conditions. Nevertheless, the degradation products, consisting of carbofuran phenol and 3-hydroxycarbofuran, were stable under anaerobic conditions, while further degradation was only observed under aerobic conditions. So, it was thought that the rapid hydrolysis of carbofuran was caused by cycles of flooding and drying that led to alternating anaerobic conditions (Venkateswarlu & Sethunathan, 1978; Ramakrishnan et al., 2019; Mohamed et al., 2021).

Hydrolysis was considered the most important degradation pathway, with a half-life ranging from 0.2 days at pH 9.5 to 1,700 days at pH 5.2. The rate of carbofuran hydrolysis in aqueous solutions accelerates significantly as the pH of the solution increases. The hydrolysis half-lives of carbofuran in water at 25°C is 690, 8.2, and 1.0 weeks at pH values of 6.0, 7.0, and 8.0, respectively. Hydrolysis reactions involve adding water to a large molecule to break it into multiple smaller molecules (Bailey et al., 1996).

The degradation pathways of carbofuran in water systems are shown in Figure 7. Hydrolysis was reported as the main mechanism leading to degradation in alkaline soils. On the other hand, acidic and neutral soils were dominated by microbial and chemical mechanisms, which led to slower degradation. Carbofuran phenol, 3-hydroxycarbofuran, and 3-ketocarbofuran are among the transformation products of carbofuran degradation that can be detected in soil (Acharid et al., 2006; Achik & Schiavon, 1989; Katagi, 2016).

One factor that reduces the effectiveness of insecticides in controlling the infestation of the brown planthopper *Nilaparvata lugens*, a significant pest of rice plants, is the rapid chemical hydrolysis of carbofuran in paddy water (Hayasaka et al., 2012). It was revealed that carbofuran phenol production from carbofuran degradation took only 5 days under alkaline paddy water conditions.

Furthermore, the effect of paddy water pH on carbofuran persistence was demonstrated in another study in which less than 5% of carbofuran remained after 5 days of application

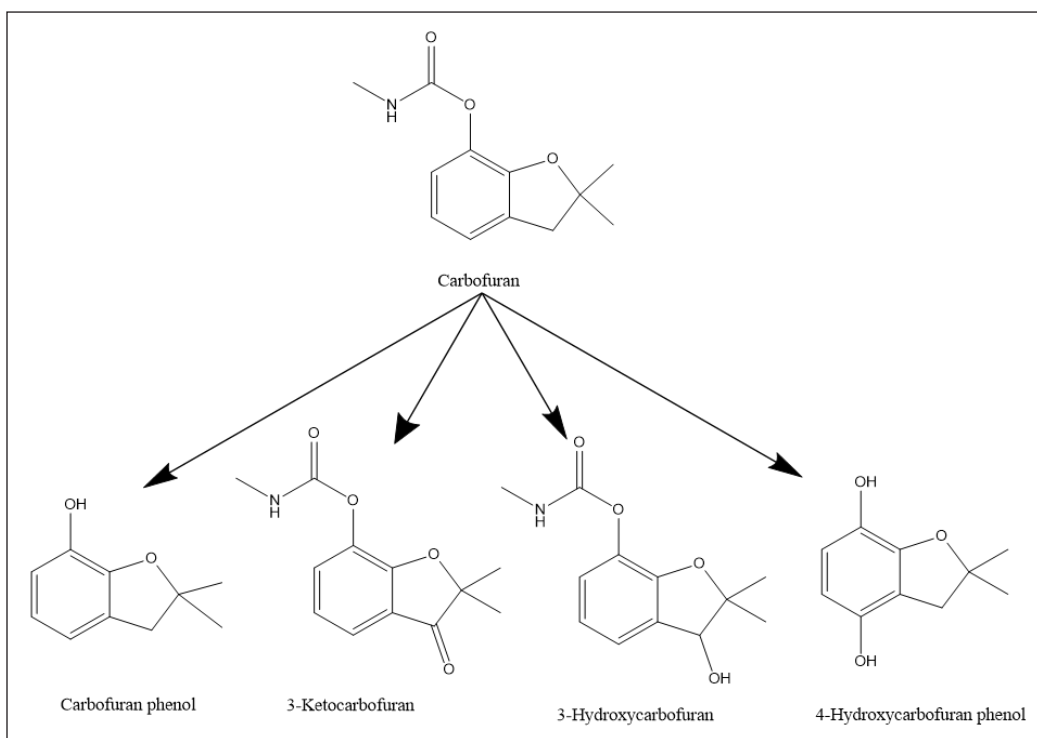


Figure 7. Environmental fate of carbofuran in water (Evert, 2002)

at pH 8. In contrast, the remaining amount of intact carbofuran after 10 days was 70% and 14% at pH 6 or lower and pH 7.1, respectively (Siddaramappa & Seiber, 1979; Roger & Bhuiyan, 2012; Xu et al., 2014). Other studies have also shown that hydrolysis and photochemical degradation are the major degradation pathways for carbofuran in water (Farahani et al., 2012; Tien et al., 2017). The biodegradation and pH of the water were among the main factors influencing the degradation of carbofuran.

### Photolysis Degradation

Photolysis of pesticides leads to chemical changes in pesticides by absorbing light radiation directly and indirectly. In direct photolysis, the pesticide molecules absorb the photons directly, whereas in indirect photolysis, a carrier and other photochemical processes (photosensitizers) are involved (Fang et al., 2022). Photolysis and hydrolysis produce more polar and stable pesticides than the parent compounds. Some pesticides persist for several months, polluting the environment (Tien et al., 2017; Sim et al., 2019). There is, thus, a need for research into the by-products of carbofuran.

A less significant degradation pathway of carbofuran is through the photolysis reaction. Carbofuran is not likely to accumulate in sunlight-exposed water and does not bioaccumulate (Baumart et al., 2011). The metabolic product of photolysis includes

2,3-dihydro-2,2 dimethyl benzofuran-4,7-diol and 2,3-dihydro-3 keto-2,2 dimethyl benzofuran-7-yl carbamate (3-ketocarbofuran) (Mahalakshmi et al., 2007; Remucal, 2014). A comparison was made between the rate of carbofuran recovery in deionized water exposed to sunlight and laboratory light for 96 hours. According to the results, the average recovery of carbofuran was 75.6% for the sunlight-exposed samples and 93.3% for the artificial light-exposed samples, respectively. The large difference suggested the significant photolytic aqueous decomposition of carbofuran (de Melo Plese et al., 2005; Gunther & Gunther, 2013).

Carbofuran photolyzes directly, producing photoproducts. At the same pH level, photolysis occurs more rapidly in pesticides in seawater, freshwater, and humic acid solutions than in distilled water. Meanwhile, although no control was used, carbofuran in freshwater photolyzes faster than seawater. In one case, adding humic acids decreased carbofuran’s direct photolysis rate (Raha & Das, 1990; Mahalakshmi et al., 2007). Figure 8 shows the metabolites’ chemical structure. The transformation of carbofuran generated

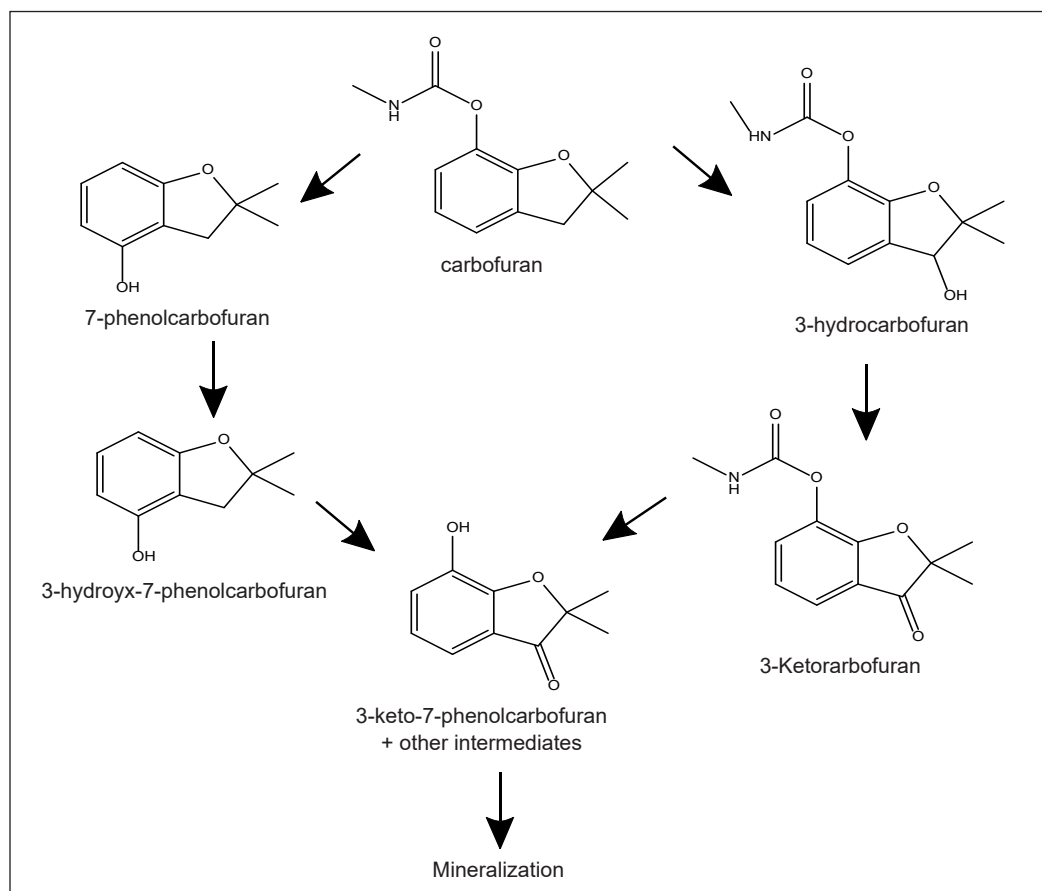


Figure 8. The chemical structure for the detected by-products of carbofuran (Mishra et al., 2020)

7-phenolcarbofuran, 3-hydroxycarbofuran, and 3-ketocarbofuran in all cases, while ZnO formed 3-keto-7-phenolcarbofuran.

Meanwhile, the vapour pressure and Henry's Law constant of carbofuran in the air are low, indicating that carbofuran has a low potential to volatilise from moist soils or water (Cid et al., 2011). Once carbofuran vaporises into the air, it is subjected to vapour-phase photooxidation by reacting with hydroxyl radicals. The half-life of this reaction at standard atmospheric conditions was estimated to be 4.6 hours (Howard, 2017). However, the oxidation and volatilisation of carbofuran in water were considered insignificant dissipation pathways (Wijesinghe et al., 2011).

As demonstrated in Table 6, the carbofuran half-lives in the Laysan and Ottawa sands were comparable (40–41 days), according to Campbell et al. (2004). In comparison, Seiber et al. (1978) found that the half-lives of carbofuran in DI water and paddy water (pH 7) were only slightly different when both were exposed to ultraviolet (UV) light and darkness in the lab. Bachman and Patterson (1999) suggested that the increase in the amount of dissolved organic matter (DOM) inversely reduced the photolysis rate of carbofuran. The assumption was made that the level of inhibition is proportional to the binding capacity of a specific DOM sample. Therefore, the binding of DOM reduced the rate of photolysis and slightly improved the persistence of carbofuran (Bolan et al., 2011).

It is interesting to highlight that the rate constant for the degradation of carbofuran in sunlight-exposed seawater was higher than in sunlight-exposed fresh water. Consequently, the half-life of carbofuran in sunlight-exposed fresh water was longer than that in sunlight-exposed seawater, in line with a similar trend observed by others (Farahani et al., 2012). Campbell et al. (2004) observed that the half-life of carbofuran in sunlight-exposed seawater (7.5 hours) was approximately 6 times faster than in sunlight-exposed DI water (41.6 hours), with the same rate of degradation in seawater (0.1 hours) as in DI water (3.1 hours) under a light intensity of 300 nm. This difference (6–31 folds) and the differences

Table 6  
*Carbofuran half-lives and photodegradation constants, k*

Matrix	Light source	-k	Half-life
Ottawa sand	Dark	0.0151/day	39.7 day
Laysan sand	Dark	0.0182/day	40.9 day
Distilled deionised water	300 nm	0.1941/h	3.1 h
Seawater	300 nm	0.328/h	0.1 h
Distilled deionised water	Sunlight	0.0162/h	41.6 h
Seawater	Sunlight	0.1284/h	7.5 h
Distilled deionised water	Ultraviolet (UV) light	$0.105 \times 10^{-2}$	27.5 day
Paddy water	Ultraviolet (UV) light	0.04	7.2 day
Distilled deionised water	Dark	0.092	31.4 day
Paddy water	Dark	0.41	9.3 day

observed in the present research were presumably due to differences in water composition (Campbell et al., 2004; Remucal, 2014; Mishra et al., 2020; Maksuk, 2021). Carbofuran's photodegradation may be aided by the existence of organic compounds, ions, and chemical sensitizers (Wang et al., 1998; Huang & Mabury, 2000; Seo et al., 2012; Wols & Hofman-Caris, 2012).

The degradation rate of pesticides also varies between countries with tropical and temperate climates. Among the main climate elements that affect the rate of pesticide degradation are the mean temperature in the region, the mean amount of rainfall, the average amount of sunlight emitted, and the existence of microbes (Daam et al., 2008). Tropical countries generally experience more rainfall than temperate-climate countries. According to past studies, the pesticide contamination of soil and water can be influenced by rainfalls through the dilution of the concentration of pesticides in the environment or by leading to pesticide runoffs that could flow into the watersheds and irrigation channels (Abdullah et al., 1997; Daam & Van den Brink, 2010).

### **Microbial Degradation**

Microbial degradation mainly occurs in water and soil when microorganisms are present. Microbial degradation includes complex biological processes. Therefore, it requires ideal conditions such as pH, temperature, and nutrient availability (Grandclément et al., 2017). The pH level and microorganism types in the water determine the rate of microbial degradation in a flooded rice field (Hayasaka et al., 2012). Depending on the metabolic pathways of microbes, microbial degradation can occur under either aerobic or anaerobic conditions, depending on oxygen availability.

Aerobic metabolism is the process of converting organic compounds into energy using oxygen. This reaction occurs in the presence of oxygen, and the end products are typically carbon dioxide and water. Aerobic microorganisms are commonly used to degrade organic pollutants in the environment in bioremediation processes. These microorganisms use enzymes such as monooxygenases and dioxygenases to oxidise and degrade organic compounds (Diale et al., 2022).

Breaking down organic compounds without using oxygen is known as anaerobic metabolism. This process occurs without oxygen and can be carried out by microorganisms such as bacteria and fungi. In terms of energy production, anaerobic metabolism is less efficient than aerobic metabolism, and the end products vary depending on the metabolic pathway. Some anaerobic bacteria, for example, use sulphate or nitrate as a final electron acceptor, producing hydrogen sulphide or nitrogen gas.

Fungi are also essential in microbial degradation, particularly in the breakdown of lignin, a complex polymer found in plant cell walls. Fungi use enzymes like lignin peroxidase and manganese peroxidase to break down lignin into simpler compounds that

other microorganisms in the environment can use. The fungal metabolism of carbofuran is a new area of study, and it has been shown that different fungal strains from the genera *Trichoderma*, *Pichia*, *Trametes*, *Aspergillus*, *Aschochyta*, *Xylaria*, *Acremonium*, *Gliocladium*, and *Mucor* can degrade carbamate pesticides (Mustapha et al., 2019).

Carbofuran was metabolised by *Mucor ramannian* to produce carbofuran phenol, which was then metabolised to yield 2-hydroxy-3-(3-methylpropan-2-ol) phenol, also known as 7a-(hydroxymethyl)-2,2-dimethylhexahydro-6H-furo[2,3-b]pyran-6-one and 3-hydroxy carbofuran-7-phenol (Sun et al., 2022). Among the pyrethroid pesticides, *Trametes versicolor* has been found to degrade carbofuran, imiprothrin, and cypermethrin. In the spent media analysis, 3-hydroxycarbofuran was the only intermediate of carbofuran degradation. It has been reported that the non-specific *monooxygenase cytochrome-P450* plays a significant role in carbofuran metabolism in this strain. (Mir-Tutusaus et al., 2014).

Pesticides degraded by microorganisms serve as a food supply for growth and energy. Based on recent studies, microbes are responsible for the breakdown of natural organic materials and the degradation of numerous xenobiotic chemicals such as petroleum hydrocarbons, pesticides, and organic solvents (Ariffin & Rahman, 2020). Strachan et al. (2018) reported that microorganisms used s-triazine herbicide as an energy source during their degradation. The conversion of contaminants into less harmful substances via microbial degradation is likely in the case of mineralisation, while in certain situations, mostly under anaerobic transformation, toxic products can also be produced (Onunga et al., 2015).

Bacteria capable of efficiently degrading carbofuran have been discovered in soil, sludge, and water contaminated with carbofuran (Mishra et al., 2020). *Pseudomonas*, *Flavobacterium*, *Achromobacterium* sp., *Sphingomonas* sp., *Arthrobacter* sp., *Enterobacter* sp., *Burkholderia* sp. PLC3, *Cupriavidus* sp. ISTL7 and *Bacillus* sp. were isolated and characterised to determine their mechanisms for degrading pesticides. *Sphingomonads* (*Sphingomonas*, *Novosphingobium*, *Sphingopyxis*, and *Sphingobium*) may be the ideal carbofuran-degrading bacteria. These bacteria degrade carbofuran and carbofuran-phenol better than others (Nguyen et al., 2015; Yan et al., 2007). So far, numerous bacterial strains from the genera *Stenotrophomonas*, *Achromobacter*, *Flavobacterium*, *Pseudomonas*, *Sphingomonas*, *Novosphingobium*, *Paracoccus*, *Aminobacter*, and *Cupriavidus* can degrade carbamate insecticides (Gupta et al., 2019; Jiang, Zhang, et al., 2020).

Microbial degradation can benefit the environment by using microorganisms to break down and transform pollutants into less harmful substances. Microbes play key roles in biodegradation, which recycles biologically essential elements within the earth's biogeochemical cycles (Jørgensen, 2008). Microbial degradation is a possible and environmentally friendly way to degrade carbofuran. However, most prior studies on carbofuran degradation have focused on bacterial degradation, and the degradation rates remain slow (Mishra et al., 2020).

In general, the use of microbial degradation for environmental remediation can have several advantages over other methods, such as physical or chemical treatments. For example, microbial degradation can be a cost-effective and environmentally friendly approach, as it can occur naturally and may not require additional chemicals or energy (Liu et al., 2021). Continuous exposure to pollutants and synergy with other strains make microbial consortiums more suitable for environmental remediation. Microbial consortiums or mixed cultures degrade carbofuran residues on a large scale. Tondon et al. (2018) discovered four carbofuran-degrading bacterial strains from contaminated soil samples: *Arthrobacter globiformis*, *Bhargavaea indica*, *Bacillus beijingensis*, and *Streptomyces* sp. Table 7 shows an overview of the microbial species studied for carbofuran biodegradation. Several strains of bacteria and fungi from different sources have been studied.

Microbial degradation completely mineralizes carbofuran into non-toxic substances and CO<sub>2</sub>. As a result, microbial carbofuran degradation is an effective and practical method. Microbes mostly use oxidative and hydrolytic mechanisms to eliminate carbofuran. Figure 9 illustrates carbofuran's comprehensive microbial degradation pathway and describes carbofuran's microbial degradation into its significant metabolites: carbofuran phenol, methylamine, and CO<sub>2</sub>. Carbofuran and carbofuran phenol are formed in *Sphingomonas* sp, but the bacterium cannot grow on them. The hydroxylase-reductase pair CfdCX and *Novosphingobium* sp conduct carbofuran phenol production. (Malhotra et al., 2021).

Table 7  
Microbial degradation of carbofuran in soil by different microbes

Microbial species, Strain or Community	Sample sources	Reference
<i>Sphingomonas</i> sp.	Soil	Park et al. (2006)
<i>Mucor ramannianus</i>	Soil	Seo et al. (2007)
<i>Novosphingobium</i> sp. FND-3	Sludge	Yan et al. (2007)
<i>Paracoccus</i> sp. YM3	Sludge	Peng et al. (2008)
<i>Enterobacter</i> sp.	Soil	Mohanta et al. (2012)
<i>Pseudomonas</i> and <i>Alcaligenes</i>	Soil	Omolo et al. (2012)
<i>Burkholderia</i> sp. PLC3	Soil	Plangklang and Reungsang (2012)
<i>Novosphingobium</i> sp	Soil	Nguyen et al. (2014)
<i>Bacillus</i> sp.	Soil	Onunga et al. (2015)
<i>Consortia</i>	Soil	Tien et al. (2017)
<i>Pseudomonas</i> sp.	Soil	Devi et al. (2017)
<i>Aspergillus</i> sp.	Soil	Devi et al. (2017)
<i>Enterobacter cloacae</i> strain TA7	Soil	Fareed et al. (2017)
<i>Arthobater globiformes</i>	Soil	Tondon et al. (2018)
<i>Streptomyces</i> sp.	Soil	Tondon et al. (2018)
<i>Cupriavidus</i> sp. ISTL7	Landfills	Gupta et al. (2019)
<i>Sphingbium</i> sp. CFD-1	Sludge	Jiang et al. (2020)



Carbofuran phenol is further degraded into novel metabolites by carbofuran-degrading microorganisms and their hydrolase enzymes. In *Novosphingobium* sp. In KN65.2, the initial step involves hydrolysis of the carbamate linkage to form carbofuran phenol, methylamine, and CO<sub>2</sub> (Duc, 2022; Yan et al., 2018).

The hydrolytic pathway is more uncomplicated and produces fewer toxic by-products than the oxidative pathway (Kim et al., 2004; Nguyen et al., 2014). Hydrolase enzymes break methyl and N-methyl carbamic acid's amide and ester bonds to yield carbofuran 7-phenol, CO<sub>2</sub>, and methylamine. The inducible oxidative enzyme hydroxylase transformed carbofuran to 4-hydroxycarbofuran. (Malhotra et al., 2021).

According to Garcia-Saavedra et al. (2018), a bacterial consortium degraded 34.99% of carbofuran in soil samples. Advances in biochemical, molecular, and biotechnological techniques have increased the potential of carbofuran-degrading microorganisms. Catabolic genes, enzymes, plasmids, mobile elements, and transposons all play essential roles in catalysing novel biochemical pathways during microbial degradation (Mishra et al., 2020). However, dehydrogenase and alkaline phosphatase activities in fertilised soil were greatly decreased by increasing carbofuran dosage (0.02–1.0 kg/ha to 5.0 kg/ha) compared to control soils (Singh et al., 2012).

The degradation process involves microbes (biodegradation), chemical reactions (chemical degradation), or light (photodegradation) (Tiryaki & Temur, 2010). The process relies on the pesticide's chemical composition and environmental conditions. Degradation of herbicides by soil bacteria and fungi can take hours to years. The degradation rate by 11 microorganisms increased under warm temperatures, ideal soil pH, appropriate oxygen, soil moisture, and good soil fertility and decreased under less favourable conditions (Meftaul et al., 2020). Pesticides degrade when they react with soil chemicals like water or oxygen. As soil pH becomes more alkaline or acidic, chemical degradation increases because microbial activity decreases (Tiryaki, 2017). In aerobic conditions, microorganisms convert organic contaminants to CO<sub>2</sub> and H<sub>2</sub>O. On the contrary, anaerobic metabolic reactions occur without molecular oxygen (Furmanczyk et al., 2018; Kaida et al., 2018; Chaudhary & Kim, 2019).

### **Oxidation and Volatilization**

Carbofuran is a commonly used carbamate derivative pesticide and a priority pollutant. Several studies have been conducted on the degradation of carbofuran by oxidation and volatilization processes. Considering that biological treatment methods involve lengthy residence times and thermal treatment can generate secondary pollutants, oxidative reactions are favoured for the rapid degradation of contaminants (Zheng et al., 2022). Therefore, a variety of effective treatment techniques, such as ultrasonic process, UV + Fe(III), UV/TiO<sub>2</sub> and UV/O<sub>3</sub>, have been applied for carbofuran removal from aqueous samples (S. Lu et al., 2021; Oliveira et al., 2006).

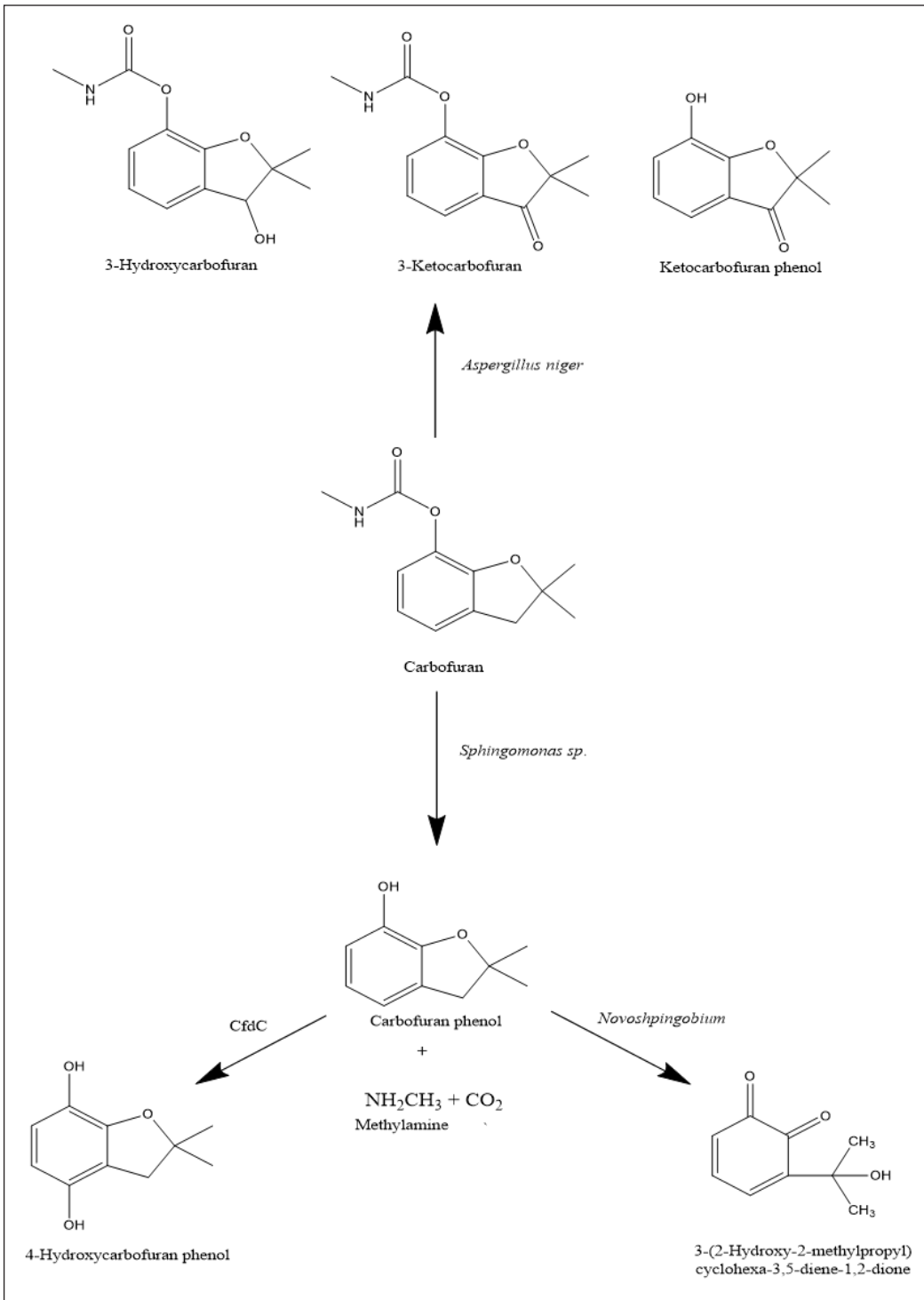


Figure 9. The microbial metabolic mechanisms for the carbofuran degradation in *Aspergillus niger*, *Sphingomonas sp.*, and *Novosphingobium sp.* (Jiang, Gao et al., 2020; Mishra et al., 2021)

Ma et al. (2009) analysed four degradation products that produced three intermediates, including 1,4-benzenedicarboxaldehyde anion, 7-benzofuranol, 2,3-dihydro-2,2-dimethyl (also known as 2,3-dihydro-2,2-dimethyl-benzofuran-7-ol), and 7-hydroxy-2,2-dimethyl-benzofuran-3-one, using GC/MS at retention times of 3.6 min, 7.44 min, and 13.8 min, respectively. The experimental results show that the Fenton process is an effective method for degrading carbofuran-contaminated water.

Javier Benitez et al. (2002) investigated how ozonation affects the breakdown of carbofuran. The research concluded that pseudo-first-order kinetics could be used to explain the global ozonation reaction. Temperature and partial ozone pressure significantly influence the removal rate of carbofuran. In contrast, similar degradation rates are obtained for pH 2 and 9 experiments, indicating a negligible contribution of the reaction with hydroxyl radicals generated by ozone self-decomposition at pH 9. The apparent pseudo-first-order rate constants for the ozonation reaction are calculated and found to be in the  $5.1 \times 10^{-4}$  to  $19.5 \times 10^{-4} \text{ s}^{-1}$  range. A kinetic competition model was used to calculate the actual rate constant for the reaction between ozone and carbofuran at 20°C, which was 1108 l/mols (Bhat & Gogate, 2021; Javier Benitez et al., 2002).

In addition to these studies, advanced oxidation processes (AOP)s, which include ultraviolet radiation, ozone, and hydrogen peroxide, are the most effective at degrading carbofuran. Several studies have proven that AOPs are promising and attractive alternatives in treating organic pollutants that are either toxic or refractory to biological treatments (Pignatello et al., 2006; M. Zhang et al., 2019). AOPs mainly generate highly oxidative free radicals, mostly the hydroxyl radical ( $\bullet\text{OH}$ ) with an  $E^\circ$  of 2.8 V/SHE (L.-A. Lu et al., 2011; Thomas et al., 2021).

In a separate investigation by Ma et al. (2010), the degrading of carbofuran was investigated using three different processes: ultrasound, the Fenton reaction, and a combination of the two. The researchers monitored the effects of various factors on the kinetics of carbofuran's degradation, including the dosages of  $\text{H}_2\text{O}_2$  and  $\text{Fe}_2^+$  as well as the initial concentrations of carbofuran. After 30 minutes of reaction time with an initial carbofuran concentration of 20 mg/L and  $\text{H}_2\text{O}_2$  and  $\text{Fe}_2^+$  dosages of 100 mg/L and 20 mg/L, more than 99% carbofuran degradation efficiency and 46% mineralization were achieved, all at pH 3, respectively. Increases in the Fenton reagents  $\text{H}_2\text{O}_2$  and  $\text{Fe}_2^+$  improved degradation efficiency, but increases in the initial carbofuran concentration had the opposite effect (Anandan et al., 2020; Bhat & Gogate, 2021; Ma et al., 2010).

## CONCLUSION

This review focused on carbofuran usage, toxicity, and degradation pathways for agricultural purposes. Hydrolysis and photolysis were suggested as the major degradation routes of carbofuran, followed by microbial degradation and oxidation. This process

generally reduces aqueous carbofuran concentrations when combined with microbial activity.

Several metabolites are formed during the breakdown of carbofuran. These include 3-hydroxy-carbofuran, 3-ketocarbofuran, carbofuran-phenol, and 3-hydroxy-5-nitrophenol. These metabolites vary in toxicity and persistence in the environment, with some being more lethal than the parent compound. The detailed pathways and mechanisms of their formation are determined by the degradation method involved. For example, microbial degradation can involve hydrolysis, oxidation, or reduction pathways, while photolysis can generate various metabolites via direct or indirect light reactions.

Based on the current research on carbofuran degradation, there are a few recommendations for future studies, such as bioremediation, advanced oxidation processes (AOPs), enzymes, and phytoremediation. Combining these methods can be an effective strategy for remediating soil and water bodies contaminated with carbofuran. Further studies are needed to enhance the conditions and increase the efficacy of these methods, not only on a laboratory level but also on a wide scale in field applications to benefit the community.

## ACKNOWLEDGMENTS

The authors acknowledge the Malaysian Ministry of Education and Universiti Sains Malaysia (USM) for awarding the Fundamental Research Grant Scheme (FRGS) to conduct this research (FRGS/1/2019/WAB05/USM/02/3)

## REFERENCES

- Abdullah, A. R., Bajet, C. M., Matin, M. A., Nhan, D. D., & Sulaiman, A. H. (1997). Ecotoxicology of pesticides in the tropical paddy field ecosystem. *Environmental Toxicology and Chemistry*, 16(1), 59-70. <https://doi.org/https://doi.org/10.1002/etc.5620160106>
- Acharid, A., Quentel, F., Elléouet, C., Olier, R., & Privat, M. (2006). Coadsorption of carbofuran and lead at the air/water interface. Possible occurrence of non-volatile pollutant cotransfer to the atmosphere. *Chemosphere*, 62(6), 989-997.
- Achik, J., & Schiavon, M. (1989). Carbofuran transfer and persistence in drained agricultural soils related to their structure and adsorption properties. *Ecotoxicology and Environmental Safety*, 18(1), 83-92. [https://doi.org/https://doi.org/10.1016/0147-6513\(89\)90094-8](https://doi.org/https://doi.org/10.1016/0147-6513(89)90094-8)
- Al-Ahmadi, M. S. (2019). Pesticides, anthropogenic activities, and the health of our environment safety. In M. Larramendy & S. Soloneski (Eds.), *Pesticides - Use and Misuse and their Impact in the Environment* (pp. 73-95). IntechOpen. <https://doi.org/10.5772/intechopen.84161>
- Fan, A. M., & Alexeeff, G. V. (2000). *Public Health Goals for Chemicals in Drinking Water*. California Environmental Protection Agency. <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=58eef5d5e0102c8e7ea84da41ddf095d545ff50a>

- Anandan, S., Ponnusamy, V. K., & Ashokkumar, M. (2020). A review on hybrid techniques for the degradation of organic pollutants in aqueous environment. *Ultrasonics Sonochemistry*, 67, Article 105130.
- Ariffin, F., & Rahman, S. A. (2020). Biodegradation of carbofuran: A review. *Journal of Environmental Microbiology and Toxicology*, 8(1), 50-57.
- Ayilara, M. S., Adeleke, B. S., Akinola, S. A., Fayose, C. A., Adeyemi, U. T., Gbadegesin, L. A., Omole, R. K., Johnson, R. M., Uthman, Q. O., & Babalola, O. O. (2023). Biopesticides as a promising alternative to synthetic pesticides: A case for microbial pesticides, phytopesticides, and nanobiopesticides. *Frontiers in Microbiology*, 14, 1-16. <https://doi.org/10.3389/fmicb.2023.1040901>
- Azam, S. M. R., Ma, H., Xu, B., Devi, S., Siddique, M. A. B., Stanley, S. L., Bhandari, B., & Zhu, J. (2020). Efficacy of ultrasound treatment in the removal of pesticide residues from fresh vegetables: A review. *Trends in Food Science & Technology*, 97, 417-432. <https://doi.org/10.1016/j.tifs.2020.01.028>
- Bachman, J., & Patterson, H. H. (1999). Photodecomposition of the carbamate pesticide carbofuran: Kinetics and the influence of dissolved organic matter. *Environmental Science & Technology*, 33(6), 874-881. <https://doi.org/10.1021/es9802652>
- Bailey, H. C., Digiorgio, C., Kroll, K., Hinton, D. E., Miller, J. L., & Starrett, G. (1996). Development of procedures for identifying pesticide toxicity in ambient waters: Carbofuran, diazinon, chlorpyrifos. *Environmental Toxicology and Chemistry: An International Journal*, 15(6), 837-845. <https://doi.org/10.1002/etc.5620150604>
- Baumart, J., Dalosto, M., & Santos, S. (2011). Effects of carbofuran and metsulfuron-methyl on the benthic macroinvertebrate community in flooded ricefields. *Acta Limnologica Brasiliensia*, 23(2), 138-144. <https://doi.org/10.1590/S2179-975X2011000200004>
- Bedair, H., Rady, H. A., Hussien, A. M., Pandey, M., Apollon, W., AlKafaas, S. S., & Ghosh, S. (2022). Pesticide detection in vegetable crops using enzyme inhibition methods: A comprehensive review. *Food Analytical Methods*, 15(7), 1979-2000. <https://doi.org/10.1007/s12161-022-02254-x>
- Benitez, F. J., Acero, J. L., & Real, F. J. (2002). Degradation of carbofuran by using ozone, UV radiation and advanced oxidation processes. *Journal of Hazardous Materials*, 89(1), 51-65. [https://doi.org/https://doi.org/10.1016/S0304-3894\(01\)00300-4](https://doi.org/https://doi.org/10.1016/S0304-3894(01)00300-4)
- Bhat, A. P., & Gogate, P. R. (2021). Degradation of nitrogen-containing hazardous compounds using advanced oxidation processes: A review on aliphatic and aromatic amines, dyes, and pesticides. *Journal of Hazardous Materials*, 403, Article 123657. <https://doi.org/10.1016/j.jhazmat.2020.123657>
- Boh, C. U., & Yeang, L. K. (2002). *Pesticide residues and microbial contamination of water resources in the MUDA rice agroecosystem*. Malaysian Institute for Nuclear Technology Research MINT. [http://inis.iaea.org/search/search.aspx?orig\\_q=RN:37065380](http://inis.iaea.org/search/search.aspx?orig_q=RN:37065380)
- Bolan, N. S., Adriano, D. C., Kunhikrishnan, A., James, T., McDowell, R., & Senesi, N. (2011). Dissolved organic matter: Biogeochemistry, dynamics, and environmental significance in soils. *Advances in Agronomy*, 110, 1-75. <https://doi.org/10.1016/B978-0-12-385531-2.00001-3>
- Bose, S., Kumar, P. S., Vo, D. V. N., Rajamohan, N., & Saravanan, R. (2021). Microbial degradation of recalcitrant pesticides: A review. *Environmental Chemistry Letters*, 19, 3209-3228. <https://doi.org/10.1007/s10311-021-01236-5>

- Brasel, J. M., Collier, A. C., & Pritsos, C. A. (2007). Differential toxic effects of carbofuran and diazinon on time of flight in pigeons (*Columba livia*): Potential for pesticide effects on migration. *Toxicology and Applied Pharmacology*, 219(2-3), 241-246. <https://doi.org/10.1016/j.taap.2006.11.028>
- Cáceres, T., Maestroni, B., Islam, M., & Cannavan, A. (2019). Sorption of 14C-carbofuran in Austrian soils: Evaluation of fate and transport of carbofuran in temperate regions. *Environmental Science and Pollution Research*, 26(1), 986-990. <https://doi.org/10.1007/s11356-018-3730-3>
- Campbell, S., David, M. D., Woodward, L. A., & Li, Q. X. (2004). Persistence of carbofuran in marine sand and water. *Chemosphere*, 54(8), 1155-1161. <https://doi.org/10.1016/j.chemosphere.2003.09.018>
- Marine, N. R. G. (1999). *Canadian Water Quality Guidelines for the Protection of Aquatic Life*. Canadian Council of Ministers of the Environment. [https://mde.maryland.gov/programs/Marylander/Documents/canadian\\_WQ\\_Chloride\\_\(en\)\[1\].pdf](https://mde.maryland.gov/programs/Marylander/Documents/canadian_WQ_Chloride_(en)[1].pdf)
- Chaudhary, D. K., & Kim, J. (2019). New insights into bioremediation strategies for oil-contaminated soil in cold environments. *International Biodeterioration & Biodegradation*, 142, 58-72. <https://doi.org/10.1016/j.ibiod.2019.05.001>
- Choez, A., & Evaristo, R. R. (2018). Acute intoxication by carbofuran in cattle. *Journal of Veterinary Medicine*, 2(1), Article 20. <https://doi.org/10.4172/2574-2868.100020>
- Chowdhury, A. Z., Jahan, S. A., Islam, M. N., Moniruzzaman, M., Alam, M. K., Zaman, M. A., Karim, N., & Gan, S. H. (2012). Occurrence of organophosphorus and carbamate pesticide residues in surface water samples from the Rangpur District of Bangladesh. *Bulletin of Environmental Contamination and Toxicology*, 89(1), 202-207. <https://doi.org/10.1007/s00128-012-0641-8>
- Cid, A., Mejuto, J. C., Moldes, Ó. A., & Morales, J. (2011). Basic degradation of 3-keto-carbofuran in the presence of non-ionic self-assembly colloids. *Fresenius Environmental Bulletin*, 20(2), 354-357
- Clasen, B., Leitemperger, J., Murussi, C., Pretto, A., Menezes, C., Dalabona, F., Marchezan, E., Adaime, M. B., Zanella, R., & Loro, V. L. (2014). Carbofuran promotes biochemical changes in carp exposed to rice field and laboratory conditions. *Ecotoxicology and Environmental Safety*, 101, 77-82. <https://doi.org/10.1016/j.ecoenv.2013.12.012>
- Conway, G. R., & Pretty, J. N. (2013). *Unwelcome Harvest: Agriculture and Pollution*. Routledge.
- Cossi, P. F., Herbert, L. T., Yusseppone, M. S., Pérez, A. F., & Kristoff, G. (2020). Toxicity evaluation of the active ingredient acetamiprid and a commercial formulation (Assail® 70) on the non-target gastropod *Biomphalaria straminea* (Mollusca: Planorbidae). *Ecotoxicology and Environmental Safety*, 192, Article 110248. <https://doi.org/10.1016/j.ecoenv.2020.110248>
- Ćwieląg-Piasecka, I., Debicka, M., & Medyńska-Juraszek, A. (2021). Effectiveness of carbaryl, carbofuran and metolachlor retention in soils under the influence of different colloid. *Minerals*, 11(9), Article 924. <https://doi.org/10.3390/min11090924>
- Cycoń, M., Mroziak, A., & Piotrowska-Seget, Z. (2017). Bioaugmentation as a strategy for the remediation of pesticide-polluted soil: A review. *Chemosphere*, 172, 52-71. <https://doi.org/10.1016/j.chemosphere.2016.12.129>

- Daam, M. A., Crum, S. J. H., Van den Brink, P. J., & Nogueira, A. J. A. (2008). Fate and effects of the insecticide chlorpyrifos in outdoor plankton-dominated microcosms in Thailand. *Environmental Toxicology and Chemistry*, 27(12), 2530-2538. <https://doi.org/10.1897/07-628.1>
- Daam, M. A., & Van den Brink, P. J. (2010). Implications of differences between temperate and tropical freshwater ecosystems for the ecological risk assessment of pesticides. *Ecotoxicology*, 19(1), 24-37. <https://doi.org/10.1007/s10646-009-0402-6>
- Daliya, P., Ljungqvist, O., Brindle, M. E., & Lobo, D. N. (2020). Guidelines for guidelines. In O. Ljungqvist, N. K. Francis & R. D. Urman (Eds.), *Enhanced Recovery After Surgery* (pp. 23-28). Springer. [https://doi.org/10.1007/978-3-030-33443-7\\_3](https://doi.org/10.1007/978-3-030-33443-7_3)
- De Melo Plese, L. P., Paraiba, L. C., FOLONI, L. L., & Trevizan, L. R. P. (2005). Kinetics of carbosulfan hydrolysis to carbofuran and the subsequent degradation of this last compound in irrigated rice fields. *Chemosphere*, 60(2), 149-156. <https://doi.org/10.1016/j.chemosphere.2005.02.049>
- Diale, M. O., Abrahams, A., & Serepa-Dlamini, M. H. (2022). *Isolation and Characterization of Bacteriocin-Like Substances from Bacillus paranthracis Strain MHSD3, a Potential Probiotic*. Research Square. <https://doi.org/10.21203/rs.3.rs-1553916/v1>
- Duc, H. D. (2022). Enhancement of carbofuran degradation by immobilized *Bacillus* sp. strain DT1. *Environmental Engineering Research*, 27(4), Article 210158. <https://doi.org/10.4491/eer.2021.158>
- Daul, J. A., Seng, M. M. K. B., Ikeda, M. M., Ombajo, M. P. S. O., Goji, M. I., Khan, M. M. B., & Magdalena, M. (2012, March 19-23). Rotterdam convention on the prior informed consent procedure for certain hazardous chemicals and pesticides in international trade. In Chemical Review Committee Eighth meeting (pp. 1-39). Geneva, Switzerland
- Edzwald, J. (2011). *Water Quality & Treatment: A Handbook on Drinking Water*. McGraw-Hill Education.
- EFSA. (2014). Reasoned opinion on the review of the existing MRLs for carbofuran, carbosulfan, benfuracarb and furathiocarb and the setting of an import tolerance for carbofuran in cultivated mushrooms. *EFSA Journal*, 12(2), Article 3559. <https://doi.org/10.2903/j.efsa.2014.3559>
- Eisler, R. (1985). *Carbofuran Hazards to Fish, Wildlife, and Invertebrates: A Synoptic Review*. U.S. Department of the Interior, Fish and Wildlife Service
- Elbashir, A. A., & Aboul-Enein, H. Y. (2015). Separation and analysis of triazine herbicide residues by capillary electrophoresis. *Biomedical Chromatography*, 29(6), 835-842. <https://doi.org/10.1002/bmc.3381>
- Evert, S. (2002). *Environmental Fate of Carbofuran*. California Environmental Protection Agency, Department of Pesticide Regulation, Sacramento. <https://www.cdpr.ca.gov/docs/emon/pubs/envfate.html>
- Fang, N., Lu, Z., Hou, Z., Zhang, C., & Zhao, X. (2022). Hydrolysis and photolysis of flupyradifurone in aqueous solution and natural water: Degradation kinetics and pathway. *Chemosphere*, 298, Article 134294. <https://doi.org/10.1016/j.chemosphere.2022.134294>
- Farahani, G. H. N., Zakaria, Z., Kuntom, A., Omar, D., & Ismail, B. S. (2007). Adsorption and desorption of carbofuran in Malaysian soils. *Advances in Environmental Biology*, 1(1), 20-26.
- Farahani, G. H. N., Zakaria, Z., Kuntom, A., Omar, D., & Ismail, B. S. (2008). Persistence of carbofuran in two Malaysian soils. *Plant Protection Quarterly*, 23(4), 179-183.



- Farahani, G. H. N., Zuriati, Z., Aini, K., & Ismail, B. S. (2012). Persistence of carbofuran in Malaysian waters. *American-Eurasian Journal of Agricultural & Environmental Sciences*, 12(5), 616-623.
- Fatah, F. A., Yaakub, N., Ridzuan, R. M., & Ahmad, A. R. (2017). The study on the economic fertilizer requirement for paddy production on a Malaysian soil. *Journal of Fundamental and Applied Sciences*, 9(2S), 777-798. <https://doi.org/10.4314/jfas.v9i2s.48>
- Fenner, K., Canonica, S., Wackett, L. P., & Elsner, M. (2013). Evaluating pesticide degradation in the environment: Blind spots and emerging opportunities. *Science*, 341(6147), 752-758.
- Filik, H., & Çekiç, S. D. (2011). Cloud point extraction of pesticide residues. In M. Stoytcheva (Ed.), *Pesticides in the Modern World-Trends in Pesticides Analysis* (pp. 247-280). IntechOpen.
- FOA. (2020). *The State of World Fisheries and Aquaculture 2020: Sustainability in Action*. Food and Agriculture Organization. <https://doi.org/10.4060/ca9229en>
- Furmanczyk, E. M., Kaminski, M. A., Lipinski, L., Dziembowski, A., & Sobczak, A. (2018). *Pseudomonas laurylsulfatovorans* sp. nov., sodium dodecyl sulfate degrading bacteria, isolated from the peaty soil of a wastewater treatment plant. *Systematic and Applied Microbiology*, 41(4), 348-354. <https://doi.org/10.1016/j.syapm.2018.03.009>
- Gammon, D. W., Liu, Z., & Becker, J. M. (2012). Carbofuran occupational dermal toxicity, exposure and risk assessment. *Pest Management Science*, 68(3), 362-370. <https://doi.org/10.1002/ps.2270>
- Garcia-Saavedra, Y., Rivera, A., Hernandez-Aladana, F., Romero-Arenas, O., Sanchez-Morales, P., & Gionocerzo, S. (2018). Carbofuran, malathion and 2, 4-D degradation by bacterial activity. *Journal of Pure and Applied Microbiology*, 12(3), 1331-1335. <http://dx.doi.org/10.22207/JPAM.12.3.35>
- Ge, Y. D., Palikhe, B. R., Gu, B., & Beatrice, G. (2021). Status of highly hazardous pesticides and their mitigation measures in Asia. *Advances in Entomology*, 10(1), 14-33. <https://doi.org/10.4236/ae.2022.101002>
- Grandclément, C., Seyssiecq, I., Piram, A., Wong-Wah-Chung, P., Vanot, G., Tiliacos, N., Roche, N., & Doumenq, P. (2017). From the conventional biological wastewater treatment to hybrid processes, the evaluation of organic micropollutant removal: A review. *Water Research*, 111, 297-317. <https://doi.org/10.1016/j.watres.2017.01.005>
- Gunther, F. A., & Gunther, J. D. (2013). *Residue Reviews: Residues Of Pesticides And Other Foreign Chemicals In Foods And Feeds* (Vol. 40). Springer Science & Business Media.
- Gupta, J., Rathour, R., Singh, R., & Thakur, I. S. (2019). Production and characterization of extracellular polymeric substances (EPS) generated by a carbofuran degrading strain *Cupriavidus* sp. ISTL7. *Bioresource Technology*, 282, 417-424. <https://doi.org/10.1016/j.biortech.2019.03.054>
- Halimah, M., Tan, Y. A., & Ismail, B. S. (2005). The fate of fluroxypyr in the soil in an oil palm agroecosystem. *Weed Biology and Management*, 5(4), 184-189. <https://doi.org/10.1111/j.1445-6664.2005.00179.x>
- Vincent, K. (2014). Reforming wildlife law: Proposals by the Law Commission for England and Wales. *International Journal for Crime, Justice and Social Democracy*, 3(2), 67-80.
- Hayasaka, D., Korenaga, T., Suzuki, K., Saito, F., Sánchez-Bayo, F., & Goka, K. (2012). Cumulative ecological impacts of two successive annual treatments of imidacloprid and fipronil on aquatic communities of

- paddy mesocosms. *Ecotoxicology and Environmental Safety*, 80, 355-362. <https://doi.org/10.1016/j.ecoenv.2012.04.004>
- Health Canada. (2020). *Guidelines for Canadian Drinking Water Quality Summary Table*. Water and Air Quality Bureau, Healthy Environments and Consumer Safety Branch, Health Canada, Ottawa, Ontario. [https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt\\_formats/pdf/pubs/water-eau/sum\\_guide-res\\_recom/summary-table-EN-2020-02-11.pdf](https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/ewh-semt/alt_formats/pdf/pubs/water-eau/sum_guide-res_recom/summary-table-EN-2020-02-11.pdf)
- Howard, P. (2017). *Handbook of Environmental Fate and Exposure Data: For Organic Chemicals, Volume III Pesticides*. Routledge.
- Huang, J., & Mabury, S. A. (2000). A new method for measuring carbonate radical reactivity toward pesticides. *Environmental Toxicology and Chemistry: An International Journal*, 19(6), 1501–1507.
- Ibitoye, O. O., Nawi, N. M., Kamarulzaman, N. H., & Man, N. (2014). Consumers' awareness towards organic rice in Malaysia. *International Food Research Journal*, 21(5), 1711-1718.
- Idayu, M., Anis, N., Akmal, S., & Esa, N. (2014). A survey on use, hazards and potential risks of rice farming pesticides in Permatang Keriang, Pulau Pinang (Malaysia). *International Journal of Scientific and Research Publications*, 4(10), 1-11.
- Islam, A., Naskar, S., Mazumder, U. K., Gupta, M., & Ghosal, S. (2008). Estrogenic properties of phyllanthin and hypophyllanthin from *Phyllanthus amarus* against carbofuran induced toxicity in female rats. *Pharmacologyonline*, 3, 1006-1016.
- Ismail, B. S., & Kalithasan, K. (2004). Adsorption, desorption, and mobility of permethrin in Malaysian soils. *Journal of Environmental Science and Health, Part B*, 39(3), 419-429. <https://doi.org/10.1081/PFC-120035927>
- Ismail, B. S., Ngan, C. K., Cheah, U. B., & Abdullah, W. Y. W. (2004). Leaching potential of pesticides in a vegetable farm in the Cameron Highlands, Malaysia. *Bulletin of Environmental Contamination and Toxicology*, 72(4), 836-843. <https://doi.org/10.1007/s00128-004-0320-5>
- Ismail, B. S., Prayitno, S., & Tayeb, M. A. (2015). Contamination of rice field water with sulfonylurea and phenoxy herbicides in the Muda Irrigation Scheme, Kedah, Malaysia. *Environmental Monitoring and Assessment*, 187(7), 1-13. <https://doi.org/10.1007/s10661-015-4600-9>
- Jaiswal, D. K., Verma, J. P., & Yadav, J. (2017). Microbe induced degradation of pesticides in agricultural soils. In S. N. Singh (Ed.), *Microbe-induced degradation of pesticides* (pp. 167-189). Springer.
- Jemutai-Kimosop, S., Orata, F. O., K'Owino, I. O., & Getenga, Z. M. (2014). The dissipation of carbofuran in two soils with different pesticide application histories within Nzoia River Drainage Basin, Kenya. *Bulletin of Environmental Contamination and Toxicology*, 92(5), 616-620. <https://doi.org/10.1007/s00128-014-1234-5>
- Jiang, W., Gao, Q., Zhang, L., Wang, H., Zhang, M., Liu, X., Zhou, Y., Ke, Z., Wu, C., & Qiu, J. (2020). Identification of the key amino acid sites of the carbofuran hydrolase CehA from a newly isolated carbofuran-degrading strain *Sphingibium* sp. CFD-1. *Ecotoxicology and Environmental Safety*, 189, Article 109938. <https://doi.org/10.1016/j.ecoenv.2019.109938>

- Jiang, W., Zhang, C., Gao, Q., Zhang, M., Qiu, J., Yan, X., & Hong, Q. (2020). Carbamate CN hydrolase gene *ameH* responsible for the detoxification step of methomyl degradation in *Aminobacter aminovorans* strain MDW-2. *Applied and Environmental Microbiology*, 87(1), Article e02005-20. <https://doi.org/10.1128/AEM.02005-20>
- Jørgensen, S. E. (2008). *Biodegradation*. In S. E. Jørgensen & B. D. Fath (Eds.), *Encyclopedia of Ecology* (pp. 366–367). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-008045405-4.00260-3>
- Kabir, K. H., Abdullah, M., Prodhan, M. D. H., Ahmed, M. S., & Alam, M. N. (2007). Determination of carbofuran residue in the samples of sugarcane (*Sacharum officinarum* L) and soil of sugarcane field. *The Agriculturists*, 5(1&2), 61-66.
- Kaida, N., Habib, S., Yasid, N. A., & Shukor, M. Y. (2018). Biodegradation of Petroleum Hydrocarbons by *Bacillus* sp.: A Review. *Bioremediation Science and Technology Research*, 6(2), 14-21.
- Kanedi, M. (2017). Physiological effects of carbofuran on earthworm *Pheretima javanica* Gates. *Advances in Life Sciences*, 7(2), 21-25.
- Katagi, T. (2016). Pesticide behavior in modified water-sediment systems. *Journal of Pesticide Science*, 41(4), 121-132. <https://doi.org/10.1584/jpestics.D16-060>
- Khairatul, A. M., Ngan, C. K., & Ismail, B. S. (2013). Adsorption and leaching studies of molinate, carbofuran and propiconazole in Muda agricultural soils. *Journal of Tropical Agriculture and Food Science*, 41(1), 127-136.
- Khuntong, S., Sirivithayapakorn, S., Pakkong, P., & Sorulump, C. (2010). Adsorption kinetics of carbamate pesticide in rice field soil. *Environment Asia*, 3(2), 20-28.
- Kim, I. S., Ryu, J. Y., Hur, H. G., Gu, M. B., Kim, S. D., & Shim, J. H. (2004). *Sphingomonas* sp. strain SB5 degrades carbofuran to a new metabolite by hydrolysis at the furanyl ring. *Journal of Agricultural and Food Chemistry*, 52(8), 2309-2314. <https://doi.org/10.1021/jf0355021>
- Kumar, R., Ali, M., Kumar, A., & Gahlot, V. (2015). Comparative bioremedial effect of withania somnifera and curcuma longa on ovaries of pesticide induced mice. *European Journal of Pharmaceutical and Medical Research*, 2(7), 249-253.
- Kumar, S. A., Krishna, A. J., Suvarna, T., Balaji, P. B., & Hyndhavi, K. (2020). An experimental study on suitability on treated sewage water for agriculture. *International Research Journal of Engineering and Technology*, 7(3), 2010-2013.
- Kuswandi, B., Futra, D., & Heng, L. Y. (2017). Nanosensors for the detection of food contaminants. In A. E. Oprea & A. M. Grumezescu (Eds.), *Nanotechnology Applications in Food* (pp. 307-333). Academic Press. <https://doi.org/https://doi.org/10.1016/B978-0-12-811942-6.00015-7>
- Lan, J., Wang, M., Ding, S., Fan, Y., Diao, X., Li, Q. X., & Zhao, H. (2019). Simultaneous detection of carbofuran and 3-hydroxy-carbofuran in vegetables and fruits by broad-specific monoclonal antibody-based ELISA. *Food and Agricultural Immunology*, 30(1), 1085-1096. <https://doi.org/10.1080/09540105.2019.1664997>
- Larson, S. J., Capel, P. D., & Majewski, M. S. (2019). *Pesticides in Surface Waters: Distribution, Trends, and Governing Factors*. CRC Press.

- Latif, Y., Sherazi, S. T. H., & Bhanger, M. I. (2011). Assessment of pesticide residues in commonly used vegetables in Hyderabad, Pakistan. *Ecotoxicology and Environmental Safety*, 74(8), 2299-2303.
- Lehel, J., Laczay, P., Déri, J., Darin, E. G., & Budai, P. (2010). Model study on the clinical signs and residue concentrations of sublethal carbofuran poisoning in birds. *Journal of Wildlife Diseases*, 46(4), 1274-1278.
- Leistra, M., & Boesten, J. (2008). *Movement of Bromide-Ion and Carbofuran in the Humic Sandy Soil of a Potato Field with Ridges and Furrows: Measurements in the Field and Computations with the PEARL Model* (No. 1750). Alterra. <https://library.wur.nl/WebQuery/wurpubs/371006>
- Liong, P. C. (1988). *Toxicity of Some Pesticides Towards Freshwater Fishes*. Department of Fisheries.
- Liu, L., Helbling, D. E., Kohler, H. P. E., & Smets, B. F. (2019). Modelling carbofuran biotransformation by *Novosphingobium* sp. KN65. 2 in the presence of coincidental carbon and indigenous microbes. *Environmental Science: Water Research & Technology*, 5(4), 798-807.
- Liu, T., Xin, Y., Liu, X., Wu, B., & Xiang, M. (2021). Advances in microbial degradation of plastics. *Chinese Journal of Biotechnology*, 37(8), 2688-2702. <https://doi.org/10.13345/j.cjb.200624>
- Lu, L. A., Ma, Y. S., Kumar, M., & Lin, J. G. (2011). Photochemical degradation of carbofuran and elucidation of removal mechanism. *Chemical Engineering Journal*, 166(1), 150-156.
- Lu, S., Liu, L., Demissie, H., An, G., & Wang, D. (2021). Design and application of metal-organic frameworks and derivatives as heterogeneous Fenton-like catalysts for organic wastewater treatment: A review. *Environment International*, 146, Article 106273. <https://doi.org/10.1016/j.envint.2020.106273>
- Ma, Y. S., Kumar, M., & Lin, J. G. (2009). Degradation of carbofuran-contaminated water by the Fenton process. *Journal of Environmental Science and Health, Part A*, 44(9), 914-920. <https://doi.org/10.1080/10934520902958807>
- Ma, Y. S., Sung, C. F., & Lin, J. G. (2010). Degradation of carbofuran in aqueous solution by ultrasound and Fenton processes: Effect of system parameters and kinetic study. *Journal of Hazardous Materials*, 178(1), 320-325. <https://doi.org/https://doi.org/10.1016/j.jhazmat.2010.01.081>
- Mackay, D., Shiu, W. Y., & Lee, S. C. (2006). *Handbook of Physical-Chemical Properties and Environmental Fate for Organic Chemicals*. CRC press.
- Mahalakshmi, M., Arabindoo, B., Palanichamy, M., & Murugesan, V. (2007). Photocatalytic degradation of carbofuran using semiconductor oxides. *Journal of Hazardous Materials*, 143(1-2), 240-245.
- Mahindru, S. N. (2009). *Food Contaminants-Origin, Propagation & Analysis*. APH Publishing.
- Maksuk, M. (2021). Analisis residu pestisida karbofuran dalam air sungai di kawasan pertanian padi [Analysis of carbofuran pesticide residues in river water at paddy field area]. *Seminar Nasional Hari Air Sedunia*, 3(1), 1-5.
- Malhotra, H., Kaur, S., & Phale, P. S. (2021). Conserved metabolic and evolutionary themes in microbial degradation of carbamate pesticides. *Frontiers in Microbiology*, 12, Article 648868. <https://doi.org/10.3389/fmicb.2021.648868>
- Meftaul, I. M., Venkateswarlu, K., Dharmarajan, R., Annamalai, P., & Megharaj, M. (2020). Pesticides in the urban environment: A potential threat that knocks at the door. *Science of The Total Environment*, 711, Article 134612. <https://doi.org/10.1016/j.scitotenv.2019.134612>

- Mir-Tutusaus, J. A., Masís-Mora, M., Corcellas, C., Eljarrat, E., Barceló, D., Sarrà, M., Caminal, G., Vicent, T., & Rodríguez-Rodríguez, C. E. (2014). Degradation of selected agrochemicals by the white rot fungus *Trametes versicolor*. *Science of the Total Environment*, *500*, 235-242. <https://doi.org/10.1016/j.scitotenv.2014.08.116>
- Mishra, S., Pang, S., Zhang, W., Lin, Z., Bhatt, P., & Chen, S. (2021). Insights into the microbial degradation and biochemical mechanisms of carbamates. *Chemosphere*, *279*, Article 130500. <https://doi.org/10.1016/j.chemosphere.2021.130500>
- Mishra, S., Zhang, W., Lin, Z., Pang, S., Huang, Y., Bhatt, P., & Chen, S. (2020). Carbofuran toxicity and its microbial degradation in contaminated environments. *Chemosphere*, *259*, Article 127419. <https://doi.org/10.1016/j.chemosphere.2020.127419>
- Mohamed, B., Rachid, M., & Amina, A. (2021). Study on biodegradation and dissipation of 14 C-carbofuran in clay soil from Loukkos Perimeter, Northwestern Morocco. *New Ideas Concerning Science and Technology*, *7*, 92-103.
- Mohanta, M. K., Saha, A. K., Zamman, M. T., Ekram, A. E., Khan, A. E., Mannan, S. B., & Fakruddin, M. (2012). Isolation and characterization of carbofuran degrading bacteria from cultivated soil. *Biochemical and Cellular Archives*, *12*(2), 313-320.
- Morales, J., Manso, J. A., & Mejuto, J. C. (2012). Basic hydrolysis of carbofuran in the presence of cyclodextrins. *Supramolecular Chemistry*, *24*(6), 399-405. <https://doi.org/10.1080/10610278.2012.688121>
- Moreira, R. A., da Silva Mansano, A., & Rocha, O. (2015). The toxicity of carbofuran to the freshwater rotifer, *Philodina roseola*. *Ecotoxicology*, *24*(3), 604-615. <https://doi.org/10.1007/s10646-014-1408-2>
- Morrice, P., Barbato, F., Iacovo, R. Dello, Seccia, S., & Ungaro, F. (2001). Kinetics and mechanism of imazosulfuron hydrolysis. *Journal of Agricultural and Food Chemistry*, *49*(8), 3816-3820. <https://doi.org/10.1021/jf010088f>
- Mukaj, M., Mai, S., Cara, M., & Ruci, T. (2017). Influence of pH in concentration of persistent organic pesticides residues in agricultural soils. *Albanian Journal of Agricultural Sciences*, *2017*, 173-178.
- Mustapha, M. U., Halimoon, N., Johar, W. L. W., & Abd Shukor, M. Y. (2019). An overview on biodegradation of carbamate pesticides by soil bacteria. *Pertanika Journal of Science & Technology*, *27*(2), 547-563.
- Nafeesa, M., Ramaraju, K., Kuttalam, S., & Doraisamy, P. (2017). Studies on growth and carbofuran degradation potential of bacterial isolates. *Madras Agricultural Journal*, *104*(1-3), 58-63.
- Nashriyah, M., & Azimahtol, H. L. P. (2002). *Residue Levels of Molinate in Rice Field Soil: Their effects on Populations of Aquatic Flora and Fauna under Recycling and Non-Recycling Practices in the MUDA Area*. International Atomic Energy Agency. [https://inis.iaea.org/search/search.aspx?orig\\_q=RN:37065383](https://inis.iaea.org/search/search.aspx?orig_q=RN:37065383)
- Nguyen, T. P. O., De Mot, R., & Springael, D. (2015). Draft genome sequence of the carbofuran-mineralizing *Novosphingobium* sp. strain KN65. 2. *Genome Announcements*, *3*(4), Article e00764-15. <https://doi.org/10.1128/genomea.00764-15>
- Nguyen, T. P. O., Helbling, D. E., Bers, K., Fida, T. T., Wattiez, R., Kohler, H.-P. E., Springael, D., & De Mot, R. (2014). Genetic and metabolic analysis of the carbofuran catabolic pathway in *Novosphingobium* sp. KN65. 2. *Applied Microbiology and Biotechnology*, *98*, 8235-8252. <https://doi.org/10.1007/s00253-014-5858-5>

- Numan, A., Khan, M., Uddin, R., Rahman, M., Bhuiyan, M., & Akter, N. (2018). Risk assessment of commonly used major pesticides for tomato (*Solanum lycopersicum* L.) cultivation in Bangladesh. *Advances in Nutrition and Food Science*, 1(2), 1-24.
- Okoli, U. A., Nubila, N. I., & Okafor, M. T. (2017). Organophosphorous pesticide: An environmental pollutant perspective. *Journal of Chemical and Pharmaceutical Research*, 9(9), 126-130.
- Oliveira, R., Almeida, M. F., Santos, L., & Madeira, L. M. (2006). Experimental design of 2, 4-dichlorophenol oxidation by Fenton's reaction. *Industrial & Engineering Chemistry Research*, 45(4), 1266-1276. <https://doi.org/10.1021/ie0509544>
- Onunga, D. O., Kowino, I. O., Ngigi, A. N., Osogo, A., Orata, F., Getenga, Z. M., & Were, H. (2015). Biodegradation of carbofuran in soils within Nzoia River Basin, Kenya. *Journal of Environmental Science and Health, Part B*, 50(6), 387-397.
- Organization, W. H. (2020). *The WHO Recommended Classification of Pesticides by Hazard and Guidelines to Classification 2019*. World Health Organization. <https://www.who.int/publications/i/item/9789240005662>
- Ostergaard, G., & Knudsen, I. (1998). The applicability of the ADI (Acceptable Daily Intake) for food additives to infants and children. *Food Additives and Contaminants*, 15, 63-74. <https://doi.org/10.1080/02652039809374617>
- Otieno, P. O., Lalah, J. O., Virani, M., Jondiko, I. O., & Schramm, K. W. (2010a). Carbofuran and its toxic metabolites provide forensic evidence for Furadan exposure in vultures (*Gyps africanus*) in Kenya. *Bulletin of Environmental Contamination and Toxicology*, 84(5), 536-544.
- Otieno, P. O., Lalah, J. O., Virani, M., Jondiko, I. O., & Schramm, K. W. (2010b). Soil and water contamination with carbofuran residues in agricultural farmlands in Kenya following the application of the technical formulation Furadan. *Journal of Environmental Science and Health, Part B*, 45(2), 137-144. <https://doi.org/10.1080/03601230903472058>
- Otieno, P. O., Lalah, J. O., Virani, M., Jondiko, I. O., & Schramm, K. W. (2011). Carbofuran use and abuse in Kenya: Residues in soils, plants, water courses and the African white-backed vultures (*Gyps africanus*) found dead. *Environmentalist*, 31(4), 382-393. <https://doi.org/10.1007/s10669-011-9350-9>
- Otieno, P. O., Schramm, K.-W., Pfister, G., Lalah, J. O., Ojwach, S. O., & Virani, M. (2012). Spatial Distribution and temporal trend in concentration of carbofuran, diazinon and chlorpyrifos ethyl residues in sediment and water in Lake Naivasha, Kenya. *Bulletin of Environmental Contamination and Toxicology*, 88(4), 526-532. <https://doi.org/10.1007/s00128-012-0529-7>
- Papadakis, E. N., Tsaoulas, A., Kotopoulou, A., Kintzikoglou, K., Vryzas, Z., & Papadopoulou-Mourkidou, E. (2015). Pesticides in the surface waters of Lake Vistonis Basin, Greece: Occurrence and environmental risk assessment. *Science of the Total Environment*, 536, 793-802. <https://doi.org/10.1016/j.scitotenv.2015.07.099>
- Park, M. R., Lee, S. W., Han, T. H., Oh, B. T., Shim, J. H., & Kim, I. S. (2006). A new intermediate in the degradation of carbofuran by *Sphingomonas* sp. strain SB5. *Journal of Microbiology and Biotechnology*, 16(8), 1306-1310.
- Park, Y., Kim, Y., Kim, J., Yoon, K. S., Clark, J., Lee, J., & Park, Y. (2013). Imidacloprid, a neonicotinoid insecticide, potentiates adipogenesis in 3T3-L1 adipocytes. *Journal of Agricultural and Food Chemistry*, 61(1), 255-259. <https://doi.org/10.1021/jf3039814>



- Peng, X., Zhang, J. S., Li, Y. Y., Li, W., Xu, G. M., & Yan, Y. C. (2008). Biodegradation of insecticide carbofuran by *Paracoccus* sp. YM3. *Journal of Environmental Science and Health. Part. B, Pesticides, Food Contaminants, and Agricultural Wastes*, 43(7), 588-594. <https://doi.org/10.1080/03601230802234492>
- Pignatello, J. J., Oliveros, E., & MacKay, A. (2006). Advanced oxidation processes for organic contaminant destruction based on the Fenton reaction and related chemistry. *Critical Reviews in Environmental Science and Technology*, 36(1), 1-84. <https://doi.org/10.1080/10643380500326564>
- Pinakini, K. S., & Kumar, T. S. M. (2006). Serial cholinesterase estimation in carbamate poisoning. *Journal of Clinical Forensic Medicine*, 13(5), 274-276. <https://doi.org/10.1016/j.jcfm.2005.11.005>
- Plangklang, P., & Reungsang, A. (2012). Isolation and characterisation of a carbofuran degrading *Burkholderia* sp. PCL3 from carbofuran-phytoremediated rhizosphere soil. *Chemistry and Ecology*, 28(3), 253-266. <https://doi.org/10.1080/02757540.2011.645032>
- Plaza, P. I., Martínez-López, E., & Lambertucci, S. A. (2019). The perfect threat: Pesticides and vultures. *Science of the Total Environment*, 687, 1207-1218. <https://doi.org/10.1016/j.scitotenv.2019.06.16>
- Prosen, H. (2012). Fate and determination of triazine herbicides in soil. In M. N. Hasaneen (Ed.), *Herbicides- Properties, Synthesis and Control of Weeds* (pp. 43-58). IntechOpen.
- Pyne, E. L. (2015). *Occurrence and Distribution of Pesticide Residues in Soil as a Result of Long-Term Application* (Master thesis). Utrecht University, Netherlands. <https://studenttheses.uu.nl/handle/20.500.12932/20943>
- Racke, K. D., Skidmore, M. W., Hamilton, D. J., Unsworth, J. B., Miyamoto, J., & Cohen, S. Z. (1997). Pesticides Report 38. Pesticide fate in tropical soils (Technical Report). *Pure and Applied Chemistry*, 69(6), 1349-1372. <https://doi.org/doi:10.1351/pac199769061349>
- Raha, P., & Das, A. K. (1990). Photodegradation of carbofuran. *Chemosphere*, 21(1-2), 99-106.
- Ramakrishnan, B., Venkateswarlu, K., Sethunathan, N., & Megharaj, M. (2019). Local applications but global implications: Can pesticides drive microorganisms to develop antimicrobial resistance? *Science of the Total Environment*, 654, 177-189. <https://doi.org/10.1016/j.scitotenv.2018.11.041>
- Ramasubramanian, T., & Paramasivam, M. (2018). Persistence and metabolism of carbofuran in the soil and sugarcane plant. *Environmental Monitoring and Assessment*, 190(9), 1-9.
- Ramesh, M., Narmadha, S., & Poopal, R. K. (2015). Toxicity of furadan (carbofuran 3% g) in *Cyprinus carpio*: Haematological, biochemical and enzymological alterations and recovery response. *Beni-Suef University Journal of Basic and Applied Sciences*, 4(4), 314-326.
- Remucal, C. K. (2014). The role of indirect photochemical degradation in the environmental fate of pesticides: A review. *Environmental Science: Processes & Impacts*, 16(4), 628-653.
- Ripley, B. D., & Chau, A. S. Y. (2020). Carbamate pesticides. In *Analysis of Pesticides in Water* (pp. 1-182). CRC Press.
- Roger, P. A., & Bhuiyan, S. I. (2012). Behavior of pesticides in rice-based agroecosystems. In P. L. Pingali & P. A. Roger (Eds.), *Impact of Pesticides on Farmer Health and the Rice Environment* (pp. 111-148). Springer.



- Ruíz-Hidalgo, K., Masis-Mora, M., Barbieri, E., Carazo-Rojas, E., & Rodríguez-Rodríguez, C. E. (2016). Ecotoxicological analysis during the removal of carbofuran in fungal bioaugmented matrices. *Chemosphere*, *144*, 864-871. <https://doi.org/10.1016/j.chemosphere.2015.09.056>
- Satar, M. D. S., Satar, S., Sebe, A., & Yesilagac, H. (2005). Carbofuran poisoning among farm workers. *The Mount Sinai Journal of Medicine*, *72*(6), 389-392.
- Sandín-España, P., & Sevilla-Morán, B. (2012). Pesticide degradation in water. In *Pesticides: Avaluation of Environmental Pollution* (pp. 79-130). CRC Press.
- Sapari, P., & Ismail, B. S. (2012). Pollution levels of thiobencarb, propanil, and pretilachlor in rice fields of the muda irrigation scheme, Kedah, Malaysia. *Environmental Monitoring and Assessment*, *184*(10), 6347-6356.
- Seiber, J. N., Catahan, M. P., & Barril, C. R. (1978). Loss of carbofuran from rice paddy water: Chemical and physical factors. *Journal of Environmental Science and Health, Part B*, *13*(2), 131-148. <https://doi.org/10.1080/03601237809372083>
- Seo, J. S., Moon, J. K., & Kim, J. H. (2012). Photodegradation of pyribenzoxim in water. *Journal of the Korean Society for Applied Biological Chemistry*, *55*(3), 391-396. <https://doi.org/10.1007/s13765-012-2042-4>
- Sharma, A., Kumar, V., Shahzad, B., Tanveer, M., Sidhu, G. P. S., Handa, N., Kohli, S. K., Yadav, P., Bali, A. S., & Parihar, R. D. (2019). Worldwide pesticide usage and its impacts on ecosystem. *SN Applied Sciences*, *1*(11), 1-16.
- Siddaramappa, R., & Seiber, J. N. (1979). Persistence of carbofuran in flooded rice soils and water. *Progress in Water Technology*, *11*, 103-111.
- Sidhu, G. K., Singh, S., Kumar, V., Dhanjal, D. S., Datta, S., & Singh, J. (2019). Toxicity, monitoring and biodegradation of organophosphate pesticides: A review. *Critical Reviews in Environmental Science and Technology*, *49*(13), 1135-1187.
- Sim, S. F., Chung, L. Y., Jonip, J., & Chai, L. K. (2019). Uptake and dissipation of carbofuran and its metabolite in Chinese kale and brinjal cultivated under humid tropic climate. *Advances in Agriculture*, *2019*, Article 7937086. <https://doi.org/10.1155/2019/7937086>
- Singh, R. P., Varshney, G., & Srivastava, G. (2012). Effect of carbofuran on enzymatic activities and growth of tomato plants in natural, fertilized and vermicompost-amended soils. *Archives of Agronomy and Soil Science*, *58*(12), 1349-1364.
- Singh, S. P., Guha, S., Bose, P., & Kunnikuruvan, S. (2017). Mechanism of the hydrolysis of endosulfan isomers. *The Journal of Physical Chemistry A*, *121*(27), 5156-5163.
- Smith, G. J. (1992). *Toxicology and Pesticide use in Relation to Wildlife, Organophosphorus, and Carbamate Compounds*. CRC Press.
- Soloneski, S., Reigosa, M. A., Molinari, G., González, N. V., & Larramendy, M. L. (2008). Genotoxic and cytotoxic effects of carbofuran and furadan® on Chinese hamster ovary (CHOK1) cells. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, *656*(1-2), 68-73. <https://doi.org/10.1016/j.mrgentox.2008.07.007>

- Sparling, D. W. (2016). *Ecotoxicology Essentials: Environmental Contaminants and Their Biological Effects on Animals and Plants*. Academic Press.
- Strachan, W. M. J., Glooschenko, W. A., & Maguire, R. J. (2018). Environmental impact and significance of pesticides. In *Analysis of Pesticides in Water* (pp. 1-23). CRC Press.
- Sun, M., Xu, W., Zhang, W., Guang, C., & Mu, W. (2022). Microbial elimination of carbamate pesticides: Specific strains and promising enzymes. *Applied Microbiology and Biotechnology*, 106(18), 5973-5986. <https://doi.org/10.1007/s00253-022-12141-4>
- Syed, J. H., Alamdar, A., Mohammad, A., Ahad, K., Shabir, Z., Ahmed, H., Ali, S. M., Sani, S. G. A. S., Bokhari, H., & Gallagher, K. D. (2014). Pesticide residues in fruits and vegetables from Pakistan: A review of the occurrence and associated human health risks. *Environmental Science and Pollution Research*, 21, 13367-13393. <https://doi.org/10.1007/s11356-014-3117-z>
- Tejada, A. W. (1995). Pesticide residues in foods and the environment as a consequence of crop protection. *Philippine Agriculturist*, 78, 63-79.
- Thomas, N., Dionysiou, D. D., & Pillai, S. C. (2021). Heterogeneous Fenton catalysts: A review of recent advances. *Journal of Hazardous Materials*, 404, Article 124082. <https://doi.org/10.1016/j.jhazmat.2020.124082>
- Tien, C., Huang, H., & Chen, C. S. (2017). Accessing the carbofuran degradation ability of cultures from natural river biofilms in different environments. *CLEAN–Soil, Air, Water*, 45(5), Article 1600380. <https://doi.org/10.1002/clen.201600380>
- Tiryaki, O. (2017). Pesticide residues and organic production. *Journal of Biological and Environmental Sciences*, 11(31), 11-23.
- Tiryaki, O., & Temur, C. (2010). The fate of pesticide in the environment. *Journal of Biological and Environmental Sciences*, 4(10), 29-38.
- Tondon, S. A., Deore, R., & Parab, A. (2018). Isolation, identification and the use of carbofuran degrading microorganisms for the removal of carbofuran pesticide from contaminated waters. *Global Journal of Bio-Science and Biotechnology*, 6, 89-95.
- Torres, J. B., & Bueno, A. de F. (2018). Conservation biological control using selective insecticides - A valuable tool for IPM. *Biological Control*, 126, 53-64.
- Trevisan, M. J., Baptista, G. C. De, Trevizan, L. R. P., & Papa, G. (2004). Residues of carbosulfan and its carbofuran metabolites and 3-hydroxy-carbofuran in oranges. *Revista Brasileira de Fruticultura*, 26(2), 230-233. <https://doi.org/10.1590/S0100-29452004000200012>
- Ukalska-Jaruga, A., Smreczak, B., & Siebielec, G. (2020). Assessment of pesticide residue content in polish agricultural soils. *Molecules*, 25 (3), Article 587. <https://doi.org/10.3390/molecules25030587>
- USEPA. (2006). *Interim Reregistration Eligibility Decision (IREDD) Document for Carbofuran*. US Environmental Protection Agency. [https://archive.epa.gov/pesticides/reregistration/web/pdf/carbofuran\\_red.pdf](https://archive.epa.gov/pesticides/reregistration/web/pdf/carbofuran_red.pdf)

- Valente-Campos, S., Spry, D. J., Palhares, J. C. P., Rudez, L. M. J., & de Aragão Umbuzeiro, G. (2019). Critical issues and alternatives for the establishment of chemical water quality criteria for livestock. *Regulatory Toxicology and Pharmacology*, *104*, 108-114. <https://doi.org/10.1016/j.yrtph.2019.03.003>
- Venkateswarlu, K., & Sethunathan, N. (1978). Degradation of carbofuran in rice soils as influenced by repeated applications and exposure to aerobic conditions following anaerobiosis. *Journal of Agricultural and Food Chemistry*, *26*(5), 1148-1151. <https://doi.org/10.1021/jf60219a019>
- Villaverde, J. J., Sevilla-Morán, B., López-Goti, C., Calvo, L., Alonso-Prados, J. L., & Sandín-España, P. (2018). Photolysis of clethodim herbicide and a formulation in aquatic environments: Fate and ecotoxicity assessment of photoproducts by QSAR models. *Science of the Total Environment*, *615*, 643-651. <https://doi.org/10.1016/j.scitotenv.2017.09.300>
- Wang, L., Cai, W.-F., & Li, Q. X. (1998). Photolysis of phloxine B in water and aqueous solutions. *Archives of Environmental Contamination and Toxicology*, *35*(3), 397-403. <https://doi.org/10.1007/s002449900394>
- Wijesinghe, M. R., Jayatillake, B., & Ratnassoriya, W. D. (2011). Solar radiation alters toxicity of carbofuran: Evidence from empirical trials with *Duttaphyrnus melanostictus*. *Journal of Tropical Forestry and Environment*, *1*(1), 48-55.
- Wols, B. A., & Hofman-Caris, C. H. M. (2012). Review of photochemical reaction constants of organic micropollutants required for UV advanced oxidation processes in water. *Water Research*, *46*(9), 2815-2827. <https://doi.org/10.1016/j.watres.2012.03.036>
- World Health Organization. (2020). *The WHO recommended classification of pesticides by hazard and guidelines to classification 2019*. World Health Organization. <https://www.who.int/publications-detail-redirect/9789240005662>
- Xu, G. F., Shen, Z. X., & Guo, R. X. (2014). The kinetics studies for the adsorption of furadan from aqueous solution by orange peel. *Advanced Materials Research*, *842*, 187-191. <https://doi.org/10.4028/www.scientific.net/AMR.842.187>
- Yan, Q. X., Hong, Q., Han, P., Dong, X. J., Shen, Y. J., & Li, S. P. (2007). Isolation and characterization of a carbofuran-degrading strain *Novosphingobium* sp. FND-3. *FEMS Microbiology Letters*, *271*(2), 207-213. <https://doi.org/10.1111/j.1574-6968.2007.00718.x>
- Yan, X., Jin, W., Wu, G., Jiang, W., Yang, Z., Ji, J., Qiu, J., He, J., Jiang, J., & Hong, Q. (2018). Hydrolase CehA and monooxygenase CfdC are responsible for carbofuran degradation in *Sphingomonas* sp. strain CDS-1. *Applied and Environmental Microbiology*, *84*(16), Article e00805-18. <https://doi.org/10.1128/AEM.00805-18>
- Zhang, C. P., He, H. M., Yu, J. Z., Hu, X. Q., Zhu, Y. H., & Wang, Q. (2016). Residues of carbosulfan and its metabolites carbofuran and 3-hydroxy carbofuran in rice field ecosystem in China. *Journal of Environmental Science and Health, Part B*, *51*(6), 351-357. <https://doi.org/10.1080/03601234.2015.1120606>
- Zhang, M., Dong, H., Zhao, L., Wang, D., & Meng, D. (2019). A review on Fenton process for organic wastewater treatment based on optimization perspective. *Science of the Total Environment*, *670*, 110-121.

Zhang, Y., Zhang, W., Li, J., Pang, S., Mishra, S., Bhatt, P., Zeng, D., & Chen, S. (2021). Emerging technologies for degradation of dichlorvos: A review. *International Journal of Environmental Research and Public Health*, 18(11), Article 5789. <https://doi.org/10.3390/ijerph18115789>

Zheng, W., Cui, T., & Li, H. (2022). Combined technologies for the remediation of soils contaminated by organic pollutants. A review. *Environmental Chemistry Letters*, 20(3), 2043-2062. <https://doi.org/10.1007/s10311-022-01407-y>

## Phenolics-Enhancing *Piper sarmentosum* (Roxburgh) Extracts Pre-Treated with Supercritical Carbon Dioxide and its Correlation with Cytotoxicity and $\alpha$ -Glucosidase Inhibitory Activities

Mohd Shahrul Ridzuan Hamil<sup>1</sup>, Fauziahanim Zakaria<sup>1</sup>, Lee Chong Yew<sup>1</sup>, Norzilawati Pauzi<sup>2</sup>, Khamsah Suryati Mohd<sup>2</sup> and Noor Hafizoh Saidan<sup>3\*</sup>

<sup>1</sup>School of Pharmaceutical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia

<sup>2</sup>Faculty of Bioresources and Food Industry, Universiti Sultan Zainal Abidin, Besut Campus, 22200 UNISZA, Besut, Terengganu, Malaysia

<sup>3</sup>Faculty of Agro Based Industry, Universiti Malaysia Kelantan, 17600 UMK, Jeli, Kelantan, Malaysia

### ABSTRACT

*Piper sarmentosum* was found to have various phytochemicals such as polyphenols and flavonoids but also contains asarone isomers, which were reported as carcinogenic in several preclinical studies. Therefore, reducing asarone isomers from this functional food is important while retaining its functional properties. This study compares the total phenolics, total flavonoids, vitexin and naringenin contents, cytotoxicity, and anti-hyperglycaemic activity between various solvent extracts using a two-step extraction with SC-CO<sub>2</sub> pre-treatment and a single-step conventional solvent extraction without SC-CO<sub>2</sub> pre-treatment. The findings showed that phenolic content was significantly enhanced in ethanol, 50% ethanol, and water extracts pre-treated with SC-CO<sub>2</sub>, significantly correlated with  $\alpha$ -glucosidase inhibitory activity. SC-CO<sub>2</sub> pre-treated extracts enhanced the viability of two normal fibroblasts NIH/3T3 and CCD-18Co cell lines. It is concluded that SC-CO<sub>2</sub> extraction offers a rapid pre-treatment step to produce safer extracts with better quality and efficacy.

### ARTICLE INFO

#### Article history:

Received: 12 January 2023

Accepted: 04 April 2023

Published: 21 December 2023

DOI: <https://doi.org/10.47836/pjst.32.1.18>

#### E-mail addresses:

shahrul.ridzuan18@gmail.com (Mohd Shahrul Ridzuan Hamil)

fauziahanimz@usm.my (Fauziahanim Zakaria)

chongyew@usm.my (Lee Chong Yew)

n.zilawati.pauzi@gmail.com (Norzilawati Pauzi)

khamsahsuryati@unisza.edu.my (Khamsah Suryati Mohd)

hafizoh.s@umk.edu.my (Noor Hafizoh Saidan)

\* Corresponding author

*Keywords:* Cytotoxicity, *P. sarmentosum*, phenolics, supercritical fluid extraction,  $\alpha$ -glucosidase

### INTRODUCTION

Carbon dioxide in supercritical fluid extraction (SFE) has been used as a treatment to lessen or enhance the extraction of natural products using organic solvents (Arumugham et al., 2022; Voung & Roach,

2014). The fundamental benefit of this method is that by adjusting pressure and temperature settings, the physicochemical properties of the solvent may be changed to increase extraction selectivity (Fayaz et al., 2021). SFE was proven to be beneficial in the removal of toxic and undesirable compounds such as aristolochic acid from *Aristolochia* plants (Liang et al., 2010), phthalate plasticiser residue in *Ganoderma lucidum* (Li et al., 2018), and many other undesirable substances from food and herbal-based supplement products (Kobori et al., 2013; Shinde & Mahadik, 2019). Previously, we established a condition to reduce potentially carcinogenic asarone isomers from *Piper sarmentosum* leaves by removing 80.42%–100% of  $\alpha$ -asarone, 91.26%–100% of  $\beta$ -asarone, and 89.24%–100% of  $\gamma$ -asarone by using supercritical carbon dioxide (SC-CO<sub>2</sub>) extraction while applying the Box–Behnken experimental design at optimised pressure ( $P$ ) = 81.16 bar, temperature ( $T$ ) = 50.11°C, and time ( $t$ ) = 80.90 min (Hamil et al., 2022).

*Piper sarmentosum* Roxburgh (Piperaceae), a creeping terrestrial shrub native to Malaysia, Cambodia and Thailand, is widely used in medicinal as well as culinary applications. In traditional medicines, all plant parts are utilised in treating health problems. Leaves and roots of *P. sarmentosum* are usually used to alleviate headaches, toothache, pleurisy, flu, rheumatism, cough and asthma (Ismail et al., 2018). Apart from that, the whole *P. sarmentosum* parts are used as an antispasmodic, expectorant, antifatulence, and appetite enhancer, as well as for treating diabetes in Thailand (Rahman et al., 2016). The roots and leaves were also reported as carminative, stomachic and able to reduce ostalgia (Muhamad & Mustafa, 2010; Ridditid et al., 1998). *Areca catechu* and *P. sarmentosum* roots can be taken together to relieve asthma and cough (Seyyedani et al., 2013). Furthermore, taken together with ginger or nutmeg, it may alleviate toothache and pleurisy. Anti-pyretic and stomachic activities have also been established from the plant leaves in Chinese traditional medicine (Sun et al., 2020).

Traditional and modern medicinal knowledge of the plant has led to scientific investigations of its pharmacological potential. One of the pharmacological properties reported was in treating diabetes or hyperglycaemia. Krisanapun et al. (2012) reported that the aerial part of *P. sarmentosum* water extract (5 and 10 mg/mL) substantially reduced glucose absorption, which was similar to positive control (sodium fluoride). The improvement of glucose metabolism and reduction of its absorption were credited with the hypoglycaemic activity. In another study, oral administration of aqueous extract (125 mg/kg) significantly reduced blood sugar after 7 days of treatment (Peungvicha et al., 1998). Hussan et al. (2013) also reported that 28 days of administration of *P. sarmentosum* leaves water extract (125 mg/kg) in streptozotocin-induced diabetic rats successfully decreased blood sugar levels. In addition, there was an improvement in glomeruli, suggesting that the extract exerted a nephroprotective effect.

SC-CO<sub>2</sub> is a new, simple, yet rapid method in *P. sarmentosum* pre-treatment to remove carcinogenic asarones from the raw leaves selectively. The pre-treated leaves can be re-extracted using solvents to extract beneficial phenolics, flavonoids and other phytochemicals. However, no study is available on the phytochemical and bioactivity profiles of the SC-CO<sub>2</sub> pre-treated *P. sarmentosum* extracts. Therefore, this work focused on evaluating the phytochemical contents, cytotoxicity and anti-hyperglycaemic activity, including  $\alpha$ -glucosidase and  $\alpha$ -amylase-inhibitory assays between three solvent extracts, namely ethanol, 50% ethanol and water extracts with their extracts after being treated with SC-CO<sub>2</sub>.

## MATERIALS AND METHODS

### Plant Material

Fresh *P. sarmentosum* leaves were collected from Batu Gajah, Perak, Malaysia. The plant was identified by a botanist at the School of Biological Sciences, USM, Penang, Malaysia (voucher specimen number 11481). The plant materials were washed thoroughly and dried at 40°C for five days. The thoroughly dried leaves were ground into fine powder and stored in a 2–7°C container before further use.

### Treatment of *P. sarmentosum* Leaves Using SC-CO<sub>2</sub>

Dried powdered leaves (50 g) were extracted using a 1-litre SC-CO<sub>2</sub> extractor (Separex, France). The experimental parameters were set according to the method established by Hamil et al. (2022): pressure = 81.16 bar, temperature = 50.11°C and dynamic extraction time = 80.90 min. The flow rate was maintained at 30 g/min, and the static extraction time was fixed at 30 min. The residue of the leaves was subsequently extracted with conventional solvents for analysis.

### Conventional Extraction of *P. sarmentosum* Leaves

Three types of extraction solvents, ethanol (EM), 50% ethanol (EWM) and water (WM), were used in the present study. First, powdered material (10 g) was macerated with one of the stated solvents or solvent mixtures at 50°C for 24 hrs. The raw material to solvent ratio was fixed at (1:20). The samples were prepared in triplicate. The same extraction methods were performed on the SC-CO<sub>2</sub> treated leaves material for each type of solvent used, namely ethanol maceration residue (EM-R), 50% ethanol maceration residue (EWM-R) and water maceration residue (WM-R). The extracts were concentrated, and their yields were calculated using Equation 1:

$$\text{Yield (\%)} = \frac{\text{Dried extract weight}}{\text{Original sample weight}} \times 100 \quad [1]$$



### **Determination of Total Phenolics Content**

A colorimetric assay was performed to determine total phenolic content (TPC) as Ghasemzadeh et al. (2010) described with slight modification. *P. sarmentosum* extracts (100 µg/mL) were prepared in methanol. Next, distilled water (1 mL) was added, followed by Folin-Ciocalteu phenol reagent (100 µL). The samples were incubated for 5 min in dark conditions before adding 20% sodium carbonate (200 µL). Further incubation was performed in dark conditions for 60 min, and absorbance was measured at 750 nm. Gallic acid prepared in the range of 0.20–200 µg/mL was used to construct the standard calibration curve. Total phenolics were calculated based on the linear regression and expressed as mg of gallic acid equivalent (mg GAE/g).

### **Determination of Total Flavonoids Content**

The aluminium chloride (AlCl<sub>3</sub>) method was used to determine total flavonoid content (TFC), as described by Asha et al. (2010). *P. sarmentosum* extracts (500 µL) were mixed with 1.5 mL of methanol. Then, 10% AlCl<sub>3</sub> and 1M potassium acetate (100 µL each) were added. Absorbance was measured at 415 nm after 30 min incubation against a blank containing all reagents and samples except AlCl<sub>3</sub>. Quercetin in the 3.125–100 µg/mL range was used to construct the standard calibration curve. Total flavonoids were calculated from the linear regression and expressed as mg of quercetin equivalent (mg QE/g).

### **HPLC Analysis of Vitexin and Naringenin**

A gradient HPLC system for the separation of vitexin and naringenin was performed based on the validated method described by Hamil et al. (2016) using an Agilent 1100 HPLC system and ZORBAX Eclipse Plus C-18 column (Agilent, USA). The mobile phase consisted of A (0.1% formic acid) and B (0.1% formic acid in acetonitrile). The mobile phase composition (A: B) was initially set at 85:15 before gradually increasing solvent B to 75% for 10 min, remaining constant for 1 min, and then changing to 85:15 in the final 5 min. An injection volume (10 µL) and flow rate were set at 1 mL/min. The temperature and detector were operated at 30°C and 330 nm.

**Preparation of Reference Markers and Samples.** Vitexin and naringenin were dissolved in HPLC grade methanol (Merck, USA) and mixed to obtain a 1 mg/mL solution. The mixture was then serially diluted (1.56–50 µg/mL) to produce calibration curves. *P. sarmentosum* extracts (2 mg/mL) were prepared in their respective solvents and filtered through 0.45-µm syringe filters prior to analysis.

**Quantification Analysis of Vitexin and Naringenin.** 10 µL of each extract was injected, and peak areas corresponding to the reference markers were recorded. The linear equations

for each marker were used to calculate the concentration of vitexin and naringenin in the samples, and the findings were given as the average of wt/wt% using Equation 2 (n=3):

$$\text{wt/wt\%} = \frac{\text{mass of marker}}{\text{mass of extract}} \times 100 \quad [2]$$

### Cytotoxicity Study

Cytotoxic activities of *P. sarmentosum* extracts were determined using mouse embryo fibroblast (NIH/3T3) and human colon fibroblast (CCD-18Co) normal cell lines using the method proposed by Tajudin et al. (2012). The cells were seeded in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% foetal bovine serum (FBS) with 1% penicillin/streptomycin. Following a 24-hour incubation, different concentrations of *P. sarmentosum* extracts (0–100 µg/mL) were added and further incubated for 72 hrs. Next, 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (MTT) reagent (20 µL) was added and subsequently incubated for 4 hrs. After that, the growth medium was removed, 100% of dimethyl sulfoxide (100 µL) was added and incubated for 15 min. Cell viability was quantified using an M200 ELISA microplate reader (Tecan, Switzerland) at 570 nm and 630 nm. Graphs of the cell's proliferation percentage against the concentration of the extracts were used to determine IC<sub>50</sub>. Doxorubicin was used as a positive control.

### Determination of α-Glucosidase Inhibition Activity

An α-glucosidase inhibition assay was evaluated according to Feng et al. (2011) with some modifications. Baker's Yeast α-glucosidase and P-Nitrophenyl-α-D-glucopyranoside (PNPG) were purchased from Sigma (USA). 20 µL of yeast α-glucosidase (0.8 U/mL) in 0.1 M phosphate buffer (pH 6.8) were added to 120 µL *P. sarmentosum* extracts (50–2000 µg/mL). The mixture was incubated for 15 min. Next, 20 µL of PNPG (0.005 M in 0.1M phosphate buffer at pH 6.8) was added and further incubated at 37°C for 15 min. Eighty µL of 0.05 M sodium carbonate was added to the mixture to stop the reaction, and the absorbance was read at 405 nm. Control samples were prepared without plant extracts, and acarbose was used as a positive control. The percentage of inhibition was calculated according to Equation 3 (n=3):

$$\text{Inhibition (\%)} = \frac{\text{Abs (Control )} - \text{Abs (Extract )}}{\text{Abs (Control )}} \times 100 \quad [3]$$

The IC<sub>50</sub> values were determined from plots of percentage inhibition versus log inhibitor concentration.

### Determination of $\alpha$ -Amylase Inhibition Activity

An  $\alpha$ -amylase inhibition assay was adapted from Bharathi et al. (2014) with slight modification. Different concentrations (100–5000  $\mu\text{g/mL}$ ) of *P. sarmentosum* extracts were prepared in 1 mL DMSO. 500  $\mu\text{L}$  of  $\alpha$ -amylase solution (0.5 mg/mL) in 0.02 M sodium phosphate buffer (pH 6.9) were mixed with 500  $\mu\text{L}$  of the extracts. The mixture was incubated at 25 °C for 10 min. Next, 500  $\mu\text{L}$  of 1% starch solution in 0.02 M sodium phosphate buffer (pH 6.9) was added and incubated at 25°C for 10 min. After that, 3, 5 dinitrosalicylic acid (DNSA) colour reagent (1 mL) was added to stop the reaction. After a 5-minute incubation in a boiling water bath, the tubes were cooled to room temperature. Distilled water (10 mL) was finally added, and absorbance was measured at 540 nm. Acarbose was used as a positive control. The percentage of inhibition by  $\alpha$ -amylase was calculated using Equation 4 (n=3):

$$\text{Inhibition (\%)} = \frac{\text{Abs (Control )} - \text{Abs (Extract )}}{\text{Abs (Control )}} \times 100 \quad [4]$$

The  $\text{IC}_{50}$  values were determined from plots of the percentage of inhibition versus log inhibitor concentration.

### Statistical Analysis

Statistical analyses were conducted using the SPSS 22.0 software package (IBM, USA). Analysis of variance (ANOVA) was performed to estimate significant differences between samples, and unpaired *t*-tests were used to compare two groups (SC-CO<sub>2</sub>-treated vs. untreated). Correlation coefficients between  $\alpha$ -glucosidase and cytotoxicity assay with metabolites profile including total phenolics, flavonoids, vitexin and naringenin were calculated using Pearson correlation analysis. Statistical significance was determined at  $p < 0.05$  in all the analyses.

## RESULTS AND DISCUSSION

### Enhancement of Phenolics and Flavonoids in SC-CO<sub>2</sub> Pre-Treated Extracts

The total phenolic and flavonoid content between EM, EWM, and WM and their SC-CO<sub>2</sub> pre-treated extracts (EM-R, EWM-R, and WM-R) are shown in Table 1. All extracts found total phenolics in a range of 39.29–107.91 mg/g gallic acid equivalent (GAE). Comparatively, total phenolics in EM-R, EWM-R and WM-R were significantly enhanced at  $p < 0.05$ . The highest enhanced total phenolics was obtained in EM-R with 107.91 mg/g compared to EM with 105.52 mg/g. The enhancement in EWM-R (75.56 mg/g) and WM-R (43.26 mg/g) was observed compared to non-treated EWM and WM with 68.11 and 39.29 mg/g, respectively. Total flavonoids ranged between 0.30–1.36 mg/g QE for ethanol, 50%

ethanol, and water extracts. Enhancement of total flavonoids in SC-CO<sub>2</sub> pre-treated extracts was observed in EM-R and EWM-R; however, the differences were not significant.

SC-CO<sub>2</sub> has been proven to enhance the extraction of metabolites from plant materials. It was used as a pre-treatment to remove non-polar compounds from elderberry (Seabra et al., 2010). Subsequent re-extraction of the residue material with 50% ethanol showed an enhanced concentration of phenolic contents. Comparing the two-step extraction to the standard single-step extraction, the total phenolics recovered were substantially higher. The enhanced extraction efficiency of phenolic compounds was attributed to the effective removal of non-polar components by CO<sub>2</sub> and the breakage of the cell walls in plant tissues, which was caused by high pressure (Vatai et al., 2009). Another study reported that an optimised SC-CO<sub>2</sub> pre-treatment removed 80.1% of caffeine from cocoa powder. Besides that, the residual powder re-extracted with solvent retained polyphenols and theobromine in the sample at 84.7% and 94.1%, respectively.

The study reported a significant enhancement in the antioxidant activity of SC-CO<sub>2</sub>-treated cocoa powder compared to the untreated sample (Kobori et al., 2013). Vuong and Roach (2014) also reported that functional polyphenolics such as catechins were maintained and preserved in green tea after decaffeination using SC-CO<sub>2</sub>. Our previous study successfully produced an asarones-free ( $\alpha$ -,  $\beta$ - and  $\gamma$ -asarones) ethanol extract using the two-step extraction technique (Hamil et al., 2022). The present study showed that SC-CO<sub>2</sub> pre-treatment could retain and enhance the quantity of other beneficial components in the sample material, such as polyphenols and flavonoids, apart from removing unwanted compounds.

Table 1

Total phenolics and flavonoids content in *P. sarmentosum* extracts. Results displayed as mean  $\pm$  standard deviation, ( $n = 3$ )

Sample	Total Phenolics (mg/g GAE) $\pm$ SD	Total Flavonoids (mg/g QE) $\pm$ SD
EM	105.52 $\pm$ 0.39	0.30 $\pm$ 0.11
EM-R	107.91 $\pm$ 0.08*	0.51 $\pm$ 0.23
EWM	68.11 $\pm$ 0.16	0.75 $\pm$ 0.05
EWM-R	75.56 $\pm$ 0.32*	0.78 $\pm$ 0.24
WM	39.29 $\pm$ 0.12	1.36 $\pm$ 0.28
WM-R	43.26 $\pm$ 0.19*	1.07 $\pm$ 0.24

Note. The symbol asterisk (\*) indicates statistically significant ( $p < 0.001$ ) compared to their non-treated extracts

### Quantification of Vitexin and Naringenin in *P. sarmentosum* Extracts

Figure 1 shows the concentration of reference markers in *P. sarmentosum* extracts. The peak of naringenin was eluted at 9.40 min, whereas vitexin was observed at 5.80 min. All extracts contained vitexin with a concentration ranging from 0.40%–0.58%. In 50%

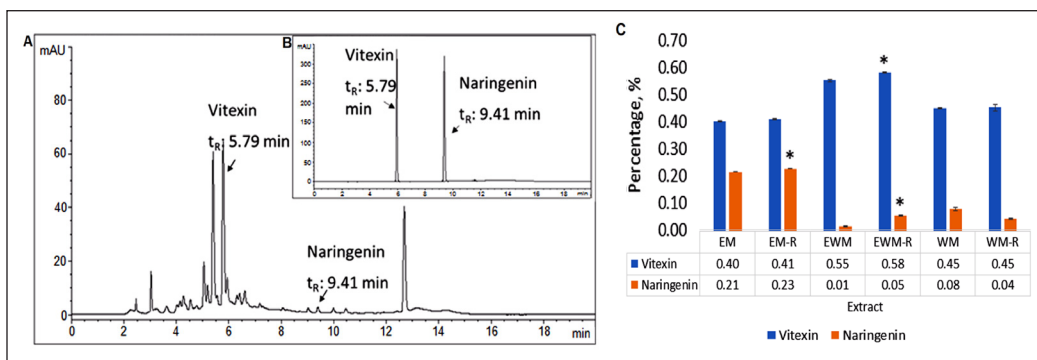


Figure 1. (A) HPLC chromatogram of *P. sarmentosum* ethanol extract, (B) chromatogram of the vitexin and naringenin standards, and (C) amount of vitexin and naringenin in SC-CO<sub>2</sub> pre-treated and non-treated extracts. The symbol asterisk (\*) indicates statistically significant ( $p < 0.01$ ) compared to their non-treated extracts

ethanol extracts, EWM-R showed a significant increase in vitexin ( $p < 0.01$ ). Vitexin was not significantly enhanced in EM-R and WM-R. Naringenin was observed in all *P. sarmentosum* extracts. It was significantly increased in EM-R and EWM-R with 0.23% and 0.05%, respectively. WM-R also shows an increase in naringenin concentration, but there was no significant difference.

Our study's findings were slightly lower than those reported by Ugusman et al. (2012). The authors reported that the vitexin concentration in *P. sarmentosum* water extract was 0.52%. They also reported that naringenin was absent in the extract. On the other hand, vitexin was reported at 0.23% in aqueous extract using a high-speed extractor machine at 80 °C for 3 hours (Azmi et al., 2021). Naringenin was previously reported from *P. sarmentosum* leaves methanol extract; however, data on the quantification of the chemical compound was not available (Subramaniam et al., 2003). *P. sarmentosum*'s flavonoids vary due to both internal and external factors. Some internal factors contributing to phytoconstituent variation are chemotypic, ecotypic, ontogenetic and genotypic. Externally, factors such as growth conditions, collection time, age of plants and storage conditions may affect their chemical compositions (Hussain et al., 2009a). Additionally, the type of solvents also contributed to the extraction efficacy of flavonoids. Generally, flavonoids have poor aqueous solubility (Zhao et al., 2019). They will likely have higher solubility in ethanol, methanol and ethyl acetate. Ethanol used in our study has the highest efficacy in extracting naringenin. Vitexin showed better extraction efficacy in 50% ethanol, possibly due to the presence of a glycoside side chain. Sugar moiety in flavonoids increases the compound's polarity, thus possessing higher solubility in polar substances (Ko et al., 2014).

## Biological Assay of *P. sarmentosum* Extracts

**Cytotoxic Activity of *P. sarmentosum* Extracts using MTT Assay.** The cytotoxic activity of *P. sarmentosum* SC-CO<sub>2</sub> pre-treated and non-treated ethanol, 50% ethanol and water

extracts were determined on normal colon fibroblast (CCD-18Co) and normal fibroblast (NIH/3T3) cell lines (Figure 2). From the experiment, it was found that the percentage viability of both cell lines was enhanced with SC-CO<sub>2</sub> pre-treated *P. sarmentosum* extracts. At 100 µg/mL, all SC-CO<sub>2</sub> pre-treated extracts showed higher cell viability than the non-treated extracts. For NIH/3T3 cells, EWM-R showed significantly higher cell viability at 81.29% compared to EWM at 68.08%. As for the CCD-18Co cell, EM-R and WM-R showed significant cell viability at 80.21% and 96.85% compared to their non-treated counterparts (54.63% and 73.09%), respectively. The positive control, doxorubicin, showed IC<sub>50</sub> at 1.0 µg/mL against NIH/3T3 cell lines and 1.0 µg/mL against CCD18-Co and cell lines, respectively.

Cytotoxicity screening on normal cell lines provides significant information on the substances' potential toxicity and safety profiles. According to the United States National Cancer Institute (NCI), a plant extract is regarded to have an active cytotoxic effect if its IC<sub>50</sub> value is 20 g/mL or below (Boik, 2001). In a previous study, chloroform extract of *P. sarmentosum* showed IC<sub>50</sub> at 64.43 µg/mL against normal human umbilical vascular endothelial cells (HUVEC) (Hussain et al., 2009b). The present study found that SC-CO<sub>2</sub> pre-treated extracts showed higher cell viability than non-treated extracts at different concentrations. However, all the extracts did not exhibit cytotoxic effects up to 100 µg/mL and thus can be considered safe for animal and human consumption.

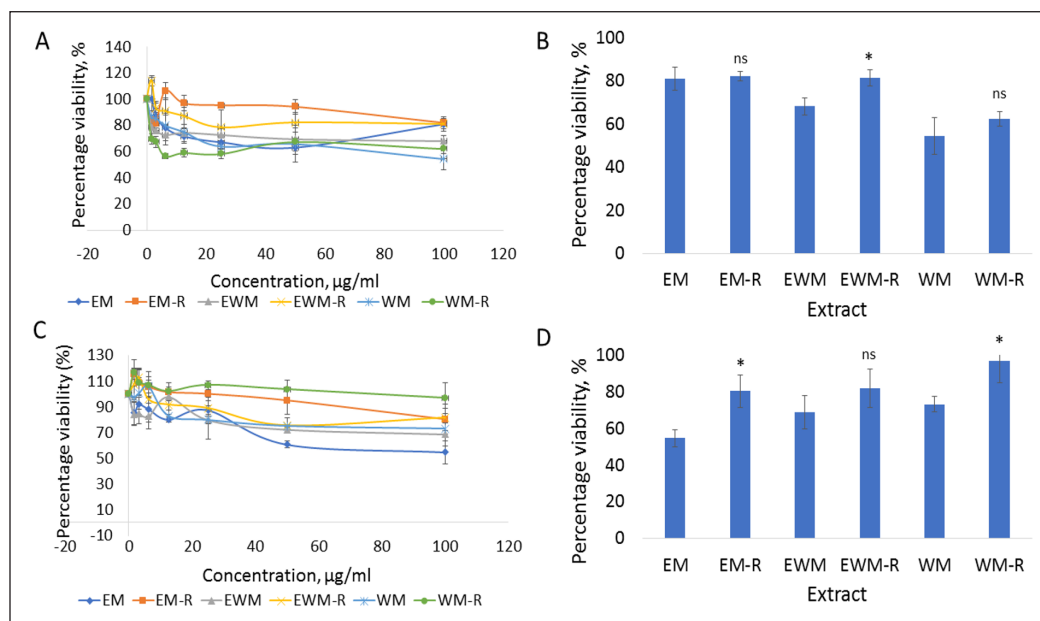


Figure 2. Percentage viability of NIH/3T3 cells treated with *P. sarmentosum* extracts in a dose-dependent manner (A), 100 µg/mL (B), percentage viability of CCD-18Co cells treated with *P. sarmentosum* extracts in a dose-dependent manner (C) and 100 µg/mL (D). The symbol asterisk (\*) indicates statistically significant (p<0.05), and ns indicates not statistically significant (p>0.05) compared to their non-treated extracts



### $\alpha$ -Glucosidase Inhibition Assay

The *in-vitro*  $\alpha$ -glucosidase inhibition assay demonstrated that the ethanol extract exhibited considerable inhibitory activity among all extracts tested. The percentage inhibition of EM-R at 1000  $\mu\text{g/mL}$  was significantly increased (73.38%) compared to EM (57.08%) at  $p < 0.001$  (Figure 3). There was an activity enhancement in EWM-R and WM-R compared to their respective non-treated extracts; however, the inhibition was lower than that of ethanol extracts. The  $\text{IC}_{50}$  values of EM and EM-R were  $1236.21 \pm 8.40 \mu\text{g/mL}$  and  $869.72 \pm 9.39 \mu\text{g/mL}$ , respectively. The  $\text{IC}_{50}$  of the positive control (acarbose) was recorded at  $181.03 \pm 2.98 \mu\text{g/mL}$ . The chemical markers naringenin and vitexin also showed significant  $\alpha$ -glucosidase inhibition activity, with  $\text{IC}_{50}$  values of  $87.35 \pm 1.05 \mu\text{g/mL}$  and  $76.49 \pm 2.32 \mu\text{g/mL}$ , respectively (Table 2).

Flavonoids are well-known as potent  $\alpha$ -glucosidase inhibitors. Chen et al. (2013) reported that a strong  $\alpha$ -glucosidase inhibitory effect ( $\text{IC}_{50}$  61.30  $\mu\text{g/mL}$ ) was obtained from *Microctis folium* extract. Isovitexin, isorhamnetin 3-O- $\beta$ -D-rutinoside and vitexin were discovered as main flavonoids in the plant, which contributed to the activity. Another study demonstrated that 25 mg/kg naringenin administration in T2D rats significantly inhibited  $\alpha$ -glucosidase activity. They mentioned that carbohydrate absorption was successfully delayed, thus decreasing the postprandial blood glucose levels (Priscilla et al., 2014). In a recent study, Daud et al. (2021) reported that phytochemical constituents, including phenolics, flavonoids, terpenoids, as well as saponins and tannins in *P. sarmentosum* extract may associated with the potent  $\alpha$ -glucosidase inhibition activity. Our study showed a strong negative correlation between  $\text{IC}_{50}$  of  $\alpha$ -glucosidase and phenolics content ( $p < 0.01$ ). Naringenin also exhibited significant correlation ( $p < 0.05$ ) with the activity (Table 3). Therefore, this class of compounds *P. sarmentosum* extracts may contribute to the extracts'  $\alpha$ -glucosidase inhibition activity.

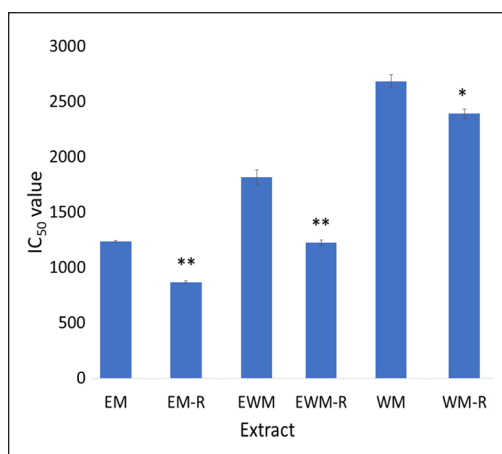


Figure 3.  $\text{IC}_{50}$  values for *in vitro*  $\alpha$ -glucosidase inhibition assay of *P. sarmentosum* extracts. The symbol asterisk (\*) indicates statistically significant ( $p < 0.01$ ), and the symbol asterisk (\*\*) indicates statistically significant ( $p < 0.001$ ) compared to their non-treated extracts

Table 2

$\text{IC}_{50}$  values for *in vitro*  $\alpha$ -glucosidase inhibition assay of chemical markers in *P. sarmentosum* and positive control acarbose. Results are displayed as mean  $\pm$  standard deviation ( $n=3$ )

Sample	$\text{IC}_{50}$ value ( $\mu\text{g/mL}$ ) $\pm$ SD
Vitexin	76.49 $\pm$ 2.32
Naringenin	87.35 $\pm$ 1.05
Acarbose	181.03 $\pm$ 2.98



Table 3

Correlation among the phytoconstituents and activities observed in *P. sarmentosum* extracts

Variables	Phenolics	Flavonoids	Vitexin	Naringenin	NIH/3T3	CCD-18Co	$\alpha$ -glucosidase
Phenolics	1.00						
Flavonoids	-0.65	1.00					
Vitexin	-0.29	0.18	1.00				
Naringenin	0.80*	-0.60	-0.76*	1.00			
NIH/3T3	0.87*	-0.72	0.02	0.55*	1.00		
CCD-18Co	-0.43	0.44	0.16	-0.36	-0.21	1.00	
$\alpha$ -glucosidase	-0.94**	0.74	0.02	-0.62*	-0.94	0.25	1.00

Note. The symbol asterisk (\*) indicates statistically significant ( $p < 0.05$ ); the symbol asterisk (\*\*) indicates statistically significant ( $p < 0.01$ )

### $\alpha$ -Amylase Inhibition Assay

In this assay, all the extracts showed no inhibition towards  $\alpha$ -amylase enzyme up to 5000  $\mu\text{g/mL}$ . The positive control (acarbose) showed inhibition with an  $\text{IC}_{50}$  value of  $325.39 \pm 1.32 \mu\text{g/mL}$ . Concentration above 5000  $\mu\text{g/mL}$  led to high turbidity and poor solubility. Hence, higher concentrations were not determined as they would not produce reliable results. According to Salehi et al. (2013),  $\alpha$ -amylase inhibition activity did not correlate with the extracts' phenolic content. In another study, a combination of quercetin-myricetin, hyperin-avicularin and kaempferol-quercetin did not exhibit synergistic effects against  $\alpha$ -amylase inhibition activities (Wang et al., 2010).

The present result was corroborated by a study conducted by Wongsu et al. (2012). Their study reported that *P. sarmentosum* extract did not inhibit  $\alpha$ -amylase but showed 70% inhibition activity against  $\alpha$ -glucosidase. Similarly, *Phyllanthus amarus*, *Euphorbia hirta*, and *Lagerstroemia speciosa* extracts also showed no inhibition against  $\alpha$ -amylase with mild to moderate  $\alpha$ -glucosidase inhibition despite high levels of tannins, terpenoids and phenolic acids (Binh et al., 2016). It can be postulated that although *P. sarmentosum* extracts showed the presence of vitexin and naringenin, these flavonoids did not contribute to the  $\alpha$ -amylase inhibition activity of the extracts. A good herbal medicine for early management and treatment of hyperglycaemia should have good inhibition activity against  $\alpha$ -glucosidase and moderate  $\alpha$ -amylase inhibitory activity. Inhibition of  $\alpha$ -amylase in excess is not favourable as it could cause upset and discomfort in the stomach (Cheplick et al., 2010).

### CONCLUSION

Using an optimised SC-CO<sub>2</sub> extraction method for the removal of toxic asarone isomers from *P. sarmentosum*, two-step extraction successfully demonstrated the feasibility of using supercritical fluid extraction as a pre-treatment step to enhance total phenolics

content significantly and produced better inhibition against  $\alpha$ -glucosidase compared to the conventional solvent extraction method. This study found that solvents that produced the total phenolic content enhancement showed a significant correlation with  $\alpha$ -glucosidase inhibition activity, with ethanol showing the best activity, followed by 50% ethanol and water extracts. SC-CO<sub>2</sub> pre-treatment also enhances the viability of NIH/3T3 and CCD-18Co cell lines, thus improving the safety profile of the extracts. It is concluded that SC-CO<sub>2</sub> extraction offers a rapid pre-treatment step to produce safer extracts with better quality and efficacy.

## ACKNOWLEDGEMENTS

This work was funded by a Bridging Research Grant from Universiti Sains Malaysia (304/PFARMASI/6316172) and an external grant from SCFE Tech Sdn. Bhd., Pulau Pinang, Malaysia (304/PFARMASI/650856/S154).

## REFERENCES

- Arumugham, T., AlYammahi, J., Rambabu, K., Hassan, S. W., & Banat, S. (2022). Supercritical CO<sub>2</sub> pretreatment of date fruit biomass for enhanced recovery of fruit sugars. *Sustainable Energy Technologies and Assessments*, 52, Article 102231. <https://doi.org/10.1016/j.seta.2022.102231>
- Asha, K., Sucheta, G., Kavita, M., Nirmala, D., & Jyoti, S. (2010). Quantification of phenolics and flavonoids by spectrophotometer from *Juglans regia*. *International Journal of Pharma and Bio Sciences*, 1(3).
- Azmi, M. F., Aminuddin, A., Jamal, J. A., Hamid, A. A., & Ugasman, A. (2021). Quantified *Piper sarmentosum* Roxb. leaves aqueous leaf extract and its antihypertensive effect in dexamethasone-induced hypertensive rats. *Sains Malaysiana*, 50(1), 171-179. <http://dx.doi.org/10.17576/jsm-2021-5001-17>
- Bharathi, A., Roopan, S. M., Vasavi, C. S., Munusami, P., Gayathri, G. A., & Gayathri, M. (2014). *In silico* molecular docking and *in vitro* antidiabetic studies of dihydropyrimido [4, 5-a] acridin-2-amines. *BioMed Research International*, 2014, Article 971569. <https://doi.org/10.1155/2014/971569>
- Binh, T. D., Trinh, D. S., & Anna, K. J. (2016). Screening for potential  $\alpha$ -glucosidase and  $\alpha$ -amylase inhibitory constituents from selected Vietnamese plants used to treat type 2 diabetes. *Journal of Ethnopharmacology*, 186, 189-195. <http://dx.doi.org/10.1016/j.jep.2016.03.060>
- Boik, J. (2001). *Natural Compound in Cancer Therapy*. Oregon Medical Press.
- Chen, Y. G., Li, P., Li, P., Yan, R., Zhang, X. Q., Wang, Y., Zhang, X. T., Ye, W. C., & Zhang, Q. W. (2013).  $\alpha$ -Glucosidase inhibitory effect and simultaneous quantification of three major flavonoid glycosides in *Microctis folium*. *Molecules*, 18(4), 4221-4232. <https://doi.org/10.3390/molecules18044221>
- Cheplick, S., Kwon, Y. I., Bhowmik, P., & Shetty, K. (2010). Phenolic-linked variation in strawberry cultivars for potential dietary management of hyperglycemia and related complications of hypertension. *Bioresource Technology*, 101(1), 404-413. <https://doi.org/10.1016/j.biortech.2009.07.068>
- Daud, D., Jalil, U. J. S. A., Sahrol, N. A. A., Zulkefli, S. Z., Adam, N. A., Ishak, N. K., Mahbob, E. N., & Tawang, A. (2021). Comparative anti-obesity and anti-diabetic properties of *Piper sarmentosum*

- and *Piper betle* aqueous extracts via *in-vitro* system. *Journal of Pharmaceutical Advanced Research*, 4(11), 1421-1427.
- Fayaz, S. M., Abdoli, M. A., Baghdadi, M., & Karbasi, A. (2021). Ag removal from e-waste using supercritical fluid: Improving efficiency and selectivity. *International Journal of Environmental Studies*, 78(3), 459-473. <https://doi.org/10.1080/00207233.2020.1834305>
- Feng, J., Yang, X. W., & Wang, R. F. (2011). Bio-assay guided isolation and identification of  $\alpha$ -glucosidase inhibitors from the leaves of *Aquilaria sinensis*. *Phytochemistry*, 72(2-3), 242-247. <https://doi.org/10.1016/j.phytochem.2010.11.025>
- Ghasemzadeh, A., Jaafar, H. Z., & Rahmat, A. (2010). Antioxidant activities, total phenolics and flavonoids content in two varieties of Malaysia young ginger (*Zingiber officinale* Roscoe). *Molecules*, 15(6), 4324-4333. <https://doi.org/10.3390/molecules15064324>
- Hamil, M. S. R., Memon, A. H., Majid, A. M. S. A., & Ismail, Z. (2016). Simultaneous determination of two isomers of asarone in *Piper sarmentosum* Roxburgh (Piperaceae) extracts using different chromatographic columns. *Tropical Journal of Pharmaceutical Research*, 15(1), 157-165. <https://doi.org/10.4314/tjpr.v15i1.22>
- Hamil, M. S. R., Mohamed, N., & Ismail, Z. (2022). Optimisation of asarone removal from *Piper sarmentosum* Roxburgh leaves using supercritical carbon dioxide extraction: The Box-Behnken design. *Brazilian Journal of Pharmaceutical Sciences*, 58, Article e19212. <https://doi.org/10.1590/s2175-97902022000X2e19212>
- Hussain, K., Ismail, Z., Sadikun, A., & Ibrahim, P. (2009a). Evaluation of metabolic changes in fruit of *Piper sarmentosum* in various seasons by metabolomics using Fourier Transform Infrared (FTIR) spectroscopy. *International Journal of Pharmaceutical and Clinical Research*, 1(2), 68-71.
- Hussain, K., Ismail, Z., Sadikun, A., Ibrahim, P., & Malik, A. (2009b). Cytotoxicity evaluation and characterization of chloroform extract of leaf of *Piper sarmentosum* possessing antiangiogenic activity. *Pharmacologyonline*, 2, 379-391.
- Hussan, F., Zin, N. N. M., Zulkefli, M. R., Yow, S. C., Abdullah, N. A., & Teoh, S. L. (2013). *Piper sarmentosum* water extract attenuates diabetic complications in streptozotocin induced sprague-dawley rats. *Sains Malaysiana*, 42(11), 1605-1612.
- Ismail, S. M., Hui, C. K., Aminuddin, A., & Ugusman, A. (2018). *Piper sarmentosum* as an antioxidant: A systematic review. *Sains Malaysiana*, 47(10), 2359-2368. <http://dx.doi.org/10.17576/jsm-2018-4710-12>
- Ko, M. J., Cheigh, C. I., & Chung, M. S. (2014). Relationship analysis between flavonoids structure and subcritical water extraction (SWE). *Food Chemistry*, 143, 147-155. <https://doi.org/10.1016/j.foodchem.2013.07.104>
- Kobori, K., Maruta, Y., Mineo, S., Shigematsu, T., & Hirayama, M. (2013). Polyphenol-retaining decaffeinated cocoa powder obtained by supercritical carbon dioxide extraction and its antioxidant activity. *Foods*, 2(4), 462-477. <https://doi.org/10.3390/foods2040462>
- Krisanapun, C., Wongkrajang, Y., Tamsiririrkkul, R., Phornchirasilp, S., & Peungvicha, P. (2012). *In vitro* evaluation of anti-diabetic potential of *Piper sarmentosum* Roxb. extract. *The FASEB Journal*, 26(S1), 686.7. [https://doi.org/10.1096/fasebj.26.1\\_supplement.686.7](https://doi.org/10.1096/fasebj.26.1_supplement.686.7)

- Li, P., Liang, Z. H., Jiang, Z., Qiu, Z., Du, B., Liu, Y. B., Li, W. Z., & Tan, L. H. (2018). Supercritical fluid extraction effectively removes phthalate plasticizers in spores of *Ganoderma lucidum*. *Food Science and Biotechnology*, 27(6), 1857-1864.
- Liang, Q., Chow, A. H., Wang, Y., Tong, H. H., & Zheng, Y. (2010). Removal of toxic aristolochic acid components from *Aristolochia* plants by supercritical fluid extraction. *Separation and Purification Technology*, 72(3), 269-274. <https://doi.org/10.1016/j.seppur.2010.02.016>
- Muhamad, Z., & Mustafa, A. M. (2010). *Traditional Malay Medicinal Plants*. Penerbit Fajar Bakti Sdn. Bhd.
- Peungvicha, P., Thirawarapan, S. S., Temsiririrkkul, R., Watanabe, H., Prasain, J. K., & Kadota, S. (1998). Hypoglycemic effect of the water extract of *Piper sarmentosum* in rats. *Journal of Ethnopharmacology*, 60(1), 27-32. [https://doi.org/10.1016/S0378-8741\(97\)00127-X](https://doi.org/10.1016/S0378-8741(97)00127-X)
- Priscilla, D. H., Roy, D., Suresh, A., Kumar, V., & Thirumurugan, K. (2014). Naringenin inhibits  $\alpha$ -glucosidase activity: A promising strategy for the regulation of postprandial hyperglycemia in high fat diet fed streptozotocin induced diabetic rats. *Chemico-Biological Interactions*, 210, 77-85. <https://doi.org/10.1016/j.cbi.2013.12.014>
- Rahman, S. F. S. A., Sijam, K., & Omar, D. (2016). *Piper sarmentosum* Roxb.: A mini review of ethnobotany, phytochemistry and pharmacology. *Journal of Analytical & Pharmaceutical Research*, 2(5), Article 00031. <https://doi.org/10.15406/japlr.2016.02.00031>
- Riditid, W., Rattanaprom, W., Thaina, P., Chittrakarn, S., & Sunbhanich, M. (1998). Neuromuscular blocking activity of methanolic extract of *Piper sarmentosum* leaves in the rat phrenic nerve-hemidiaphragm preparation. *Journal of Ethnopharmacology*, 61(2), 135-142. [https://doi.org/10.1016/S0378-8741\(98\)00025-7](https://doi.org/10.1016/S0378-8741(98)00025-7)
- Salehi, P., Asghari, B., Esmacili, M. A., Dehghan, H., & Ghazi, I. (2013).  $\alpha$ -Glucosidase and  $\alpha$ -amylase inhibitory effect and antioxidant activity of ten plant extracts traditionally used in Iran for diabetes. *Journal of Medicinal Plants Research*, 7(6), 257-266. <https://doi.org/10.5897/JMPR11.1320>
- Seabra, I. J., Braga, M. E., Batista, M. T., & de Sousa, H. C. (2010). Fractioned high pressure extraction of anthocyanins from elderberry (*Sambucus nigra* L.) pomace. *Food and Bioprocess Technology*, 3, 674-683. <https://doi.org/10.1007/s11947-008-0134-2>
- Seyyedani, A., Yahya, F., Kamarolzaman, M. F. F., Suhaili, Z., Desa, M. N. M., Khairi, H. M., Somchit, M. N., Fatimah, C. A., Teh, L. K., Salleh, M. Z., & Zakaria, Z. A. (2013). Review on the ethnomedicinal, phytochemical and pharmacological properties of *Piper sarmentosum*: Scientific justification of its traditional use. *CELLMED*, 3(3), 19-1. <http://dx.doi.org/10.5667/tang.2013.0002>
- Shinde, V., & Mahadik, K. (2019). Supercritical fluid extraction: A new technology to herbals. *International Journal of Herbal Medicine*, 7(1), 27-34.
- Subramaniam, V., Adenan, M. I., Ahmad, A. R., & Sahdan, R. (2003). Natural antioxidants: *Piper sarmentosum* (kadok) and *Morinda elliptica* (mengkudu). *Malaysian Journal of Nutrition*, 9(1), 41-51.
- Sun, X., Chen, W., Dai, W., Xin, H., Rahmand, K., Wang, Y., Zhang, J., Zhang, S., Xu, L., & Han, T. (2020). *Piper sarmentosum* Roxb.: A review on its botany, traditional uses, phytochemistry, and pharmacological activities. *Journal of Ethnopharmacology*, 263, Article 112897. <https://doi.org/10.1016/j.jep.2020.112897>

- Tajudin, T. J. S. A., Mat, N., Siti-Aishah, A. B., Yusran, A. A. M., Alwi, A., & Ali, A. M. (2012). Cytotoxicity, antiproliferative effects, and apoptosis induction of methanolic extract of *Cynometra cauliflora* Linn. whole fruit on human promyelocytic leukemia HL-60 cells. *Evidence-Based Complementary and Alternative Medicine*, 2012, Article 127373. <https://doi.org/10.1155/2012/127373>
- Ugusman, A., Zakaria, Z., Hui, C. K., Nordin, N. A. M. M., & Mahdy, Z. A. (2012). Flavonoids of *Piper sarmentosum* and its cytoprotective effects against oxidative stress. *EXCLI Journal*, 11, 705-714.
- Vatai, T., Škerget, M., & Knez, Ž. (2009). Extraction of phenolic compounds from elder berry and different grape marc varieties using organic solvents and/or supercritical carbon dioxide. *Journal of Food Engineering*, 90(2), 246-254. <https://doi.org/10.1016/j.jfoodeng.2008.06.028>
- Vuong, Q. V., & Roach, P. D. (2014). Caffeine in green tea: Its removal and isolation. *Separation & Purification Reviews*, 43(2), 155-174. <https://doi.org/10.1080/15422119.2013.771127>
- Wang, H., Du, Y. J., & Song, H. C. (2010).  $\alpha$ -Glucosidase and  $\alpha$ -amylase inhibitory activities of guava leaves. *Food Chemistry*, 123(1), 6-13. <https://doi.org/10.1016/j.foodchem.2010.03.088>
- Wongsa, P., Chaiwarit, J., & Zamaludien, A. (2012). *In vitro* screening of phenolic compounds, potential inhibition against  $\alpha$ -amylase and  $\alpha$ -glucosidase of culinary herbs in Thailand. *Food Chemistry*, 131(3), 964-971. <https://doi.org/10.1016/j.foodchem.2011.09.088>
- Zhao, J., Yang, J., & Xie, Y. (2019). Improvement strategies for the oral bioavailability of poorly water-soluble flavonoids: An overview. *International Journal of Pharmaceutics*, 570, Article 118642. <https://doi.org/10.1016/j.ijpharm.2019.118642>



## Estimation of Leachate Volume and Treatment Cost Avoidance Through Waste Segregation Programme in Malaysia

Josfirin Uding Rangga<sup>1,2</sup>, Sharifah Norkhadijah Syed Ismail<sup>1\*</sup>, Irniza Rasdi<sup>1</sup> and Karmegam Karuppiah<sup>1</sup>

<sup>1</sup>Department of Environmental and Occupational Health, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Department of Health Sciences, Faculty of Humanities and Health Sciences, Curtin University Malaysia, CDT 250, 98009 Miri, Sarawak, Malaysia

### ABSTRACT

This study aims to calculate the avoided leachate volume and treatment cost from waste segregation practices compared to the existing waste disposal methods of landfilling. The mathematical equations were used to analyse the waste segregation rate, leachate volume and treatment cost. The study findings reveal that 99.4% of mixed waste was disposed of in landfills, resulting in an annual generation of 565 thousand cubic meters (m<sup>3</sup>) of leachate with an estimated treatment cost of MYR 19.82 million (USD 4.36 million). The segregated waste, which accounts for only 0.06%, reduces 354 m<sup>3</sup> in leachate volume and a cost-saving of MYR 12.42 thousand (USD 2.73 thousand) in the treatment expenses per year. The findings concluded that waste segregation practice could reduce waste management costs by reducing leachate production and treatment costs and environmental impacts.

*Keywords:* Environment, health, leachate generation, recycling, waste

### ARTICLE INFO

#### Article history:

Received: 16 January 2023

Accepted: 14 June 2023

Published: 21 December 2023

DOI: <https://doi.org/10.47836/pjst.32.1.19>

#### E-mail addresses:

josfirin@curtin.edu.my (Josfirin Uding Rangga)

norkhadijah@upm.edu.my (Sharifah Norkhadijah Syed Ismail)

irniza@upm.edu.my (Irniza Rasdi)

megam@upm.edu.my (Karmegam Karuppiah)

\* Corresponding author

### INTRODUCTION

Globally, 2.01 billion tonnes of municipal solid waste are generated annually, and this figure is projected to increase to 3.04 billion tonnes by 2050. The worldwide waste generation per capita ranges from 0.11 to 4.54 kg per day. It is estimated that waste generation will increase by 19% in high-income countries, while low- and middle-income countries are expected to experience



a growth of over 40% by 2050 (Kaza et al., 2018). In Malaysia, the available data showed a significant increase in daily waste generation in recent years, from 36,843 tonnes in 2018 to 39,936 tonnes in 2022 (Bernama, 2022). In a year, the country generates about 12,840,000 million t/yr. Of waste (Jain, 2017). A Malaysian generates about 1.03 kg of waste per day, which is higher than Japan (0.98 kg/capita/day), China (0.63 kg/capita/day), and India (0.5 kg/capita/day) (Hoornweg & Perinaz, 2012). The Malaysian recycling rate has increased from 24.6% in 2018 to 31.5% in 2021; by 2025, the government will achieve a 40% recycling rate (Jamliah, 2022; Kamel, 2021). Other countries, such as Germany, South Korea, and Austria, have a higher recycling rate than Malaysia, which are 67.1%, 59.7%, and 59%, respectively (Tiseo, 2022).

Approximately 37% of global waste is disposed of in landfills, 31% is dumped in open sites, 19% is recycled and composted, and 11% is disposed of through incineration. The remaining 8% of waste is disposed of in sanitary landfills with landfill gas collection systems (Kaza et al., 2018). About 65% of waste in Malaysia comes from households, 28% from commercial and institutions, and 7% from industrial waste (National Solid Waste Management Department, 2013). The country disposes of 69% of waste in landfills, which does not include illegal or open dumping waste where the amount may be higher than reported (Jamliah, 2022). There are 158 active landfills throughout the country, with 141 non-sanitary and 17 sanitary landfills (National Solid Waste Management Department, 2021). Landfilling is the primary waste disposal method in the country due to its low maintenance cost and ease of operation (Imran et al., 2019). However, it is not the best solution for waste management, as landfills cause environmental degradation. One of the landfills' problems is leachate generation and uncontrolled discharging into the environment. A recent study indicated that an effective solid waste management system should encompass environmentally friendly and cost-effective solutions while requiring collaboration among government, non-governmental agencies, and the community. The study emphasised that this can be accomplished by implementing waste segregation using the 4Rs approach: reduce, reuse, recycle, and recover (Mor & Ravindra, 2023).

Landfill leachate is the liquid formed when rainwater filters through waste in a landfill. When this liquid encounters buried waste, it leaches or draws out chemicals or constituents from it (USEPA, 2023). The volume of leachate and its compositions depend on the waste compositions, moisture availability, age of landfill sites, rainfall intensity, and land area. Organic waste and construction and demolition debris contribute to organic and inorganic constituents in leachate; for example, disposal of lead batteries, cans, steel scrap, electronic devices, thermometers, fluorescent lights, and other products in a landfill will generate leachate containing heavy metals (Maiti et al., 2016). The high moisture of the waste or landfill causes a high volume of leachate due to the moisture in the stabilisation rate and supports the methanogenic fermentation of solid waste (Adhikari et al., 2014). In terms of

landfill age, leachate generation decreases over time due to waste stabilisation and reduced filtration rate (Aziz et al., 2012). Even a closed landfill produces some leachate (Ibrahim et al., 2017). Rainfall intensity influences leachate generation because intensive precipitation increases the volume of leachate (Aziz et al., 2012). A large landfill area also generates a high volume of leachate (Ibrahim et al., 2017). Its compositions are categorised into three groups: organic matters (i.e., biochemical oxygen demand, chemical oxygen demand), inorganic matters (i.e., phosphates, sulphates, ammonia, nitrate, nitrogen, and chlorides), and heavy metals (i.e., cadmium, chromium, lead, nickel, zinc, mercury, copper) (Maiti et al., 2016; Magda & Gaber, 2015).

Leachate generation in landfills has become a concern as improper landfills and leachate management will lead to environmental pollution, such as groundwater and soil pollution. For instance, a study conducted by Suratman et al. (2011) in Malaysia reported that leachate quality from most non-sanitary landfills exceeded the standard limits, and water sources were polluted. Non-sanitary landfills in Malaysia have a more severe impact on surface and groundwater contamination than sanitary landfills (Taha et al., 2011). For example, a study in Nigeria determined a high concentration of heavy metals in landfill leachate and water samples (Aderemi et al., 2011). A study by Sharifah et al. (2015) also reported a high concentration of heavy metals in Malaysia's topsoil of non-sanitary landfills. The contaminated groundwater with heavy metals that serve as a source of drinking water, as well as the accumulation of heavy metals in soil and vegetables in high amounts, can cause adverse health effects to both animals and human beings (Foufou et al., 2017; Khaled & Muhammad, 2016).

Thus, it is necessary to design the engineering control of leachate pollution properly and build treatment facilities to reduce these adverse environmental impacts. For example, properly lining landfill cells and leachate ponds can minimise groundwater contamination (Magda & Gaber, 2015). A combination of conventional (i.e., nitrification and denitrification process) and advanced technologies (i.e., integrated method of physiochemical, biological, chemical, and physical treatment) provide a sustainable, cost-effective, and environmentally friendly (Aziz & Ramli, 2018; Show et al., 2019). Despite this, developing countries like Malaysia face barriers to these good facilities, such as budget constraints and trained operators (Aziz & Ramli, 2018). On average, the Malaysian government requires USD5 billion per year to manage waste in the country (Utusan Online, 2017). The fee for leachate treatment alone is about 35 Malaysian Ringgit (MYR) per cubic meter (m<sup>3</sup>) (Ministry of Housing and Local Government, 2015).

Waste minimisation and recycling are proven to be the best methods to reduce the amount of waste being dumped in landfills as well as reduce the generation of leachate. Under the Solid Waste and Public Cleansing Management Act, 2007 (Act 672), waste segregation at the source is mandatory, as stated in Section 74 (1) and (2). Households need

to segregate their waste into food, paper, plastic, miscellaneous, and non-recyclable waste, and noncompliance may result in a fine of no more than RM 1000. It is crucial to enhance waste segregation programmes to facilitate the separation of recyclable materials, organic waste, and hazardous waste for recycling, composting, or treatment. These approaches will aid in reducing the overall volume of waste being disposed of in landfills, thus promoting resource conservation and minimising the environmental impacts associated with waste generation (Coelho & Lange, 2018; Slagstad & Brattebø, 2012).

Proper waste segregation enables effective recycling and resource recovery, which helps conserve natural resources, save energy, reduce the demand for raw materials, and mitigate emissions (Liikanen et al., 2018). The programmes also promote the separate collection and treatment of organic waste, reducing the volume of waste in landfills and contributing to renewable energy generation in the form of biogas through anaerobic digestion treatment (Liu et al., 2017). Diverting organic waste from landfills significantly reduces the generation of greenhouse gas emissions. The study by Turner et al. (2015) reports that recycling source-segregated materials can reduce 70% of GHG emissions and global warming potential. Similarly, achieving a source-separated collection efficiency of 25% for food waste and 17% for other wastes can help avoid 77% of GHG emissions (Jun et al., 2013). The government can implement this strategy as part of its efforts to mitigate climate change and achieve environmental goals in Malaysia.

Separating hazardous materials, such as batteries, chemicals, and electronic waste, at the source enables proper handling and disposal, thereby preventing environmental and human health harm (Kaya, 2016; Pathak et al., 2019). For example, increasing the source segregation intensity from 25% (74.6% residual MSW, 19.0% paper, 6.40% packaging) to 52% (51.9% residual MSW, 11.6% paper, 12.1% light packaging, 17.9% organic, 3.30% glass, 3.20% metals) could result in significant reductions in the potential for human toxicology, terrestrial eco-toxicity, GHG emissions, and leachate production (Maria & Micale, 2014). Environmental pollution, such as soil and water pollution, can be prevented if hazardous and non-hazardous are properly managed.

Promoting waste segregation practices has a positive impact on raising public awareness and education. It helps foster a sense of environmental responsibility and sustainability among individuals and communities. Increased awareness can drive long-term behaviour change and cultivate a culture of sustainable waste management practices (Debrah et al., 2021; Knickmeyer, 2020). In addition, waste segregation programmes support the principles of a circular economy by promoting the recovery of valuable resources from waste streams (Ministry of Economy, 2021). Separating recyclable materials to be recycled and reused reduces the demand for virgin resources. It conserves natural resources and reduces energy consumption and associated environmental impacts from resource extraction and manufacturing processes (Ministry of Economy, 2021). Thus, Malaysia

needs to implement waste segregation effectively to aid the country in moving towards sustainable waste management practices as well as contributing to achieving national and international environmental goals (Ministry of Economy, 2021; Ministry of Housing and Local Government, 2015).

There are limited studies done to elucidate the contributions of waste segregation to the avoided leachate generation in landfills and the cost savings to the leachate treatment, especially in Malaysia. The available study by Abunama et al. (2018) only determined the elements that contribute to leachate production, such as waste volume, rainfall level, and emanated gases, using the developed model of adaptive neural fuzzy inference system (ANFIS). A previous study by Ibrahim et al. (2016) only estimated the volume of leachate generated by closed landfills in Selangor using simple mathematical models where rainfall level and landfill size were considered. Another study by Aziz et al. (2012) also analysed the leachate generation rate by applying the field measurements and the Water Balance Method (WBM). A recent study conducted in Malaysia has shown that the commercial and industrial sectors actively reuse and recycle end-of-life vehicles (ELVs) to reduce waste disposal. It is achieved through donating, selling, trading, and sending unused vehicle parts to workshops (Othman et al., 2021). Most studies did not focus on the cost estimation for the leachate treatment. Recent studies on waste management have reported the development of composites from plastic waste and the transformation of waste into green building materials (Soni et al., 2022; Soni et al., 2023). These approaches can reduce the volume of waste as well as leachate generation when waste segregation and recovery are effectively implemented.

Thus, the present study aims to estimate the volume of leachate generated from landfilled waste and the avoided volume from segregated waste. The study intended to estimate the cost of leachate treatment through waste landfilling and cost savings through waste segregation. Waste segregation efficiency is the key factor in estimating leachate volume in a waste segregation program in Malaysia. Efficient segregation methods help divert recyclable and biodegradable waste from landfills, reducing waste volume and potential leachate generation. The second factor is the waste generation rate. The volume of waste generated is a crucial factor in estimating leachate volume. The higher the amount of waste generated, the greater the potential for leachate production. In addition, the composition of the segregated waste plays a significant role in estimating leachate volume. The estimations will aid authorities in designing and operating effective treatment systems, determining treatment requirements such as the capacity of the leachate collection system, and facilitating financial planning by providing cost estimates for maintenance and operational budgeting. The waste volume (tonnes) data was obtained from SWCorp and analysed using the adapted mathematical models. Six states of Malaysia and two Federal states were selected as the study area as the waste segregation programme under

Act 672 has been imposed in these states since 2015. The study’s findings are expected to aid the authorities in planning waste management strategies regarding waste reduction and environmental impact protection. A model of leachate quantification and treatment cost estimation could also serve as the baseline data expected to stimulate more research on waste management in Malaysia.

## METHODOLOGY

### Study Area

The states under Act 672 were selected as the study area. The selected states were Kuala Lumpur (KUL), Putrajaya (PJY), Pahang (PHG), Perlis (PLS), Kedah (KDH), Negeri Sembilan (NSN), Melaka (MLK), and Johor (JHR) as shown in Figure 1. Pahang is the largest state, with around 35,965 square kilometres (km<sup>2</sup>). The second-largest state is Johor (19,166 km<sup>2</sup>), followed by Kedah (9,492 km<sup>2</sup>), Negeri Sembilan (6,656 km<sup>2</sup>), Melaka (1,720 km<sup>2</sup>), Perlis (816 km<sup>2</sup>), Kuala Lumpur (243 km<sup>2</sup>) and Putrajaya (49 km<sup>2</sup>) (Department of Statistics Malaysia, 2019b). Johor has the largest population (3,776.0 million), followed by Kedah (2,178.7 million), Kuala Lumpur (1,796.7 million), Pahang (1,679.7 million),

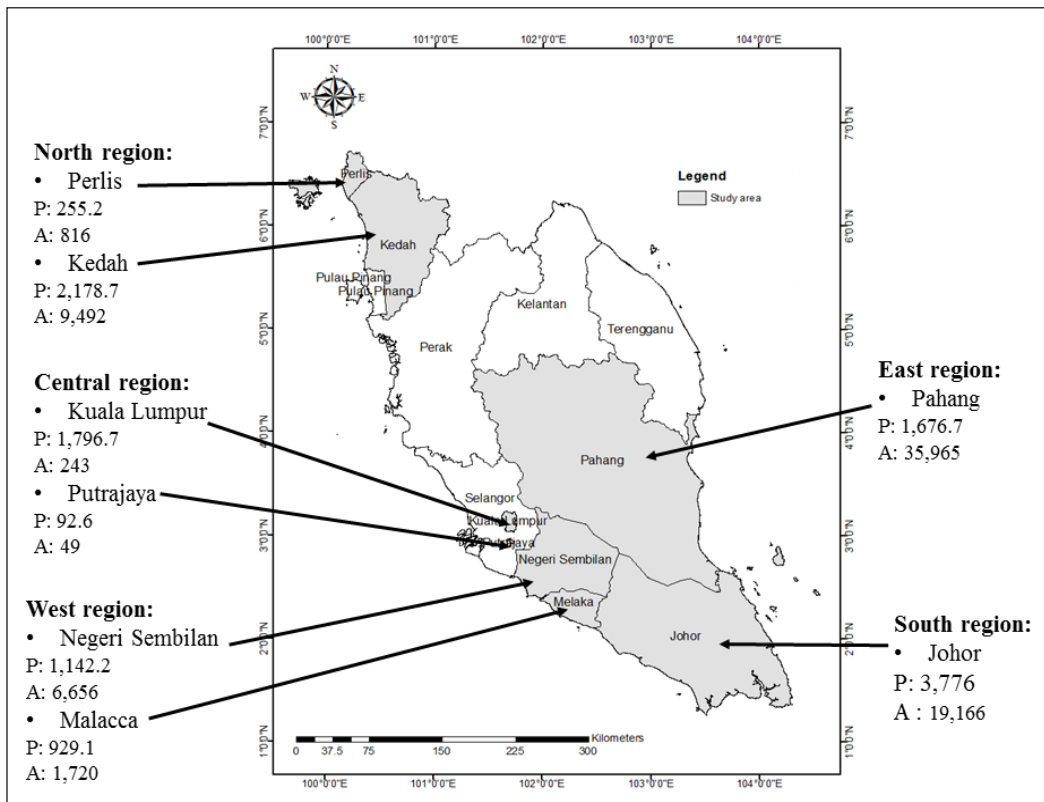


Figure 1. The selected study area (states) in west Malaysia: P: population ('000); A: area (km<sup>2</sup>).

Negeri Sembilan (1,142.2 million), Perlis (255.2 million) and Putrajaya (92.6 million) (Department of Statistics Malaysia, 2019a).

The states were chosen because household waste segregation at source is mandatory under Act 672. Selangor, Perak, Penang, Terengganu, and Kelantan were excluded as they still did not conform to the Act when the study was conducted. The Act only applied to Peninsular Malaysia. Thus, the Borneo regions of Malaysia (Sabah and Sarawak) were excluded from the study as both states have their law and regulations.

### Study Variables and Mathematical Models

**Segregation Rate (SR).** The segregation rate (%) was determined using Equation 1, adapted from the National Solid Waste Management Department (2013). Waste segregation rate (SR) is defined as the total segregated waste (SegW) volume divided by total solid waste volume (tSW) multiplied by one hundred. The tSW is the total volume of landfilled and segregated waste (tonne).

$$SR = \sum \left[ \frac{SegW}{tSW} \right] \times 100 \quad (1)$$

**The Volume of Leachate (VL).** The mathematical model of Equation 2 was used to estimate the volume of leachate generated in landfills. It was adapted based on the reported value by the KPKT (2015). Where the volume (m<sup>3</sup>) of leachate (VL) is equal to the volume (tonne) of solid waste (SW) multiplied by 0.21 cubic meters (m<sup>3</sup>t<sup>-1</sup>), a tonne (t) of waste could generate 0.21 cubic meters (m<sup>3</sup>) of leachate (KPKT, 2015).

$$VL = \sum SW \times 0.21 \quad (2)$$

**Leachate Treatment Cost (LTC).** The leachate treatment cost (LTC) was calculated using Equation 3. The costs were estimated based on the volume (m<sup>3</sup>) of leachate (VL) from Equation 2 multiplied by the treatment price (TP). In Malaysia, the TP is RM35/m<sup>3</sup> of leachate (KPKT, 2015).

$$LTC = \sum VL \times TP \quad (3)$$

The mathematical models are summarised in Table 1.

### Data Collection and Calculation

The solid waste data (tonnes) of the selected states were obtained from SWCorp as presented in Table 2, as first published in the Journal of Material Cycles and Waste Management,

Table 1  
The summary of mathematical models

Parameter	Mathematical models	Source
Segregation rates (%)	$SR = \sum \left[ \frac{SegW}{tSW} \right] \times 100 \quad (1)$ <p>SR = segregation rates (%)                      SegW = volume of segregated waste (tonne)                      tSW = total solid waste volume (tonne)                      100 = to obtain the percentage (%)</p>	(Ministry of Housing and Local Government, 2015)
Leachate production (m <sup>3</sup> )	$VL = \sum SW \times 0.21 \quad (2)$ <p>VL = volume of leachate (m<sup>3</sup>).                      SW = volume of solid waste (tonne).                      0.21 = volume of leachate per tonne of waste.</p>	(Ministry of Housing and Local Government, 2015)
Leachate treatment (MYR (USD))	$LTC = \sum VL \times TP \quad (3)$ <p>LTC = leachate treatment cost (MYR (USD)).                      VL = estimated volume of leachate (m<sup>3</sup>) using Equation 2.                      TP = leachate Treatment Price (RM35 m<sup>-3</sup>).</p>	(Ministry of Housing and Local Government, 2015)

Table 2  
The volume of landfilled (L) and segregated (S) waste

Waste (tonnes)	KUL	PJY	PHG	PLS	KDH	NSN	MLK	JHR
2014 L	680,806.63	26,193.53	302,233.41	49,440.65	309,494.57	218,617.04	360,734.76	810,829.75
2015 L	613,561.50	28,649.16	308,009.50	63,792.00	433,996.39	233,137.48	313,950.85	818,515.73
S	140.02	16.26	134.72	4.29	50.57	45.54	46.41	60.52
2016 L	614,357.93	29,360.29	247,664.08	41,269.00	424,287.45	245,802.71	214,688.00	912,033.82
S	429.38	89.79	385.50	6.25	178.07	231.41	238.55	426.11
2017 L	610,849.75	28,903.84	242,309.86	30,366.10	345,456.08	242,538.16	218,275.08	839,252.11
S	342.75	168.18	353.16	23.40	134.95	283.70	284.43	667.50
2018 L	610,068.90	29,360.29	229,811.12	30,118.21	388,397.00	252,529.82	203,308.89	869,692.81
S	272.90	94.18	212.59	28.35	131.48	216.38	261.41	795.16



2022 by Springer Nature (Rangga et al., 2023). The MSW (in tonne) was categorised as landfilled waste (i.e., MWS that had been disposed of in landfills) and segregated waste (i.e., separated waste or materials for recycling purposes). The available data at the time of the study was only from 2014 to 2018. Microsoft Excel 2016 was used to calculate the parameters of the study. The mathematical models were inserted and programmed in the Excel “function formula” to avoid any errors during the analysis.

## RESULTS AND DISCUSSION

The present study reported that the country’s overall waste segregation rate was 0.06% (Figure 2). It was analysed based on the segregated waste volume divided by the total volume of waste. The waste segregation rate increased from 0.02% in 2015 to 0.06% in 2018 after the government made it compulsory by implementing the Act 672. Still, enormous amounts of waste were deposited in landfills, averaging 99.94% annually. Based on the reported data, the amount of waste generated and disposed of in the country’s landfills increased from 33 thousand tonnes per day in 2012 to 39 thousand tonnes per day in 2022 (Hassan et al., 2022).

Low segregation practices in the country were influenced by knowledge, attitude, awareness, facilities, and incentives (Irina et al., 2014; Azilah et al., 2015; Low et al., 2016; Malik et al., 2015).

The findings in Figure 3 show that the volume of leachate was reduced from 579 thousand cubic metres ( $m^3$ ) in 2014 to 548 thousand  $m^3$  in 2018, which parallels the reduction in the volume of waste disposed of in landfills through waste segregation. About 565 thousand  $m^3$  per year ( $m^3/yr.$ ) of leachate is generated in landfills. The waste segregation practice increased the avoided volume of leachate discharge into the environment from 104  $m^3$  in 2015 to 422  $m^3$  in 2018. On average, 354  $m^3/yr$  of leachate was avoided.

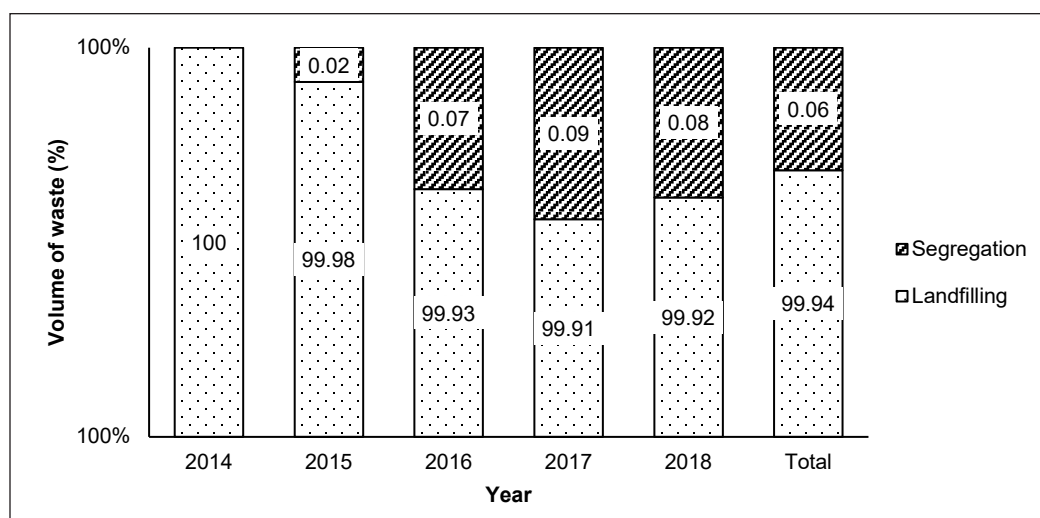
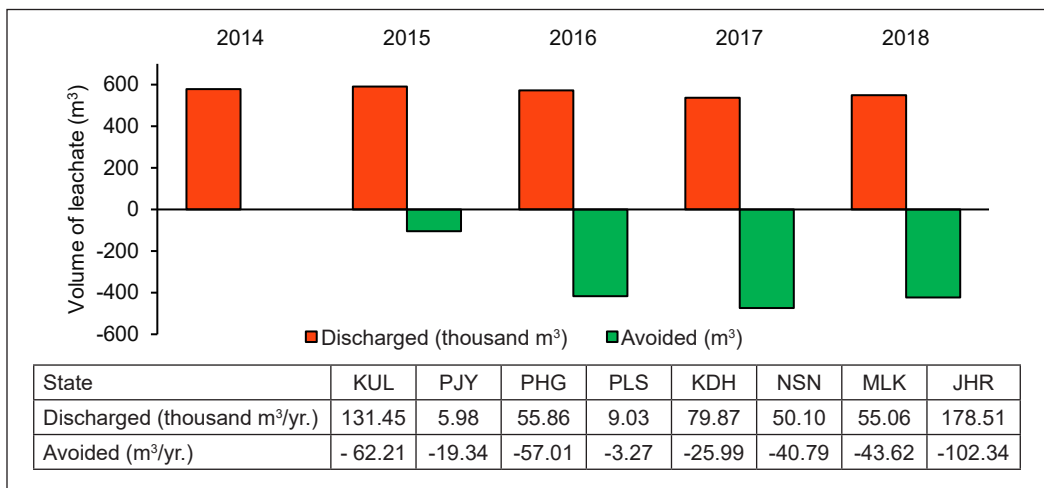
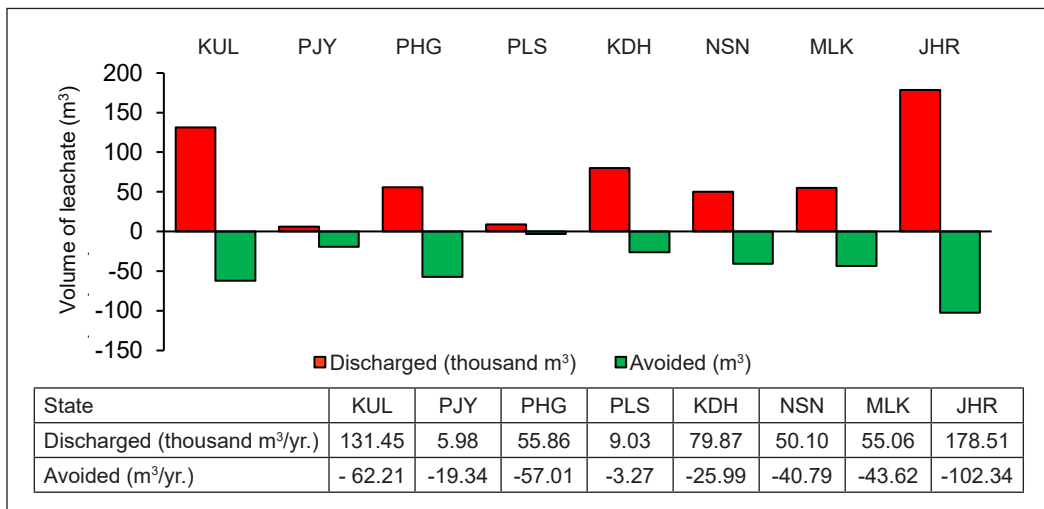


Figure 2. Waste segregation rates

Among the states under Act 672, Johor produced the highest volume of leachate (about 178 thousand m<sup>3</sup>/yr.) as the state disposed of the highest volume of waste in the landfills. It was followed by Kuala Lumpur (104 thousand m<sup>3</sup>/yr.), Kedah (79 thousand m<sup>3</sup>/yr.), Pahang (55 thousand m<sup>3</sup>/yr.), Negeri Sembilan (50 thousand m<sup>3</sup>/yr.), Melaka (55 thousand m<sup>3</sup>/yr.), Perlis (nine thousand m<sup>3</sup>/yr.) and Putrajaya (five thousand m<sup>3</sup>/yr.), as shown in Figure 3. Johor avoided a high volume of leachate discharge into the environment (102 m<sup>3</sup>/yr.). Waste segregation practices in other states also avoided the production of leachate, for example, Kuala Lumpur (62 m<sup>3</sup>/yr.), Pahang (57 m<sup>3</sup>/yr.), Melaka (43 m<sup>3</sup>/yr.), Negeri Sembilan (40 m<sup>3</sup>/yr.), Kedah (25 m<sup>3</sup>/yr.), Putrajaya (19 m<sup>3</sup>/yr.), and Perlis (3.27 m<sup>3</sup>/yr.).



(a)



(b)

Figure 3. The estimated volume of leachate (m<sup>3</sup>) generated (a) from 2014 to 2018 and (b) among the states under Act 672. Negative (-) indicates the avoided volume of leachate discharged in landfills

A previous study by Ibrahim et al. (2017) reported that six Selangor landfills produced an average of 3,936 m<sup>3</sup>/yr of leachate by considering the factors of rainfall intensity and landfill area. The present study used the mathematical models from the literature in which one tonne of waste could generate 0.21 cubic meters (m<sup>3</sup>) of leachate (Ministry of Housing and Local Government, 2015). The study estimated the volume of leachate based on the volume of waste generated. Similarly, a study in Jordan proved that 30 kilograms of waste (52% organic waste, 13% paper, 17% plastic, 3% glass, 1% metals, and 14% others) generated about 15.7 litres of leachate (i.e., equivalent to 1 tonne of waste generates 0.52 m<sup>3</sup> of leachate) after 200 days of decomposition (Aljaradin & Persson, 2016). Therefore, the present study's findings can be used as the baseline data to improve the waste management strategy. It shows that waste segregation can increase the recovery of recyclable materials and waste deviation from landfills and reduce leachate generation in landfills.

Improper waste segregation leads to a high generation of landfill leachate volume but also causes adverse impacts on the environment and human health. Landfill leachate contains heavy metals (such as cadmium, copper, lead, and zinc), organic matter, and inorganic matter, resulting from the disposal of mixed waste in landfills (Maiti et al., 2016; Magda & Gaber, 2015), which leads to contamination of soil and water bodies. A recent study reported that heavy metal concentrations in leachate from Ethiopia's dump site exceeded the permissible limits set by WHO standards, with nearby stream water highly polluted with chemicals, having chemical oxygen demand (COD) values ranging from 10 mg/l to 61 mg/l in all water samples (Mekonnen et al., 2020). A study in Saudi Arabia also revealed high concentrations of heavy metals in leachate, soil, and waste samples. They found that 59.3% of water samples were contaminated, rendering the groundwater unsafe for drinking (Alghamdi et al., 2021).

In Malaysia, a recent study reported that the leachate from the Simpang Renggang landfill had high volumes of COD (1712–1633 mg/l), biological oxygen demand (137–138 mg/l), ammonia (405 mg/l), and suspended solids (341–359 mg/l), exceeding Malaysian standards (Detho et al., 2020; Zaini et al., 2022). The disposal of a large volume of mixed waste in landfills also contributes to the emission of greenhouse gases (GHGs) and other toxic pollutants such as non-methane organic compounds (NMVOCs) (Wu et al., 2018; Cheng et al., 2019). People living near landfill sites are exposed to landfill pollutants through inhalation, ingestion, and dermal contact, resulting in a higher risk of carcinogenic and non-carcinogenic health issues, especially among children, who are a vulnerable group, compared to adults (Nai et al., 2021; Obiri-Nyarko et al., 2021; Yao et al., 2019).

Therefore, comprehensive mitigation strategies are necessary to ensure proper waste segregation practices and minimise the risks associated with waste management. One such strategy is encouraging and mandating waste segregation at the source. The community must be aware of and actively participate in waste segregation, particularly when it comes

to food waste, which accounts for 45% of the total waste generated (National Solid Waste Management Department, 2013). In Malaysia, approximately 17,000 tonnes of food waste are generated daily (Noor, 2022). Malaysian waste composition shows a high moisture content, with 59.45% being a mix of 45% food waste, 13% plastic, 9% paper, 12% diapers, 6% garden waste, and 16% other materials (e.g., glass, metal, leather, wood & rubber textiles). The high moisture content of the waste contributes to increased degradation and leachate production in landfills (Aljaradin & Persson, 2016). Waste segregation helps reduce the moisture content and the amount of waste disposed of in landfills. Waste segregation alone is insufficient; it must be integrated with other strategies to achieve sustainable waste management. An effective combination of recycling, composting, anaerobic digestion, and incineration, supported by efficient separate waste collection, is considered the best option for integrated solid waste management (Coelho & Lange, 2018). A study by Yong et al. (2015) stated that a combination of waste separation at the source with other approaches has a significant environmental impact, saving on global warming (30%), acidification (18%), nutrient enrichment (28%), and photochemical ozone formation (29%).

Another mitigation strategy to minimise the environmental impact of landfill leachate is the utilisation of effective and efficient leachate treatment methods prior to discharging it into the environment. Various leachate treatment technologies, such as photocatalysis, advanced oxidation processes (AOPs), electrooxidation (EO), electrocoagulation, and bio-granulation, can be practised, and improvement is required to increase the efficiency in leachate management (Bandala et al., 2021). Previous studies have shown that photocatalysis can remove 61% to 85% of COD and 52% of phthalates in leachate (Azadi et al., 2021; Eslami et al., 2021). AOPs, specifically using supercritical water oxidation (ScWO) as a reagent, can remove 96% of phosphorus, 70% of nitrate, 94% of total dissolved solids, and 87% of true colour (Martins et al., 2020). EO can effectively remove 68% of COD and 40% of total organic carbon (Yan et al., 2021). Electrocoagulation, as indicated in a study by Dia et al. (2018), can reduce turbidity by 82% and Zn content by 95% in leachate. Moreover, bio-granulation can eliminate nearly 100% of ammonia and 90% of COD in leachate (Moa et al., 2021). Thus, treating landfill leachate before its discharge into the environment is crucial, albeit incurring some costs to operate the treatment facilities.

Table 3 indicates the estimated costs of leachate treatment for different regions. Johor is projected to have the highest cost of leachate treatment, requiring MYR 6.25 million (USD 1.38 million) annually. Kuala Lumpur follows with MYR 4.60 million (USD 1.01 million) yearly, and Pahang with MYR 1.96 million (USD 0.43 million) yearly. On average, the total annual cost for leachate treatment amounts to approximately MYR 19.82 million (USD 4.36 million). Implementing waste segregation practices that divert waste from landfills can help avoid these leachate treatment costs, saving about MYR 12.42 thousand (USD 2.73 thousand) annually. Among the regions, Johor shows

Table 3

*The estimated cost of leachate treatment, MYR (USD)*

VARIABLE	STATE	Leachate treatment cost	
		MYR	USD
Landfilled waste (million/year)	KUL	4.60	1.01
	PJY	0.21	0.05
	PHG	1.96	0.43
	PLS	0.32	0.07
	KDH	2.80	0.62
	NSN	1.75	0.39
	MLK	1.93	0.42
	JHR	6.25	1.38
	<b>TOTAL</b>	<b>19.82</b>	<b>4.36</b>
Segregated waste (thousand /year)	KUL	2.18	0.48
	PJY	0.68	0.15
	PHG	2.00	0.44
	PLS	0.11	0.02
	KDH	0.91	0.20
	NSN	1.43	0.31
	MLK	1.53	0.34
	JHR	3.58	0.79
	<b>TOTAL</b>	<b>12.42</b>	<b>2.73</b>

the highest savings in leachate treatment costs, amounting to MYR 2.18 thousand (USD 3.58 thousand) per year, which is attributed to its high volume of waste segregation. Kuala Lumpur follows with MYR 2.18 thousand (USD 0.48 thousand) per year, and Pahang with MYR 2 thousand (USD 0.44 thousand) per year. The economic benefits of waste segregation programs include avoiding costs associated with leachate treatment, such as construction, operation, and maintenance.

The operational process involved in leachate treatment, which incurs ongoing costs like energy consumption, labour expenses, and chemical usage (Almeida et al., 2020; Cingolani et al., 2016), can be reduced by minimising leachate generation through waste segregation practices. In addition to cost savings in leachate treatment, waste segregation contributes to savings in developing new landfill sites. Treating 60% of waste can help save up to 40% of landfill area requirements (Pujara et al., 2019). Leachate contaminants can potentially degrade the quality of landfill liners (Özçoban et al., 2022). Controlling the amount of leachate discharged into landfills minimises landfill liner degradation, thus extending the lifespan of the landfills and reducing the costs of landfill construction. In Malaysia, constructing a new landfill requires approximately USD 7.75 million (Zaipul & Ahmad, 2017). Avoiding leachate treatment costs also helps to prevent or reduce clean-up and

remediation expenses associated with environmental liabilities, as the volume of leachate discharge in landfills is decreased. For instance, the costs associated with environmental clean-up resulting from leachate leakage in landfills can be avoided.

The estimated costs in the present study were based on the estimated volume of leachate generated in landfills. The study used the cost of leachate treatment in Malaysia, which was MYR 35 m<sup>-3</sup> (USD 7.41 m<sup>-3</sup>) multiplied by the estimated volume of leachate (m<sup>3</sup>). According to Almeida et al. (2020), in Brazil, the estimated cost of leachate treatment for the reverse osmosis (RO) process was USD 8.58 m<sup>-3</sup> for 20 years of the RO unit operation after landfill closure. In Italy and China, the estimated cost of leachate treatment for the RO process was USD 2.7 m<sup>-3</sup> and USD 3.86 m<sup>-3</sup>, respectively (Cingolani et al., 2016; Sato et al., 2021). It indicates that the leachate treatment cost used in this study aligns with the estimated costs from previous studies, showing little difference. Therefore, the method and estimated treatment cost based on the leachate volume presented in this study can be utilised by the authorities for conducting cost analyses and budget planning for waste management, even though specific treatment methods and other factors such as operating and capital costs were not analysed.

The treatment costs depend on the methods and technologies used to treat leachate in a landfill. For example, the conventional leachate treatment method, such as denitrification and nitrification, requires a high operation cost because the method requires high oxygen and carbon, respectively (Show et al., 2019). An integrated method of leachate treatment, which combines conventional and advanced processes, should be practised as it is cost-effective, sustainable, and acceptable to the community (Show et al., 2019). There are some constraints in implementing an integrated treatment technology, especially for developing countries like Malaysia. High costs of operation and maintenance and well-trained operators are required. As the treatment costs can be avoided through segregation and the deviation of waste from landfills, the best options for the country are waste reduction, recycling, and segregation strategies that are inexpensive and environmentally friendly.

Increasing the intensity of waste segregation can enhance the recovery of recyclables and decrease the amount of waste dumped in landfills (Maria & Micale, 2014), leading to a reduction in leachate generation and a decrease in treatment costs. Malaysia must implement the best practices for waste segregation, focusing on achieving the maximum reduction in leachate volume and treatment costs. These practices should increase segregation and recycling rates by promoting public awareness and education (Otitoju & Lau, 2014; Knickmeyer, 2020). The best practice for waste segregation programmes in Malaysia to ensure maximum reduction of leachate volume and treatment cost avoidance is through law enforcement. Enforce mandatory waste segregation, not only at the residential level but also among industrial and commercial sectors, providing clear and consistent guidelines for implementation (Fauziah & Agamuthu, 2010; Yiing & Latifah, 2017).

The Malaysian government should ensure the availability of adequate infrastructure, such as separate bins, collection trucks, and recycling centres, to promote waste segregation (Ting et al., 2016). Previous studies indicate that incentives and rewards motivate individuals to participate actively and effectively in waste segregation programmes (Ali et al., 2015; Low et al., 2016). Continuous evaluation and active monitoring should be carried out to enhance the quality of waste management and reduce waste generation and disposal in landfills (Abas & Wee, 2014). Another approach is to identify new waste management techniques and technologies that can minimise leachate generation and improve the treatment process in terms of both costs and environmental impact (Bandala et al., 2021). It can be achieved through active research and development initiatives and stakeholder collaboration (Abas & Wee, 2014; Samsudin & Don, 2013). The goal of reducing leachate volume and treatment costs can be accomplished by exploring innovative solutions.

### **Waste Segregation Challenges in Malaysia**

There are several challenges in implementing waste segregation programmes in Malaysia. One of the constraints is the level of public awareness and knowledge about waste segregation (Malik et al., 2015). Effective awareness campaigns and education programmes should be conducted through various platforms such as workshops, social media, and public outreach (Otitoju & Lau, 2014). While a high percentage of people have heard about waste segregation (86%), the actual participation in waste segregation practices is relatively low (42%), as reported in Sarawak (Otitoju & Lau, 2014). Previous studies have shown that people possess a high level of knowledge but only a moderate attitude and practice towards waste segregation (Al-Naggar et al., 2019; Mahat et al., 2019). Individual behaviour is one of the factors that influences waste segregation practices. For example, a study conducted in Putrajaya reported that 52% of the community practised waste segregation at the source, while 47% did not (Malik et al., 2015). Similarly, a study in Johor found that only 37% of the respondents claimed to practice waste segregation and recycling actively (Azilah et al., 2015). Thus, behaviour change strategies can be implemented in addition to awareness campaigns to overcome these challenges. For example, applying social norms strategies in the community, led by leaders and local influencers, can positively influence others and attract more people to participate (Knickmeyer, 2020).

The limited availability of segregation facilities, collection systems, and recycling centres for the public to participate in waste segregation programmes poses a challenge to good waste segregation practices in Malaysia. It has been reported that 48.37% of respondents in Selangor agreed to practice waste segregation and recycling if they had access to recycling bins (Ting et al., 2016). Similarly, in another state, 90% of households in Kuala Lumpur were willing to practice waste segregation if infrastructure and facilities were provided (Irina & Chamburi, 2015). Thus, the respective agencies of the local authority



need to overcome this problem by investing in the development and expansion of waste management infrastructure. Previous studies have also reported that socioeconomic factors impact households' waste segregation practices. In Malaysia, most recyclers (84.2%) had a family income above MYR4000, higher education levels, lived in two-storey terraced houses, and most were Chinese. They were actively involved in recycling compared to those with lower income, lower education levels, and living in one-storey terraced houses (Irina et al., 2014).

The authorities need to ensure that waste management services are accessible and affordable for all segments of society and provide incentives to promote participation. People tend to refuse to participate in waste segregation and recycling programmes when they have limited access to them and no financial incentives, such as rewards or penalties, are provided (Ali et al., 2015; Low et al., 2016). The lack of stakeholder collaboration is also a challenge in implementing good waste segregation practices in Malaysia. More intensive collaboration between the government and non-governmental organisations (NGOs) is crucial in promoting and inspiring public participation in waste segregation activities (Abas & Wee, 2014; Mor & Ravindra, 2023; Samsudin & Don, 2013). Another challenge is Malaysia's lack of enforcement and monitoring of waste management. For example, although Act 672 was introduced in 2007, it was officially implemented in 2011 to improve the quality of waste management in the country (Abas & Wee, 2014). The impact of the Act on current waste management practices is still limited (Abas & Wee, 2014; Yahaya & Larsen, 2010) due to several factors, such as a lack of manpower and financial resources to effectively conduct enforcement and monitoring activities (Abas & Wee, 2014; Manaf et al., 2009). To mitigate regulations, fines for noncompliance and strict adherence to the law should be strengthened and implemented with a robust monitoring system.

### **The Role of Government Policies and Regulations**

Government policies and regulations are crucial in promoting waste segregation programmes and reducing leachate volume in Malaysia. Their key role encompasses various aspects, such as establishing and developing legislation and frameworks, implementing waste management plans, enforcing mandatory waste segregation, developing infrastructure, fostering community awareness and education, facilitating enforcement measures, promoting collaboration and partnerships, and supporting research and development. For instance, in the 8<sup>th</sup> Malaysia Plan (2001-2005), the emphasis was placed on the 3Rs program (reduce, reuse, and recycle) to encourage participation. The program was relaunched, and the National Recycling Day was announced to be celebrated annually on November 11<sup>th</sup>.

The National Strategic Plan for Solid Waste Management (NSP) was also implemented in 2005 to enhance the efficiency of non-sanitary landfills and construct new sanitary landfills, promoting environmental sustainability in Malaysia (Sreenivasan et al.,

2012). During the 9<sup>th</sup> Malaysia Plan (2006-2010), the Master Plan on National Waste Minimization was implemented to provide a clear vision, strategies, and stakeholder roles in minimising solid waste disposal (Yiing & Latifah, 2017). Establishing a national solid waste management policy aimed at building an integrated system that is comprehensive, sustainable, cost-effective, and protects the environment and human health (Agamuthu & Victor, 2010).

Act 672 and the Solid Waste and Public Cleansing Management Corporation Act (Act 673) were gazetted as part of these efforts on August 30th, 2007. These acts aimed to improve service quality in waste collection, transportation, treatment, and disposal, protect the environment and human well-being, standardise solid waste management and public cleansing services, and ensure proper waste management throughout the country. The JPSPN and SWCorp were introduced to enhance waste management. JPSPN, established under Section 6(1) of Act 672, is responsible for formulating policies, plans, and strategies for waste management and public cleansing. SWCorp, under Section 17(1) of Act 673, is responsible for recommending and implementing policies, plans, and strategies for waste management and public cleansing (Anuar & Wahab, 2015).

Enforcement of Act 672 commenced on September 1<sup>st</sup>, 2011, during the 10<sup>th</sup> Malaysia Plan (2010-2015), as the federal government assumed full responsibility for solid waste management and privatisation of concessionaires from local authorities, aiming to improve waste management quality (Yiing & Latifah, 2017). Mandatory solid waste separation at source was introduced under Section 74(1) of Act 672, with a fine of up to one thousand ringgits imposed on individuals who fail to comply. Starting in 2016, solid waste segregation at source became mandatory in states governed by Act 672, aiming to increase the recovery of recyclables, reduce waste disposal in landfills, and ensure their long-term sustainability (Fauziah & Agamuthu, 2010; Yiing & Latifah, 2017). Part XI of Act 672 establishes the “Solid Waste and Public Cleansing Management Fund,” which is used to finance solid waste and public cleansing activities, services, and research and development related to waste management in Malaysia (Fauziah & Agamuthu, 2010). The government also plans to implement an integrated waste management system nationwide. This system has already been initiated in Johor as a “test bed” and involves the combination of a sanitary landfill, waste-to-energy (WTE) facilities, anaerobic digestion, and a materials recycling facility (MRF) (Ministry of Housing and Local Government, 2015).

In the 11<sup>th</sup> Malaysia Plan (2016–2020), the government aims to manage various types of waste, including solid, agricultural, construction, radioactive, mining, sewage, and scheduled waste, more comprehensively through a life cycle approach. As part of this strategy, the government encourages the adoption of 3R practices among the community to increase recycling rates and invests in converting waste into electricity through energy recovery. Since 2016, other strategies have been implemented, such as enforcing Government Green

Procurement (GGP) to promote product recycling in public procurement, implementing a media strategy to raise awareness and provide public education, establishing a code of practice for construction and demolition (C&D) waste management, and fostering smart partnerships between NGOs, private sectors, academic institutions, and communities in 3R activities (Ministry of Housing and Local Government, 2015).

The pay-polluter principle (PPP), also known as Pay-As-You-Throw (PAYT), was introduced in 2018 as part of the key initiative under Thrust 1. It aims to enforce waste reduction and separation at the source, encourage recycling activities among the community, and promote cost control, equality, and environmental quality. Polluters will be charged based on the volume, weight, or hybrid waste disposal system. The Take Back System, launched in 2018, aims to boost recycling demand among manufacturers, producers, importers, and retailers to increase the recycling rate in the industrial sector (Ministry of Housing and Local Government, 2015). Under the 12<sup>th</sup> Malaysia Plan (2021–2025), the government targets a recycling rate of 40% by 2025. Therefore, there will be intensified enforcement of waste segregation at the source and the implementation of 3R initiatives, particularly in urban areas. More waste collection, segregation, and recycling facilities will be introduced, especially in residential, institutional, and commercial areas.

The government has also introduced a sustainable economy cycle that emphasises the restoration of resources through reuse and recycling initiatives. This approach follows the principles of the 3Rs, aiming to minimise resource consumption, maximise product reuse, and recycle materials back into the manufacturing cycle. The economic cycle is expected to reduce waste generation and disposal, as well as minimise the environmental and health impacts associated with solid waste management in Malaysia (Ministry of Economy, 2021). Thus, the effective implementation of these approaches is expected to encourage waste segregation practices, reduce waste volume, and minimise leachate generation in Malaysia's landfills.

## CONCLUSION

The present study reports that many mixed wastes are sent to landfills, resulting in a high volume of landfill leachate generated in a year. Discharging untreated leachate can cause environmental problems such as soil pollution and surface and groundwater contamination, as well as impact human health. The treatment is crucial, with the high costs required to operate the treatment facilities. The waste segregation practices showed they potentially avoid leachate production and benefit leachate treatment cost savings. The savings cost can be used and allocated for sustainable approaches such as conducting awareness programmes and improving recycling facilities in the country. It is expected that the government in Malaysia can use the model described in this study to plan waste management strategies, create legislation, and support segregation and recycling efforts. The findings of the study

are summarised as follows:

- 99.94% of waste is disposed of in the landfills.
- 0.06% waste segregation rate in the country.
- 565 thousand m<sup>3</sup> of leachate generated by landfilled waste per year.
- 354 m<sup>3</sup> per year of leachate can be avoided through waste segregation practices.
- USD 4.36 million per year is required for landfill leachate treatment.
- Waste segregation practices can prevent USD 2.73 thousand per year leachate treatment costs.

However, the waste segregation rate was reported as low. Thus, promoting community engagement and education is important to encourage waste segregation practices and reduce landfill leachate in Malaysia. One approach that can be taken to raise community awareness is through an educational campaign that utilises various mediums such as social media, posters, pamphlets, and community meetings to disseminate information effectively. At the school level, for example, waste management and segregation can be introduced in the curriculum as an early step to educate and encourage students to practice waste segregation. Authorities should also establish community waste segregation centres or recycling stations that provide convenience for households to drop off their segregated waste. It can be achieved by ensuring the availability of proper waste segregation infrastructure, such as separate bins for different waste. Another approach to encourage community engagement and education is the establishment of public-private partnerships, which can contribute resources, expertise, and infrastructure to strengthen waste segregation initiatives. By combining these approaches, community engagement and education can effectively promote waste segregation practices and minimise leachate volume in landfills.

### **The Limitations of the Study**

There are a few limitations to this study. First, the study's findings only represent the states implementing Act 672. However, other states can use this result as a baseline data for improving their waste management strategies. Second, the secondary data were analysed using mathematical models adapted from the literature where the estimated value may be under- or over-estimated. Nonetheless, the values used in the equations of the volume of leachate per tonne of waste and leachate treatment price were compared and were within the values from previous field studies. The study recommends that some factors in the analysis must be considered to estimate the quantity of the leachate. For example, the analysis of leachate production must consider other factors such as waste composition, waste moisture availability, landfills' age, rainfall intensity, and land area in a region. Capital expenses (CAPEX) and operational expenses (OPEX) should be included in estimating leachate treatment costs by determining economical methods or processes based on the country settings that should be performed in future studies.

## ACKNOWLEDGEMENT

The authors sincerely appreciate the financial support provided by the Graduate Research Fellowship (GRF) of Universiti Putra Malaysia (UPM) 2018/2019, as well as the datasets provided by Solid Waste Management and Public Cleansing Corporation (SWCorp), Cyberjaya, Selangor, Malaysia.

## REFERENCES

- Abas, M. A., & Wee, S. T. (2014). The issues of policy implementation on solid waste management in Malaysia. *International Journal of Conceptions on Management and Social Sciences*, 2(3), 12-17.
- Abunama, T., Othman, F., & Younes, M. K. (2018). Predicting sanitary landfill leachate generation in humid regions using ANFIS modeling. *Environmental Monitoring and Assessment*, 190(10), 1-15. <https://doi.org/https://doi.org/10.1007/s10661-018-6966-y>
- Aderemi, A. O., Oriaku, A. V, Adewumi, G. A., & Otitolaju, A. A. (2011). Assessment of groundwater contamination by leachate near a municipal solid waste landfill. *African Journal of Environmental Science and Technology*, 5(11), 933-940.
- Adhikari, B., Dahal, K. R., & Khanal, S. N. (2014). A review of factors affecting the composition of municipal solid waste landfill leachate. *International Journal of Engineering Science and Innovative Technology (IJESIT)*, 3(5), 273-281.
- Agamuthu, P., & Victor, D. (2010). *Policy evolution of solid waste management in Malaysia*. In *International Solid Waste Association Congress 2011* (pp. 916-924). UM Research Repository. [http://eprints.um.edu.my/12896/1/2.Policy\\_evolution\\_of\\_Solid\\_Waste\\_Management\\_in\\_Malaysia.pdf](http://eprints.um.edu.my/12896/1/2.Policy_evolution_of_Solid_Waste_Management_in_Malaysia.pdf)
- Al-Naggar, R. A., Abdulghani, M. A. M., & Al-Areefi, M. A. (2019). Effects of inappropriate waste management on health: Knowledge, attitude and practice among Malaysian population. *Malaysian Journal of Public Health Medicine*, 19(1), 101-109.
- Alghamdi, A., Aly, A. A., & Ibrahim, H. M. (2021). Assessing the environmental impacts of municipal solid waste landfill leachate on groundwater and soil contamination in western Saudi Arabia. *Arabian Journal of Geosciences*, 14, Article 350. <https://doi.org/10.1007/s12517-021-06583-9>
- Ali, A. B., Alavi, N., Goudarzi, G., Teymouri, P., Kambiz, A., & Mohammad, R. (2015). Household recycling knowledge, attitudes and practices towards solid waste management. *Resources, Conservation and Recycling*, 102, 94-100. <https://doi.org/10.1016/j.resconrec.2015.06.014>
- Aljaradin, M., & Persson, K. M. (2016). The emission potential from municipal solid waste landfill in Jordan. *Journal of Ecological Engineering*, 17(1), 38-48. <https://doi.org/10.12911/22998993/61188>
- Almeida, R. de, Bila, D. M., Quintaes, B. R., & Campos, J. C. (2020). Cost estimation of landfill leachate treatment by reverse osmosis in a Brazilian landfill. *Waste Management and Research*, 38(10), 1087-1092. <https://doi.org/10.1177/0734242X20928411>
- Anuar, H. M., & Wahab, H. A. (2015). Sisa pepejal dan pembersihan awam: Pengurusan dan perundangan [Waste management and public cleansing: Management and legislation]. *Solid Waste Solution Journal*, 1, 1-19.

- Azadi, S., Karimi-Jashni, A., Javadpour, S., & Mahmoudian-Boroujerd, L. (2021). Photocatalytic landfill leachate treatment using P-type TiO<sub>2</sub> nanoparticles under visible light irradiation. *Environment, Development and Sustainability*, 23, 6047-6065. <https://doi.org/10.1007/s10668-020-00861-4>
- Azilah, M. A., Foziah, J., & Ho, C. . (2015). The effects of socio-economic influences on households recycling behaviour in Iskandar Malaysia. *Social and Behavioral Sciences*, 202, 124-134. <https://doi.org/10.1016/j.sbspro.2015.08.215>
- Aziz, H. A., Adlan, M. N., Amilin, K., Yusoff, M. S., Ramly, N. H., & Umar, M. (2012). Quantification of leachate generation rate from a semi-aerobic landfill in Malaysia. *Environmental Engineering and Management Journal*, 11(9), 1581-1585.
- Aziz, H. A., & Ramli, S. F. (2018). Recent development in sanitary landfilling and landfill leachate treatment in Malaysia. *International Journal of Environmental Engineering*, 9(3-4), 201-229. <https://doi.org/10.1504/IJEE.2018.10018737>
- Bandala, E. R., Liu, A., Wijesiri, B., Zeidman, A. B., & Goonetilleke, A. (2021). Emerging materials and technologies for landfill leachate treatment: A critical review. *Environmental Pollution*, 291, Article 118133. <https://doi.org/10.1016/j.envpol.2021.118133>
- Bernama. (2022, March 10). Can Malaysia achieve 40 per cent recycling rate by 2025? *New Straits Times*. <https://www.nst.com.my/news/nation/2022/03/778625/can-malaysia-achieve-40-cent-recycling-rate-2025>
- Cheng, Z., Sun, Z., Zhu, S., Lou, Z., Zhu, N., & Feng, L. (2019). The identification and health risk assessment of odor emissions from waste landfilling and composting. *Science of the Total Environment*, 649, 1038-1044. <https://doi.org/10.1016/j.scitotenv.2018.08.230>
- Cingolani, D., Eusebi, A. L., & Battistoni, P. (2016). Osmosis process for leachate treatment in industrial platform: Economic and performances evaluations to zero liquid discharge. *Journal of Environmental Management*, 203, 782-790. <https://doi.org/10.1016/j.jenvman.2016.05.012>
- Coelho, L. M. G., & Lange, L. C. (2018). Applying life cycle assessment to support environmentally sustainable waste management strategies in Brazil. *Resources, Conservation and Recycling*, 128, 438-450. <https://doi.org/10.1016/j.resconrec.2016.09.026>
- Debrah, J. K., Vidal, D. G., & Dinis, M. A. P. (2021). Raising awareness on solid waste management through formal education for sustainability: A developing countries evidence review. *Recycling*, 6(6), 1-21. <https://doi.org/10.3390/recycling6010006>
- Department of Statistics Malaysia. (2019a). *Demographic Statistics First Quarter 2019 , Malaysia*. [https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=430&bul\\_id=Mno2WGQ3QUdmaUM3c3l0NzN0aW9tZz09&menu\\_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09](https://www.dosm.gov.my/v1/index.php?r=column/cthemByCat&cat=430&bul_id=Mno2WGQ3QUdmaUM3c3l0NzN0aW9tZz09&menu_id=L0pheU43NWJwRWVSZklWdzQ4TlhUUT09)
- Department of Statistics Malaysia. (2019b). *Statistic*. [https://www.dosm.gov.my/v1/index.php?r=column/cone&menu\\_id=d1dTR0JMK2hUUUFnTnp5WUR2d3VBQT09](https://www.dosm.gov.my/v1/index.php?r=column/cone&menu_id=d1dTR0JMK2hUUUFnTnp5WUR2d3VBQT09)
- Detho, A., Daud, Z., Samo, S. R., Khan, Z., Memon, D. A., & Awang, H. (2020). Physicochemical characteristics of landfill leachate from Simpang Renggam landfill site, Johor, Malaysia. *Quest Research Journal*, 18(2), 82-88. <https://doi.org/10.52584/QRJ.1802.13>



- Dia, O., Drogui, P., Buelna, G., & Dubé, R. (2018). Hybrid process, electrocoagulation-biofiltration for landfill leachate treatment. *Waste Management*, 75, 391-399. <https://doi.org/10.1016/j.wasman.2018.02.016>
- Eslami, A., Saghi, M. H., Akbari-adergani, B., Sadeghi, S., Ghaderpoori, M., Rabbani, M., & Alinejad, A. (2021). Synthesis of modified ZnO nanorods and investigation of its application for removal of phthalate from landfill leachate: A case study in Aradkouh landfill site. *Journal of Environmental Health Science and Engineering*, 19, 133-142. <https://doi.org/10.1007/s40201-020-00587-5>
- Fauziah, S. H., & Agamuthu, P. (2010). Landfills in Malaysia: Past, present and future. In *International Conference on Final Sinks* (pp. 1-9). UM Research Repository. <https://eprints.um.edu.my/11314/>
- Foufou, A., Djorfi, S., Haied, N., Kechiched, R., Azlaoui, M., & Hani, A. (2017). Water pollution diagnosis and risk assessment of Wadi Zied plain aquifer caused by the leachates of Annaba landfill (N-E Algeria). *Energy Procedia*, 119, 393-406. <https://doi.org/10.1016/j.egypro.2017.07.123>
- Hassan, M. S., Hayin, N. A. M., Ismail, M. R., & Mahmud, N. H. (2022). 14 juta tan sampah pada 2022 [14 million tons of waste in 2022]. *Metro*. <https://www.hmetro.com.my/utama/2021/07/732770/14-juta-tan-sampah-pada-2022>
- Hoornweg, D., & Perinaz, B. T. (2012). *What a Waste: A global Review of Solid Waste Management*. The World Bank. <http://hdl.handle.net/10986/17388>
- Ibrahim, T. N. T., Mahmood, N. Z., & Othman, F. (2016). Estimation of leachate generation from MSW landfills in Selangor. *Environmental Science*, 19(1), 43-48.
- Ibrahim, T. N. T., Mahmood, N. Z., & Othman, F. (2017). Estimation of leachate generation from MSW landfills in Selangor. *Asian Journal of Microbiology, Biotechnology & Environmental Sciences*, 19(1), 43-48.
- Imran, A., Shreeshivadasan, C., Norhayati, A., & Mohd Danish, A. (2019). Sanitary landfill is a solution in solid waste management or a silent threat to environment: Malaysian scenario article history. *Open International Journal of Informatics*, 7, 135-146.
- Irina, S. Z., & Chamburi, S. (2015). An analysis of household acceptance of curbside recycling scheme in Kuala Lumpur, Malaysia. *Habitat International*, 47, 248-255. <https://doi.org/10.1016/j.habitatint.2015.01.014>
- Irina, S. Z., Zainura, Z. N., & Rafiu, O. Y. (2014). The profiles of household solid waste recyclers and non-recyclers in Kuala Lumpur, Malaysia. *Habitat International*, 42, 83-89. <https://doi.org/10.1016/j.habitatint.2013.10.010>
- Jain, A. (2017). *Waste Management in ASEAN countries*. United Nations Environment Programme. <http://www.unenvironment.org/resources/report/waste-management-asean-countries-summary-report>
- Jamliah, A. (2022, February 21). Kadar kitar semula kebangsaan meningkat [The national recycling rate has increased]. *Utusan Malaysia*. <https://www.utusan.com.my/nasional/2022/02/kitar-semula/>
- Jun, D., Mingjiang, N., Yong, C., Daoan, Z., & Chao, F. (2013). Life cycle and economic assessment of source-separated MSW collection with regard to greenhouse gas emissions: A case study in China. *Environmental Science and Pollution Research*, 20(8), 5512-5524. <https://doi.org/10.1007/s11356-013-1569-1>
- Kamel, H. (2021, October 18). Recycling is key to waste recovery's future. *The Malaysia Reserve*. <https://themalaysianreserve.com/2021/10/18/recycling-is-key-to-waste-recovery-s-future/>



- Kaya, M. (2016). Recovery of metals and nonmetals from electronic waste by physical and chemical recycling processes. *Waste Management*, 57, 64-90. <https://doi.org/10.1016/j.wasman.2016.08.004>
- Kaza, S., Yao, L., Bhada-Tata, P., & Van, W. F. (2018). *What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050*. World Bank Publications. [https://datatopics.worldbank.org/what-a-waste/trends\\_in\\_solid\\_waste\\_management.html](https://datatopics.worldbank.org/what-a-waste/trends_in_solid_waste_management.html)
- Khaled, S. B., & Muhammad, A. A. (2016). Field accumulation risks of heavy metals in soil and vegetable crop irrigated with sewage water in western region of Saudi Arabia. *Saudi Journal of Biological Sciences*, 23(1), S32-S44. <https://doi.org/10.1016/j.sjbs.2015.09.023>
- Knickmeyer, D. (2020). Social factors influencing household waste separation: A literature review on good practices to improve the recycling performance of urban areas. *Journal of Cleaner Production*, 245, Article 118605. <https://doi.org/10.1016/j.jclepro.2019.118605>
- Liikanen, M., Havukainen, J., Viana, E., & Horttanainen, M. (2018). Steps towards more environmentally sustainable municipal solid waste management - A life cycle assessment study of São Paulo, Brazil. *Journal of Cleaner Production*, 196, 150-162. <https://doi.org/10.1016/j.jclepro.2018.06.005>
- Liu, Y., Ni, Z., Kong, X., & Liu, J. (2017). Greenhouse gas emissions from municipal solid waste with a high organic fraction under different management scenarios. *Journal of Cleaner Production*, 147, 451-457. <https://doi.org/10.1016/j.jclepro.2017.01.135>
- Low, S. T., Tee, S. Y., & Choong, W. W. (2016). Preferred attributes of waste separation behaviour: An empirical study. *International Conference on Sustainable Design, Engineering and Construction*, 145, 738-745. <https://doi.org/10.1016/j.proeng.2016.04.094>
- Magda, M. A. E. S., & Gaber, I. A. Z. (2015). Impact of landfill leachate on the groundwater quality: A case study in Egypt. *Journal of Advanced Research*, 6(4), 579-586. <https://doi.org/10.1016/j.jare.2014.02.003>
- Mahat, H., Hashim, M., Nayan, N., Saleh, Y., & Norkhaidi, S. B. (2019). E-waste disposal awareness among the Malaysian community. *Knowledge Management and E-Learning*, 11(3), 393-408. <https://doi.org/10.34105/j.kmel.2019.11.021>
- Maiti, S. K., De, S., Hazra, T., Debsarkar, A., & Dutta, A. (2016). Characterization of leachate and its impact on surface and groundwater quality of a closed dumpsite - A case study at Dhapa, Kolkata, India. *International Conference on Solid Waste Management*, 35, 391-399. <https://doi.org/10.1016/j.proenv.2016.07.019>
- Malik, N. K. A., Abdullah, S. H., & Manaf, L. A. (2015). Community participation on solid waste segregation through recycling programmes in Putrajaya. *Procedia Environmental Sciences*, 30, 10-14. <https://doi.org/10.1016/j.proenv.2015.10.002>
- Manaf, L. A., Samah, M. A. A., & Zukki, N. I. M. (2009). Municipal solid waste management in Malaysia: Practices and challenges. *Waste Management*, 29(11), 2902-2906. <https://doi.org/https://doi.org/10.1016/j.wasman.2008.07.015>
- Maria, F. D., & Micale, C. (2014). A holistic life cycle analysis of waste management scenarios at increasing source segregation intensity: The case of an Italian urban area. *Waste Management*, 34(11), 2382-2392. <https://doi.org/10.1016/j.wasman.2014.06.007>

- Martins, D. C. C., Scandelay, A. P. J., Cardozo-Filho, L., & Tavares, C. R. G. (2020). Supercritical water oxidation treatment of humic acid as a model organic compound of landfill leachate. *The Canadian Journal of Chemical Engineering*, 98, 868-878. <https://doi.org/https://doi.org/10.1002/cjce.23691>
- Mekonnen, B., Haddis, A., & Zeine, W. (2020). Assessment of the effect of solid waste dump site on surrounding soil and river water quality in Tepi Town, Southwest Ethiopia. *Journal of Environmental and Public Health*, 2020, Article 5157046. <https://doi.org/10.1155/2020/5157046>
- Ministry of Economy. (2021). *Twelfth Malaysia Plan (2021-2025)*. <https://rmke12.epu.gov.my/en/documents/twelfth-plan>
- Ministry of Housing and Local Government. (2015). *Solid Waste Management Lab 2015*. [http://www.kpkt.gov.my/resources/index/user\\_1/Attachments/hebahan\\_slider/slaid\\_dapatan\\_makmal.pdf](http://www.kpkt.gov.my/resources/index/user_1/Attachments/hebahan_slider/slaid_dapatan_makmal.pdf)
- Moa, S., Wang, Y., & Chen, F. (2021). Stability and microbial diversity of the aerobic granular sludge under low carbon to nitrogen ratio. *Environmental Chemistry*, 3, 904-913. <https://doi.org/https://doi.org/10.7524/j.issn.0254-6108.2019101105>
- Mor, S., & Ravindra, K. (2023). Municipal solid waste landfills in lower- and middle-income countries: Environmental impacts, challenges and sustainable management practices. *Process Safety and Environmental Protection*, 174, 510-530. <https://doi.org/10.1016/j.psep.2023.04.014>
- Nai, C., Tang, M., Liu, Y., Xu, Y., Dong, L., Liu, J., & Huang, Q. (2021). Potentially contamination and health risk to shallow groundwater caused by closed industrial solid waste landfills: Site reclamation evaluation strategies. *Journal of Cleaner Production*, 286, Article 125402. <https://doi.org/10.1016/j.jclepro.2020.125402>
- National Solid Waste Management Department. (2013). *Survey on Solid Waste Composition, Characteristics & Existing Practice of Solid Waste Recycling in Malaysia*. Ministry of Housing and Local Government. [http://jpspn.kpkt.gov.my/resources/index/user\\_1/Sumber\\_Rujukan/kajian/Final\\_Report\\_REVz.pdf](http://jpspn.kpkt.gov.my/resources/index/user_1/Sumber_Rujukan/kajian/Final_Report_REVz.pdf)
- National Solid Waste Management Department. (2021). *Bilangan Tapak Pelupusan Mengikut Negeri [The Number of Disposal Sites According to State]*. Ministry of Housing and Local Government. <http://jpspn.kpkt.gov.my/index.php/pages/view/69%0Ahttp://jpspn.kpkt.gov.my/index.php/pages/view/134>
- Noor, A. S. (2022, April 25). 17,000 tan sisa makanan setiap hari [17,000 tonnes of food waste every day]. *Berita Harian Online*. <https://www.bharian.com.my/berita/nasional/2022/04/949530/17000-tan-sisa-makanan-setiap-hari>
- Obiri-Nyarko, F., Duah, A. A., Karikari, A. Y., Agyekum, W. A., Manu, E., & Tagoe, R. (2021). Assessment of heavy metal contamination in soils at the Kpone landfill site, Ghana: Implication for ecological and health risk assessment. *Chemosphere*, 282, Article 131007. <https://doi.org/10.1016/j.chemosphere.2021.131007>
- Othman, N., Razali, A., Chelliapan, S., Mohammad, R., & Kamyab, H. (2021). Chapter 18 - A design framework for an integrated end-of-life vehicle waste management system in Malaysia. In *Soft Computing Techniques in Solid Waste and Wastewater Management* (pp. 305-319). Elsevier. <https://doi.org/10.1016/B978-0-12-824463-0.00021-5>
- Ototoju, T. A., & Lau, S. (2014). Municipal solid waste management: Household waste segregation in Kuching South City, Sarawak, Malaysia. *American Journal of Engineering Research*, 03(06), 82-91.

- Özçoban, M. Ş., Acarer, S., & Tüfekci, N. (2022). Effect of solid waste landfill leachate contaminants on hydraulic conductivity of landfill liners. *Water Science and Technology*, 85(5), 1581-1599. <https://doi.org/10.2166/wst.2022.033>
- Pathak, P., Srivastava, R. R., & Ojasvi. (2019). Chapter 5 - Environmental management of e-waste. *Electronic Waste Management and Treatment Technology, 2019*, 103-132. <https://doi.org/10.1016/B978-0-12-816190-6.00005-4>
- Pujara, Y., Pathak, P., Sharma, A., & Govani, J. (2019). Review on Indian municipal solid waste management practices for reduction of environmental impacts to achieve sustainable development goals. *Journal of Environmental Management*, 248, Article 109238. <https://doi.org/10.1016/j.jenvman.2019.07.009>
- Rangga, J. U., Ismail, S. N. S., Rasdi, I., & Karuppiah, K. (2023). Contribution of waste segregation to GHG emissions, land usage, and health risk of NMVOC exposure: A study in Malaysia. *Journal of Material Cycles and Waste Management*, 25, 181-197. <https://doi.org/https://doi.org/10.1007/s10163-022-01530-6>
- Samsudin, M. D. M., & Don, M. M. (2013). Municipal solid waste management in Malaysia: Current practices, challenges and prospect. *Jurnal Teknologi (Sciences and Engineering)*, 62(1), 95-101. <https://doi.org/10.11113/jt.v62.1293>
- Sato, Y., Zeng, Q., Meng, L., & Chen, G. (2021). Importance of combined electrochemical process sequence and electrode arrangements: A lab-scale trial of real reverse osmosis landfill leachate concentrate. *Water Research*, 192, Article 116849. <https://doi.org/10.1016/j.watres.2021.116849>
- Sharifah, N. S. I., Che, F. I., Armi, A. S. M., Erneeza, M. H., & Aini, S. A. W. (2015). Soil contamination from non-sanitary waste landfill in Langat water catchment area, Malaysia. *Journal of Scientific Research & Reports*, 7(6), 480-493. <https://doi.org/10.9734/JSRR/2015/15102>
- Show, P. L., Pal, P., Leong, H. Y., Juan, J. C., & Ling, T. C. (2019). A review on the advanced leachate treatment technologies and their performance comparison: An opportunity to keep the environment safe. *Environmental Monitoring and Assessment*, 191(4), 1-28.
- Slagstad, H., & Brattebø, H. (2012). LCA for household waste management when planning a new urban settlement. *Waste Management*, 32(7), 1482-1490. <https://doi.org/10.1016/j.wasman.2012.03.018>
- Soni, A., Das, P. K., Hashmi, A. W., Yusuf, M., Kamyab, H., & Chelliapan, S. (2022). Challenges and opportunities of utilizing municipal solid waste as alternative building materials for sustainable development goals: A review. *Sustainable Chemistry and Pharmacy*, 27, Article 100706. <https://doi.org/10.1016/j.scp.2022.100706>
- Soni, A., Das, P. K., Yusuf, M., Ridha, S., Kamyab, H., Chelliapan, S., Kirpichnikova, I., & Mussa, Z. H. (2023). Valorization of post-consumers plastics and agro-waste in sustainable polymeric composites for tribological applications. *Waste and Biomass Valorization*, 2023, 1-17. <https://doi.org/10.1007/s12649-023-02103-w>
- Sreenivasan, J., Govindan, M., Chinnasami, M., & Kadiresu, I. (2012). Solid Waste Management in Malaysia - A Move Towards Sustainability. In L. F. M. Rebellon (Ed.), *Waste Management: An Integrated Vision* (pp. 55-70). InTech.
- Suratman, S., Tawnie, I., & Sefei, A. (2011). Impact of landfills on groundwater in Selangor, Malaysia. *ASM Science Journal*, 5(2), 101-107.

- Taha, M. R., Yaacob, W. Z. W., Samsudin, A. R., & Yaakob, J. (2011). Groundwater quality at two landfill sites in Selangor, Malaysia. *Bulletin of the Geological Society of Malaysia*, 57(57), 13-18. <https://doi.org/10.7186/bgsm57201103>
- Ting, L. S., Yee, T. S., & Wai, C. W. (2016). Preferred attributes of waste separation behaviour: An empirical study. *International Conference on Sustainable Design, Engineering and Construction*, 145, 738-745. <https://doi.org/10.1016/j.proeng.2016.04.094>
- Tiseo, I. (2022). *Global recycling rates of municipal solid waste 2020, by select country*. Statista. <https://www.statista.com/statistics/1052439/rate-of-msw-recycling-worldwide-by-key-country/>
- Turner, D. A., Williams, I. D., & Kemp, S. (2015). Greenhouse gas emission factors for recycling of source-segregated waste materials. *Resources, Conservation & Recycling*, 105, 186-197. <https://doi.org/10.1016/j.resconrec.2015.10.026>
- USEPA. (2023). *What is a Municipal Solid Waste Landfill?* United States Environmental Protection Agency. <https://www.epa.gov/landfills/municipal-solid-waste-landfills#whatis>
- Utusan Online. (2017). 13.5 juta tan sampah setiap tahun. *Utusan Online*. [https://upm.edu.my/akhbar/135\\_juta\\_tan\\_sampah\\_setiap\\_tahun-126](https://upm.edu.my/akhbar/135_juta_tan_sampah_setiap_tahun-126)
- Wu, C., Liu, J., Liu, S., Li, W., Yan, L., Shu, M., Zhao, P., Zhou, P., & Cao, W. (2018). Assessment of the health risks and odor concentration of volatile compounds from a municipal solid waste landfill in China. *Chemosphere*, 202, 1-8. <https://doi.org/10.1016/j.chemosphere.2018.03.068>
- Yahaya, N., & Larsen, I. (2010). *Federalising Solid Waste Management in Peninsular Malaysia*. International Solid Waste Association (ISWA).
- Yan, C., Cheng, Z., Quan, X., Feng, C., & Cheng, G. (2021). Electrochemical pretreatment of landfill leachate RO concentrate with multi-channel mesh electrodes. *Environmental Chemistry*, 2, 603-613. <https://doi.org/10.7524/j.issn.0254-6108.2020030801>
- Yao, X. Z., Ma, R. C., Li, H. J., Wang, C., Zhang, C., Yin, S. S., Wu, D., He, X. Y., Wang, J., Zhan, L. T., & He, R. (2019). Assessment of the major odor contributors and health risks of volatile compounds in three disposal technologies for municipal solid waste. *Waste Management*, 91, 128-138. <https://doi.org/10.1016/j.wasman.2019.05.009>
- Yiing, C. M., & Latifah, A. M. (2017). Solid waste management transformation and future challenges of source separation and recycling practice in Malaysia. *Resources, Conservation & Recycling*, 116, 1-14. <https://doi.org/10.1016/j.resconrec.2016.09.012>
- Yong, C., Jun, D., Yuanjun, T., Qunxing, H., & Mingjiang, N. (2015). Life cycle assessment of municipal solid waste source-separated collection and integrated waste management systems in Hangzhou, China. *Journal of Material Cycles and Waste Management*, 17(4), 695-706. <https://doi.org/10.1007/s10163-014-0300-8>
- Zaini, M. S. I., Hasan, M., & Zolkepli, M. F. (2022). Urban landfills investigation for leachate assessment using electrical resistivity imaging in Johor, Malaysia. *Environmental Challenges*, 6, Article 100415. <https://doi.org/10.1016/j.envc.2021.100415>
- Zaipul, A. Z., & Ahmad, R. S. (2017). Policies, challenges and strategies for municipal waste management in Malaysia. *Journal of Science, Technology and Innovation Policy*, 3(1), 18-22.

## Comparison Using Intelligent Systems for Data Prediction and Near Miss Detection Techniques

Lek Ming Lim<sup>1</sup>, Saratha Sathasivam<sup>1</sup>, Mohd. Tahir Ismail<sup>1</sup>, Ahmad Sufriil Azlan Mohamed<sup>2</sup>, Olayemi Joshua Ibidoja<sup>1</sup> and Majid Khan Majahar Ali<sup>1\*</sup>

<sup>1</sup>*School of Mathematical Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*

<sup>2</sup>*School of Computer Sciences, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*

### ABSTRACT

Malaysia ranks third among ASEAN countries in terms of deaths due to accidents, with an alarming increase in the number of fatalities each year. Road conditions contribute significantly to near-miss incidents, while the inefficiency of installed CCTVs and the lack of monitoring system algorithms worsen the situation. The objective of this research is to address the issue of increasing accidents and fatalities on Malaysian roads. Specifically, the study aims to investigate the use of video technology and machine learning algorithms for the car detection and analysis of near-miss accidents. To achieve this goal, the researchers focused on Penang, where the MBPP has deployed 1841 CCTV cameras to monitor traffic and document near-miss accidents. The study utilised the YOLOv3, YOLOv4, and Faster RCNN algorithms for vehicle detection. Additionally, the study employed image processing techniques such as Bird's Eye View and Social Distancing Monitoring to detect and analyse how near misses occur. Various video lengths (20s, 40s, 60s and 80s) were tested

to compare the algorithms' error detection percentage and test duration. The results indicate that Faster RCNN beats YOLOv3 and YOLOv4 in car detection with low error detection, whereas YOLOv3 and YOLOv4 outperform near-miss detection, while Faster RCNN does not perform it. Overall, this study demonstrates the potential of video technology and machine learning algorithms in near-miss accident detection and analysis. Transportation authorities can better understand the causes of accidents and take appropriate measures to improve road

### ARTICLE INFO

#### Article history:

Received: 13 February 2023

Accepted: 21 July 2023

Published: 21 December 2023

DOI: <https://doi.org/10.47836/pjst.32.1.20>

#### E-mail addresses:

[limlekming@gmail.com](mailto:limlekming@gmail.com) (Lek Ming Lim)

[m.tahir@usm.my](mailto:m.tahir@usm.my) (Saratha Sathasivam)

[saratha@usm.my](mailto:saratha@usm.my) (Mohd. Tahir Ismail)

[sufriil@usm.my](mailto:sufriil@usm.my) (Ahmad Sufriil Azlan Mohamed)

[ojibidoja@fugusau.edu.ng](mailto:ojibidoja@fugusau.edu.ng) (Olayemi Joshua Ibidoja)

[majidkhanmajaharali@usm.my](mailto:majidkhanmajaharali@usm.my) (Majid Khan Majahar Ali)

\* Corresponding author

safety using these models. This research can be a foundation for further traffic safety and accident prevention studies.

*Keywords:* Bird's Eye View, intelligent systems, machine learning, near-miss, object detection, Social Distancing Monitoring, vehicle detection

---

## INTRODUCTION

According to the World Health Organization (2020), the death rate from transportation ranked fifth in 2019. The majority of those killed in traffic accidents are teenagers. In 2030, the fatality rate from transportation is predicted to rise.

The main issue with transportation is a lack of accurate and reliable data. The manual data is the police data, also known as POL 37 data, collected after accidents, but some data would not be recorded. Therefore, the information could not be utilised to predict traffic conditions due to missing and inaccurate data in the statistical record. Majlis Bandar Pulau Pinang (MBPP) deployed 1841 Closed-circuit television (CCTV) cameras all around Pulau Pinang in 2019. CCTV is difficult to use to its full potential due to the lack of an algorithm capable of calculating and detecting vehicles, as well as the limited storage capacity, which can only store videos for one and a half months.

Furthermore, POL data is a hardcopy report. It cannot calculate vehicles automatically and does not record near misses due to the limitation of converting the manual report into a visual report. Aside from that, no autonomous algorithm can be utilised for near-miss counting. It is hard to measure near misses simultaneously in CCTV videos since no previous research exists.

Near misses are one of the transportation issues that must be addressed to reduce the likelihood of fatalities and accidents and meet the goals of the Penang 2030 mission. Since near misses cause accidents, near-miss reports are investigated to enhance road safety. Heinrich's 300-29-1 model demonstrated that the probability of 300 near misses can result in 29 minor injuries, and then these probabilities of minor injuries can result in one major injury. If 300 near-miss probabilities can cause a fatality or accident, it will emit carbon, harming the environment and contributing to atmospheric pollution. Therefore, reporting near misses is important to reduce the probability of accidents on specific roads and find out the root cause.

Near misses result from dangerous activities caused by human mistakes and situations produced by malfunctioning procedures or systems in Malaysian road traffic. According to Aldred (2016), cyclist behaviours influence near misses. A mixed traffic flow scenario contains a wide range of collision types, which contribute to near misses, too. According to Wang et al. (2020), the weather, which causes low-vision conditions, is another aspect that might lead to near misses in traffic flow. As a result, in the research, vision-based detection is used to detect objects and crash types in the monitoring system.



Vehicle detection is an algorithm-based computer vision technique used to detect vehicles, count vehicles and estimate the average speed of vehicles (Meng et al., 2020). Arinaldi et al. (2018) used visualisation in traffic monitoring systems. The research collected statistics on vehicle counts, vehicle type, predicted vehicle speed, and vehicle lane shift. These are all factors to consider in their study. Researchers who conducted prior investigations only used images and a linked-in algorithm and software method to conduct vehicle detection. The main reason is that the monitoring systems will show all details and information through visualisation reporting.

Appendix A summarises previous research and highlights gaps in vehicle detection, particularly in Penang. Limited studies employ various methods for detecting vehicles and identifying near misses, focusing on software and algorithm development in engineering. While image processing methods are commonly used, this study aims to improve accuracy by implementing Convolutional Neural Network (CNN) and Fast Region-based Convolutional Neural Network (RCNN). Many researchers connect their model or algorithm to software to enable visualisation in monitoring systems. Such systems aid in data analysis from image and video processing. In this study, Social Distancing Monitoring and Bird's Eye View are used in vehicle identification to analyse images and identify near misses.

Models or software that recognise objects or track vehicles include the Faster RCNN, and You Only Look Once (YOLO). These models are not well-known from near-miss studies. Huang et al. (2020) detected vehicles in traffic utilising Faster RCNN, and You Only Look Once version 3 (YOLOv3) on 40-second basis CCTV recordings. Kumar et al. (2020) employed YOLOv3, and You Only Look Once version 4 (YOLOv4) for surveillance in traffic monitoring systems. Ammar et al. (2021) compare Faster RCNN, YOLOv3, and YOLOv4 for vehicle recognition from images. Faster RCNN, YOLOv3, and YOLOv4 were used by Sowmya and Radha (2021) to develop vehicle identification and classification algorithms for effective heavy vehicle traffic monitoring. Lim, Ali et al. (2022) experimented with different video quality in vehicle recognition using YOLOv3 and Faster RCNN.

In this experiment, YOLOv3 and YOLOv4 perform image processing on vehicle recognition in CCTV recordings using the approaches of Bird's Eye View and Social Distancing Monitoring. The road condition may be monitored using vehicle detection to study the process of near misses and accidents. The data acquired may be utilised to forecast near misses in traffic flow and determine the causes of the issues.

## **MATERIALS AND METHODS**

### **YOLO**

Figure 1 shows the flow chart of YOLO in vehicle detection.



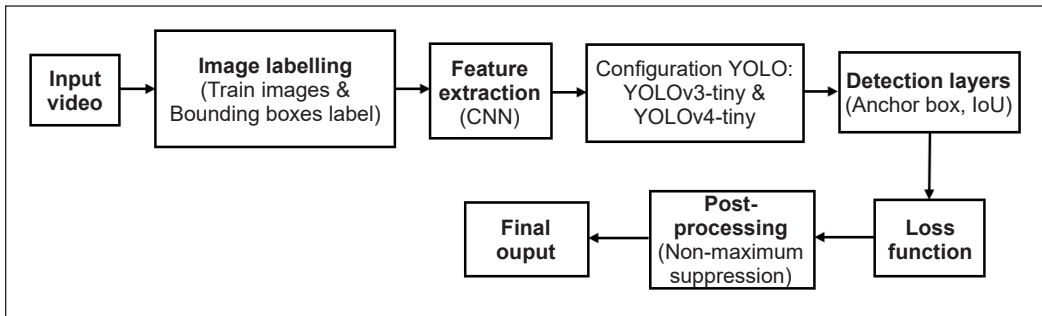


Figure 1. YOLO flow chart

### Image Labelling

Training YOLO to identify an object involves the training image and the bounding box label. Collecting training images is the first step in initiating the object detector. Bounding box annotations are used to aid the object detector’s learning. Each object detected by the detector is enclosed by a box and labelled with the object class forecasted by the detector. Image training in YOLO is depicted in Figure 2.

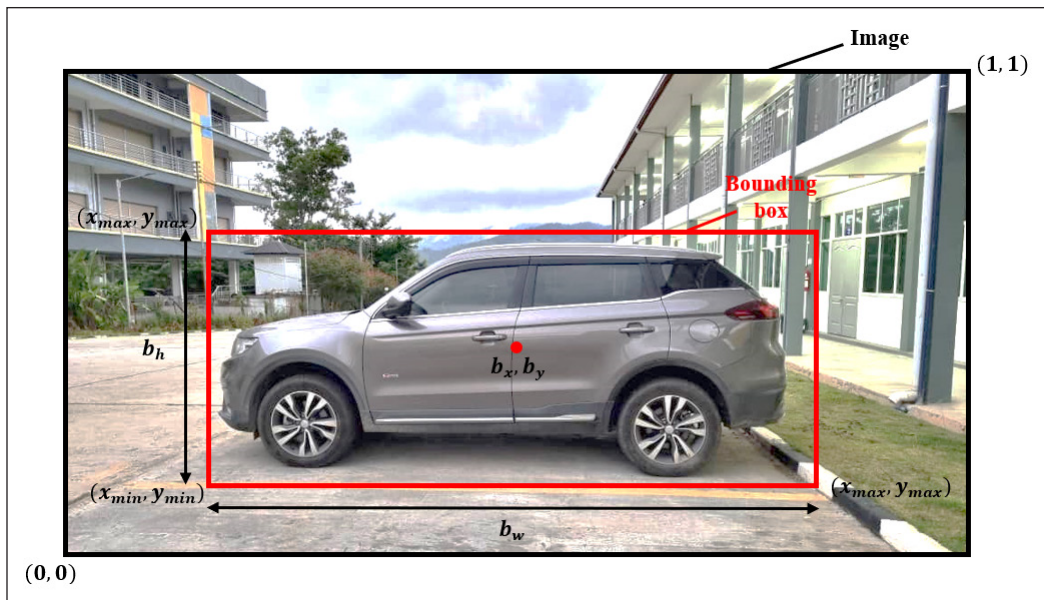


Figure 2. Train image in YOLO

Define  $W, H$  is the dimension of the original image (Equation 1) (AlKishri & Al-Bahri, 2021).

$$b_x = \frac{(x_{min} + x_{max})}{2 * W}, \quad b_y = \frac{(y_{min} + y_{max})}{2 * H}$$

$$b_w = \frac{(x_{max} - x_{min})}{W}, \quad b_h = \frac{(y_{max} - y_{min})}{H} \quad [1]$$

where:  $b_x, b_y$  is the box locations;  $b_x, b_y, b_w, b_h$  is the width and height of the entire image;  $b_x, b_y, b_w, b_h$  is the bounding box prediction;  $x_{min}, y_{min}, x_{max}, y_{max}$  is the image bounding box coordinate; and  $W, H$  is the width and height of the bounding box.

## Feature Extraction

Gai et al. (2021) stated that other deep learning methods for object detection typically involve merging the feature extraction process with the classification prediction. It involves using a convolutional network to extract features and predict the objects' locations and classifications.

## Detection Layers

Silva et al. (2020) stated that during model training, the convolutional network layer was configured to detect only one class: the car. The number of filters used in the layer was directly set in the configuration. The formula for determining the number of filters is in Equation 2:

$$\begin{aligned} Filters &= (classes + w + h + x + y + Confidence\ score) \times num \\ &= (classes + 1 + 1 + 1 + 1 + 1) \times num \\ &= (classes + 5) \times num \end{aligned} \quad [2]$$

where 5 represents the conditions ( $x, y$  - position of boundary box,  $w, h$  - width and height of the image and  $C$  - the Confidence score), and  $c$  is the class probability. YOLO detect 3 boxes per grid cell (Equation 3), so the:

$$Filters = (classes + 5) \times 3 \quad [3]$$

Therefore, there are 18 filters for 1 class. These 18 filters will be placed before two YOLO layers.

Figure 3 illustrates how YOLO performs detection. Initially, the input image is divided into  $S \times S$  grids. If the centre of a ground truth object lies within a grid cell, the grid is assigned to detect that object. Each grid cell has  $B$  bounding boxes and corresponding confidence scores for those boxes. The bounding boxes and confidence scores are then multiplied with class probability maps. Finally, the results of the final detection are obtained.

The final detection takes place during inference using the formula derived from Figure 4, a tensor of size  $S \times S * (B * 5 + c)$ . In this formula,  $S$  represents the number of grids,  $B$

is the bounding box, 5 denotes the four bounding box attributes and one object confidence per cell,  $c$  is the class probability, and tensor refers to the network output. Redmon et al. (2016) employed the PASCAL VOC dataset, which consists of 20 object classes, resulting in  $S = 7, B = 2, c = 20$ . As a result, the YOLO model network structure produces a  $7 \times 7 \times 30$  tensor. YOLOv3-tiny and YOLOv4-tiny model network structures are used for feature extraction in the detection layers.

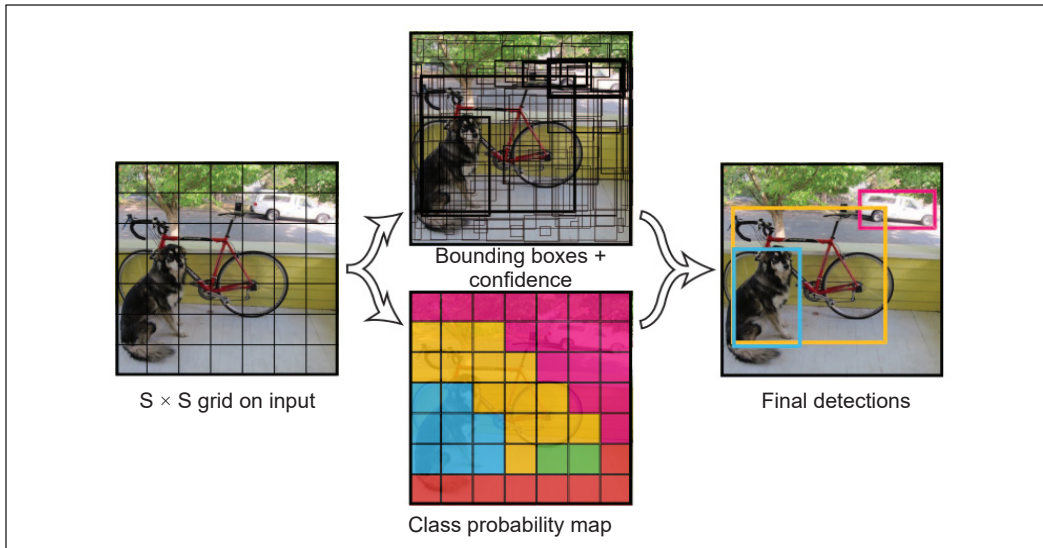


Figure 3. YOLO algorithm detection calculation (Redmon et al., 2016)

### YOLO Model Training

The network was set up in model training to identify a car class. The number of filters is set directly in the convolutional network layer (Silva et al., 2020). The formula of the filter number is in Equation 4.

$$\begin{aligned}
 \text{Filters} &= (\text{classes} + w + h + x + y + \text{Confidence score}) \times \text{num} \\
 &= (\text{classes} + 1 + 1 + 1 + 1 + 1) \times \text{num} \\
 &= (\text{classes} + 5) \times \text{num}
 \end{aligned}
 \tag{4}$$

where 5 represents the conditions ( $x, y$  - position of boundary box,  $w, h$  - width and height of the image and  $C$  - the Confidence score), and  $c$  is the class probability. YOLO detect 3 boxes per grid cell (Equation 5), so the

$$\text{Filters} = (\text{classes} + 5) \times 3
 \tag{5}$$

Therefore, there are 18 filters for 1 class. These 18 filters will be placed before two YOLO layers.

**YOLOv3.** The object detection algorithm YOLOv3 is in its third iteration of You Only Look Once. The accuracy has been greatly improved over earlier methods (Wang, 2021). Among these single-stage detectors, the YOLO family model may have the fastest object-detecting algorithm with the greatest achieved accuracy rate (Redmon et al., 2016). The real-time performance of YOLO series models reported in the literature is evaluated using a graphics processor unit (GPU) card with high-performance computational power (Redmon & Farhadi, 2018). YOLO predicts the class and object position using a forward convolutional network. With the creation of the basic network, YOLOv3, which uses Darknet-53 to replace the backbone network and uses multi-scale characteristics to determine the target, was proposed (Redmon & Farhadi, 2018).

YOLOv3 improves the network by introducing a residual module based on Darknet-19, which is You Only Look Once version 2's (YOLOv2) backbone and expands the network (Wang, 2021). Darknet-53, the enhanced network, contains 53 convolution layers. Table 1 shows the YOLOv3 network structure (Ezat et al., 2021).

Table 1  
*YOLOv3 network structure*

	Type	Filters	Size/Stride	Output
	Convolutional	32	$3 \times 3/1$	$256 \times 256$
	Convolutional	64	$3 \times 3/2$	$128 \times 128$
1×	Convolutional	32	$3 \times 3/1$	
	Convolutional	64	$3 \times 3/1$	
	Residual			$128 \times 128$
	Convolutional	128	$3 \times 3/2$	$64 \times 64$
2×	Convolutional	64	$3 \times 3/1$	
	Convolutional	128	$3 \times 3/1$	
	Residual			$64 \times 64$
	Convolutional	256	$3 \times 3/2$	$32 \times 32$
8×	Convolutional	128	$1 \times 1/1$	
	Convolutional	256	$3 \times 3/1$	
	Residual			$32 \times 32$
	Convolutional	512	$3 \times 3/2$	$16 \times 16$
8×	Convolutional	256	$1 \times 1/1$	
	Convolutional	512	$3 \times 3/1$	
	Residual			$16 \times 16$
	Convolutional	1024	$3 \times 3/2$	$8 \times 8$
4×	Convolutional	512	$1 \times 1/1$	
	Convolutional	1024	$3 \times 3/1$	
	Residual			$8 \times 8$
	Avgpool		Global	
	Convolutional	1000	$1 \times 1/1$	$8 \times 8$
	Softmax			

**YOLOv4.** YOLOv4 is a one-stage object detection algorithm that builds on YOLOv3 with various additional techniques and features. The YOLOv4 network structure is constructed of the Cross Stage Partial Darknet-53 (CSPDarknet53) backbone, the Spatial pyramid pooling (SPP) additional module, the Path Aggregation Network (PANet) path-aggregation neck, and the anchor-based YOLOv3 head (Bochkovskiy et al., 2020). Figure 4 shows the network structure of YOLOv4 (Abdurahman et al., 2021).

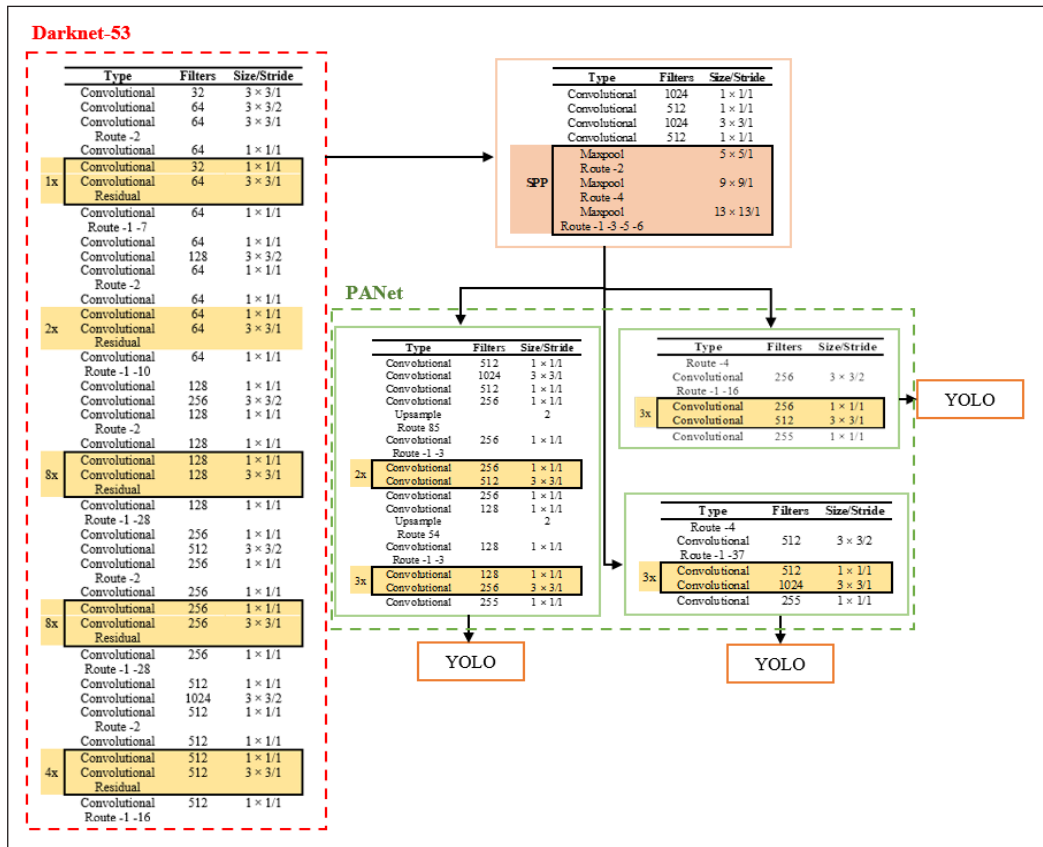


Figure 4. YOLOv4 network structure (Abdurahman et al., 2021)

### Anchor Box and Intersection over Union

According to the study by Ezat et al. (2021), the Intersection over Union (IoU) metric is utilised to evaluate object detection accuracy by comparing the overlap of two boxes. The calculation of IoU for various configurations of bounding boxes is illustrated in Figure 5.

### Loss Function

According to the study of Cepni et al. (2020), the loss function is calculated using a sum squared error. The loss function consists of three parts: localisation error, confidence

error and classification error (Equation 6).

$$\begin{aligned} \text{loss} &= \sum_{i=0}^{S^2} \text{coordError} + \text{iouError} \\ &+ \text{classError} \end{aligned} \quad [6]$$

where  $S$  is the number of grids,  $\text{coordError}$  is the localisation error,  $\text{iouError}$  is the confidence error, and  $\text{classError}$  is the classification error.

**Localisation Error.** In the equation, localisation error is represented by  $\text{coordError}$ . The localisation error is used to calculate the accuracy of the estimated bounding boxes. Equation 7 uses error to measure the position and dimensions of bounding boxes (Zhang et al., 2019).

$$\begin{aligned} \text{coordError} &= \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} [(x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2] \\ &+ \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{\text{obj}} \left[ (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right] \end{aligned} \quad [7]$$

where  $S$  is the number of grids,  $B$  is the boundary box,  $x$  and  $y$  are the centre of the box relative to the bounds of the grid cell (coordinate system), and  $w$  and  $h$  are the width and height of the entire image (scale of the original image),  $\lambda_{\text{coord}}$  increase the weight of loss in the coordinate of the box and its value is fixed as  $\lambda_{\text{coord}} = 5$ ,  $1_{ij}^{\text{obj}} = 1$  if the  $j$ th boundary box in cell  $i$  is responsible for detecting the object, otherwise 0.

**Confidence Error.** The confidence error is represented by  $\text{iouError}$ . It is used to measure differently by the presence or absence of the object and bounding box (Jiang et al., 2020). The first line of Equation 8 represents the object's presence and is detected by the bounding box, while the second line of Equation 8 represents the object's absence and is not detected by the bounding box.

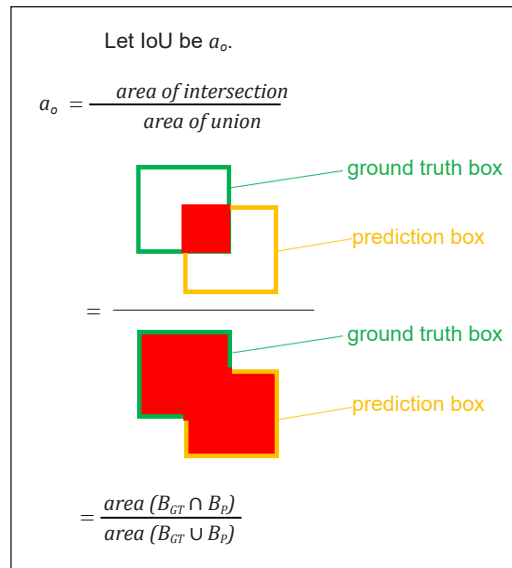


Figure 5. IoU calculation (which  $B_{CT}$  refers to the ground truth box represented by a green square, while  $B_P$  refers to the prediction box represented by an orange square. The IoU value can be calculated as the area of overlap between these two boxes, which is denoted by  $a_o$ )

$$iouError = \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{obj} (C_i - \hat{C}_i)^2 + \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B 1_{ij}^{noobj} (C_i - \hat{C}_i)^2 \quad [8]$$

where  $S$  is the number of grids,  $B$  is the boundary box,  $\hat{C}_i$  is the prediction confidence score,  $\lambda_{noobj}$  is the weights to prevent unbalanced classes because most of the bounding boxes do not detect objects or is a background, and its value is fixed as  $\lambda_{noobj} = 0.5$  to prevent the loss be tilted toward negative values or objects,  $1_{ij}^{noobj}$  is the supplement of  $1_{ij}^{obj}$ .

**Classification Error.** The classification error is represented by *classError*. It calculates the correctness of the estimated object (Equation 9) (Jiang et al., 2020).

$$classError = \sum_{i=0}^{S^2} 1_{ij}^{obj} \sum_{c \in classes} (p_i(c) - \hat{p}_i(c))^2 \quad [9]$$

where  $S$  is the number of grids,  $1_{ij}^{obj}$  if an object is present, otherwise 0,  $\hat{p}_i(c)$  is the prediction class probability for class  $c$ .

### Post Process

According to Adarsh et al. (2020), Non-Maximum Suppression (NMS) is an additional step after object detection to remove redundant predictions and choose the most accurate prediction for each object in an image. This technique involves computing the IoU between every pair of predictions and eliminating the prediction with the lower confidence score.

### Final Output

The monitoring system will identify the object and display it in the image. The final output is shown in the final detection in Figure 6.

### Faster RCNN

In the research of Girshick (2015), Faster RCNN was developed. Figure 12 depicts the two stages in faster RCNN. It is a hybrid of Fast RCNN and the Region Proposal Network (RPN). Figure 7 shows the architecture of Faster RCNN.

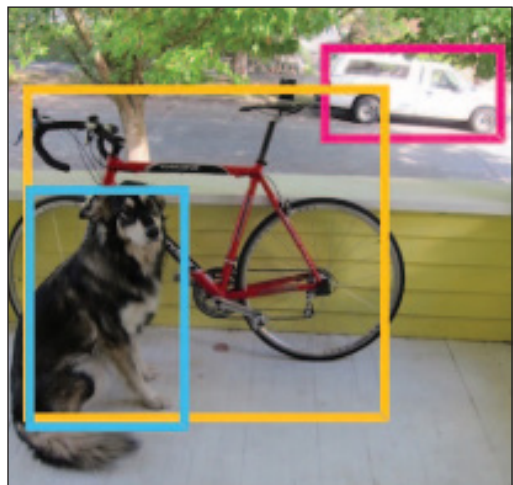


Figure 6. YOLO final detection output (Redmon et al., 2016)



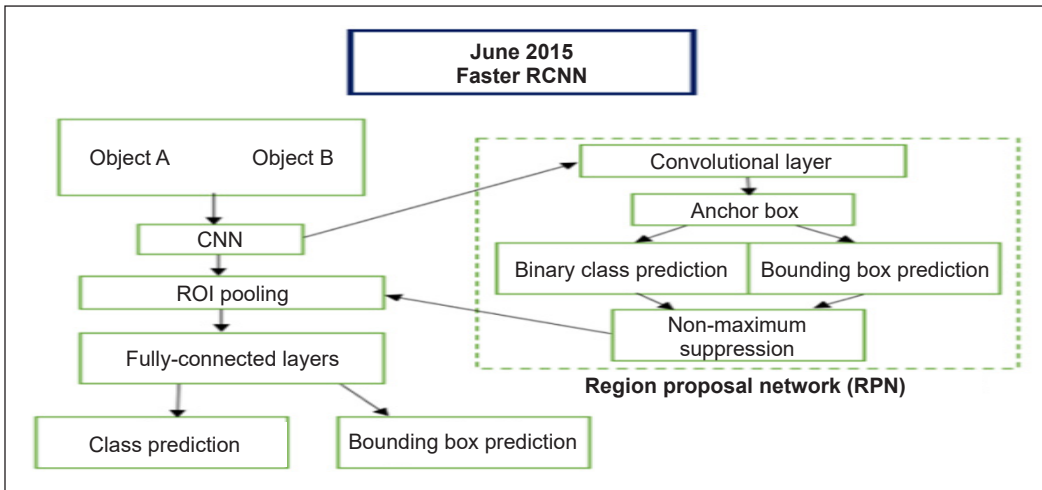


Figure 7. Faster RCNN architecture

**RPN.** Figure 8 shows the first stage of Faster RCNN, known as RPN. Faster RCNN will replace the external algorithm, selective search, with an RPN, which will share many parameters with Fast RCNN. So, to share the convolutional layers, utilise this concept of parameter sharing, a region proposal network, and Fast RCNN around the network (Zhao et al., 2020).

When an image is captured, it is passed through convolutional layers, beginning and ending at a particular feature map (Ren et al., 2017). Due to the external algorithm’s need to avoid providing object proposals, all of these actions are shared. It should be replaced rather than removed entirely. Proposals with dense anchor boxes are replaced for each

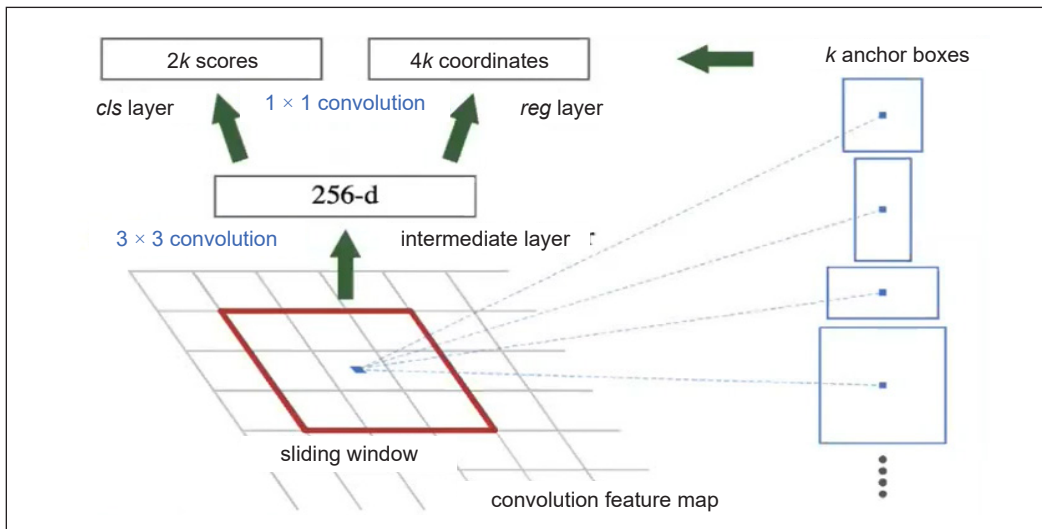


Figure 8. RPN (Ren et al., 2017)

pixel in the convolutional feature map, which considers  $k$  anchor boxes. Then, there are several anchor boxes in the pixels.

Following the study of Wang et al. (2019),  $w$  times  $h$ , which is the feature map's resolution, multiplied by  $k$  of those anchor boxes. Then, make any necessary adjustments to the anchor boxes. The Region Proposal Network will detect the object's existence. So, for each anchor, there are two classes: one for adjusting the anchor boxes. According to the research of Gou et al. (2019), the Region Proposal Network is led by two heads.  $2k$  scores represent the object's existence. Then,  $4k$  coordinates denote the top, bottom, left, and right from the centre point to the bounding box. The pixels are denoted by  $W$  and  $H$ . The red square in the illustration above is the  $3 \times 3$  convolution in the section of the translation-invariant anchor. As a result, it contains nine anchors with three scales and three aspect ratios. It uses three sizes for anchors, with box areas of  $128 \times 128$ ,  $256 \times 256$ , and  $512 \times 512$  pixels and aspect ratios of 1:1, 1:2, and 2:1.

The Region Proposal Network is trained on a good loss function about classification and regression. Then, there is the problem of properly labelling data and observing the intersection over the union with ground truth. The equation below shows the loss function in the Region Proposal (Sekar & Perumal, 2021). The loss function comprises the loss function for the classifier,  $L(C_i)$  and the loss function for the regressor,  $L(R_i)$  (Equation 10).

$$L\{p_i, t_i\} = [L(C_i) + L(R_i)] \quad [10]$$

where:  $i$  represents the selected proposal region index;  $p_i$  is the probability that the candidate box  $i$  is the object;  $t_i$  is the position coordinates of the predicted box.

The loss function of the classifier,  $L(C_i)$  (Equation 11) (Jiang & Shi, 2021).

$$L(C_i) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) \quad [11]$$

where:  $C_i$  is the classifier in candidate box  $i$ ;  $N_{cls}$  is the number of anchors available in mini-batch, is normalised and weighted by a parameter  $\lambda$ ;  $i$  represents the selected proposal region index;  $L_{cls}$  is classification loss, which is a log function for classifying object or not object;  $p_i$  is the value of predicted probability in the candidate box  $i$ ;  $p_i^*$ , also called a ground truth label, which is used to mark the anchor positive when  $p_i^* = 1$  or negative when  $p_i^* = 0$ ;

The classification loss,  $L_{cls}$ , is the logarithm loss for two categories which are object and non-object (Equation 12) (Wang et al., 2019).

$$L_{cls}(p_i, p_i^*) = \begin{cases} -\log p_i & \text{if } p_i^* = 1 \\ -\log(1 - p_i) & \text{if } p_i^* = 0 \end{cases} \quad [12]$$

The loss function of the regressor,  $L(R_i)$  (Equation 13) (Jiang & Shi, 2021).

$$L(R_i) = \lambda \frac{1}{N_{reg}} \sum_i L_{reg}(t_i, t_i^*) \quad [13]$$

where:  $R_i$  is the regressor in candidate box  $i$ ;  $\lambda$  is a constant value or a balancing factor;  $N_{reg}$  is the number of anchor locations normalised and weighted by a parameter  $\lambda$ ;  $i$  represents the selected proposal region index;  $L_{reg}$  is regression loss;  $t_i$  is the position coordinates of the predicted box;  $t_i^*$  is the coordinate vector of the corresponding object bounding box.

The regression loss,  $L_{reg}$ , is the loss of the  $i$ th region given by the Equations 14 and 15 (Wang et al., 2019).

$$L_{reg}(t_i, t_i^*) = smooth_{L1}(t_i, t_i^*) \quad [14]$$

$$smooth_{L1}(m) = \begin{cases} 0.5m^2, & |m| < 1 \\ |m| - 0.5, & |m| \geq 1 \end{cases} \quad [15]$$

where  $smooth_{L1}(m)$  is a smooth function of function  $m$ .

The parameter of regression expression,  $\{t_i\}$  and the region values of  $x, y, w$  and  $h$  are shown in Equation 16 (Jiang & Shi, 2021).

$$\begin{aligned} t_x &= \frac{(x-x_a)}{w_a}, & t_y &= \frac{(y-y_a)}{h_a} \\ t_w &= \log\left(\frac{w}{w_a}\right), & t_h &= \log\left(\frac{h}{h_a}\right) \\ t_x^* &= \frac{(x^*-x_a)}{w_a}, & t_y^* &= \frac{(y^*-y_a)}{h_a} \\ t_w^* &= \log\left(\frac{w^*}{w_a}\right), & t_h^* &= \log\left(\frac{h^*}{h_a}\right) \end{aligned} \quad [16]$$

where  $x, y$  are the predicted bounding box coordinates;  $x^*, y^*$  is the ground truth bounding box coordinates;  $w, h$  is the width and height of the predicted bounding box;  $w^*, h^*$  is the width and height of the ground truth bounding box;  $t_x, t_y$  is the regression value of  $x$  and  $y$  of the predicted bounding box coordinates;  $t_x^*, t_y^*$  is the regression value of  $x$  and  $y$  of the ground truth bounding box coordinates;  $t_w, t_h$  is the regression value of the width and height of the predicted bounding box;  $t_w^*, t_h^*$  is the regression value of the width and height of the ground truth bounding box;  $t_i = \{t_x, t_y, t_w, t_h\}$  is a vector-prediction parametrised candidate frame coordinate;  $t_i^* = \{t_x^*, t_y^*, t_w^*, t_h^*\}$  is the coordinate vector of real boundaries.

The following step is an alternate training process (Ren et al., 2017). To begin, transfer learning from an ImageNet pre-trained network to the proposed network for the region. Second, repeat the transfer learning from ImageNet to Fast RCNN. The Region Proposal Network proposes objects for Fast RCNN to consider for object locations. This detector network is then used to initialise RPN. Before fine-tuning the final layers, transfer learning

from Fast RCNN to an RPN network and fix the convolutional layers. The neural network's head then performs the same thing. Finally, fine-tune the faster scene while keeping the shared convolutional layers alone. It is how things will eventually become completely convolutional.

**Fast RCNN.** The second stage contains a detection network that employs Fast RCNN (Ren et al., 2017). The RPN and shared convolutional features produce proposals, which are then input into the Region of Interest pooling (RoI) pooling layer, followed by the remaining layers of the backbone CNN to predict the class and class-specific box refinement for each proposal (Lokanath et al., 2017).

### Social Distancing Monitoring

Social Distancing Monitoring is an approach that employs CCTV footage to observe and calculate the distance between vehicles. Figure 9 depicts the flowchart for Social Distancing Monitoring that recognises each vehicle in videos and shows the distance between vehicles through different coloured bounding boxes (Lim, Sadullah et al., 2022).

### Bird's Eye View

Bird's Eye View is a method for detecting vehicles and converting them into points. It shows the distance between the points in the specifically drawn area. The various coloured points show the distance between vehicles. Figure 10 illustrates the flowchart of the Bird's Eye View on vehicle detection (Lim, Ali et al., 2022).

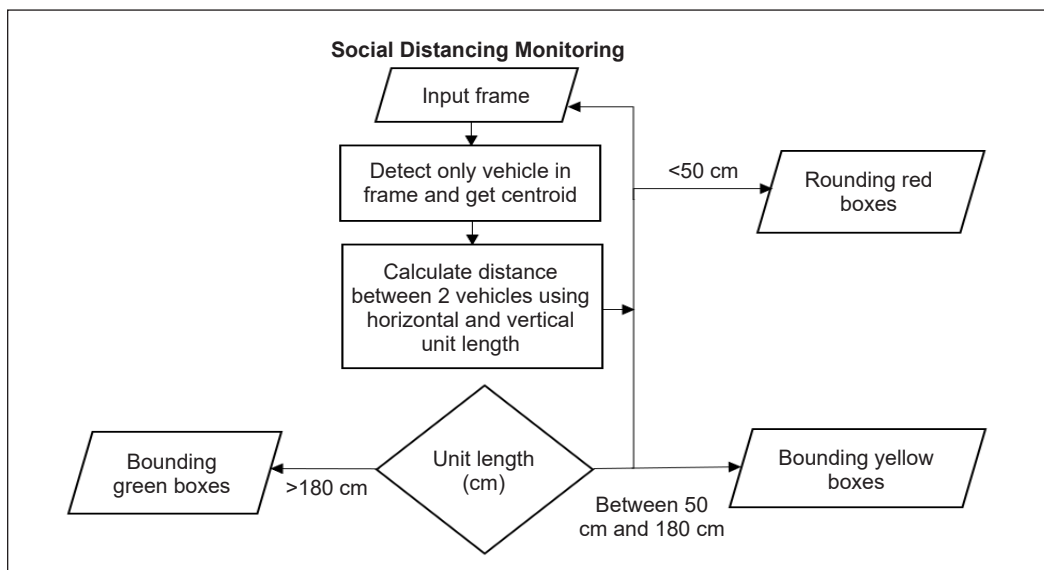


Figure 9. Social Distancing Monitoring Algorithm (Lim, Ali et al., 2022)

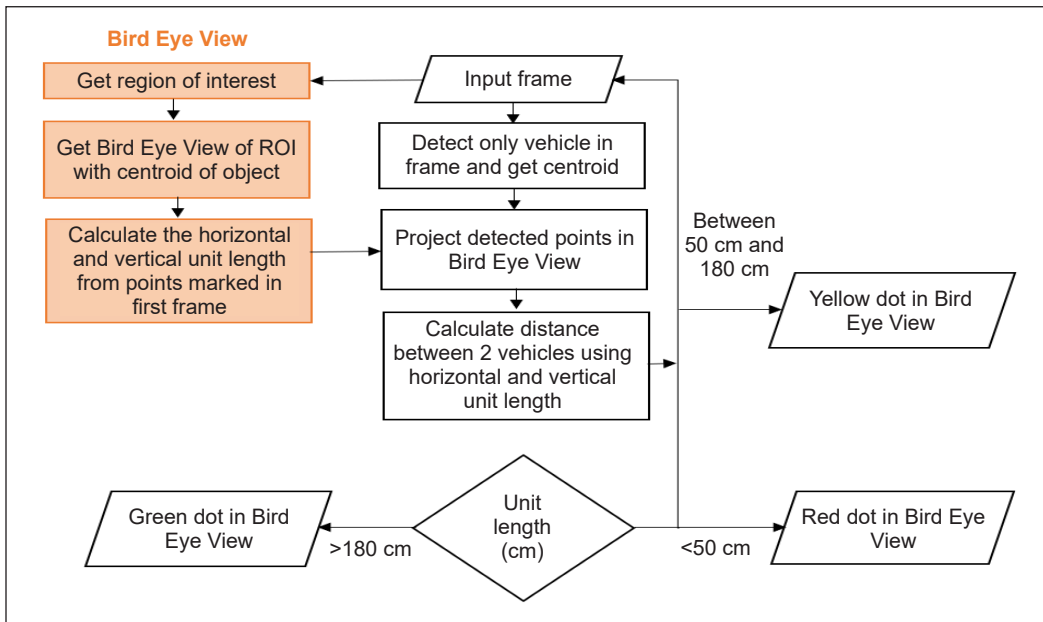


Figure 10. Bird's Eye View algorithm (Lim, Sadullah et al., 2022)

## RESULTS AND DISCUSSION

### YOLOv3 Results

In YOLOv3, vehicle and near-miss detection were conducted together in Social Distancing Monitoring and Bird's Eye View. The Social Distancing Monitoring technique displays vehicle detection in monitoring and closeness in videos (Vinitha & Velantina, 2020), whereas the Bird's Eye View presents cars as points in a specified drawn box and displays the risk level. The algorithms described above are employed in YOLOv3 to observe and calculate real-world near misses (Ong, 2020). Only cars will be detected in this study.

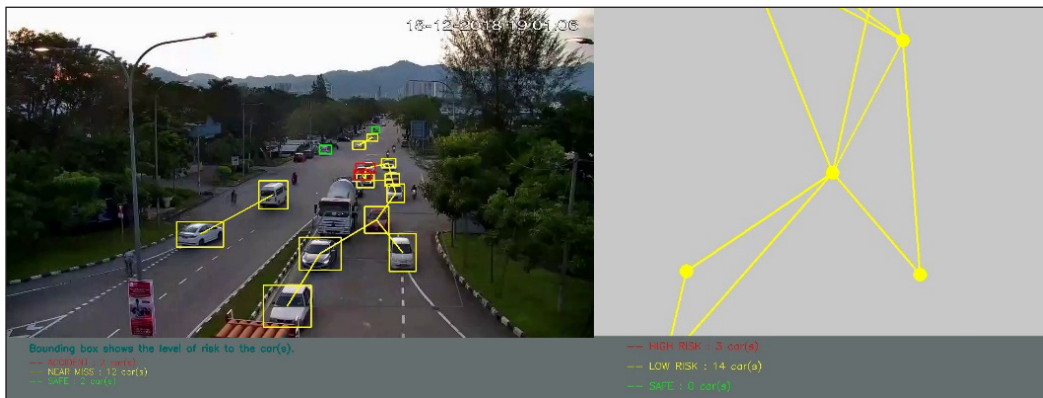


Figure 11. Near miss detection result in YOLOv3

Figure 11 exclusively demonstrates that YOLOv3 conducts vehicle detection to detect the likelihood of accidents and near misses between cars. The Lebuhraya Tun Dr Lim Chong Eu video was taken from 18/12/2018, 19:00:00 to 18/12/2018, 19:01:19. Only cars within the specified rectangle area will be displayed using the Bird’s Eye View method. Because it is in a poor perspective, the results outside the rectangle region will be ignored, and only cars within the drawn rectangle zone will be counted.

Table 2 compares different durations of the same video (in the 20s, 40s, 60s, and 80s). In the research of Sonnleitner et al. (2020), three videos with a 10-minute duration and 30 frames per second were acquired online to carry out the study. When the length of the video is increased, the running time also increases. It is a reason for this study’s choice of a video duration of 20 seconds since the computing time is too long for the computational procedure when conducting the image process using YOLOv3 compared to an 80-second video. Even though the data displayed in the 80-second video is more dependable than the data shown in the 20-second video, counting data manually is a huge undertaking for recording data. The longer the video length, the higher the number of frames obtained.

The videos show that the percentage of near-miss detection is high. It demonstrates that the vehicles in the videos are too near to one other while coming to a halt in front of a traffic signal or when stuck in traffic. So, the percentage of near-miss detection increased from 35.56% to 53.27%.

In the paper by Huang et al. (2020), the researchers gathered data from three distinct videos, each lasting 40 seconds. The videos contain different weather and scenarios. The YOLOv3 algorithm is used for video traffic monitoring. They also mention the vision field and the location of the video collection. The research shows missed detection when there is a large video data collection. If the field of view is too large, it will cause error detection. In

Table 2  
Comparison of various length videos in YOLOv3

CCTV video time		19:00:00 – 19:00:19	19:00:19 – 19:00:39	19:00:39 – 19:00:59	19:00:59 – 19:01:19
Duration of video (s)		20	40	60	80
Computational time (s)		612	1223	1835	2447
Total number of frames, D		599	1199	1799	2399
Error detection	Number of frames, A	6	166	315	316
	Percentage, A/D x 100%	1 %	13.85 %	17.51 %	13.17 %
	Object detected	Motorcycle	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck
Near miss detection	Number of frames, B	213	478	826	1278
	Percentage, B/D x 100%	35.56 %	39.87 %	45.91 %	53.27 %

this study, the duration of the 20s, 40s, 60s, and 80s videos were tested in the experiment. The vision-blocked view and the place of cameras might be factors in future work.

## YOLOv4 Results

YOLOv4 is applied to conduct Social Distancing Monitoring and Bird's Eye View to simultaneously conduct vehicle and near-miss detection. Object detection and near-miss detection are displayed in the monitoring system.

Figure 12 displays car detection and near-miss detection by using YOLOv4. The cameras were set up at Lebuhraya Tun Dr Lim Chong Eu, and the footage was captured between 18/12/2018, 19:00:00, 18/12/2018, and 19:01:19. The image on the right resulted from using Bird's Eye View. The vehicles in the drawn boxes, or RoI, are displayed in the results.

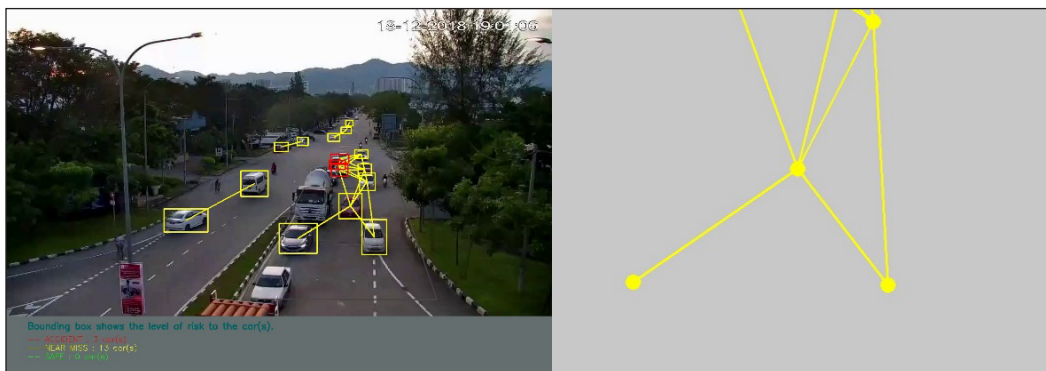


Figure 12. Near miss detection result in YOLOv4

Table 3 analyses the duration of 20s, 40s, 60s, and 80s videos by using the YOLOv4. In this experiment, a 20-second video was selected since the computation for the image process with YOLOv4 was much too long compared to an 80-second video. Even though the data in the 80-second video is more reliable than in the 20-second video, physically gathering data is a significant amount of labour for data recording. As the duration of the video rises, so does the total frame rate.

Sowmya and Radha (2021) proposed an algorithm's classifier (YOLOv4) that is tested using a computer vision approach. It uses a custom vehicle dataset that includes 3,500 images of trucks and buses. According to experimental data, the proposed technique has the best detection accuracy among various YOLO applications. Kumar et al. (2020) show that YOLOv4 investigates object detection in video and object detection in grayscale video. Objects are detected frame by frame in a video for various object categories in the sample image of video with a duration of 2 minutes and 22 seconds for multiple object detection and the sample image of grayscale video with a duration of 9 seconds for multiple object detection, which is input for multiple object detection.



Table 3  
Comparison of various length videos in YOLOv4

CCTV video time		19:00:00 – 19:00:19	19:00:19 – 19:00:39	19:00:39 – 19:00:59	19:00:59 – 19:01:19
Duration of video (s)		20	40	60	80
Computational time (s)		671	1343	2015	2687
Total number of frames, D		599	1199	1799	2399
Error detection	Number of frames, A	11	174	326	434
	Percentage, A/D x 100%	1.87 %	14.51 %	18.12 %	18.09 %
	Object detected	Motorcycle	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck
Near miss detection	Number of frames, B	354	835	1353	1913
	Percentage, B/D x 100%	59.1 %	69.64 %	75.21 %	79.74 %

Table 3  
Comparison of various length videos in YOLOv4

CCTV video time		19:00:00 – 19:00:19	19:00:19 – 19:00:39	19:00:39 – 19:00:59	19:00:59 – 19:01:19
Duration of video (s)		20	40	60	80
Computational time (s)		671	1343	2015	2687
Total number of frames, D		599	1199	1799	2399
Error detection	Number of frames, A	11	174	326	434
	Percentage, A/D x 100%	1.87 %	14.51 %	18.12 %	18.09 %
	Object detected	Motorcycle	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck
Near miss detection	Number of frames, B	354	835	1353	1913
	Percentage, B/D x 100%	59.1 %	69.64 %	75.21 %	79.74 %

### Faster RCNN Results

The Faster RCNN only detected cars in the experiment data. There are four different lengths of footage: 20s, 40s, 60s, and 80s. Figure 13 only shows cars identified by the Faster RCNN. The green boxes represent spotted vehicles and display the percentage similarity of vehicles in training datasets.

Table 4 shows the results of comparing four different video durations. The time required to compute the algorithm increases as the video length increases. Due to the time taken by Faster RCNN being too long, this study used a 20-second video instead of an 80-second

video. The number of frames taken increases, as does the video length. As a result, Faster RCNN takes longer to compute than YOLOv3 and YOLOv4 (Alganci et al., 2020). The accuracy results in Faster RCNN are higher than YOLO (Dixit et al., 2019). The research of Dixit et al. (2019) shows that the precision of Faster RCNN is the highest, but the speed of object detection is the slowest compared to other object detection models.

Faster RCNN has error-detected motorcycles, a lorry and a cement truck as cars. The percentage of error detection increased from 1.34% to 59.29%, correspondingly to 20s video until 80s video. It also does not detect small objects (cars) in further places.

Hariato et al.'s (2021) research employed the Google Open Image Dataset v6 and four different classes of vehicles. The study collected 1000 images from each class, with 800 photos utilised as training data and the rest as research data. The vehicle detection was then performed using Faster RCNN on the Tesla K80 GPU. According to Tariq et al. (2021), the data was collected independently. The collected dataset contains five different colours of cars and four different views of cars. Five hundred images were distributed to each class. Therefore, the researchers obtained 2500 images from 15 high-quality camera streams with a 20-minute duration.

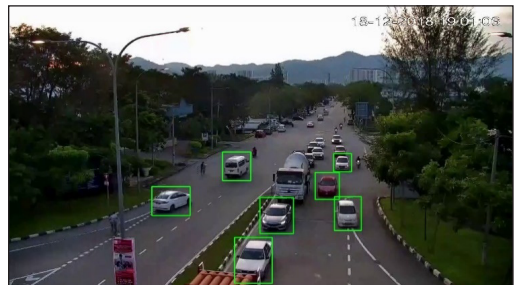


Figure 13. Result of vehicle detection in Faster RCNN

Table 4  
Comparison of various length videos in Faster RCNN

CCTV video time		19:00:00 – 19:00:19	19:00:19 – 19:00:39	19:00:39 – 19:00:59	19:00:59 – 19:01:19
Duration of video (s)		20	40	60	80
Computational time (s)		1534	3072	4608	6144
Total number of frames, D		599	1200	1800	2400
Error detection	Number of frames, A	8	253	903	1423
	Percentage, A/D x 100%	1.34 %	21.08 %	50.17 %	59.29 %
	Object detected	Motorcycle	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck	Motorcycle, lorry, cement truck

### Comparison Between YOLOv3, YOLOv4 and Faster RCNN

The videos have been experimented with in YOLOv3, YOLOv4 and Faster RCNN. Table 5 compares YOLOv3, YOLOv4 and Faster RCNN in the videos. The following comparison is based on Tables 2, 3 and 4.

Table 5  
Comparison detections in three systems

Difference	YOLOv3	YOLOv4	Faster RCNN
Method	Social Distancing Monitoring and Bird's Eye View	Social Distancing Monitoring and Bird's Eye View	Vehicle detection
Speed	Fast	Faster	Slow
Error detection	Higher than Faster RCNN	Higher than Faster RCNN	Low
Near miss detection	Very high	Lower than YOLOv3	null
Accident detection	Low	Lower than YOLOv3	null

The quality of the footage given by MBPP varies based on their position. They are not in the same place and have various levels of video quality. In 2015, MBPP deployed 534 CCTV cameras and another 1841 in 2019. Some places already installed CCTV cameras in 2015, although some were of poor quality compared to other locations.

Finally, it may be concluded that the comparison between Faster RCNN, YOLOv3 and YOLOv4. It requires more computing time to run Faster RCNN compared to YOLOv3 and YOLOv4 in videos because YOLOv3 and YOLOv4 have a simplified construction and Faster RCNN has been trained to do categorisation and regression of bounding boxes at the same time.

Faster RCNN displayed more precise results when compared to YOLOv3 and YOLOv4 in the footage of Lebuhraya Lim Chong Eu. Faster RCNN required more dataset samples depending on images or videos to train the algorithm, while YOLOv3 and YOLOv4 did not need to train the algorithm because their dataset trained it. Both algorithms have detection errors in vehicle detection. YOLOv3, YOLOv4 and Faster RCNN detected motorcycles, trucks, and lorries as vehicles in the Lebuhraya Lim Chong Eu videos.

From Table 5, the Faster RCNN does not have data for detecting near misses and accidents. Social Distancing Monitoring and Bird's Eye View techniques are unavailable at Faster RCNN.

Figure 14 depicts the test time in the videos utilising YOLOv3, YOLOv4 and Faster RCNN. YOLOv4 uses more computational time than YOLOv3. As revealed by Wang (2021), the detection efficiency of YOLOv3, YOLOv3-tiny, YOLOv3-SPP3, YOLOv4 and YOLOv4-tiny was comparable. When compared to YOLOv4 models, YOLOv3 models require less testing time. Faster RCNN has the highest computational time compared to YOLOv3 and YOLOv4 in both quality videos because it is a two-stage detector (Soviany & Ionescu, 2018).

Figure 15 displays the detection of errors in the videos using YOLOv3, YOLOv4 and Faster RCNN. Error detection happened in several cases, such as identifying motorcycles, cement trucks, and lorries as cars in the videos. When the cars are too close to one another, and at the traffic signal or junction, YOLOv4 detects more errors than YOLOv3.

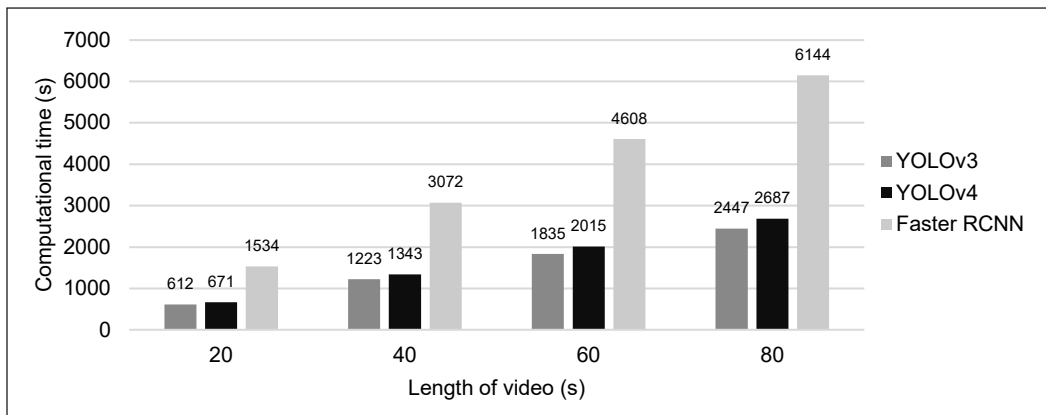


Figure 14. Comparison of computational time in three systems

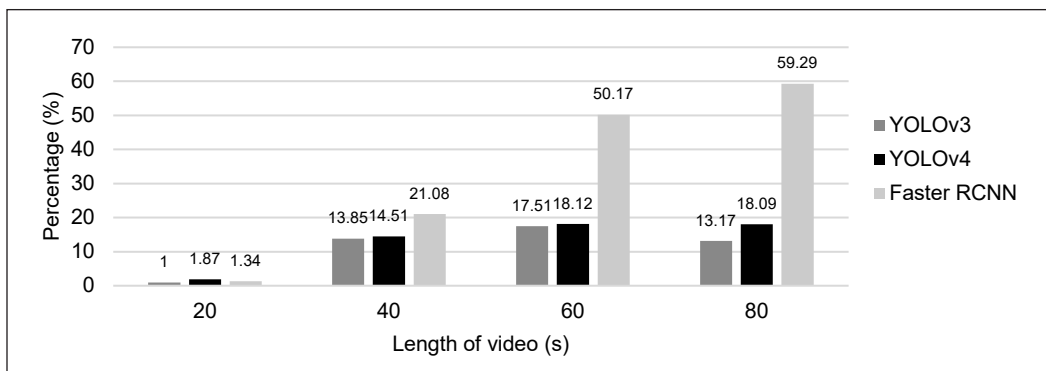


Figure 15. Comparison of error detection in three systems

Through the experiment, YOLOv3 shows higher accuracy than YOLOv4 and Faster RCNN. YOLOv3 and YOLOv4 can detect the car in the further part of the videos. Besides that, the experimental performance of YOLOv3 and YOLOv4 in identifying small and large vehicles is demonstrated by Nepal and Eslamiat (2020). In the research of Sowmya and Radha (2021), Faster R-CNN is not suited for usage in real-time videos since the system depends on the precision ratio to ensure detection speed.

Figure 16 displays the near-miss detection in the videos by applying YOLOv3, YOLOv4 and Faster RCNN. YOLOv4 shows a higher percentage of near misses than YOLOv3 in the videos. Since Faster RCNN does not have an algorithm to conduct the near-miss detection, there is no data for it.

This data lends weight to the idea that near misses occur during the busy hours in the black spot. Besides driving behaviour, the primary cause of near misses is also a major problem. According to Matsui et al. (2013), to reduce the number of deaths and serious injuries in Japan, the construction of driving safety systems requires precise functioning of the interaction scenario between the automobile and the pedestrian.

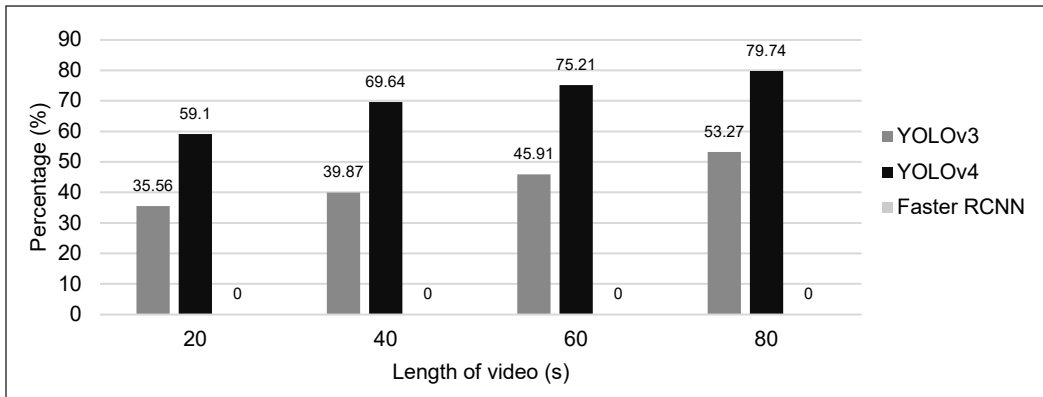


Figure 16. Comparison of near miss detection in three systems

Due to real-world accidents and a shortage of statistics, the researchers focused on near misses. As an outcome, the near miss is spotted utilising the methods of Bird's Eye View and Social Distancing Monitoring in the video recorded from the black spot position. Faster RCNN cannot detect near misses because of the failure of the Social Distancing Monitoring technique and Bird's Eye View method.

The reliability of this data is impacted by the video given by the Penang state government. The video is reliable and can be recorded as historical data for future usage. According to Calles et al. (2017), near misses warn drivers and prevent accidents. Driver behaviours and drivers' experience can contribute to accidents among young people. Future studies should collaborate with hospitals and insurance companies for more complete data and reduce the data flaws.

## CONCLUSION

The ultimate purpose of this inquiry is to learn about near misses in Penang Island so that fatalities and air pollution can be reduced in the city. This project can achieve the goals mentioned in the Sustainable Development Goals (SDGs) and is concerned with the Penang 2030 mission. The time (18/12/2018, 7:00:00 p.m. to 7:01:19 p.m.) and Lebuhraya Tun Dr Lim Chong Eu black spot location are chosen as the experiment data to apply vehicle detection techniques YOLOv3, YOLOv4, and Faster RCNN. YOLOv3 and YOLOv4 were faster and more accurate than the Faster RCNN in vehicle detection. YOLOv3 performed better than YOLOv4 in precision and speed. The probability of near misses is high, as demonstrated by Lebuhraya Tun Dr Lim Chong Eu, which YOLOv3 and YOLOv4 calculated. The quality of CCTV video and the angle of cameras need to be enhanced in future work. The length of videos can be longer if the algorithm can calculate it automatically. Near misses and accidents can be predicted by traffic flow. It can also forecast future seasonal variations, assuring the success of the Penang 2030 objective.

## ACKNOWLEDGEMENTS

The authors express their gratitude to the team of MBPP CCTV, Pulau Pinang, Malaysia, for their assistance in supplying important data, as well as the School of Mathematical Sciences, Universiti Science Malaysia, for the financing.

## REFERENCES

- Abdullah, A., & Oothariasamy, J. (2020). Vehicle counting using deep learning models: A comparative study. *International Journal of Advanced Computer Science and Applications*, 11(7), 697-703. <https://dx.doi.org/10.14569/IJACSA.2020.0110784>
- Abdurahman, F., Fante, K. A., & Aliy, M. (2021). Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models. *BMC Bioinformatics*, 22(1), Article 112. <https://doi.org/10.1186/s12859-021-04036-4>
- Aldred, R. (2016). Cycling near misses: Their frequency, impact and prevention. *Transport Research Part A*, 90(1), 69-83. <https://doi.org/10.1016/j.tra.2016.04.016>
- Adarsh, P., Rathi, P., & Kumar, M. (2020, March). YOLO v3-Tiny: Object Detection and Recognition using one stage improved model. In *2020 6th international conference on advanced computing and communication systems (ICACCS)* (pp. 687-694). IEEE Publishing. <https://doi.org/10.1109/ICACCS48705.2020.9074315>
- Alganci, U., Soydas, M., & Sertel, E. (2020). Comparative research on deep learning approaches for airplane detection from very high-resolution satellite images. *Remote Sensing*, 12(3), Article 458. <https://doi.org/10.3390/rs12030458>
- AlKishri, W., & Al-Bahri, M. (2021). Object recognition for organizing the movement of self-driving car. *International Journal of Computation and Applied Sciences*, 10(1), 1-8. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=3828692](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3828692)
- Ammar, A., Koubaa, A., Ahmed, M., Saad, A., & Benjdira, B. (2021). Vehicle detection from aerial images using deep learning: A comparative study. *Electronics*, 10(7), Article 820. <https://doi.org/10.3390/electronics10070820>
- Arinaldi, A., Pradana, J., A., & Gurusniaga, A., A. (2018). Detection and classification of vehicles for traffic video analytics. *Procedia Computer Science*, 144, 259-268. <https://doi.org/10.1016/j.procs.2018.10.527>
- Bochkovskiy, A., Wang, C., Y., & Liao, H., Y., M. (2020). *YOLOv4: Optimal speed and accuracy of object detection*. ArXiv Preprint. <https://doi.org/10.48550/arXiv.2004.10934>
- Calles, M., B., Nelson, T., & Winters, M. (2017). Comparing crowdsourced near-miss and collision cycling data and official bike safety reporting. *Transportation Research Record: Journal of the Transportation Research Board*, 2662(1), 1-11. <https://doi.org/10.3141/2662-01>
- Cepni, S., Atik, M. E., & Druan, Z. (2020). Vehicle detection using different deep learning algorithms from image sequence. *Baltic Journal of Modern Computing*, 8(2), 347-358. <https://doi.org/10.22364/bjmc.2020.8.2.10>
- Ciberlin, J., Grbic, R., Teslic, N. and Pilipovic, M. (2019). Object detection and object tracking in front of the vehicle using front view camera. In *IEEE 2019 Zooming Innovation in Consumer Technologies Conference (ZINC)* (pp. 27-32). IEEE Publishing. <https://doi.org/10.1109/ZINC.2019.8769367>

- Ding, X., & Yang, R. (2019). Vehicle and parking space detection based on improved YOLO network model. *Journal of Physics: Conference Series*, 1325, 1-7. <https://doi.org/10.1088/1742-6596/1325/1/012084>
- Dixit, K. S., Chadaga, M. G., Savalgimath, S. S., Rakshith, G. R., & Kumar, M. N. (2019). Evaluation and evolution of object detection techniques YOLO and R-CNN. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2S3), 824-829. <https://doi.org/10.35940/ijrte.B1154070782S319>
- Ezat, W., A., Dessouky, M., M., & Ismail, N., A. (2021). Evaluation of deep learning YOLOv3 algorithm for object detection and classification. *Menoufia Journal of Electronic Engineering Research (MJEER)*, 30(1), 52-57. <http://dx.doi.org/10.21608/mjeer.2021.146237>
- Gai, J., Zhong, K., Du, X., Yan, K., & Shen, J. (2021). Detection of gear fault severity based on parameter-optimized deep belief network using sparrow search algorithm. *Measurement*, 185, Article 110079. <https://doi.org/10.1016/j.measurement.2021.110079>
- Girshick, R. (2015). *Fast R-CNN*. ArXiv Preprint. <https://doi.org/10.48550/arXiv.1504.08083>
- Gou, C., Peng, B., Li, T., & Gao, Z. (2019). Pavement crack detection based on the improved faster-RCNN. In *2019 IEEE 14th International Conference on Intelligent Systems and Knowledge Engineering (ISKE)* (pp. 962-967). IEEE Publishing. <https://doi.org/10.1109/iske47853.2019.9170456>
- Gulati, I., & Srinivasan, R. (2019). Image processing in intelligent traffic management. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2S4), 213-218.
- Hariato, R. A., Pranoto, Y. M., & Gunawan, T. P. (2021). Data augmentation and faster RCNN improve vehicle detection and recognition. In *2021 3rd East Indonesia Conference on Computer and Information Technology (EICONCIT)* (pp. 128-133). IEEE Publishing. <https://doi.org/10.1109/eiconcit50028.2021.9431863>
- Huang, Y. Q., Zheng, J. C., Sun, S. D., Yang, C. F., & Liu, J. (2020). Optimized YOLOv3 algorithm and its application in traffic flow detections. *Applied Sciences*, 10(9), Article 3079. <https://doi.org/10.3390/app10093079>
- Jiang, Z., G., & Shi, X., T. (2021). Application research of key frames extraction technology combined with optimized faster R-CNN algorithm in traffic video analysis. *Complexity*, 2021, Article 6620425. <https://doi.org/10.1155/2021/6620425>
- Jiang, Z., Zhao, L., Li, S., & Jia, Y. (2020). *Real-time object detection method for embedded devices*. ArXiv Preprint. <https://doi.org/10.48550/arXiv.2011.04244>
- Kumar, B. C., Punitha, R., & Mohana. (2020). YOLOv3 and YOLOv4: Multiple object detection for surveillance applications. In *2020 Third international conference on smart systems and inventive technology (ICSSIT)* (pp. 1316-1321). IEEE Publishing. <https://doi.org/10.1109/icssit48917.2020.9214094>
- Lim, L. M., Ali, M. K. M., Ismail, M. T., & Mohamed, A. S. A. (2022). Data safety prediction using bird's eye view and social distancing monitoring for Penang roads. *Pertanika Journal of Science & Technology*, 30(4), 2563-2587. <https://doi.org/10.47836/pjst.30.4.15>
- Lim, L. M., Sadullah, A. F. M., Ismail, M. T., Awang, N., & Ali, M. K. M. (2022). Penang data safety prediction using database development and alternative road identification using Dijkstra approach. *AIP Conference Proceedings*, 2465, Article 040007. <https://doi.org/10.1063/5.0078252>



- Lokanath, M., Kumar, K. S., & Keerthi, E. S. (2017). Accurate object classification and detection by faster-RCNN. *IOP Conference Series: Materials Science and Engineering*, 263, Article 052028. <https://doi.org/10.1088/1757-899x/263/5/052028>
- Mahto, P., Garg, P., Seth, P., & Panda, J. (2020). Refining Yolov4 for vehicle detection. *International Journal of Advanced Research in Engineering and Technology (IJARET)*, 11(5), 409-419.
- Manne, G., & Raghuvanshi, N. (2017). Review on vehicle detection techniques. *International Journal of Science, Engineering and Technology Research (IJSETR)*, 6(4), 482-486.
- Maqbool, S., Khan, M., Tahir, J., Jalil, A., Ali, A., & Ahmad, J. (2018). Vehicle detection, tracking and counting. In *2018 IEEE 3rd International Conference on Signal and Image Processing (ICSIP)* (pp. 126-132). IEEE Publishing. <https://doi.org/10.1109/SIPROCESS.2018.8600460>
- Matsui, Y., Hitosugi, M., Doi, T., Oikawa, S., Takahashi, K., & Ando, K. (2013). Features of pedestrian behavior in car-to-pedestrian contact situations in Near-Miss incidents in Japan. *Traffic Injury Prevention*, 14(1), 58-63. <https://doi.org/10.1080/15389588.2013.796372>
- Meng, C., Bao, H., & Ma, Y. (2020). Vehicle detection: A review. *Journal of Physics: Conference Series CISAT 2020*, 1634, Article 012107. <https://doi.org/10.1088/1742-6596/1634/1/012107>
- Moutakki, Z., Ouloul, I. M., Afdel, K., & Amghar, A. (2018). Real-time system based on feature extraction for vehicle detection and classification. *Transport and Telecommunication*, 19(2), Article 93. <https://sciendo.com/article/10.2478/ttj-2018-0008>
- Nepal, U., & Eslamiat, H. (2020). Comparing YOLOv3, YOLOv4 and YOLOv5 for autonomous landing spot detection in faulty UAVs. *Sensors*, 22(464) 1-15. <https://doi.org/10.3390/s22020464>
- Nousi, P., Triantafyllidou, D., Tefas, A., & Pitas, I. (2019). Joint lightweight object tracking and detection for unmanned vehicles. In *2019 IEEE International Conference on Image Processing (ICIP)* (pp. 160-164). IEEE Publishing. <https://doi.org/10.1109/ICIP.2019.8802988>
- Ong, Y. (2020). *Near Miss Vehicle Collisions Estimation using YOLO*. (Unpublished master's thesis). Universiti Sains Malaysia, Malaysia.
- Pawar, B., Humbe, V., & Kundnani, L. R. (2017). Morphological approach for moving vehicle detection. *IOSR Journal of Computer Engineering (IOSR-JCE)*, 3(1), 73-80.
- Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. ArXiv Preprint. <https://doi.org/10.48550/arXiv.1804.02767>
- Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). IEEE Publishing. <https://doi.org/10.1109/cvpr.2016.91>
- Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137-1149. <https://doi.org/10.1109/tpami.2016.2577031>
- Rin, V., & Nuthong, C. (2019). Front moving vehicle detection and tracking with Kalman Filter. In *IEEE 4th International Conference on Computer and Communication Systems* (pp. 304-310). IEEE Publishing. [10.1109/CCOMS.2019.8821772](https://doi.org/10.1109/CCOMS.2019.8821772)

- Sekar, A., & Perumal, V. (2021). Automatic road crack detection and classification using multi-tasking faster RCNN. *Journal of Intelligent & Fuzzy Systems*, 41(6), 6615-6628. <https://doi.org/10.3233/jifs-210475>
- Silva, L. A., Sanchez San Blas, H., Peral García, D., Sales Mendes, A., & Villarubia González, G. (2020). An architectural multi-agent system for a pavement monitoring system with pothole recognition in UAV images. *Sensors*, 20(21), Article 6205. <https://doi.org/10.3390/s20216205>
- Sonnleitner, E., Barth, O., Palmanshofer, A., & Kurz, M. (2020). Traffic measurement and congestion detection based on real-time highway video data. *Applied Sciences*, 10(18), Article 6270. <https://doi.org/10.3390/app10186270>
- Soviany, P., & Ionescu, R., T. (2018). *Optimizing the trade-off between single-stage and two-stage deep object detectors using image difficulty prediction*. ArXiv Preprint. <https://doi.org/10.48550/arXiv.1803.08707>
- Sowmya, V., & Radha, R. (2021). Comparative analysis on deep learning approaches for heavy-vehicle detection based on data augmentation and transfer-learning techniques. *Journal of Scientific Research*, 13(3), 809-820. <https://doi.org/10.3329/jsr.v13i3.52332>
- Tariq, A., Khan, M. Z., & Khan, M. U. G. (2021). Real time vehicle detection and colour recognition using tuned features of faster-RCNN. In *2021 1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)* (pp. 262-267). IEEE Publishing. <https://doi.org/10.1109/caida51941.2021.9425106>
- Uus, J., & Krilavičius, T. (2018). Vehicle detection and classification in aerial imagery. *IEEE International Conference on Image Processing*, 2410(1), 80-85.
- Vançin, S., & Erdem, E. (2018). Detection of the vehicle direction with adaptive threshold algorithm using magnetic sensor nodes. *Journal of Polytechnic*, 21(2), 333-340.
- Vinitha, V., & Velantina, V. (2020). Social distancing detection system with artificial intelligence using computer vision and deep learning. *International Research Journal of Engineering and Technology (IRJET)*, 7(8), 4049-4053.
- Wang, C., Dai, Y., Zhou, W., & Geng, Y. (2020). A vision-based video crash detection framework for mixed traffic flow environment considering low-visibility condition. *Hindawi, Journal of Advanced Transportation*, 2020, Article 9194028. <https://doi.org/10.1155/2020/9194028>
- Wang, G., Guo, J. M., Chen, Y., Li, Y., & Xu, Q. (2019). A PSO and BFO-based learning strategy applied to faster R-CNN for object detection in autonomous driving. *IEEE Access*, 7, 18840-18859. <https://doi.org/10.1109/access.2019.2897283>
- Wang, S. (2021). Research towards yolo-series algorithms: Comparison and analysis of object detection models for real-time UAV applications. *Journal of Physics: Conference Series*, 1948(1), Article 012021. <https://doi.org/10.1088/1742-6596/1948/1/012021>
- World Health Organization. (2020, December 9). *The Top 10 Causes of Death*. World Health Organization. <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>
- Zhang, Y., Shen, Y., & Zhang, J. (2019). An improved tiny-YOLOv3 pedestrian detection algorithm. *Optik*, 183, 17-23. <https://doi.org/10.1016/j.ijleo.2019.02.038>

- Zhang, Z., Trivedi, C., & Liu, X. (2018). Automated detection of grade-crossing-trespassing near misses based on computer vision analysis of surveillance video data. *Safety Science, 110*(Part B), 276-285. <https://doi.org/10.1016/j.ssci.2017.11.023>
- Zhao, J., Wei, S., Xie, H., & Zhong, H. (2020). Image recognition of small UAVs based on faster RCNN. In *Proceedings of the 2020 4th International Conference on Vision, Image and Signal Processing* (pp. 1-6). ACM Publishing. <https://doi.org/10.1145/3448823.3448844>
- Zhao, X., Pu, F., Wang, H., Chen, H. and Xu, Z. (2019). Detection, tracking, geolocation of moving vehicle from UAV using monocular camera. *IEEE Access, 7*(1), 101160-101170.
- Zohra, A., F., Kamilia, S., & Souad, S. (2018). Detection and classification of vehicles using deep learning. *International Journal of Computer Science Trends and Technology (IJCSST), 6*(3), 23-29.

**APPENDIX A**

No.	Researchers	Type of paper	Application data			Type of method			Objects detected			Remarks
			Research	Image process	Empirical data	Survey/ Questionnaire	Model	Software	Others	Vehicles	Pedestrians	
1	Manne and Raghuvanshi (2017)	✓	✓			✓			✓			- MATLAB - India - Edge Detection - Top-Hat Processing - Masking Operation - India
2	Pawar et al. (2017)	✓	✓			✓						- Region Convolutional Neural Networks (RCNN) - Support Vector Machine (SVM) - Indonesia - Gaussian Mixture Model - Blob detection - Hungarian Algorithm - Kalman Filter - Pakistan
3	Arinaldi et al. (2018)	✓	✓			✓			✓	✓		- Background subtraction - Filtering and contour detection - North Africa
4	Maqbool et al. (2018)	✓	✓			✓			✓			- Adaptive threshold detection algorithm (ATDA) - Turkey
5	Moutakki et al. (2018)	✓	✓			✓			✓	✓		
6	Vançin and Erdem (2018)	✓	✓			✓			✓	✓		

No.	Researchers	Type of paper	Application data			Type of method			Objects detected			Remarks
			Research	Image process	Empirical data	Survey/Questionnaire	Model	Software	Others	Vehicles	Pedestrians	
7	Zhang et al. (2018)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- Histograms of Oriented Gradients</li> <li>- AdaBoost algorithm</li> <li>- Image segmentation</li> <li>- Edge detection</li> <li>- Kalman filter</li> <li>- China</li> </ul>
8	Zohra et al. (2018)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- CNN</li> <li>- North Africa</li> </ul>
9	Ciberlin et al. (2019)	✓	✓			✓		✓	✓			<ul style="list-style-type: none"> <li>- Viola-Jones algorithm</li> <li>- YOLOv3</li> <li>- Median Flow tracking</li> <li>- Correlation tracking method</li> <li>- Croatia</li> </ul>
10	Ding and Yang (2019)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- YOLOv3</li> <li>- China</li> </ul>
11	Gulati & Srinivasan (2019)	✓	✓					✓				<ul style="list-style-type: none"> <li>- Comparison between detector systems and algorithms of image processing</li> <li>- India</li> </ul>
12	Nousi et al. (2019)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- Kernelised Correlation Filter</li> <li>- YOLOv2</li> <li>- Greece</li> </ul>
13	Rin and Nuthong (2019)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- Kalman Filter</li> <li>- Thailand</li> </ul>
14	Uus and Krilavičius (2019)	✓	✓			✓			✓			<ul style="list-style-type: none"> <li>- YOLOv3</li> <li>- Lithuania</li> </ul>

No.	Researchers	Type of paper	Application data		Type of method				Objects detected			Remarks	
			Research	Image process	Empirical data	Survey/ Questionnaire	Model	Software	Others	Vehicles	Pedestrians		Motorcycles/ Bicycles
15	Zhao et al. (2019)	✓	✓			✓		✓	✓				- YOLOv3 - Kernelised Correlation Filter - China
16	Abdullah and Oothariasamy (2020)	✓	✓			✓		✓	✓				- YOLOv3 - KL, Malaysia
17	Cepni et al. (2020)	✓	✓			✓		✓	✓				- YOLOv3 - Turkey - CNN
18	Mahto et al. (2020)	✓	✓			✓		✓	✓				- YOLOv4 - India

## Development of GIS-based Ground Flash Density and its Statistical Analysis for Lightning Performance Evaluation of Transmission Lines in Peninsular Malaysia

Nurzanariah Roslan<sup>1\*</sup>, Ungku Anisa Ungku Amirulddin<sup>1</sup>, Mohd Zainal Abidin Ab. Kadir<sup>2</sup> and Noradlina Abdullah<sup>3</sup>

<sup>1</sup>*Institute of Power Engineering, Department of Electrical & Electronics Engineering, College of Engineering, Universiti Tenaga Nasional, 43000 Selangor, Malaysia*

<sup>2</sup>*Advanced Lightning, Power and Energy Research Centre (ALPER), Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

<sup>3</sup>*TNB Research Sdn Bhd, No. 1, Kawasan Institusi Penyelidikan, Jalan Ayer Itam, Kajang, 43000 Selangor, Malaysia*

### ABSTRACT

Malaysia is one of the world's highest lightning regions, making it an ideal location for studying lightning activities, as they cause many power outages on overhead transmission lines. This study presents ground flash density (GFD) mapping and statistical analysis of lightning flash data in Peninsular Malaysia, which will be used to evaluate the lightning performance of transmission lines. Using Geographical Information System (GIS) software, the GFD map and lightning flash data for statistical analysis were extracted. MATLAB was then used to perform statistical analysis and obtain the probability of peak lightning current using the generalized extreme value (GEV) distribution. This study analyzed six years of lightning flash data from 2012 to 2017 recorded by the Lightning Location System (LLS) and used the Peninsular Malaysia base map from the Department of Survey and Mapping Malaysia (JUPEM). Results show that the GFD mapping approach effectively classifies GFD distribution and identifies areas with high lightning activity.

81% of 4,536,380 lightning flashes were negative polarity, with a higher mean peak current magnitude than positive ones. More lightning activity was observed during the Southwest Monsoon (June-September) and the first Inter-Monsoon season (April-May). Pahang had the most lightning flashes due to its large land area. The GFD map overlaid on the transmission line demonstrated how lightning performance on the transmission

### ARTICLE INFO

#### Article history:

Received: 15 February 2023

Accepted: 14 June 2023

Published: 15 January 2024

DOI: <https://doi.org/10.47836/pjst.32.1.21>

#### E-mail addresses:

[nurzanariah@uniten.edu.my](mailto:nurzanariah@uniten.edu.my) (Nurzanariah Roslan)

[anisa@uniten.edu.my](mailto:anisa@uniten.edu.my) (Ungku Anisa Ungku Amirulddin)

[mzk@upm.edu.my](mailto:mzk@upm.edu.my) (Mohd Zainal Abidin Ab. Kadir)

[noradlina.abdullah@tnb.com.my](mailto:noradlina.abdullah@tnb.com.my) (Noradlina Abdullah)

\*Corresponding author



line can be assessed. These findings are useful for utility and protection engineers to improve the performance of transmission lines.

*Keywords:* Energy, GIS, ground flash density, lightning, statistical, transmission line

## INTRODUCTION

Malaysia is a tropical country with a unique location and topography and is known for having high lightning activity, with about 40 strikes per square kilometer per year (Ab-Kadir, 2016; Islam et al., 2019; Rawi et al., 2017). It makes lightning-related outages a major concern for Tenaga Nasional Berhad (TNB), the main electricity provider. TNB has found that 50 – 60% of power outages in their system are caused by lightning (Abdullah et al., 2008). Lightning strikes on phase or shield wires or towers can result in high overvoltage and electrical discharges across insulator strings, causing faults, equipment damage, energy losses, and maintenance costs, especially on 500 kV transmission lines.

To mitigate the impact of lightning on the electricity supply, TNB Research Sdn. Bhd. installed a Lightning Location System (LLS), also known as a Lightning Detection Network (LDN), in Peninsular Malaysia in 1994 (Wooi et al., 2016). The LLS determines the coordinates of cloud-to-ground lightning strikes using the Time of Arrival (TOA) and Magnetic Direction Finding (MDF) principles (Abdullah et al., 2008). The original LLS had 7 lightning sensors but has been upgraded several times (Abdullah et al., 2008; Wooi et al., 2016), as shown in Figure 1. Figure 2 shows the location of the latest upgraded lightning sensors, denoted as 1, 2, 3, 4, and 5, in Peninsular Malaysia as of 2015.

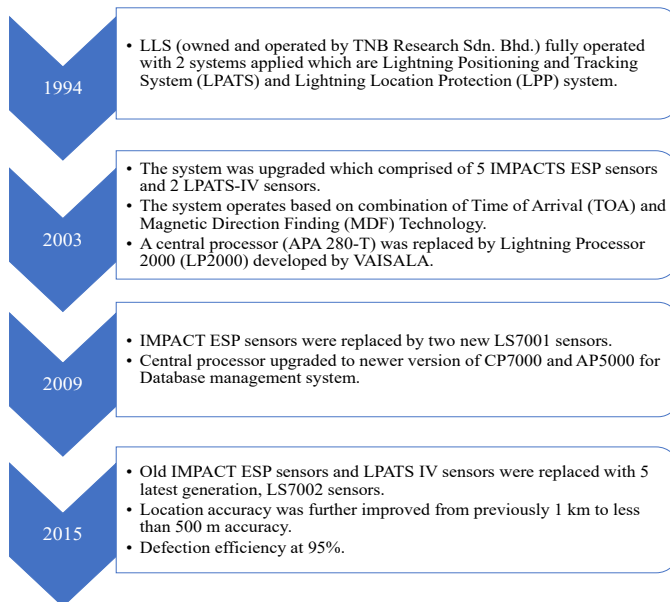


Figure 1. History of lightning location system in Peninsular Malaysia

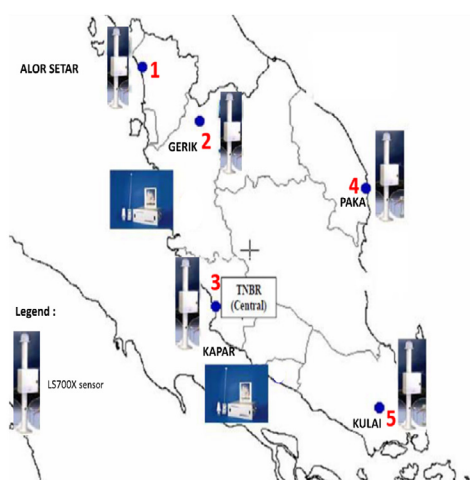


Figure 2. Location of lightning sensors in Peninsular Malaysia (Rawi et al., 2018)

Based on the information shown in Figure 1, the IMPACT ESP lightning sensors in the LLS use TOA and MDF methods to achieve the accuracy and efficiency requirements. With a location accuracy of less than 500 m and a detection efficiency of 95%, The system can detect lightning strikes up to 600 km. In addition, the sensors can reliably differentiate intra-cloud flashes and identify stroke polarity for flashes (Abdullah et al., 2008). The LLS equipment has undergone progressive upgrades to enhance its reliability, which has resulted in improved detection efficiency and location accuracy. Specifically, after a

major upgrade in 2015, the LLS has been able to achieve a detection efficiency of more than 95% for cloud-to-ground (CG) lightning activities, with a location accuracy of up to 250 m (Abdullah, 2019).

The LLS is a tool to enhance power line performance and detect electrical failures in the power network. The information from the LLS is leveraged through the Fault Analysis and Lightning Location System (FALLS) software, which the LLS manufacturer developed.

A geographic information system (GIS) is a powerful tool for storing, manipulating, managing, processing, and visualizing geospatial data. Its analytical capabilities, including statistical analysis summaries, calculations, data interrelationships, buffer generation, and overlay function (ESRI, 2010; Javor et al., 2018), make it an even more powerful tool. This technology is widely used in science and industry, providing a foundation for mapping and analysis. With GIS, users can better understand patterns, relationships, and the context of geographic information, leading to better communication and enhanced efficiency, management, and decision-making. It is also observed that there is a growing trend in the use of GIS within the electrical sector (ESRI, 2010; Husain et al., 2012; Kakumoto et al., 2016; Kezunovic et al., 2015; Korir & Ngigi, 2015; Leite et al., 2019; Li et al., 2014; Lin & Xu, 2016; Rahman et al., 2020; Yatim et al., 2019; Yan et al., 2016; Zheng et al., 2021).

Several studies have shown the widespread use and popularity of GIS applications with lightning data (Ali et al., 2018; Biswas et al., 2020; Edirisinghe & Maduranga, 2021; Farukh et al., 2017; Hodanish et al., 2019; Javor et al., 2018; Mammadov et al., 2021; Mishra et al., 2022; Misztal & Siłuch, 2021). However, due to the limitation of information provided by the commercially available LLS, utility found it difficult to work, especially on finding the root cause and correlation of tripping with some other factors such as topography and

weather data, thus limiting the research on utilizing GIS for transmission lines' lightning performance evaluation. Our research leverages GIS capabilities to create a GFD map and perform a statistical analysis of lightning flash data to evaluate lightning performance on transmission lines in Peninsular Malaysia and address the gap. The GFD map produced by GIS can also be utilized to analyze past lightning occurrences, known as historical attributes, to assess their impact on TNB assets.

### **Lightning Performance of Transmission Line**

A transmission line is vital in an electricity power system network since it transmits and distributes electrical power along the line. A direct or indirect lightning strike on the transmission line may disrupt supply. TNB's statistics from 2001 to 2013 showed that between 39% and 61% of transmission line outages are caused by lightning every year in Peninsular Malaysia (Rawi et al., 2018).

Similar scenarios were also highlighted in other countries such as Brazil, Indonesia, and China, where line outages due to lightning were reported to be 50 to 70% for 230 kV lines, 66% for 150 kV lines, and 46% for 500 kV lines, respectively (Rawi et al., 2018; Warmi & Michishita, 2018). High outage rates due to lightning activities were also reported in Russia (He & Zeng, 2010). Thus, numerous studies have been conducted on improving the lightning performance of transmission lines. According to Rawi et al. (2018) and Warmi and Michishita (2018), lightning strikes on overhead transmission lines are affected by several factors such as lightning density, tower footing resistance (TFR), soil resistivity (SR), span length of transmission line, terrain characteristics, and monsoon seasons.

The lightning performance of a transmission line is often evaluated based on the flashover rate, expressed as the number of flashovers by 100 km per year. This flashover rate of the transmission line can be further broken down into two important parameters: back flashover rate (BFR) and shielding failure flashover rate (SFFOR). The BFR refers to the annual outage rate on a tower line length basis caused by a back flashover on a transmission line, while the SFFOR is the annual number of flashovers on a tower line length basis caused by shielding failures (IEEE Power Engineering Society, 1997).

GFD is one of the most important parameters that play a significant role in evaluating the exposure of a transmission line to lightning strikes and identifying potential areas for improvement in the line's design or maintenance. This parameter, coupled with others, will contribute towards the overall assessment of the lightning performance of the transmission line and reduce the risk of outages or damage caused by lightning strike failures, as highlighted in this discussion.

### **Ground Flash Density**

The IEC 62858 standard defines ground flash density,  $N_g$ , as the number of cloud-to-ground

flashes per kilometer squared per year (International Electrotechnical Commission, 2015). This term is often viewed as the primary descriptor of lightning activity (Ab-Kadir, 2016; Bouquegneau, 2014; Phillips, 2004). Ground flash density is one of the fundamental lightning parameters that provide the basis for estimating the frequency of lightning effects on the electrical system. This parameter can be measured from records of lightning flash counters (LFC) or LLS, and it can also be estimated using thunderstorm day or hourly records. Peninsular Malaysia was reported to experience GFD values as high as 28 flashes/km<sup>2</sup>/year for about six years between 2004 and 2010 (Abdullah & Hatta, 2012).

### Back Flashover Rate (BFR)

As highlighted earlier, GFD is an important parameter for evaluating the lightning performance of transmission lines. Back flashover is a well-known dominant cause of line tripping, and it occurs when lightning strikes the transmission tower or shield wire as opposed to the shielding failure, typically with the lightning current of less than 20 kA, which normally occurs due to failure of the shield wire to protect the phase conductor (Sardi et al., 2008). Therefore, analyzing back flashover is crucial to determine the effectiveness of a transmission line's design and maintenance in withstanding lightning strikes. It can be achieved by examining various parameters, such as the Back Flashover Rate (BFR).

The BFR of the line can be calculated using the Equation 1 (IEEE Power Engineering Society, 1997; Sardi et al., 2008),

$$BFR = 0.6N_s P(I_c > I_f) \quad [1]$$

where,

$N_s$  is the number of flashes to the line per 100 km per year and is given by Equation 2:

$$N_s = N_g (28h_t^{0.6} + b) / 10 \quad [2]$$

$P(I_c > I_f)$  is the cumulative probability of the critical back flashover current,  $I_c$  exceeding  $I_f$  is then given by Equation 3:

$$P(I_c > I_f) = 1 / [1 + (\frac{I_c}{I_f})^{2.6}] \quad [3]$$

where  $N_g$  is GFD,  $h_t$  is the tower height,  $b$  is the distance between shield wires, and  $I_f$  is the median current of 31 kA.

### Monsoon Seasonal Variation of Lightning Activity

Lightning activity in tropical countries is reported to be related to the monsoon season (Dewan et al., 2022; Isa et al., 2021; Kamra & Kumar, 2021; Rawi et al., 2017, 2018; Wooi et al., 2016). Peninsular Malaysia has four monsoon seasons: the Northeast Monsoon (December to March), the Southwest Monsoon (June to September), and the two shorter

periods of inter-monsoon seasons (April to May and October to November). According to Rawi et al. (2017), the most active periods of lightning are during the inter-monsoon season, which is from April to May and October to November.

### Use Cases of Area and Lightning Flashes Data

Peninsular Malaysia is in Southeast Asia, between 1° to 7° North latitude and 99° to 105° East longitude, with a total area of 131,802 km<sup>2</sup>. It comprises 11 states and 2 federal territories, grouped into four regions: Northern, East Coast, Central, and Southern. These regions are represented by labels (N1-N4, E1-E3, C1-C3, and S1-S3) in a map shown in Figure 3, and the land areas of each state and federal territory are listed in Table 1.

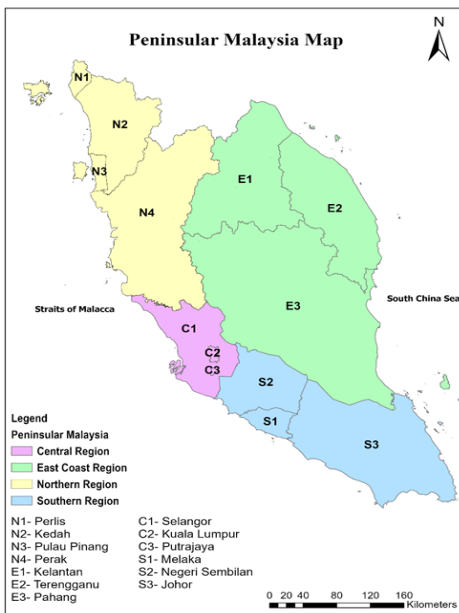


Figure 3. Map of study area replotted using ArcGIS software

Table 1  
List of regions in Peninsular Malaysia with respective land areas of state and federal territories

Region	State	Area (km <sup>2</sup> )
Northern Region	Perlis, N1	813
	Kedah, N2	9,471
	Pulau Pinang, N3	1,050
	Perak, N4	20,912
East Coast Region	Kelantan, E1	15,030
	Terengganu, E2	12,956
	Pahang, E3	35,944
Central Region	Selangor, C1	7,920
	Kuala Lumpur, C2	243
	Putrajaya, C3	49
Southern Region	Melaka, S1	1,655
	Negeri Sembilan, S2	6,653
	Johor, S3	19,106

This study utilized lightning flashes data provided by the LLS from the period of 2012 up to the year 2017. Although lightning flash data beyond 2018 was not accessible for this study, the data from this 6-year period, which comprises over 4 million records, is still valuable for analyzing lightning patterns and trends. This analysis is essential for developing an effective approach to mitigate lightning damage. Table 2 presents lightning flash data, including date and time, latitude and longitude in decimal degrees, discrimination (type of flash with polarity), and the peak current in kA.

Table 2

*Sample of original lightning data*

Date and Time	Lat (°)	Lon (°)	Discrimination	Peak Current (kA)
1/1/2017 2:27:17 AM	2.53043	103.686874	CG-	-13
1/1/2017 8:10:12 AM	5.69484	102.118904	CG-	-12
1/1/2017 8:38:46 AM	5.642838	102.334763	CG-	-40
1/1/2017 3:31:31 PM	2.689012	101.49189	CG-	-31
1/1/2017 3:33:55 PM	2.691144	101.486023	CG-	-54

Note. CG- means negative polarity cloud-to-ground flash

## METHODOLOGY

The study used ArcGIS software version 10.8.2 to process the data collected from the LLS and the base map from the Department of Survey and Mapping Malaysia (JUPEM). The data was used for statistical analysis and to create a GFD map. The selected transmission line was then assessed using the created GFD map. The procedure for the study is summarized in Figure 4.

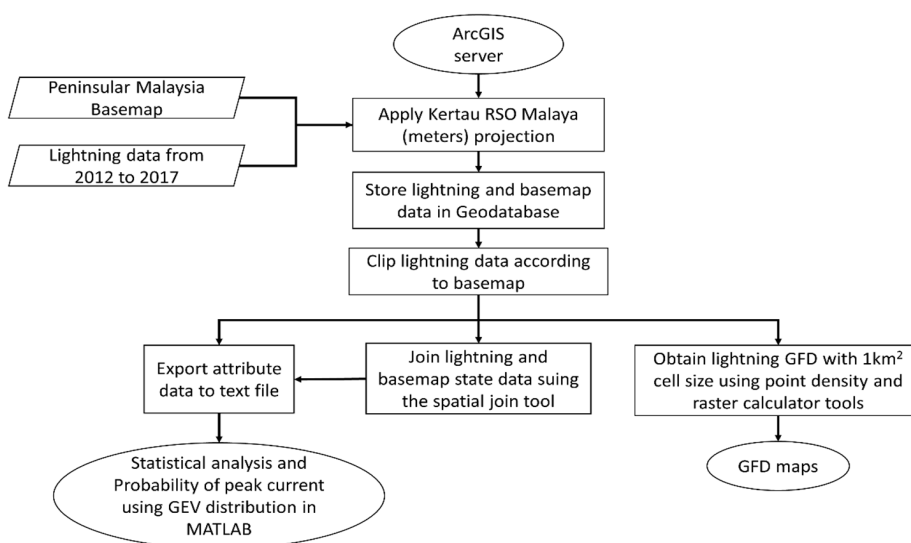


Figure 4. Procedure for statistical analysis and GFD maps

List of definitions for the terms used in Figure 4 are stated as the following (Chang, 2008; Xie et al., 2000; Yao, 2020):

- ArcGIS server serves GIS services such as maps, geodata, and image services. All services provide remote access to a geodatabase through the LAN, WAN, or Internet using an ArcGIS Server.

- Kertau RSO Malaya (Meters) is a projected coordinate system in Malaysia.
- A geodatabase is a set of structured information that can store, manage, and access data.
- Clip is a geoprocessing tool to extract input features overlaid on the clip features.
- Spatial data describes the location of spatial features on the Earth's surface based on a geographic coordinate system with longitude and latitude values.
- Attribute data describes an item or spatial feature in more detail and is stored in tabular format in the geodatabase.
- Spatial join is an analysis to append additional data based on the relative locations of the features in the two layers.
- Point density is a geoprocessing tool to calculate the density of point features around each output grid cell.
- A raster calculator is a geoprocessing tool for performing raster analysis using a Map Algebra expression.

The ArcGIS desktop software was required to create a geodatabase of the lightning and Peninsular Malaysia base map data imported into the software. This geodatabase, or “gdb.” enabled users to design, store, and manage spatial and non-spatial data in a single environment. The raw lightning flash data were then clipped to the base map of Peninsular Malaysia to remove all lightning flash data that fell outside the study area.

### **Statistical Analysis and Peak Current Probability**

As shown in Figure 4, the statistical analysis of the clipped lightning flashes data was conducted based on yearly and monthly frequency and seasonal variation during the monsoon. The analysis was performed separately for negative and positive flashes datasets with distinct characteristics. The statistical results for each dataset were calculated, including the maximum and minimum peak value (kA), median, mode, mean, and standard deviation. Additionally, the distribution of lightning flashes by state and region was included in the analysis using the spatial join tool available in ArcGIS software.

The lightning peak current probability density function (pdf) for each data set was also computed using the General Extreme Value (GEV) distribution to describe how often random lightning events happen. GEV distribution was used to represent the probability distribution of lightning flashes' peak current data, as shown in Figure 5. GEV is a family of continuous probability distributions developed within extreme values theory to combine Gumbel, Frechet, and Weibull families or specified as type I, II, and III extreme value



distributions, respectively. The GEV distribution is commonly used for representing large-range data such as extreme weather. Previous researchers discovered that using GEV distribution for extreme data was practical, popular, and satisfactory for estimating

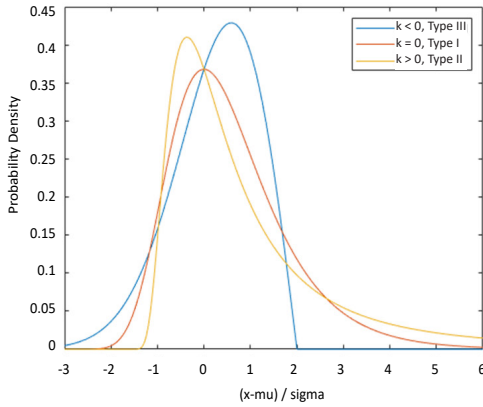


Figure 5. The generalized extreme value distribution (The MathWorks Inc., 2021)

extreme weather, such as the subtropical monsoon, extreme daily temperatures, and extreme rainfall (Shukla et al., 2010; Yendra et al., 2021).

GEV distribution combines the type I, II, and III extreme value distributions into a single family to allow for a continuous range of possible shapes. Type I distribution is normally a distributed curve without skewing, while type II and III are left-skewed and right-skewed, respectively. The curve distribution is filled with a location parameter,  $\mu$  ( $\mu$ ), a scale parameter,  $\sigma$  ( $\sigma$ ), and a shape parameter,  $k$ . When  $k < 0$ ,

the GEV is equivalent to the type III extreme value. When  $k > 0$ , the GEV changes to the type II. As  $k$  approaches 0, the GEV becomes type I in the limit.

The probability density function of the GEV distribution can be expressed as Equation 4 (Provost et al., 2018).

$$f(x; \mu, \sigma, k) = \frac{1}{\sigma} \left[ 1 + k \left( \frac{x - \mu}{\sigma} \right) \right]^{(-\frac{1}{k})-1} \exp \left\{ - \left[ 1 + k \left( \frac{x - \mu}{\sigma} \right) \right]^{-1/k} \right\} \quad [4]$$

Based on GEV distribution for the lightning data sets, the lightning peak current at 5%, 50%, and 95% occurrence probability for the entire Peninsular Malaysia were determined from 2012 to 2017. The MATLAB software implemented all the statistical analyses and probability peak currents of the lightning flashes method presented here.

### Ground Flash Density Map of Peninsular Malaysia

The lightning data for the period 2012 to 2017, imported into the ArcGIS software, were separated into negative and positive polarity data sets. The GFD maps for Peninsular Malaysia were developed using these data sets and the ArcGIS tools of point density and raster calculator, with a 1 km x 1 km grid cell.

## RESULTS AND DISCUSSION

### Annual Variation and Statistical Analysis for Peninsular Malaysia

According to historical data of lightning events in Peninsular Malaysia, 4,629,881 lightning

flashes (including both negative and positive polarities) were recorded by the LLS between 2012 and 2017. However, only 4,536,380 lightning flashes that landed on the land of Peninsular Malaysia were used in this study (i.e., removing flashes that landed in the sea). The summary of findings for the lightning flashes from 2012 to 2017 is shown in Figure 6, with ranges from 371,773 in 2014 to 830,568 in 2017 for negative polarity and from 27,219 in 2014 to 434,416 in 2017 for positive polarity.

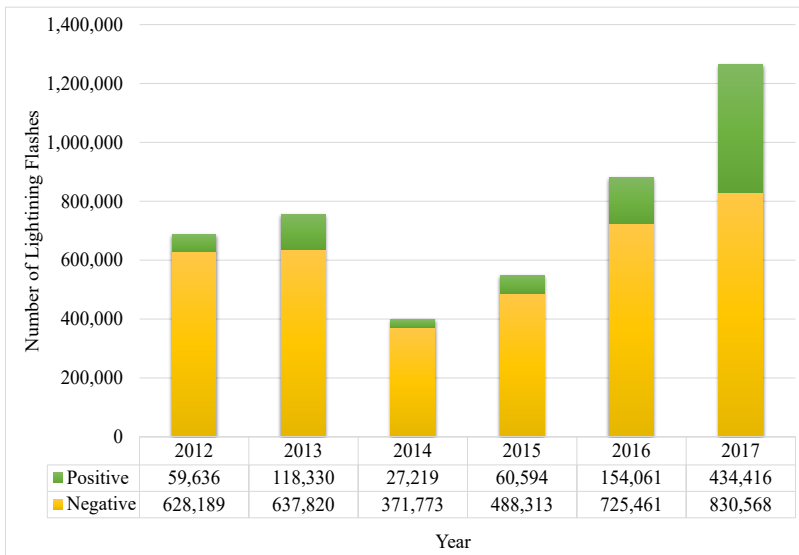


Figure 6. Annual variation of lightning distribution based on polarity from 2012 to 2017

The data shows that negative polarity lightning flashes contributed about 81% of the total flashes from 2012 to 2017, while the remaining were positive. Hence, it is evident that most lightning flashes in Peninsular Malaysia are negative. According to the yearly trend of positive polarity lightning flashes, the percentage of positive polarity lightning flashes has shown an increasing trend from 2014 to 2017, with the recorded percentages being 7%, 11%, 18%, and 34%, respectively. Furthermore, the results showed that there have been more lightning flashes over the last three years studied (i.e., 2015, 2016, and 2017). It could be attributed to the upgraded LLS system, where the detection efficiency and location accuracy improved to 95% and 500 m, respectively. Notably, the system can detect lightning strikes as far as 600 km from the sensors.

A comprehensive statistical analysis and peak current probability were carried out to present the lightning flashes from 2012 to 2017. Tables 3 and 4 show the statistical analysis and peak current probability of negative and positive lightning flashes. Within the 6 years, the highest magnitude of maximum peak currents observed for negative and positive lightning flashes were 508 kA (in 2016) and 427 kA (in 2017), respectively. In contrast, the

Table 3

*Statistical analysis and lightning peak current probability of negative lightning flashes*

Year	No. of Flashes	Min (kA)	Max (kA)	Median (kA)	Mode (kA)	Mean (kA)	Standard Deviation $\sigma$	Peak current Probability		
								95%	50%	5%
2012	628,189	3	290	18	12	21	15.67	6	18	52
2013	637,820	3	444	17	9	21	16.62	6	17	56
2014	371,773	3	221	20	13	24	16.06	8	20	56
2015	488,313	1	322	19	12	24	17.86	7	19	59
2016	725,461	1	508	18	6	22	18.03	5	18	57
2017	830,568	1	426	19	10	25	21.96	5	19	68
2012 - 2017	3,682,124	1	508	18	12	23	18.25	6	18	58

Table 4

*Statistical analysis and lightning peak current probability of positive lightning flashes*

Year	No. of Flashes	Min (kA)	Max (kA)	Median (kA)	Mode (kA)	Mean (kA)	Standard Deviation $\sigma$	Peak current Probability		
								95%	50%	5%
2012	59,636	3	186	8	6	11	10.47	4	8	25
2013	118,330	3	214	8	6	10	9.26	4	8	23
2014	27,219	3	245	9	7	13	13.40	5	9	31
2015	60,594	1	234	8	6	11	10.64	4	8	23
2016	154,061	1	253	6	5	9	8.99	3	6	19
2017	434,416	1	427	7	5	9	8.96	3	7	21
2012 - 2017	854,256	1	427	7	5	9	9.48	3	7	21

lowest magnitude of maximum peak currents observed for negative and positive lightning flashes were 221 kA (2014) and 186 kA (2012), respectively. The median peak current recorded for negative lightning flashes was 17 to 20 kA and 6 to 9 kA for positive lightning flashes. At the same time, the mean peak current recorded for negative lightning flashes was in the range of 21 to 25 kA and 9 to 13 kA for positive lightning flashes.

In previous studies conducted at various locations in Malaysia, several findings were reported, but the available information is limited for comparison. However, our study recorded lightning peak currents comparable to previous studies, as listed in Table 5. It is worth noting that these studies were based on different LLS, with detailed information on the SAFIR 3000 LLS (Chan & Mohamed, 2018; Johari et al., 2021). To further validate these findings, previous studies conducted along the same case line in Peninsular Malaysia using the same LLS system between 2004 and 2015 reported similar results (Rawi et al., 2018). These studies found median peak first stroke currents of 18 kA for negative flashes and slightly higher at 14 kA for positive flashes. Additionally, a high peak first stroke current of 378 kA was reported. It is important to note that the flash peak current corresponds to the peak current of the first stroke in the case of multi-stroke flashes (Diendorfer et al., 2014).

Table 5  
 Comparison of peak currents reported in previous studies for negative and positive lightning flashes

Year	LLS - reported Peak Currents								No. of Flash	Source
	Negative lightning flash				Positive lightning flash					
	Min (kA)	Max (kA)	Mean (kA)	Median (kA)	Min (kA)	Max (kA)	Mean (kA)	Median (kA)		
2012 – 2017	1	508	23	18	1	427	9	7	4,536,380	
2013 – 2015	1.9	139.6	13.6	13.6	10	86.7	14.2	12.8	572,282	Johari et al. (2021)
2015	1.80	140			1.9	100.5			201,296	Chan and Bin Mohamed (2018)

The analysis of the probability of lightning peak current for the same data set showed that 95% of the time, the peak current occurred in the range of 5 to 8 kA for negative lightning flashes and 3 to 5 kA for positive lightning flashes. Meanwhile, 50% of the time, the peak current was within 17 to 20 kA for negative lightning flashes and 6 to 9 kA for positive ones. Additionally, 5% of the time, the peak current was 52 to 68 kA for negative lightning flashes and 19 to 31 kA for positive ones.

**Monthly and Monsoon Seasonal Variation for Peninsular Malaysia**

Figure 7 reveals the monthly variation of lightning flashes during 6 years. The highest occurrences of negative lightning, with 616,860 flashes, were recorded in May, while the lowest negative lightning flashes, with 75,602 flashes, were observed in February, and for positive lightning flashes, the month of April had the most lightning, with 156,788 flashes while February had the fewest with 11,928 flashes.

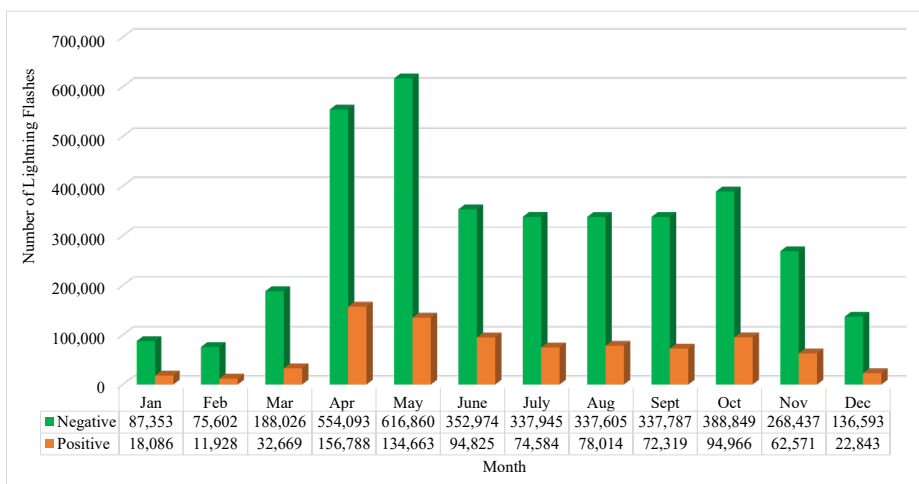


Figure 7. Monthly variation of lightning distribution based on polarity from 2012 to 2017

A statistical analysis of the lightning flashes from 2012 to 2017 for Peninsular Malaysia was carried out according to the monsoon season, as shown in Figure 8 and Table 6, respectively. It is observed that most lightning occurred during the Southwest Monsoon season from May to September, while the lowest was recorded during the Northeast Monsoon season from December to March. During the two Inter-Monsoon periods, from April to May and October to November, there were also higher lightning occurrences compared to the Northeast Monsoon season. These findings are consistent with previous studies that showed higher lightning flashes occurring during the two inter-monsoon periods (Rawi et al., 2017).

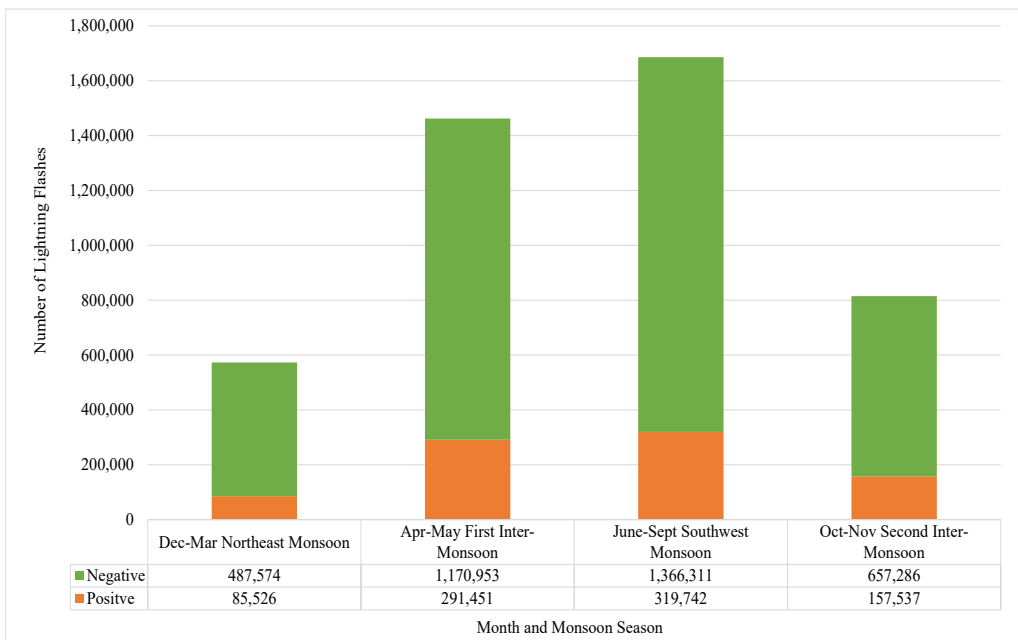


Figure 8. Monsoon season variation of lightning distribution based on polarity from 2012 to 2017

Table 6  
Statistical analysis of lightning flashes and peak current probability for Peninsular Malaysia based on monsoon seasons

		Negative Polarity							Peak Current Probability		
Monsoon	Period	No. of Flashes	Min (kA)	Max (kA)	Median (kA)	Mode (kA)	Mean (kA)	Standard Deviation $\sigma$	95%	50%	5%
Northeast	December – March	487,574	1	424	18	12	23.09	17.90	6	18	57
First Inter-Monsoon	April – May	1,170,953	1	348	18	10	23.09	17.90	5	18	56

Table 6 (Continue)

Negative Polarity											
Monsoon	Period	No. of Flashes	Min (kA)	Max (kA)	Median (kA)	Mode (kA)	Mean (kA)	Standard Deviation $\sigma$	Peak Current Probability		
									95%	50%	5%
Southwest	June – September	1,366,311	1	508	17	11	21.78	17.18	6	17	55
Second Inter-Monsoon	October – November	657,286	1	444	19	11	24.85	20.05	6	19	64
Positive Polarity											
Northeast	December – March	85,526	1	287	8	6	11.38	11.77	4	8	28
First Inter-Monsoon	April – May	291,451	1	427	7	5	9.11	9.11	3	7	21
Southwest	June – September	319,742	1	265	7	5	8.90	8.73	3	7	20
Second Inter-Monsoon	October – November	157,537	1	299	7	5	9.38	9.40	3	7	21

According to the monsoon seasonal variation, the highest maximum peak current of negative and positive lightning flashes observed were 508 kA (during the Southwest Monsoon) and 427 kA (during the first Inter-Monsoon), respectively. Table 6 shows little variation in the lightning peak current magnitude at 95%, 50%, and 5% occurrence probability when analyzed according to the monsoon seasons.

### Lightning Occurrence Distribution by State in Peninsular Malaysia

According to the detected latitude and longitude, the frequency of lightning flashes was plotted on a graph based on state and region, as depicted in Figure 9. The state with the highest negative and positive lightning flashes was E3 (Pahang), with 916,286 and 267,255, respectively. In contrast, C3 (Putrajaya) has the least lightning flashes, with 4,609 negative and 376 positive lightning flashes. In Peninsular Malaysia, the state of Pahang has the largest area, as shown in Table 1, which results in a higher total number of lightning flashes, compared to Putrajaya state, which has the smallest area and the lowest number of lightning flashes.

### Ground Flash Density Maps

The GFD maps for Peninsular Malaysia from 2012 to 2017 were produced separately for negative and positive polarity lightning. Figure 10 and Figure 11 compare the negative and positive GFD maps between 2012 and 2017, respectively. These results show that the hotspot areas for negative GFD are within the central region of Peninsular Malaysia (Kuala Lumpur and Selangor). From 2012 to 2015, the highest negative GFD was 45 flashes/km<sup>2</sup>/

year. However, in 2016, the highest count increased to 62 flashes/km<sup>2</sup>/year, and 73 flashes/km<sup>2</sup>/year were reported to be the highest negative GFD in 2017.

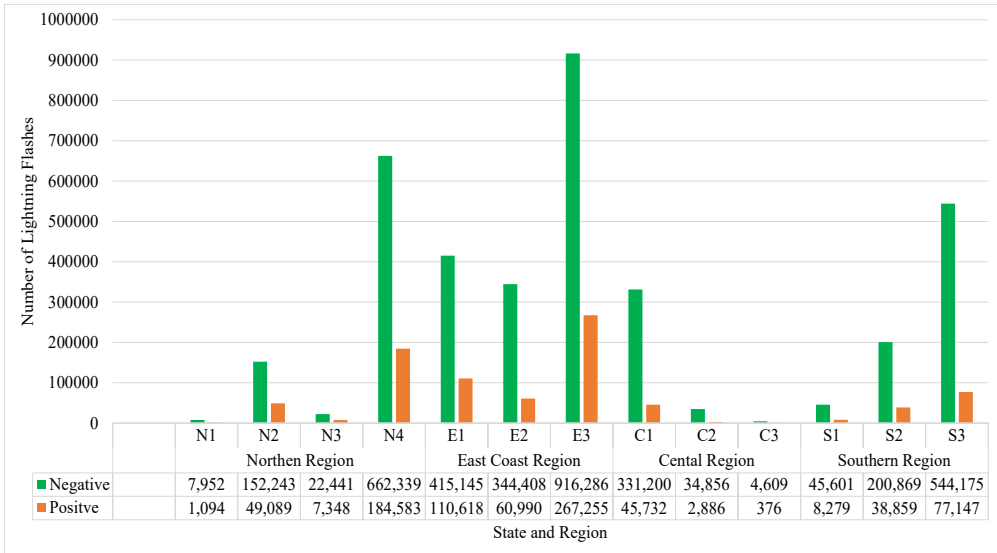


Figure 9. Lightning flash distribution in Peninsular Malaysia from 2012–2017 based on polarity and state

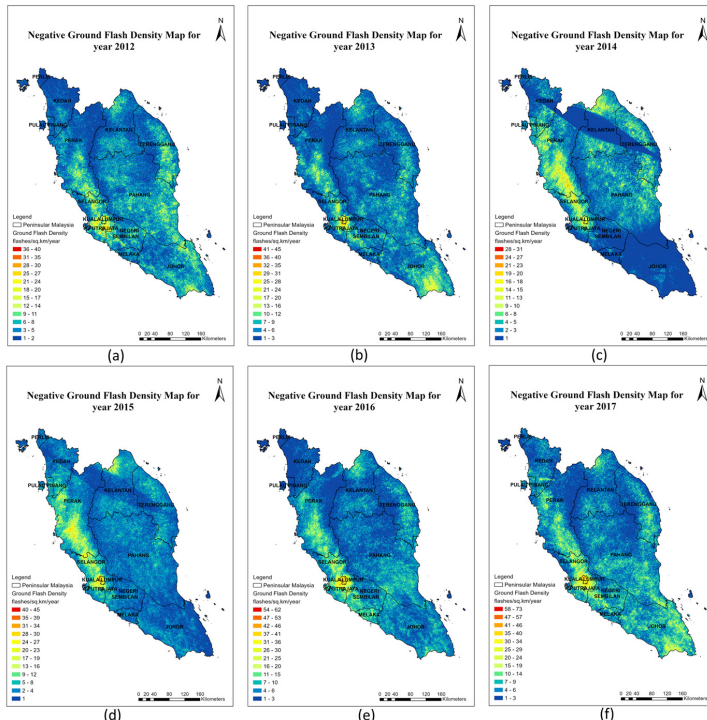


Figure 10. Negative ground flash density in Peninsular Malaysia. (a) 2012; (b) 2013; (c) 2014; (d) 2015; (e) 2016; (f) 2017



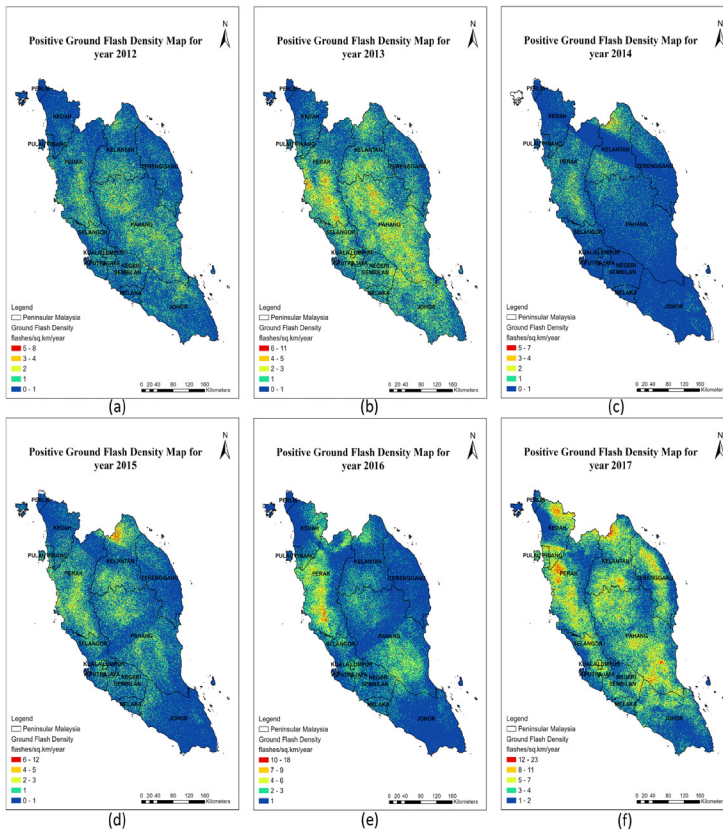


Figure 11. Positive ground flash density in Peninsular Malaysia. (a) 2012; (b) 2013; (c) 2014; (d) 2015; (e) 2016; (f) 2017

Due to the lower occurrences, the maximum recorded value of positive GFD is lower compared to the negative GFD. The positive GFD maps show that the hotspot area is within the Northern region (Perak) and East Coast regions (Pahang and Kelantan). The highest positive GFD of 23 flashes/km<sup>2</sup>/year was observed in 2017, 28% more than the highest positive GFD recorded in 2016 and approximately double the highest positive GFD recorded in 2015. Based on the combined lightning flashes from 2012 to 2017, Figure 12 shows the average GFD map for Peninsular Malaysia. It is seen that the Central region is the main hotspot for negative GFD, with the highest average GFD of 38 flashes/km<sup>2</sup>/year. Meanwhile, the Northern and East Coast regions are the hotspots for positive GFD, with the highest average GFD of 6 flashes/km<sup>2</sup>/year.

Table 7 provides information on the specific locations in Peninsular Malaysia with the highest count of lightning flash density (GFD), determined from the GFD map created in ArcGIS. For instance, the highest count of negative GFD between 2012 and 2017 was recorded in Kuala Lumpur, with a value of 73 flashes/km<sup>2</sup>/year in 2017. These 73 negative lightning flashes occurred within a 1 km square area, centered at 3° 04' 02.31" N latitude

and 101° 43' 38.44" E longitude. The developed GFD map enables efficient geolocation of lightning incidents within Peninsular Malaysia, which is important for lightning protection design.

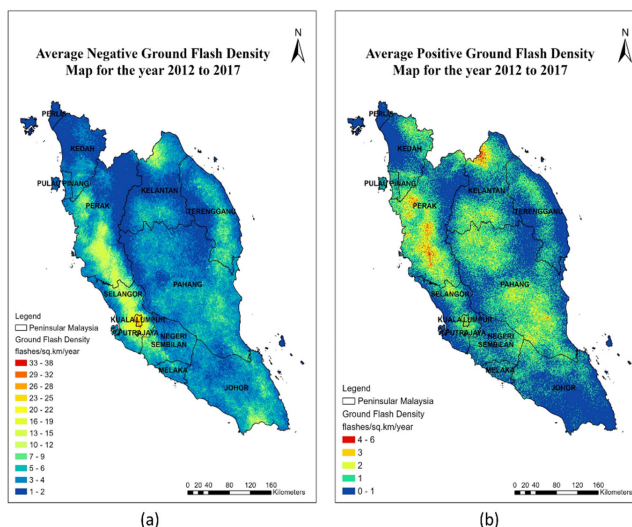


Figure 12. Average ground flash density in Peninsular Malaysia from 2012 to 2017. (a) Negative polarity; (b) Positive polarity

Table 7

Maximum GFD count and the location from the year 2012 to 2017 for both polarities

Year	Coordinates of the central point of a 1 km <sup>2</sup> grid cell (Latitude/Longitude)	State	Max Negative GFD	Coordinates of the central point of a 1 km <sup>2</sup> grid cell (Latitude/Longitude)	State	Max Positive GFD
2012-2017	N 3° 09' 27.67"/ E 101° 42' 00.50"	Kuala Lumpur	38	N 5°52' 18.40"/ E 101° 58' 24.05015489	Kelantan	6
2012	N 2° 57' 31.89"/ E 101° 45' 48.88	Selangor	40	N 4° 07' 32.91"/ E 101° 49' 58.44" & N 2° 24' 01.64"/ E 103° 30' 33.81"	Pahang & Johor	8
2013	N 3° 10' 32.48"/ E 101° 39' 50.77"	Kuala Lumpur	45	N 4° 31' 12.37"/E 100° 40' 41.80"	Perak	11
2014	N 4° 19' 20.96"/ E 101° 2' 22.12" & N 3° 57' 41.33"/ E 101° 16' 29.53"	Perak	31	N 5° 47' 57.39"/ E 101° 54' 04.49" & N 4° 20' 58.76 / E101° 02' 54.22"	Perak & Kelantan	7
2015	N 3° 13' 47.51"/ E101° 37' 40.72"	Selangor	45	N 5° 48' 30.88"/ E 102° 01' 39.57"	Kelantan	12
2016	N 3° 09' 27.44"/ E 101° 40' 23.32"	Kuala Lumpur	62	N 4° 03' 04.84"/ E 101° 05' 40.02"	Perak	18

Table 7 (Continue)

Year	Coordinates of the central point of a 1 km <sup>2</sup> grid cell (Latitude/Longitude)	State	Max Negative GFD	Coordinates of the central point of a 1 km <sup>2</sup> grid cell (Latitude/Longitude)	State	Max Positive GFD
2017	N 3° 04' 02.31"/ E 101° 43' 38.44"	Kuala Lumpur	73	N 5° 04' 52.38"/ E 100° 45' 58.27"/ & N 5° 50' 07.90"/ E 101° 56' 14.26"	Perak & Kelantan	23

### Application of GFD Map on Transmission Line Performance Evaluation

The lightning exposure of a selected transmission line can be determined by overlaying the GFD map onto the transmission line and focusing on the lightning flashes data within a 5-km buffer radius of the transmission line's centerline. Figure 13 shows an example of the lightning exposure on a 500 kV transmission line A-B.

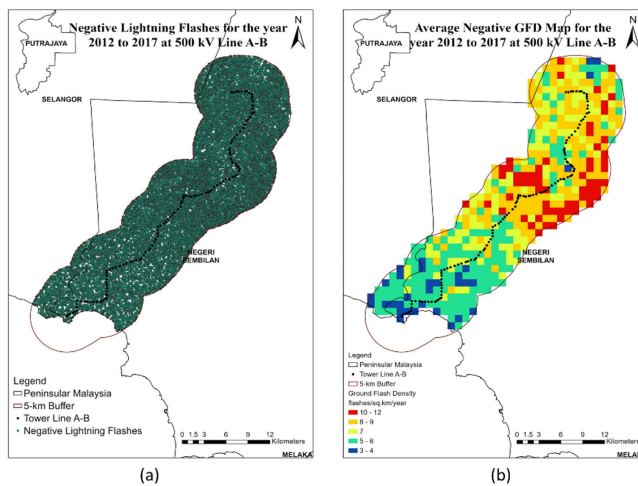


Figure 13. Lightning exposure within a 5-km buffer from the 500 kV Line A-B from 2012 to 2017. (a) Negative lightning flashes recorded data; (b) Negative GFD map

Between 2012 and 2017, the transmission line was exposed to 23,529 negative lightning flashes within a 5-km buffer along the line. The average negative GFD within this buffer was recorded between 3 and 12 flashes/km<sup>2</sup>/year, with peak current ranging from 2 kA to 278 kA. As demonstrated in Figure 13 (b), the towers located in areas with higher lightning GFD values pose a higher risk of lightning related faults. Therefore, it is crucial to prioritize on monitoring and maintaining these areas. This concern has been previously addressed by other researchers (Hatta et al., 2019; Rawi et al., 2018; Tofani et al., 2018).

Details related to these towers, such as TFR and SR values, ground elevation, and the historical lightning-related tripping incidents, can be extracted by utilizing the GFD map

developed in ArcGIS software, as demonstrated in Figure 14. The analysis of the BFR value involved applying Equation (1) to (3). Notably, Figure 14 reveals a crucial parameter for BFR calculation, specifically the average negative GFD,  $N_g$ , value of 10 flashes/km<sup>2</sup>/year. This information is essential for calculating the BFR value and serves as key metric for evaluating the lightning performance of transmission lines, as highlighted by Ardila et al. (2023).

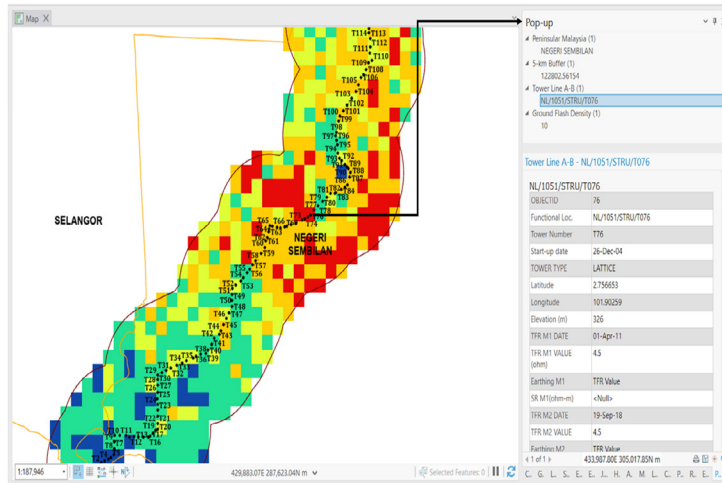


Figure 14. Negative GFD and tower details at one of the towers along the 500 kV Line A-B

By assuming  $I_c = 248 \text{ kA}$  and given  $h_t = 45 \text{ m}$   $b = 2 \text{ m}$

$$P(I_c > I_f) = 1/[1 + (\frac{248}{31})^{2.6}] = 4.4671 \times 10^{-3}$$

$$N_s = N_g(28h_t^{0.6} + b)/10 = 276.84$$

$$BFR = 0.742 \text{ flashovers/100 km/year}$$

It means that 0.742 back flashovers are to be expected in a year for 100 km.

Therefore, the GFD map developed for transmission lines can evaluate their performance, monitor them, and design appropriate lightning protection systems. This information can help identify high-risk areas and design lightning protection systems, such as installing shield wire, reducing tower footing resistance, increasing insulation levels, and installing arresters to reduce the impact of lightning strikes (Ahmed et al., 2023; IEEE Power Engineering Society, 1997). By implementing lightning mitigation measures, routine maintenance, and inspections of transmission lines, their reliability can be significantly improved.

## CONCLUSION

In this study, GFD mapping and statistical analysis were used to assess the impact of lightning on transmission lines. The findings provide valuable information for power companies and researchers. Results show that 81% of the 4,536,380 lightning flashes were caused by negative polarity lightning, which severely impacts transmission lines. The annual lightning flashes increased from 2015 to 2017 due to improved detection by the LLS network after an upgrade in 2015.

The mean peak current of negative lightning flashes was higher than that of positive flashes, in the range of 21 to 25 kA and 9 to 13 kA, respectively. The 95% probability of occurrence for negative and positive lightning flashes was 5 to 8 kA and 3 to 5 kA, respectively. More lightning activity was recorded during the Southwest Monsoon (June to September) and first Inter-Monsoon (April to May) compared to other seasons. The state of Pahang had the most lightning flashes due to its large land area.

The results from the developed GFD map indicate that the Central region was the major hotspot for negative GFD and that the hotspots for positive GFD were in the Northern and East Coast regions. The highest negative and positive GFD were recorded in 2017 at 73 flashes/km<sup>2</sup>/year and 23 flashes/km<sup>2</sup>/year, respectively. Over the six years studied (2012–2017), the highest average negative and positive GFD were 38 flashes/km<sup>2</sup>/year and 6 flashes/km<sup>2</sup>/year, respectively.

To summarize, this paper outlines a method for conducting a statistical analysis of lightning flashes and creating GFD maps using GIS to assess lightning performance on transmission lines. For future work, it is recommended to incorporate meteorological data better to understand correlations between lightning distribution and transmission line performance.

## ACKNOWLEDGMENTS

The authors thank and acknowledge Tenaga Nasional Berhad and UNITEN R&D Sdn. Bhd., located in Malaysia for the financial support of this project. TNB and URND financially support this research under the TNB Seeding Fund with project code U-TS-RD-19-30 and by Geran Putra Berimpak Universiti Putra Malaysia GP-GPB/2022/9708800 under the Research University Initiative for publication support. We extend our appreciation to TNB Research Sdn. Bhd. Malaysia for providing lightning flash data and the Department of Survey and Mapping Malaysia (JUPEM) for providing the Peninsular Malaysia base map.

## REFERENCES

- Abdullah, N. (2019). *Technical report for technical assessment of lightning protection technical expert grade E16(S)*. Lightning & Earthing Unit, TNB Research Sdn. Bhd.

- Abdullah, N., & Hatta, N. M. (2012, December 2-5). *Cloud-to-ground lightning occurrences in Peninsular Malaysia and its use in improvement of distribution line lightning performances*. [Paper presentation]. IEEE International Conference on Power and Energy (PECon), Kota Kinabalu, Malaysia. <https://doi.org/10.1109/PECon.2012.6450330>
- Abdullah, N., Yahaya, M. P., & Hudi, N. S. (2008, December 1-3). *Implementation and use of lightning detection network in Malaysia*. [Paper presentation]. IEEE 2<sup>nd</sup> International Power and Energy Conference, Johor Bahru, Malaysia. <https://doi.org/10.1109/PECON.2008.4762504>
- Ab-Kadir, M. Z. A. (2016). Lightning severity in Malaysia and some parameters of interest for engineering applications. *Thermal Science*, 20(2), 437-450. <https://doi.org/10.2298/TSCI151026028A>
- Ahmed, N. Y., Illias, H. A., & Mokhlis, H. (2023). A protocol for selecting viable transmission line arrester for optimal lightning protection. *Electric Power Systems Research*, 221, 109489. <https://doi.org/10.1016/j.epsr.2023.109489>
- Ali, H. M. S. A., Aman, M. M., Ahmed, S. J., Shaikh, S., & Munir, A. (2018, June 12-15). *Lightning flash density map of Pakistan on ARC-GIS® software - An empirical approach*. [Paper presentation]. IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe (EEEIC / I&CPS Europe), Palermo, Italy. <https://doi.org/10.1109/EEEIC.2018.8494579>
- Ardila, B., Soto, E., & Zamora, J. (2023). Modeling lightning flashes in transmission structures. *Electric Power Systems Research*, 223, 109636. <https://doi.org/10.1016/j.epsr.2023.109636>
- Biswas, R. N., Islam, M. N., Mia, M. J., & Islam, M. N. (2020). Modeling on the spatial vulnerability of lightning disaster in Bangladesh using GIS and IDW techniques. *Spatial Information Research*, 28(5), 507-521. <https://doi.org/10.1007/s41324-019-00311-y>
- Bouquegneau, C. (2014, October 11-18). *Lightning density based on lightning location systems*. [Paper presentation]. International Conference on Lightning Protection (ICLP), Shanghai, China. <https://doi.org/10.1109/ICLP.2014.6973446>
- Chan, H. G., & Mohamed, A. I. B. (2018). Investigation on the occurrence of positive cloud to ground (+CG) lightning in UMP Pekan. *Journal of Atmospheric and Solar-Terrestrial Physics*, 179, 206-213. <https://doi.org/10.1016/j.jastp.2018.07.016>
- Chang, K. T. (2008). *Introduction to Geographic Information Systems* (Vol. 4). McGraw-Hill.
- Dewan, A., Islam, K. M. A., Fariha, T. R., Murshed, M. M., Ishtiaque, A., Adnan, M. S. G., Kabir, Z., & Chowdhury, M. B. H. (2022). Spatial pattern and land surface features associated with cloud-to-ground lightning in Bangladesh: An exploratory study. *Earth Systems and Environment*, 6(2), 437-451. <https://doi.org/10.1007/s41748-022-00310-4>
- Diendorfer, G., Pichler, H., Achleitner, G., & Broneder, M. (2014, October 11-18). *Lightning caused outages in the Austrian Power Grid transmission line network*. [Paper presentation]. International Conference on Lightning Protection (ICLP), Shanghai, China. <https://doi.org/10.1109/ICLP.2014.6973112>
- Edirisinghe, M., & Maduranga, U. G. D. (2021). Distribution of lightning accidents in Sri Lanka from 1974 to 2019 using the DesInventar database. *ISPRS International Journal of Geo-Information*, 10(3), Article 117. <https://doi.org/10.3390/ijgi10030117>



- ESRI. (2010). *The geographic approach™ to electric distribution*. Esri. <https://www.esri.com>
- Farukh, M. A., Ahmed, S. U., Islam, M. A., & Baten, M. A. (2017). Spatial vulnerability assessment of extreme lightning events in Bangladesh using GIS. *Journal of Environmental Science and Natural Resources*, 10(2), 11-18. <https://doi.org/10.3329/jesnr.v10i2.39008>
- Hatta, N. M., Abdullah, N., Osman, M., & Abidin Ab Kadir, M. Z. (2019). Statistical Lightning Study for 33kV Overhead Line in Peninsular Malaysia. [Paper presentation]. *2019 11th Asia-Pacific International Conference on Lightning (APL)*, 1–3. <https://doi.org/10.1109/APL.2019.8816020>
- He, J. L., & Zeng, R. (2010). *Lightning Shielding Failure Analysis of 1000 kV Ultra-High Voltage AC Transmission Line*. CIGRE. [https://cigreindia.org/CIGRE%20Lib/CIGRE%20Session%202010%20paper/C4\\_201\\_2010.pdf](https://cigreindia.org/CIGRE%20Lib/CIGRE%20Session%202010%20paper/C4_201_2010.pdf)
- Hodanish, S. J., Vogt, B. J., & Wolyn, P. (2019). Colorado lightning climatology. *Journal of Operational Meteorology*, 7(4), 45-60. <https://doi.org/10.15191/nwajom.2019.0704>
- Husain, F., Sulaiman, N. A., Hashim, K. A., & Samad, A. M. (2012, July 16-17). *A study on TNB transmission line route sustainability and suitability using GIS-AHP*. [Paper presentation]. IEEE Control and System Graduate Research Colloquium, Shah Alam, Malaysia. <https://doi.org/10.1109/ICSGRC.2012.6287193>
- International Electrotechnical Commission. (2015). *Lightning Density based on Lightning Location Systems (LLS) - General principles*. ICE. <https://webstore.iec.ch/publication/23064>
- IEEE Power Engineering Society (1997). *IEEE Std 1243-1997: IEEE guide for improving the lightning performance of transmission lines*. IEEE Xplore. <https://ieeexplore.ieee.org/document/640303>
- Isa, N. A., Baharudin, Z. A., Zainuddin, H., Sutikno, T., Zainon, M., & Zulkefle, A. A. (2021). Distribution of attempted leader with monsoon seasons and negative cloud-to-ground flashes in Melaka, Malaysia. *Indonesian Journal of Electrical Engineering and Computer Science*, 23(3), 1324-1330. <https://doi.org/10.11591/ijeecs.v23.i3.pp1324-1330>
- Islam, M. A., Chan, A., Khaw, S. K. K., Salleh, S. A., Zakaria, N. H., Isa, N. A., Gee, M. O. C., & Azari, M. (2019). Importance of lightning detection network development in Peninsular Malaysia. *IOP Conference Series: Earth and Environmental Science*, 385(1), Article 012036. <https://doi.org/10.1088/1755-1315/385/1/012036>
- Javor, V., Stoimenov, L., Dzakovic, N., Dinkic, N., Javor, D., & Betz, H. D. (2018). LINETGIS analysis of lightning flash density in Serbia based on ten years data. *Serbian Journal of Electrical Engineering*, 15(2), 201-211. <https://doi.org/10.2298/SJEE1802201J>
- Johari, D., Aiman Misri Amir, M. F., Hashim, N., Baharom, R., & Haris, F. A. (2021, March 8-9). *Positive cloud-to-ground lightning observed in Shah Alam, Malaysia based on SAFIR 3000 lightning location system*. [Paper presentation]. IEEE International Conference in Power Engineering Application (ICPEA), Shah Alam, Malaysia. <https://doi.org/10.1109/ICPEA51500.2021.9417761>
- Kakumoto, Y., Koyamatsu, Y., Shiota, A., Qudaih, Y., & Mitani, Y. (2016). Application of geographic information system to power distribution system analysis. *Energy Procedia*, 100, 360-365. <https://doi.org/10.1016/j.egypro.2016.10.189>
- Kamra, A. K., & Kumar, P. R. (2021). Regional variability in lightning activity over South Asia. *International Journal of Climatology*, 41(1), 625-646. <https://doi.org/10.1002/joc.6641>



- Kezunovic, M., Djokic, T., Chen, P. C., & Malbasa, V. (2015, January 5-8). *Improved transmission line fault location using automated correlation of big data from lightning strikes and fault-induced traveling waves*. [Paper presentation]. 48<sup>th</sup> Hawaii International Conference on System Sciences, Kauai, USA. <https://doi.org/10.1109/HICSS.2015.328>
- Korir, J. K., & Ngigi, M. M. (2015). The use of GIS in high voltage transmission line routing in Kenya: A case study of 132 kV Kilimambogo-Thika-Kiganjo Line. *African Journal of Geography and Regional Planning*, 2(1), 207-217.
- Leite, J. B., Mantovani, J. R. S., Dokic, T., Yan, Q., Chen, P. C., & Kezunovic, M. (2019). Resiliency assessment in distribution networks using GIS-based predictive risk analytics. *IEEE Transactions on Power Systems*, 34(6), 4249-4257. <https://doi.org/10.1109/TPWRS.2019.2913090>
- Li, L., Feng, C., Wei-Ning, S., Yao-Qiang, X., & Xing-Zhi, W. (2014, October 20-22). *Study on evaluation system of meteorological hazards for power grid based on Cloud GIS*. [Paper presentation]. International Conference on Power System Technology, Chengdu, China. <https://doi.org/10.1109/POWERCON.2014.6993685>
- Lin, Z., & Xu, Z. (2016). A meteorological risk assessment method for power lines based on GIS and multi-sensor integration. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 3(8), 19-26. <https://doi.org/10.5194/isprsannals-III-8-19-2016>
- Mammadov, R., Safarov, S., Safarov, E., & Rasouli, A. A. (2021). Visualizing the lightning flashes over the Republic of Azerbaijan by analyzing the lightning imaging sensor data. *ANAS Transactions, Earth Sciences*, 1, 93-105. <https://doi.org/10.33677/ggianas20210100058>
- Mishra, M., Acharyya, T., Santos, C. A. G., Silva, R. M. da, Chand, P., Bhattacharyya, D., Srivastava, S., & Singh, O. (2022). Mapping main risk areas of lightning fatalities between 2000 and 2020 over Odisha state (India): A diagnostic approach to reduce lightning fatalities using statistical and spatiotemporal analyses. *International Journal of Disaster Risk Reduction*, 79, Article 103145. <https://doi.org/10.1016/j.ijdrr.2022.103145>
- Misztal, K., & Siłuch, M. (2021). Spatial analysis of atmospheric discharges in Lubelszczyzna in 2018. *Applied Geomatics*, 13(4), 969-979. <https://doi.org/10.1007/s12518-021-00378-z>
- Phillips, A. (2004). *Handbook for Improving Overhead Transmission Line Lightning Performance*. (Technical Report No. 1002019). EPRI.
- Provost, S. B., Saboor, A., Cordeiro, G. M., & Mansoor, M. (2018). On the q-generalized extreme value distribution. *REVSTAT-Statistical Journal*, 16(1), 45-70. <https://doi.org/https://doi.org/10.57805/revstat.v16i1.232>
- Rahman, M. A., Maulud, K. A., Bahri, M. S., Hussain, M., Oon, A. R., Suhatdi, S., Hashim, C. C., & Mohd, F. (2020). Development of GIS database for infrastructure management: Power distribution network system. *IOP Conference Series: Earth and Environmental Science*, 540(1), Article 012067. <https://doi.org/10.1088/1755-1315/540/1/012067>
- Rawi, I. M., Abidin Ab Kadir, M. Z., Gomes, C., & Azis, N. (2018). A case study on 500 kV line performance related to lightning in Malaysia. *IEEE Transactions on Power Delivery*, 33(5), 2180-2186. <https://doi.org/10.1109/TPWRD.2017.2787660>

- Rawi, I. M., Kadir, M. Z. A. A., & Izadi, M. (2017). Seasonal variation of transmission line outages in Peninsular Malaysia. *Pertanika Journal Science & Technology*, 25(S), 213-220.
- Sardi, J., Kadir, A. M. Z. A., Ahmad, W. W. F., Hizam, H., Rawi, I. M., & Ahmad, A. (2008, December 1-3). *Backflashover analysis for 132 kV Kuala Krai-Gua Musang transmission line*. [Paper presentation]. IEEE 2nd International Power and Energy Conference, Johor Bahru, Malaysia. <https://doi.org/10.1109/PECON.2008.4762716>
- Shukla, R. K., Trivedi, M., & Kumar, M. (2010). On the proficient use of GEV distribution: A case study of subtropical monsoon region in India. *Computer Science Series*, 8(1), 81-92.
- The MathWorks Inc. (2021). *MATLAB (Version R2021a)*. MathWorks. <https://www.mathworks.com/>
- Tofani, K. M., Jintaka, D. R., Hartono, J., Munir, B. S., & Bramantyo S. D. A. Harsono, B. (2018). *Lightning activity evaluation in South Sumatera Region*. 2018 2nd Borneo International Conference on Applied Mathematics and Engineering (BICAME) (pp. 313–316). IEEE. <https://doi.org/10.1109/BICAME45512.2018.1570505913>
- Warmi, Y., & Michishita, K. (2018). Tower-footing resistance and lightning trip-outs of 150 kV transmission lines in West Sumatra in Indonesia. *MATEC Web of Conferences*, 215, Article 01022. <https://doi.org/10.1051/mateconf/201821501022>
- Wooi, C. L., Abdul-Malek, Z., Ahmad, N. A., Mokhtari, M., & Khavari, A. H. (2016). Cloud-to-ground lightning in Malaysia: A review study. *Applied Mechanics and Materials*, 818, 140-145. <https://doi.org/10.4028/www.scientific.net/AMM.818.140>
- Xie, M., Zhou, G., Li, D., & Gong, J. (2000). Design and implementation of attribute database management system in a GIS system: GeoStar. *Geographic Information Sciences*, 6(2), 170-180. <https://doi.org/10.1080/10824000009480547>
- Yan, Q., Dokic, T., & Kezunovic, M. (2016, April 18-20). *GIS-based risk assessment for electric power consumers under severe weather conditions*. [Paper presentation]. 18th Mediterranean Electrotechnical Conference (MELECON), Lemesos, Cyprus. <https://doi.org/10.1109/MELCON.2016.7495473>
- Yatim, F. M., Majid, Z., & Amerudin, S. (2019). The key success factor of geographic information system re-implementation in Tenaga Nasional Berhad Distribution, Malaysia: A case study of Cheras Tenaga Nasional Berhad Distribution station. *International Journal of Built Environment and Sustainability*, 6(1-2), 111-119. <https://doi.org/10.11113/ijbes.v6.n1-2.390>
- Yao, X. A. (2020). Georeferencing and geocoding. In A. Kobayashi (Ed.), *International encyclopedia of human geography* (pp.111-117). Elsevier. <https://doi.org/10.1016/B978-0-08-102295-5.10548-7>
- Yendra, R., Hanaish, I. S., & Fudholi, A. (2021). Power bayesian markov chain monte carlo (MCMC) for modelling extreme temperatures in Sumatra Island using generalised extreme value (GEV) and generalised logistic (GLO) distributions. *Mathematical Modelling of Engineering Problems*, 8(3), 365-376. <https://doi.org/10.18280/mmep.080305>
- Zheng, Y., Wu, Z., Shu, S., Xu, J., Fang, C., & Xie, W. (2021). Lightning risk assessment model for transmission lines with lift-based improved analytic hierarchy process. *IET Generation, Transmission & Distribution*, 15(20), 2905-2914. <https://doi.org/10.1049/gtd2.12227>

## Preparation of Activated Carbon from Sugarcane Bagasse Using Microwave-assisted $ZnCl_2$ Chemical Activation: Optimization and Characterization Study

Atiqa Rahmawati<sup>1\*</sup>, Fadzkurisma Robbika<sup>1</sup> and Yuafni<sup>2</sup>

<sup>1</sup>Department of Leather Processing Technology, Politeknik ATK Yogyakarta, Special District of Yogyakarta, Yogyakarta, Indonesia

<sup>2</sup>Department of Leather Product Technology, Politeknik ATK Yogyakarta, Special District of Yogyakarta, Yogyakarta, Indonesia

### ABSTRACT

Microwave-assisted activation is a green technology technique that can synthesize activated carbon from bagasse. In this study, microwave-assisted  $ZnCl_2$  chemical activation was applied to convert bagasse to activated carbon (BAC). Activating activated carbon was optimized using surface response methodology (RSM). The Box-Behnken (BBD) design was used to assist in the optimum synthesis condition, with the loading of  $ZnCl_2$  concentration (A: 10–50% w/v), heating time (B: 2–12 min), and microwave power (C: 150–800 W). The BAC was characterized using Brunauer-Emmett-Teller (BET), Scanning Electron Microscopy (SEM), Fourier Transform Infrared (FT-IR), and moisture content. The findings of the BAC optimization were at a  $ZnCl_2$  concentration of 24.281 (% w/v), 12 min of heating time, and 800 W of microwave power. The characterization result shows that  $BAC_{op}$  has mesoporous carbon and a desirable surface area of 446.874  $m^2/g$ , average pore size of 4.071 nm, pore volume of 0.054 cc/g, and total pore volume of 0.2531 cc/g. SEM analysis revealed that microwave-generated pore structures lead to the  $ZnCl_2$  activation process. The pore structures of the raw material and activated carbon were different. The FT-IR analysis shows the existence of functional groups as well as changes in functional groups from raw material to activated carbon. The moisture content study findings are 5.51 to 9.21% in accordance with the Indonesian National Standard (SNI) 06-3730-1995.

### ARTICLE INFO

#### Article history:

Received: 20 February 2023

Accepted: 20 July 2023

Published: 15 January 2024

DOI: <https://doi.org/10.47836/pjst.32.1.22>

#### E-mail addresses:

tiqa054@gmail.com (Atiqa Rahmawati)

Fadzkurisma.risma@gmail.com (Fadzkurisma Robbika)

yuafni@atk.ac.id (Yuafni)

\*Corresponding author

The isothermal adsorption-desorption process is classified as type IV adsorption with hysteresis loop H4, suggesting that the pore distribution in activated carbon is mesoporous with a tiny pore width and slit-shape pore materials.

**Keywords:** Activated carbon, box-behnken, microwave-assisted activation, optimization, sugarcane bagasse,  $ZnCl_2$

## INTRODUCTION

Biomass waste as a feedstock for renewable energy or the development of material innovation is a challenge in research (Daochalermwong et al., 2020). The significant increase in biomass waste is getting much attention due to environmentally friendly aspects and the availability of raw materials, which is also one aspect of sustainable development (Ummartyotin & Pechyen, 2016). Using biomass waste as raw material for material innovation or renewable energy has developed methods and is still being developed to meet industrial needs. One of the developments focuses on the use of biomass waste as an adsorption material known as an adsorbent (Ummartyotin & Pechyen, 2016). Activated carbon is an amorphous compound produced from materials with a high carbon level and is specially modified to produce a larger surface area (Arnell et al., 2019). On an industrial scale, activated carbon is utilized as an adsorbent for purification and separation processes (Ozdemir et al., 2014). Hence, it is utilized in wastewater treatment, metal recovery, solvent recovery, food processing, and improving odor and flavor (Ao et al., 2018).

Several studies have produced activated carbon from various biomass wastes such as sugarcane bagasse (el Nemr et al., 2022; Foo et al., 2013; Jiang et al., 2021; Karri et al., 2020; Kaushik et al., 2017; Salihi et al., 2017), macadamia nut endocarp (Junior et al., 2014), pineapple leaf (Daochalermwong et al., 2020), rice husk (Arnell et al., 2019), grape stalk (Ozdemir et al., 2014), cotton stalk (Deng et al., 2009), sludge waste (Bian et al., 2018), rubber fruit shell (Suhdi & Wang, 2021) and many more. Based on the study's findings, biomass waste can be converted into low-cost adsorbents (Buthiyappan et al., 2019). Activated carbon preparation from lignocellulose often uses physical or chemical activation (Abdulhameed et al., 2021). The physical activation process involves the carbonization of raw materials under inert conditions, followed by the activation of the resultant carbon at high temperatures (Mahmood et al., 2017). In contrast, chemical activation consists of a single step. It is conducted in the presence of dehydrating reagents such as KOH,  $K_2CO_3$ , NaOH,  $ZnCl_2$ , and  $H_3PO_3$ , which inhibit tar formation and enhance pyrolytic decomposition (Deng et al., 2009). In this study, we used a combination of chemical and physical activation to improve activated carbon characteristics. Lam et al. (2017) state that combining chemical and physical activation will produce activated carbon with highly accessible pores.  $ZnCl_2$ , as an activator in chemical activation, was used for its advantages in generating activated carbon with small pore volume, pore size, and higher surface area (Luo et al., 2019). Furthermore, microwaves can affect activated carbon's pore structure and surface area (Arnell et al., 2019).

Sugarcane Bagasse is one of the biomass wastes that can be developed into activated carbon. Sugarcane bagasse comprises approximately 50% cellulose, 25% hemicellulose, and 25% lignin, each reacting vigorously with dyestuff and heavy metals (Buthiyappan et al., 2019). However, unmodified sugarcane bagasse has limited absorption efficiency, long

contact time, and low pH sensitivity. Sugarcane bagasse can be chemically and physically modified to generate low-cost activated carbon and efficiently remove heavy metals, dyestuffs, and phenols (Buthiyappan et al., 2019). The availability of sugarcane bagasse based on Indonesian Plantation Research Center (P3GI) data, from 62 sugar factories, 29,911 million tonnes of sugarcane can be milled per year, resulting in 2,991 million tonnes of bagasse will be produced (Hidayati et al., 2016). Meanwhile, up to 3,500 tonnes of sugarcane is used daily at the Madukismo Sugar Factory in Yogyakarta, and 1,400 tonnes of bagasse are generated daily. 50% of the bagasse produced by the milling process will be used as boiler feed, the others being stored as waste with no economic value. Bagasse can be used as a raw material for producing activated carbon for adsorption due to the high carbon concentration of the lignocellulose in bagasse (Hidayati et al., 2016).

Activated carbon is composed of functional groups bound to fused aromatic rings, which are predicted to have comparable chemical characteristics to aromatic hydrocarbons. Thermal, hydrothermal, or chemical treatment can alter and adjust surface functional groups in carbon matrices for specific applications (Ao et al., 2018). To optimize the activated carbon preparation from bagasse, we used Response Surface Methodology (RSM). RSM was chosen in some recent studies to optimize the preparation parameters and further analyze potential interaction between various parameters (Yuan et al., 2018). Box-Behnken Design (BBD) is one of the RSM commonly used in food processing, analytical chemistry, and AC production. BBD provides the following advantages: BBD assists in optimizing variables with the fewest number of experiments, contributes further analysis interaction between variables, can effectively estimate factors of the quadratic model, and avoids combination factors at an extreme range (Yuan et al., 2018).

This study converted bagasse from the Madukismo sugar factory in Yogyakarta, Indonesia into activated carbon. Chemical and physical activation were combined to prepare activated carbon. The activation was performed using microwave-assisted  $ZnCl_2$  chemical activation. This study aims to optimize activated carbon (AC) production from sugarcane bagasse using  $ZnCl_2$  as a precursor in chemical activation and microwave-assisted activation as physical activation. Optimization of activated carbon used RSM. The BBD design was approached in the optimum synthesis condition. The effect of microwave power, heating time, and  $ZnCl_2$  concentration on activated carbon was investigated. In addition, BET surface area, iodine number, surface functional group analysis by Boehm titration, FT-IR analysis, and SEM-EDX were used to characterize the bagasse-activated carbon.

## METHODOLOGY

### Activated Carbon (AC) Preparation

**Raw Material.** Sugarcane Bagasse (SB) was obtained from Madukismo Sugar Factory, Yogyakarta, Indonesia. Sugarcane bagasse was washed with distilled water several times

to remove the dirt and dried in an oven at 110°C for 24 hrs. The material was subsequently pulverized and granulometrically into particles with sizes 177 µm. The proximate examination of the raw material was performed at Pusat Studi Pangan dan Gizi Gajah Mada University, Indonesia to determine the concentration of hemicellulose, cellulose, and lignin.

**Synthesis of Activated Carbon.** Conventional pyrolysis and microwave-assisted ZnCl<sub>2</sub> chemical activation were utilized to prepare the activated carbon. Sugarcane bagasse was placed in a furnace and heated at a rate of 15°C/min from 30 to 500°C for one hour. The generated char is properly preserved in a plastic container for further activation. The activation process for activated carbon consists of two steps: chemical and physical. The first step is chemical activation. In 50 ml of ZnCl<sub>2</sub> solution, the carbon is submerged for 12 hrs at impregnation ratios of 10% w/v (low factor level), 30% w/v (midpoint), and 50% w/v (high factor level). The carbon was subsequently reactivated (second step) using a conventional microwave oven (Samsung ME731K 2.45 GHz) and subjected to different heating times (min) and microwave radiation (W). Table 1 shows the combination of variables for the optimization process. The optimization process of activated carbon was performed in duplicate for each combination variable. The resulting material was repeatedly rinsed with 0.1 M HCl and distilled water until the pH of the solution dropped to 7.0. The resulting AC was separated, dried at 110°C for 24 hrs, and stored for further analysis.

Table 1  
Real variables and levels in BBD experimental design

Factors	Coded value		
	-1	0	+1
ZnCl <sub>2</sub> concentration, A (% w/v)	10	30	50
Heating time, B (min)	2	7	12
Microwave power, C (W)	150	450	800

**Box-Behnken Experimental Design.** Response surface methodology with Box Behnken Design (BBD) was used to examine the interaction between variables and response in the activated carbon activation (Junior et al., 2014). In this study, independent variables A, B, and C were altered (-1, 0, +1), with the iodine number of activated carbon (mg/g) as the dependent variable. Table 1 shows the variable combinations utilizing the BBD. The BBD experimental design employs 15 experiments with 12 factorial and 3 central points. Response prediction (Y) can be determined using a second-order polynomial regression model (Equation 1) (Junior et al., 2014; Yuan et al., 2018).

$$Y = b_0 + \sum_{i=1}^k b_i x_i + \sum_{i=1}^k b_{ii} x_i^2 + \sum_{i=1}^{k-1} \sum_{j=2}^k b_{ij} x_i x_j \quad [1]$$



Where  $Y$  is the predicted response value,  $x_i$  and  $x_j$  are real or code variables,  $b_0$  represents the constant term,  $b_i$  represents the linear effect term,  $b_{ii}$  represents the quadratic effect term, and  $b_{ij}$  represents the interaction effect term (Yuan et al., 2018). The experimental data regression model and statistical significance model were assessed using analysis of variance (ANOVA). Design Expert is utilized for design experiment development, model evaluation, and optimization.

### **Characterization of Activated Carbon**

Moisture Meter MA 45 Sartorius obtained the water content of activated carbon.  $1 \pm 0.2$  gram activated carbon was placed in an aluminum container, then heated at  $105^\circ\text{C}$  until the constant weight reached. Moisture content was displayed when the process finished. The iodine number was determined using the sodium thiosulfate volumetric method (Siragi et al., 2021). The iodine number is a relative indicator of porosity in activated carbon and can be used to estimate the surface area of activated carbon. The pore size of activated carbon can be approximated by analyzing the iodine number  $>1.0$  nm (Özhan et al., 2014). The surface area of activated carbon was determined by Brunauer-Emmet-Teller (BET). Pore structure and pore distribution were evaluated by nitrogen adsorption at 77 K using NOVA 2000 Quantachrome. Scanning electron microscopy (SEM) was performed using Jeol JSM-6510 LA, surface morphology analysis for sugarcane bagasse, carbon, and activated carbon. The functional group of raw material, carbon, and activated carbon was analyzed and identified by Fourier transform infrared spectrometer (FTIR-Perkin Elmer frontier) in the scanning range  $4000\text{--}700$   $\text{cm}^{-1}$ .

## **RESULTS AND DISCUSSION**

### **Carbonization of Sugarcane Bagasse**

Table 2 shows the results of sugarcane bagasse's hemicellulose, cellulose, and lignin. Important factors in the carbonization process are carbonation temperature and biomass composition, such as lignin, hemicellulose, and cellulose (Pratama et al., 2018). Cellulose and hemicellulose are chemically similar, the main difference being the number of saccharide units. Meanwhile, lignin is a hydrophobic compound and can inhibit water penetration. The cellulose content in bagasse will be more easily decomposed due to the thermal breakdown of the sugar in the bagasse. In addition, the decomposition of cellulose will produce volatile components (Pratama et al., 2018). Thermal decomposition of hemicellulose occurs at a temperature of  $200^\circ\text{C}$ , followed by the decomposition of cellulose at a temperature range of  $250\text{--}400^\circ\text{C}$  (Ukanwa et al., 2019), while the minimum temperature of lignin is decomposed at  $500^\circ\text{C}$  (Hidayati et al., 2016). In this study, the carbonation was carried out at a temperature of  $500^\circ\text{C}$  so that the volatile components of the biomass would evaporate, and pore formation occurred in the carbon. In the process, the carbonation and



Table 2

*Composition of sugarcane bagasse from Madukismo Sugar Factory, Yogyakarta, Indonesia*

Sample	Content		
	Hemicelulose (%)	Celulose (%)	Lignin (%)
Sugarcane Bagasse	24.81	39.67	19.61
	24.18	40.99	18.93
Average	24.295	40.330	19.270

activation of volatile compounds evaporate from cellulose and hemicellulose. The more volatile compounds evaporate, the more pores are formed. It will result in a decrease in yield and an increase in adsorption power (Elsayed & Zalut, 2015). The yield from the carbonation process is 26%.

### Response Surface Methodology and Statistical Analysis of Activated Carbon

Figures 1, 2, and 3 showed three-dimensional response surfaces for two factors. The effect of heating time on the iodine number is shown in Figures 1 and 2. Figures 1 and 2 show that the iodine number increases with heating time. Activated carbon in the sample has become dehydrated, contributing to the rise in iodine concentration (Baytar, Şahin, Saka, & Ağrak, 2018). The amount of energy transmitted to the precursor of activated carbon rises as the heating time of the microwave increases (Junior et al., 2014). In addition, prolonged microwave radiation promotes the formation of activated carbon, which has a large number of pores and active sites and improves its capability to adsorb chemicals. Furthermore, when the radiation time reaches a certain value, it can reduce the value of

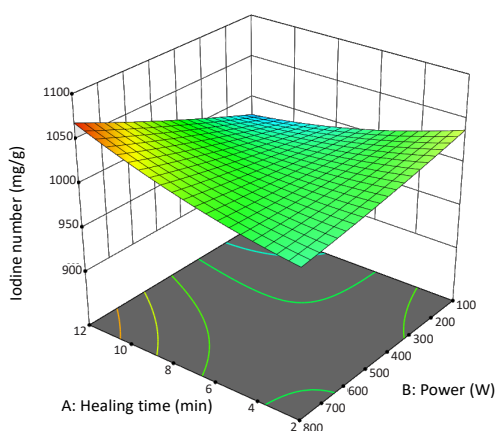


Figure 1. Three-dimensional response surface of iodine number for interaction between heating time and microwave power

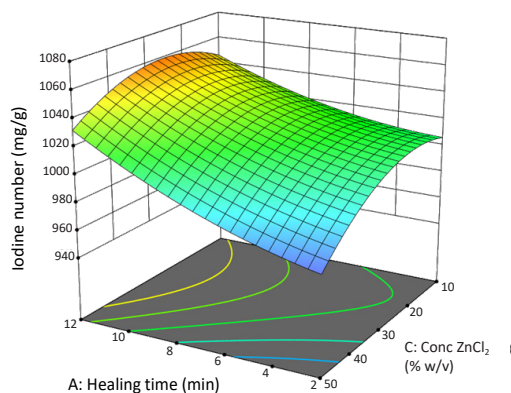


Figure 2. Three-dimensional response surface of iodine number for interaction between heating time and ZnCl<sub>2</sub> concentration

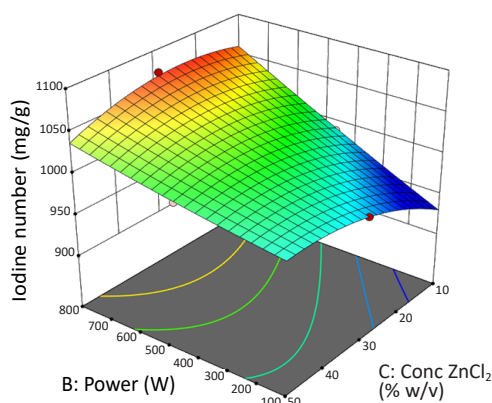


Figure 3. Three-dimensional response surface of iodine number for interaction between power microwave and  $\text{ZnCl}_2$  concentration

when the power was increased from 180 to 800 W. The increased iodine number value is due to activated carbon's ability to absorb microwave energy and modify pore structure. Furthermore, the formation of pore structure is enhanced by the interaction of  $\text{ZnCl}_2$  and sample at high microwave energies (İzgi et al., 2019). The effect of microwave power on activating activated carbon is an increase in system temperature that promotes the formation of pores in the material (Junior et al., 2014). The system temperature and the sample temperature directly affect the microwave power. Activated carbon's removal of tar and volatile compounds is enhanced by increased temperatures (İzgi et al., 2019).

In addition, sample mass and microwave power are closely related. A higher microwave power level allows more energy to enter the pores, resulting in the development of the active side and larger pores in activated carbon (Ao et al., 2018). However, when microwave power reaches a limit, too much energy will cause the amount of carbon to burn and destroy the pore structure (Ao et al., 2018). It was proven in research conducted by (İzgi et al., 2019), where the value of the iodine number decreased at a microwave power of more than 800 W. According to other studies, the value of the iodine number decreases when the microwave power approaches 500 W (Baytar, Şahin, & Saka, 2018). Overheating activated carbon can diminish the number of mesopores and micropores in the material (Ayu, 2017).

The impact of  $\text{ZnCl}_2$  concentration on the findings can be observed from Figures 2 and 3. The iodine number rises from 10 to 30% w/v in  $\text{ZnCl}_2$  concentration. In comparison, activator concentrations beyond 30 to 50% w/v tend to decline the iodine number because low  $\text{ZnCl}_2$  concentration cannot optimally activate raw material. In contrast, when the activator concentration is high, activated carbon develops a macroporous structure (Teğin et al., 2020). In addition, alterations in the pore structure of the activated carbon promote the decrease in iodine number. The interaction between the sample and excess

the iodine number because the increase in heating time can damage the pore walls of the activated carbon (İzgi et al., 2019); the external shrinkage and carbon framework will collapse (Baytar, Şahin, & Saka, 2018). A prolonged heating time will decrease its absorption capability (Ao et al., 2018).

Figures 1 and 3 show the effect of microwave power on the iodine number. Figures 1 and 3 show that the result iodine number increases as microwave power increases. The study's findings showed that the iodine number increased from around 900 mg/g to approximately 1000 mg/g

ZnCl<sub>2</sub> destroys activated carbon's micropores and forms more mesopores and macropores (Baytar, Şahin, & Saka, 2018). High concentrations of ZnCl<sub>2</sub> will promote pore collapse and destroy activated carbon's micropores (Özhan et al., 2014). Furthermore, an increase in the impregnation ratio that is too high has a negative effect, which causes combustion and clogs the pores due to an excessive amount of reagent, which reduces the access area of carbon (Ao et al., 2018).

The experimental design, as well as the responses and the predictions that were obtained, are presented in Table 3. In this study, responses were provided as iodine numbers. The iodine number is the defining characteristic of activated carbon. The iodine number is determined by the surface area of pores having a radius greater than 1 nm. Iodine adsorption was used to investigate the pore structure of activated carbon (Teğın et al., 2020). The statistical significance and components of the model were evaluated using ANOVA with a 95% confidence level. Table 4 presents the findings of the ANOVA analysis. Table 4 shows that model terms B, AB, BC, A<sup>2</sup>, and C<sup>2</sup> are statistically significant. Based on Table 4, BBD terms are considered statistically significant when the *p*-value <0.05 under the selection condition, while the *p*-value >0.05 in the BBD model is disregarded (Jawad et al., 2020). The correlation of a second-order polynomial equation between input parameters and the response is given in Equation 2.

$$Y=1012.45+12.32B+30.91AB-21.40BC+8.14A^2-19.56C^2 \quad [2]$$

The BBD model can also be validated graphically by analyzing the residual distribution's characteristics and the correlation between the actual and predicted iodine number values (Jawad et al., 2020). The normal probability versus externally studentized residuals are shown in Figure 4a. This plot determines that the residuals are normally distributed. The points in a normal distribution should be in an approximately straight line. Since the points in Figure 4a are in a straight line, it can be inferred that they are normally distributed. The residuals' normal distributions represent both the residuals' independence and the validity of the assumptions (Reghioua et al., 2021). Figure 4b shows the correlation between the iodine number's actual and predicted values. According to Figure 4b, the predicted and actual points were mostly close. This finding indicates that the experimental results of this study are acceptable.

Therefore, the Model *F*-value of 34.49 implies that the model is significant. There is only a 0.06% chance that an *F*-value this large could occur due to noise. The Lack of Fit *F*-value of 0.69 implies that the Lack of Fit is not significant relative to the pure error. There is a 63.84% chance that a Lack of Fit *F*-value this large could occur due to noise. A non-significant lack of fit is good, meaning the model is significant. The model has the adjusted R<sup>2</sup> of 0.9556, and the predicted R<sup>2</sup> of 0.8537 are reasonably in agreement. The significant model requires a difference of less than 0.2. A high correlation coefficient (R<sup>2</sup> =

0.9556) indicated the high combability between expected and experimental values (Jawad et al., 2021).

Table 3  
*Experimental design of BBD with 3-variables and experimental data for iodine number*

Run	Variable			Responses Y = iodine number (mg/g)	
	A	B	C	Actual	Predicted
1	7	450	30	1015.89	1012.45
2	12	100	30	981.87	981.79
3	12	800	30	1072.31	1068.24
4	2	800	30	1002.40	1002.48
5	7	100	50	1008.78	1007.74
6	2	450	10	1003.70	1002.58
7	2	450	50	998.56	995.53
8	2	100	30	1035.60	1039.66
9	7	450	30	1015.89	1012.45
10	12	450	10	989.82	992.85
11	7	100	10	961.27	958.32
12	7	800	50	986.65	989.59
13	7	450	30	1005.56	1012.45
14	12	450	50	1012.03	1013.15
15	7	800	10	1024.72	1025.75

Table 4  
*Analysis of variance (ANOVA) for the iodine number*

Source	Sum of Squares	df	Mean Square	F-value	p-value
<b>Model</b>	8969.67	9	996.63	34.49	0.0006
A-Heating time	31.06	1	31.06	1.07	0.3473
B-Power	1214.09	1	1214.09	42.02	0.0013
C-Conc ZnCl <sub>2</sub>	87.85	1	87.85	3.04	0.1417
AB	3821.40	1	3821.40	132.26	< 0.0001
AC	186.98	1	186.98	6.47	0.0517
BC	1831.11	1	1831.11	63.37	0.0005
A <sup>2</sup>	244.38	1	244.38	8.46	0.0335
B <sup>2</sup>	22.37	1	22.37	0.7744	0.4191
C <sup>2</sup>	1411.94	1	1411.94	48.87	0.0009
<b>Residual</b>	144.47	5	28.89		
Lack of Fit	73.33	3	24.44	0.6872	0.6384
Pure Error	71.14	2	35.57		
<b>Cor Total</b>	9114.14	14			

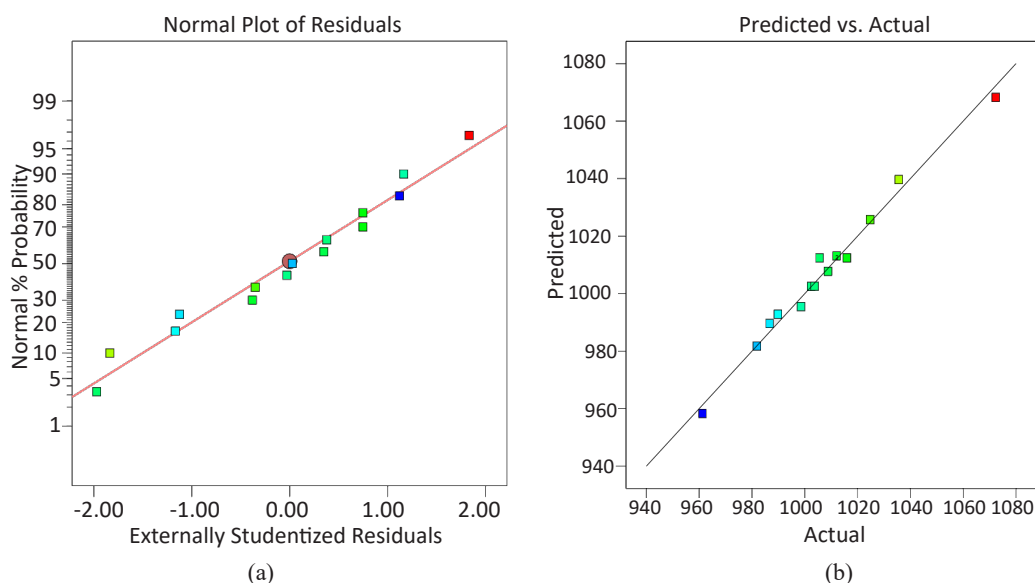


Figure 4. Normal probability plot of residuals for (a) BBD model (a), (b) plot of the relationship between the predicted and actual value of iodine number (mg/g)

## Optimization Process

The optimum conditions for synthesizing activated carbon or process variables can be determined using the iodine number. Design Experts determined the optimal conditions for the synthesis of activated carbon. The process conditions were optimized if the desirability was close to 1. (Junior et al., 2014). A response with a desirable value of 1 indicates that it was received. In contrast, a desirability rating of 0 indicates that the response value exceeds the acceptable threshold (Jawad et al., 2021). The main objective is determining the point at which desirability is at its maximum level. A desirability function is thus a well-established method for determining the simultaneous optimization of process variables (A:  $\text{ZnCl}_2$  concentration, B: Heating Time, C: Microwave Power) that produces the best performance level for the response (Iodine Number mg/g).

The numerical optimization showed the best iodine number was produced at 24.281%  $\text{ZnCl}_2$ , 12 minutes of heating time, and 800 W of microwave power; at these operating parameters, the iodine number is 1068.24 mg/g with a desirability value of 0.968 as shown in Figure 5. Overall, there was a good correlation between the results from numerical optimization utilizing desirability functions and the results from actual data. The actual iodine number is 1072.31 mg/g. The predicted and actual results have a relatively low error value of 2.035%, which means there is little difference between the actual and predicted values. The study shows that the BBD model with desirability functions may be successfully utilized to optimize the experimental condition for synthesizing activated carbon using microwave-assisted chemical activation (Jawad et al., 2021).

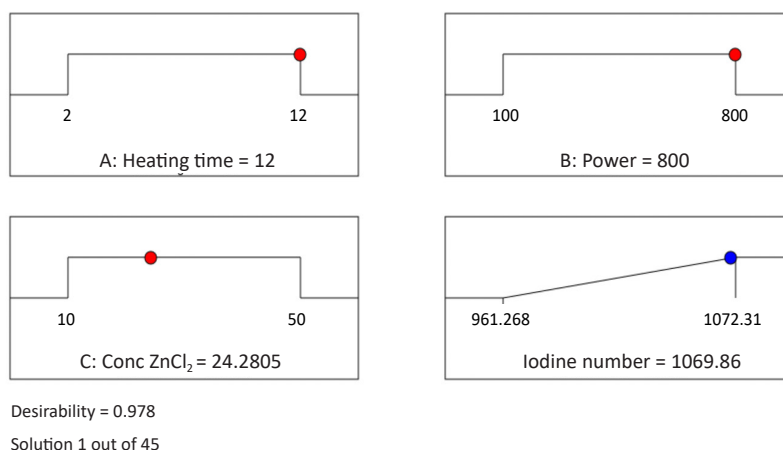


Figure 5. Desirability ramps for the optimization of preparation activated carbon input parameters for iodine number by Microwave-assisted  $\text{ZnCl}_2$  chemical activation

## Characterization of Activated Carbon

**Moisture Content.** The moisture content of activated carbon can affect the adsorption process. The lower the moisture content of activated carbon, the emptier pores can be filled by the adsorbate, allowing for an optimum adsorption process. The relatively low moisture content of activated carbon implies that water has evaporated during the carbonation and activation processes (Tasanif et al., 2020). In this study, the moisture content of activated carbon ranged from 5.51 to 9.21%. The study showed that the moisture content of activated carbon agreed with Indonesian National Standard (SNI) 06-3730-1995, which stated that the water content of powdered activated carbon should not exceed 15%.

**FT-IR Analysis.** Figure 6 shows the FTIR test results for bagasse, carbon, and activated carbon. FTIR tests on bagasse showed a band of hydroxyl groups (O-H) at wave numbers around  $3400\text{--}3500\text{ cm}^{-1}$ , C-H bonds at  $2800\text{--}2900\text{ cm}^{-1}$ , and a C=C bond at  $1500\text{--}1600\text{ cm}^{-1}$ . It is similar to a study that reported most agricultural waste, including cellulose, hemicellulose, and lignin, had wave numbers between  $3412$  and  $3392\text{ cm}^{-1}$ , indicating the presence of O-H groups (Vo et al., 2019). O-H functional groups ( $3590\text{--}3650\text{ cm}^{-1}$ ), hydrogen bonds, and C-H bonds ( $2850\text{--}2970\text{ cm}^{-1}$ ) derived from cellulose, as well as C=C bonds ( $1500\text{--}1600\text{ cm}^{-1}$ ) derived from lignin aromatic rings, are also present in bagasse (Pratama et al., 2018) The results of the analysis of carbon and activated carbon are similar in Figure 6. The lack of peaks at wave numbers  $3400\text{--}3500\text{ cm}^{-1}$  and  $2800\text{--}2900\text{ cm}^{-1}$  in the results of the FTIR analysis of carbon and activated carbon indicates that the O-H bond and the C-H bond have degraded. A monomer group replaces the O-H group, formerly a hydrogen bond. The O-H groups will attach to aromatic compounds as a result of the

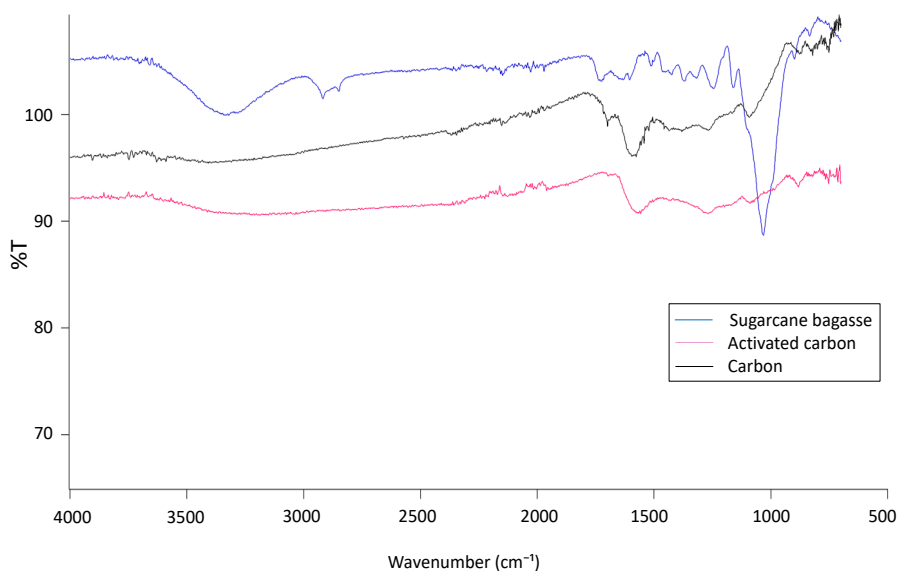


Figure 6. FT-IR analysis for sugarcane bagasse, carbon, and activated carbon

cellulose molecules' aromatization into polyaromatic structures during the carbonation and activation processes (Pratama et al., 2018). A C=O (carbonyl compound) bond is indicated at wave numbers between  $1870\text{ cm}^{-1}$  and  $1650\text{ cm}^{-1}$  in carbon and activated carbon. The presence of the C=O group, which denotes carbonyl acid stretching and is a characteristic group found in activated carbon, indicates that bagasse has generated a carbon-active material (Ayu, 2017). The difference between carbon and activated carbon can be seen in the wave numbers between  $800\text{ cm}^{-1}$  and  $1000\text{ cm}^{-1}$ , which indicate the presence of C-Cl bonds as a result of the  $\text{ZnCl}_2$  activation process (Vo et al., 2019).

**Morphological structure.** SEM analysis has been performed on the surface morphology of activated carbon. Figure 7 shows the 1000x SEM results on sugarcane bagasse, carbon, and activated carbon. Figure 7(a) shows the scanning electron microscopy (SEM) results of untreated bagasse; the bagasse's surface is smooth and has slight cracks and voids. Figure 7(b) shows the scanning electron microscope (SEM) analysis of carbon bagasse, showing that a porosity structure with an irregular pore structure has formed. While Figure 7(c) shows the scanning electron microscopy (SEM) study results on activated carbon following the activation process, the pores generated are more numerous and not uniform besides the surface of the activated carbon developing cavities.

The heating method with a furnace and microwave promotes the formation of irregular pores and cavities on the surface of activated carbon (İzgi et al., 2019). Heating in furnaces and microwaves leads to the evaporation of carbon components like lignin and organic matter, forming pores in activated carbon [Figure 7(c)]. These pores can be mesoporous



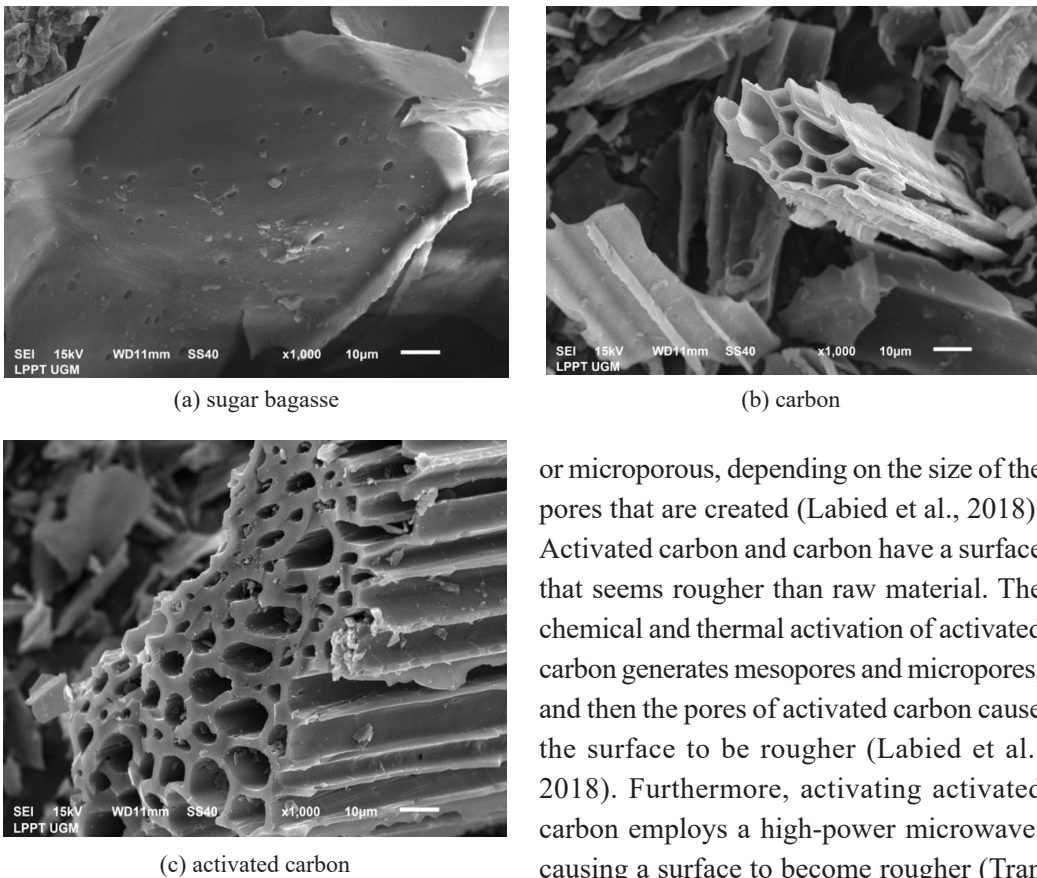


Figure 7. Scanning electron microscope (SEM) analysis of (a) sugarcane bagasse, (b) carbon, and (c) activated carbon at optimum conditions

or microporous, depending on the size of the pores that are created (Labied et al., 2018). Activated carbon and carbon have a surface that seems rougher than raw material. The chemical and thermal activation of activated carbon generates mesopores and micropores, and then the pores of activated carbon cause the surface to be rougher (Labied et al., 2018). Furthermore, activating activated carbon employs a high-power microwave, causing a surface to become rougher (Tran et al., 2016).

Specific surface area and pore structure. The  $N_2$  adsorption-desorption isotherm and pore size distribution are shown in Figure

8. Figure 8(a) shows the adsorption-desorption isotherm of an activated carbon with a type IV isotherm according to IUPAC standards. Isotherms type IV show that the pores generated are mesopores with small pore diameters (Horvat et al., 2022). Thommes et al. (2015) stated that type IV isotherms generate mesopore adsorbents with cylinder and conical geometries that are closed at the ends. The adsorption-desorption isotherm of activated carbon has a hysterical loop, as shown in Figure 8(a). Since the amount of  $N_2$  gas adsorbed is not equal to that of adsorbed, hysteresis indicates that pore condensation occurred during adsorption (Luo et al., 2019). According to Thommes et al. (2015), loop hysteresis on  $BAC_{op}$  is classified as hysterical H4, which exhibits slit-shaped pore material and is frequently observed on micro-mesoporous carbon. Hysteresis Loop H4 indicates a complex interaction with micro and mesopore (Horvat et al., 2022).

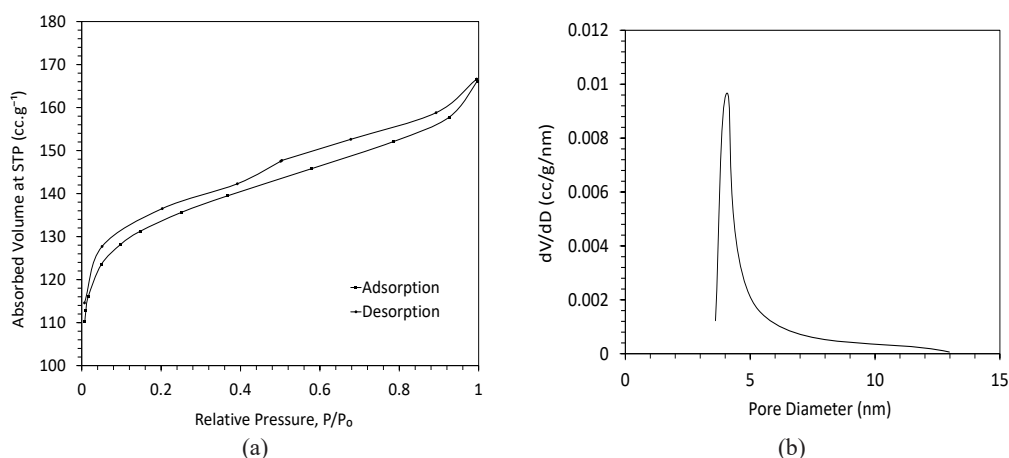


Figure 8. (a) N<sub>2</sub> Adsorption-desorption isotherm of AC<sub>op</sub> and (b) Barret-Joyner-Halenda BJH Corresponding Pore size distribution of AC<sub>op</sub>

Figure 8 shows the Barret-Joyner-Halenda (BJH) pore size distribution of BAC<sub>op</sub>. Most BAC<sub>op</sub> pores have a diameter between 3.5 and 5.4 nm, indicating that BAC<sub>op</sub> pores are part of the mesopore. The pore volume of BAC<sub>op</sub> is 0.054 cc/g, the average pore diameter is 4.071 nm, and the total pore volume is 0.2531 cc/g. It has a surface area of 446.874 m<sup>2</sup>/g. Table 5 compares the surface area of activated carbon with several previous studies. Based on the BET study, the average pore width of BAC<sub>op</sub> is 4.071 nm, with pore sizes ranging from 3.5 to 5.4 nm. Based on the average pore width and the pore size distribution, BAC<sub>op</sub> may adsorb chromium ions with a molecular size of less than 0.088 nm (Labied et al., 2018).

## CONCLUSION

This study reported the synthesis of activated carbon from bagasse by ZnCl<sub>2</sub> using microwave-assisted activation. Optimizing the condition of the AC synthesis process has been accomplished with effectiveness utilizing the response surface methodology approach. The BAC optimization (BAC<sub>op</sub>) was at a ZnCl<sub>2</sub> concentration of 24.281 (%w/v), 12 minutes of heating time, and 800 W of microwave power. The characterization results show that BAC<sub>op</sub> has mesoporous carbon and a desirable surface area of 446.874 m<sup>2</sup>/g, average pore size of 4.071 nm, pore volume of 0.054 cc/g, and total pore volume of 0.2531 cc/g. SEM analysis revealed that microwave-generated pore structures lead to the ZnCl<sub>2</sub> activation process. The FT-IR analysis shows the existence of functional groups as well as changes in functional groups from raw material to activated carbon. The moisture content study findings are 5.51 to 9.21% in accordance with the Indonesian National Standard (SNI) 06-3730-1995. The isothermal adsorption-desorption process is classified as type IV adsorption with hysteresis loop H4, suggesting that the pore distribution in activated carbon is mesoporous with a tiny pore width and slit-shape pore materials.

## ACKNOWLEDGEMENTS

We thank Badan Pemberdayaan Sumber Daya Manusia Industri Industry Ministry, Indonesia for financial support. We are also thankful to the Laboratory of Chemistry, Department of Leather Processing Technology, Politeknik ATK Yogyakarta, Indonesia for providing the facilities for our study.

## REFERENCES

- Abdulhameed, A. S., Firdaus Hum, N. N. M., Rangabhashiyam, S., Jawad, A. H., Wilson, L. D., Yaseen, Z. M., Al-Kahtani, A. A., & Alothman, Z. A. (2021). Statistical modeling and mechanistic pathway for methylene blue dye removal by high surface area and mesoporous grass-based activated carbon using  $K_2CO_3$  activator. *Journal of Environmental Chemical Engineering*, 9(4), Article 105530. <https://doi.org/10.1016/j.jece.2021.105530>
- Ao, W., Fu, J., Mao, X., Kang, Q., Ran, C., Liu, Y., Zhang, H., Gao, Z., Li, J., Liu, G., & Dai, J. (2018). Microwave assisted preparation of activated carbon from biomass: A review. *Renewable and Sustainable Energy Reviews*, 92, 958-979. <https://doi.org/10.1016/j.rser.2018.04.051>
- Arnelli, A., Putri, U. H. H., Cholis, F. N., & Astuti, Y. (2019). Use of microwave radiation for activating carbon from rice husk using  $ZnCl_2$  activator. *Jurnal Kimia Sains dan Aplikasi*, 22(6), 283-291. <https://doi.org/10.14710/jksa.22.6.283-291>
- Ayu, G. E. (2017). *Pembuatan karbon aktif dari tempurung kelapa (coconut shell) dengan proses pengaktifan kimia  $ZnCl_2$  menggunakan microwave* [Preparation of activated carbon from coconut shell using Microwave-assisted  $ZnCl_2$  Chemical activation]. [Doctoral dissertation]. Universitas Sumatera Utara, Indonesia. <http://repositori.usu.ac.id/handle/123456789/20774>
- Baytar, O., Şahin, Ö., & Saka, C. (2018). Sequential application of microwave and conventional heating methods for preparation of activated carbon from biomass and its methylene blue adsorption. *Applied Thermal Engineering*, 138, 542-551. <https://doi.org/10.1016/j.applthermaleng.2018.04.039>
- Baytar, O., Şahin, Ö., Saka, C., & Ağrak, S. (2018). Characterization of microwave and conventional heating on the pyrolysis of pistachio shells for the adsorption of methylene blue and iodine. *Analytical Letters*, 51(14), 2205-2220. <https://doi.org/10.1080/00032719.2017.1415920>
- Bian, Y., Yuan, Q., Zhu, G., Ren, B., Hursthouse, A., & Zhang, P. (2018). Recycling of waste sludge: Preparation and application of sludge-based activated carbon. *International Journal of Polymer Science*, Article 8320609. <https://doi.org/10.1155/2018/8320609>
- Buthiyappan, A., Gopalan, J., & Raman, A. A. A. (2019). Synthesis of iron oxides impregnated green adsorbent from sugarcane bagasse: Characterization and evaluation of adsorption efficiency. *Journal of Environmental Management*, 249, Article 109323. <https://doi.org/10.1016/j.jenvman.2019.109323>
- Daochalermwong, A., Chanka, N., Songsrirote, K., Dittanet, P., Niamnuy, C., & Seubsai, A. (2020). Removal of heavy metal ions using modified celluloses prepared from pineapple leaf fiber. *ACS Omega*, 5(10), 5285-5296. <https://doi.org/10.1021/acsomega.9b04326>
- Deng, H., Yang, L., Tao, G., & Dai, J. (2009). Preparation and characterization of activated carbon from cotton stalk by microwave assisted chemical activation-application in methylene blue adsorption from aqueous solution. *Journal of Hazardous Materials*, 166(2-3), 1514-1521. <https://doi.org/10.1016/j.jhazmat.2008.12.080>

- Elsayed, M. A., & Zalal, O. A. (2015). Factor affecting microwave assisted preparation of activated carbon from local raw materials. *International Letters of Chemistry, Physics and Astronomy*, 47, 15-23. <https://doi.org/10.18052/www.scipress.com/ilcpa.47.15>
- el Nemr, A., Aboughaly, R. M., el Sikaily, A., Masoud, M. S., Ramadan, M. S., & Ragab, S. (2022). Microporous-activated carbons of type I adsorption isotherm derived from sugarcane bagasse impregnated with zinc chloride. *Carbon Letters*, 32(1), 229-249. <https://doi.org/10.1007/s42823-021-00270-1>
- Foo, K. Y., Lee, L. K., & Hameed, B. H. (2013). Preparation of activated carbon from sugarcane bagasse by microwave assisted activation for the remediation of semi-aerobic landfill leachate. *Bioresource Technology*, 134, 166-172. <https://doi.org/10.1016/j.biortech.2013.01.139>
- Hidayati, A. S. D. S. N., Kurniawan, S., Restu, N. W., & Ismuyanto, B. (2016). Potensi ampas tebu sebagai alternatif bahan baku pembuatan karbon aktif [Potential sugarcane bagasse as an alternative raw material for activated carbon production]. *Natural B*, 3(4), 311-317.
- Horvat, G., Pantić, M., Knez, Ž., & Novak, Z. (2022). A brief evaluation of pore structure determination for bioaerogels. *Gels*, 8(7), Article 438. <https://doi.org/10.3390/gels8070438>
- İzgi, M. S., Saka, C., Baytar, O., Saraçoğlu, G., & Şahin, Ö. (2019). Preparation and characterization of activated carbon from microwave and conventional heated almond shells using phosphoric acid activation. *Analytical Letters*, 52(5), 772-789. <https://doi.org/10.1080/00032719.2018.1495223>
- Jawad, A. H., Abdulhameed, A. S., Hanafiah, M. A. K. M., ALothman, Z. A., Khan, M. R., & Surip, S. N. (2021). Numerical desirability function for adsorption of methylene blue dye by sulfonated pomegranate peel biochar: Modeling, kinetic, isotherm, thermodynamic, and mechanism study. *Korean Journal of Chemical Engineering*, 38(7), 1499-1509. <https://doi.org/10.1007/s11814-021-0801-9>
- Jawad, A. H., Mohammed, I. A., & Abdulhameed, A. S. (2020). Tuning of fly ash loading into chitosan-ethylene glycol diglycidyl ether composite for enhanced removal of reactive red 120 dye: Optimization using the box-behnken design. *Journal of Polymers and the Environment*, 28(10), 2720-2733. <https://doi.org/10.1007/s10924-020-01804-w>
- Jiang, W., Zhang, L., Guo, X., Yang, M., Lu, Y., Wang, Y., Zheng, Y., & Wei, G. (2021). Adsorption of cationic dye from water using an iron oxide/activated carbon magnetic composites prepared from sugarcane bagasse by microwave method. *Environmental Technology*, 42(3), 337-350. <https://doi.org/10.1080/09593330.2019.1627425>
- Junior, O. P., Cazetta, A. L., Gomes, R. C., Barizão, É. O., Souza, I. P. A. F., Martins, A. C., Asefa, T., & Almeida, V. C. (2014). Synthesis of ZnCl<sub>2</sub>-activated carbon from macadamia nut endocarp (*Macadamia integrifolia*) by microwave-assisted pyrolysis: Optimization using RSM and methylene blue adsorption. *Journal of Analytical and Applied Pyrolysis*, 105, 166-176. <https://doi.org/10.1016/j.jaap.2013.10.015>
- Karri, R. R., Sahu, J. N., & Meikap, B. C. (2020). Improving efficacy of Cr (VI) adsorption process on sustainable adsorbent derived from waste biomass (sugarcane bagasse) with help of ant colony optimization. *Industrial Crops and Products*, 143, Article 111927. <https://doi.org/10.1016/j.indcrop.2019.111927>
- Kaushik, A., Basu, S., Singh, K., Batra, V. S., & Balakrishnan, M. (2017). Activated carbon from sugarcane bagasse ash for melanoidins recovery. *Journal of Environmental Management*, 200, 29-34. <https://doi.org/10.1016/j.jenvman.2017.05.060>
- Labied, R., Benturki, O., Hamitouche, A. Y. E., & Donnot, A. (2018). Adsorption of hexavalent chromium by activated carbon obtained from a waste lignocellulosic material (*Ziziphus jujuba* cores): Kinetic, equilibrium, and thermodynamic study. *Adsorption Science and Technology*, 36(3-4), 1066-1099. <https://doi.org/10.1177/0263617417750739>

- Lam, S. S., Liew, R. K., Wong, Y. M., Yek, P. N. Y., Ma, N. L., Lee, C. L., & Chase, H. A. (2017). Microwave-assisted pyrolysis with chemical activation, an innovative method to convert orange peel into activated carbon with improved properties as dye adsorbent. *Journal of Cleaner Production*, *162*, 1376-1387. <https://doi.org/10.1016/j.jclepro.2017.06.131>
- Luo, X., Cai, Y., Liu, L., & Zeng, J. (2019). Cr(VI) adsorption performance and mechanism of an effective activated carbon prepared from bagasse with a one-step pyrolysis and ZnCl<sub>2</sub> activation method. *Cellulose*, *26*(8), 4921-4934. <https://doi.org/10.1007/s10570-019-02418-9>
- Mahmood, T., Ali, R., Naeem, A., Hamayun, M., & Aslam, M. (2017). Potential of used *Camellia sinensis* leaves as precursor for activated carbon preparation by chemical activation with H<sub>3</sub>PO<sub>4</sub>; optimization using response surface methodology. *Process Safety and Environmental Protection*, *109*, 548-563. <https://doi.org/10.1016/j.psep.2017.04.024>
- Ozdemir, I., Şahin, M., Orhan, R., & Erdem, M. (2014). Preparation and characterization of activated carbon from grape stalk by zinc chloride activation. *Fuel Processing Technology*, *125*, 200-206. <https://doi.org/10.1016/j.fuproc.2014.04.002>
- Özhan, A., Şahin, Ö., Küçük, M. M., & Saka, C. (2014). Preparation and characterization of activated carbon from pine cone by microwave-induced ZnCl<sub>2</sub> activation and its effects on the adsorption of methylene blue. *Cellulose*, *21*(4), 2457-2467. <https://doi.org/10.1007/s10570-014-0299-y>
- Pratama, B. S., Aldriana, P., Bambang, Ismuyanto., & Saptati, A. S. D. (2018). Konversi ampas tebu menjadi biochar dan karbon aktif untuk penyisihan Cr (VI) [Conversion of sugarcane bagasse to biochar and activated carbon for Cr(VI) removal]. *Jurnal Rekayasa Bahan Alam Dan Energi Berkelanjutan*, *2*(1), 7-12. <http://rbaet.ub.ac.id/index.php/rbaet/article/view/45>
- Puchana-Rosero, M. J., Adebayo, M. A., Lima, E. C., Machado, F. M., Thue, P. S., Vaghetti, J. C. P., Umpierrez, C. S., & Gutierrez, M. (2016). Microwave-assisted activated carbon obtained from the sludge of tannery-treatment effluent plant for removal of leather dyes. *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, *504*, 105-115. <https://doi.org/10.1016/j.colsurfa.2016.05.059>
- Reghioua, A., Barkat, D., Jawad, A. H., Abdulhameed, A. S., Al-Kahtani, A. A., & Allothman, Z. A. (2021). Parametric optimization by Box-Behnken design for synthesis of magnetic chitosan-benzil/ZnO/Fe<sub>3</sub>O<sub>4</sub> nanocomposite and textile dye removal. *Journal of Environmental Chemical Engineering*, *9*(3), Article 105166. <https://doi.org/10.1016/j.jece.2021.105166>
- Salihi, I. U., Kutty, S. R. M., & Isa, M. H. (2017). Adsorption of lead ions onto activated carbon derived from sugarcane bagasse. *IOP Conference Series: Materials Science and Engineering*, *201*(1), Article 012034. <https://doi.org/10.1088/1757-899X/201/1/012034>
- Sinyoung, S., Chaiwat, W., & Kunchariyakun, K. (2021). Preparation of activated carbon from bagasse by microwave-assisted phosphoric acid activation. *Walailak Journal of Science and Technology*, *18*(16), Article 22796. <https://doi.org/10.48048/wjst.2021.22796>
- Siragi D. B. M., Desmecht, D., Hima, H. I., Mamane, O. S., & Natatou, I. (2021). Optimization of activated carbons prepared from *Parinari macrophylla* shells. *Materials Sciences and Applications*, *12*(05), 207-222. <https://doi.org/10.4236/msa.2021.125014>
- Suhdi, & Wang, S. C. (2021). Fine activated carbon from rubber fruit shell prepared by using ZnCl<sub>2</sub> and KOH activation. *Applied Sciences*, *11*(9), Article 3994. <https://doi.org/10.3390/app11093994>
- Tasanif, R., Isa, I., & Kunusa, W. R. (2020). Potensi ampas tebu sebagai adsorben logam berat Cd, Cu dan Cr [Potential of sugarcane bagasse as an adsorbent for heavy metals Cd, Cu and Cr]. *Journal of Chemistry*, *2*(1), 34-44.

- Teğin, Ş. Ö., Şahin, Ö., Baytar, O., & İzgi, M. S. (2020). Preparation and characterization of activated carbon from almond shell by microwave-assisted using  $ZnCl_2$  activator. *International Journal of Chemistry and Technology*, 4(2), 130-137. <https://doi.org/10.32571/ijct.747943>
- Thommes, M., Kaneko, K., Neimark, A. V., Olivier, J. P., Rodriguez-Reinoso, F., Rouquerol, J., & Sing, K. S. W. (2015). Physisorption of gases, with special reference to the evaluation of surface area and pore size distribution (IUPAC Technical Report). *Pure and Applied Chemistry*, 87(9-10), 1051-1069. <https://doi.org/10.1515/pac-2014-1117>
- Tran, T. V., Pham, V. T., Quynh, N. T. P., Cong, H. T., Tam, D. T. T., Thuan, V. N., & Bach, L. G. (2016). Production of activated carbon from sugarcane bagasse by chemical activation with  $ZnCl_2$ : Preparation and characterization study. *Research Journal of Chemical Sciences*, 6(5), 42-47.
- Ummartyotin, S., & Pechyen, C. (2016). Strategies for development and implementation of bio-based materials as effective renewable resources of energy: A comprehensive review on adsorbent technology. *Renewable and Sustainable Energy Reviews*, 62, 654-664. <https://doi.org/10.1016/j.rser.2016.04.066>
- Vo, A. T., Nguyen, V. P., Ouakouak, A., Nieva, A., Doma, B. T., Tran, H. N., & Chao, H. P. (2019). Efficient removal of Cr(VI) from water by biochar and activated carbon prepared through hydrothermal carbonization and pyrolysis: Adsorption-coupled reduction mechanism. *Water*, 11(6), Article 1164. <https://doi.org/10.3390/w11061164>
- Yuan, Z., Xu, Z., Zhang, D., Chen, W., Zhang, T., Huang, Y., Gu, L., Deng, H., & Tian, D. (2018). Box-Behnken design approach towards optimization of activated carbon synthesized by co-pyrolysis of waste polyester textiles and  $MgCl_2$ . *Applied Surface Science*, 427, 340-348. <https://doi.org/10.1016/j.apsusc.2017.08.241>



# Conceptual Design and Selection of Natural Fibre Reinforced Composite Cyclist Helmet Liner Using an Integrated Approach

Nurul Ain Maidin<sup>1,2</sup>, Mohd Sapuan Salit<sup>1\*</sup>, Mastura Mohammad Taha<sup>2</sup> and Mohd Zuhri Mohamed Yusoff<sup>1</sup>

<sup>1</sup>Department of Mechanical and Manufacturing, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

<sup>2</sup>Department of Manufacturing Engineering Technology, Faculty of Industrial & Manufacturing Technology & Engineering (FTKIP), Universiti Teknikal Malaysia Melaka, 76100, Melaka, Malaysia

## ABSTRACT

This paper describes the conceptual design phase in the product development of a natural fibre composites cyclist helmet liner, beginning with idea generation and ending with selecting the best design concept. The integrated Theory of Inventive Problem Solving (TRIZ), Biomimetic methods, and the Grey Relational Analysis (GRA) method are demonstrated in this paper. This work aims to produce nature-inspired design concepts and determine the best design concept for the composite cyclist helmet liner. Following that, four design concepts were generated using the TRIZ-Biomimetic method, and the variance of concepts was developed using a morphological chart. The GRA method was chosen as the multiple criteria decision-making tool to compare their cost and weight criteria. The design concept C1 was selected as the best design concept for the natural fibre composites of cyclist helmet liner conceptual design when the highest grey relational grade (GRG) value and rank with a value of 1.0000 satisfied the GRA method conditions. This paper demonstrates how the integrated method of TRIZ-Biomimetics-Morphological Chart and GRA helps researchers and engineers develop designs inspired by nature and select the

best design concept during the conceptual design stage using a systematic strategy and justified solutions.

## ARTICLE INFO

### Article history:

Received: 14 March 2023

Accepted: 20 July 2023

Published: 15 January 2024

DOI: <https://doi.org/10.47836/pjst.32.1.23>

### E-mail addresses:

ain2513@gmail.com (Nurul Ain Maidin)

sapuan@upm.edu.my (Mohd Sapuan Salit)

mastura.taha@utem.edu.my (Mastura Mohammad Taha)

zuhri@upm.edu.my (Mohd Zuhri Mohamed Yusoff)

\*Corresponding author

*Keywords:* Biomimetics, concept selection, conceptual design, cyclist helmet liner, GRA, integrated approach, morphological chart, TRIZ

## INTRODUCTION

The recreational, fitness, transportation, and environmentally beneficial benefits



of cycling have made it a worldwide phenomenon. Common bicycle-friendly cities like Copenhagen, Denmark, and Amsterdam, Netherlands, have cycling participation rates of about 36 % and 34 %, respectively (Pucher et al., 2010; The Worldwatch Institute, 2013), and local governments in various other countries have been encouraging cycling (Leng et al., 2022) due to the health benefits, traffic relief, and environmental benefits. For instance, between 2000 and 2012, the percentage of Americans who rode their bikes to work increased by 61 % (Bliven et al., 2019b; McKenzie, 2014). The Malaysian transport ministry has noticed a rise in cyclists using Malaysian roads for recreational and sporting purposes (<https://www.mot.gov.my/en/land/safety/cyclist-safety>). Despite its widespread acceptance, cycling is frequently risky. One-third of all visits to the emergency room and three-quarters of all deaths (Høye, 2018, Monique et al., 2023) are due to head injuries (Baker et al., 1993; Bliven et al., 2019a; Rivara et al., 1998).

In particular, traumatic brain injury (TBI) has been labelled the "silent epidemic" because of its potential for delayed effects on human cognition, emotion, and sensitivity (Takhounts et al., 2013). A helmet cuts the risk of fatal brain trauma in a bicycle accident by around half. This rule should be enforced even for short travels (<https://www.freemalaysiatoday.com/category/leisure/health/2022/08/23/cycling-an-easy-and-beneficial-activity-for-all-ages/>). The shell, liner, and retention straps make up the standard bicycle helmet. The liner is essential because it takes the brunt of the impact, protecting the wearer's head from harm. Liner materials like expanded polystyrene foam (EPS) and expanded polypropylene (EPP) are often utilised (Bliven et al., 2019a). While EPS's low manufacturing cost and high crushing stress-to-weight ratio support its widespread use, it does have some downsides (Blanco et al., 2014).

The benefits of bio-inspired designs have attracted the attention of many researchers working to incorporate them into engineering products. Because of their one-of-a-kind properties, bio-inspired structures can absorb more energy than their artificial counterparts can (Kassar et al., 2016). There has been a recent uptick in academic and engineering interest in eco-friendly product development. Polymer composites can replace traditional petroleum-based plastics to achieve this goal. Polymer composites are used primarily to lessen the vehicle's overall mass, reducing emissions. More stringent regulations and standards for reduced emissions and biodegradability (Ahmad et al., 2015) have led to a steady transition from synthetic fibre composites to natural fibre composites (NFCs). Given their low price, high specific properties, and low density, natural fibre composites have been frequently used to replace synthetic fibres (Al-Ghazali et al., 2022; Pickering et al., 2016; Sapuan & Mansor, 2015). Due to their low cost, lightweight, environmental friendliness, recyclability, outstanding specific strength and stiffness, and image as a natural product (Sapuan & Mansor, 2015), researchers are currently interested in using natural fibre composites (NFC) in various industries. NFC has been used for internal and external parts to replace more expensive materials, including carbon, aramid, and glass fibres.

While bicycle helmets made from synthetic composites are commonplace, attempts to use composites made from natural fibres are far rarer. Bharath et al. (2016) analysed the uses of renewable resource-based natural fibre biocomposite materials, focusing on developing coir-polyester composites for use in helmets. Biocomposite helmet manufacturing methods and mechanical characterisation were presented at an exhibition (Bharath et al., 2018). These studies are among the few that have focused on biocomposites in helmets in recent years. To the authors' knowledge, no natural fibre has been used as a reinforcing component in biocomposites, particularly for creating bicycle helmet liners.

The development process for composite products, encompassing materials, design concepts, manufacturing process choices, and life cycle analysis, must be studied early on in the concurrent engineering process when the product is still in the conceptual design phase (Sapuan et al., 2014). Conceptual design is crucial when making the calls to determine the following stages of product development (Ulrich & Eppinger, 2012; Mansor et al., 2016). The Grey Relational Analysis (GRA) approach was used to select the composite materials for the bicycle helmet. The previous study by the same researchers showed that the pineapple/polyethylene composite was the best natural fibre polymer-reinforced composite, with the natural fibre cycling helmet receiving the top ranking, according to the results. This study suggested thermoplastic polyethylene as a particularly ideal matrix in composite cyclist helmets during the selection process for the best thermoplastic matrix material using the  $\sigma_6$  technique, with the decision based on the highest performance, the lightest weight, and the most environmentally friendly criteria (Maidin et al., 2023).

The GRA approach was used while evaluating the right natural fibre to fulfil consumer and environmental needs. From the results, the GRA method was utilised and revealed that pineapple was the best-ranked natural fibre among the ten (10) natural fibre alternatives (Maidin et al., 2022). The three phases of conceptual design are ideation (pre-concept), creation (concept design), and evaluation (concept design). Designers and engineers use the ideation approach to develop potential design concepts during this phase. Some ways to create new ideas include brainstorming, blue ocean strategies, biomimicry, and the theory of inventive problem-solving (TRIZ) (Sapuan, 2017). In the early 2000s, researchers started using TRIZ tools with other approaches, such as Quality Function Deployment (QFD) (Yamashina et al., 2002). *Teoriya Resheniya Izobretatelskikh Zadach* (or TRIZ for short) is a Russian methodology developed by scientist and engineer Genrich Asthuller, who studied over 400 000 patents (Ilevbare et al., 2013). Via its initial problem description and identification, problem-solving methods, and specific solutions (Li et al., 2013), TRIZ aids scientists, engineers, and other industries in addressing difficulties.

At the conceptual design stage, TRIZ can be utilised alone or in conjunction with different approaches to produce ideas for solutions to scientific, technical, and technological problems. In recent years, academics' interest in solving engineering problems through

a combination of TRIZ and biomimetics has increased. Scientists believe creating new products can achieve positive environmental outcomes when nature is integrated into engineering technology. Eract Vincent (2009) suggested the TRIZ to understand better how nature and technology interact. In discussing the hybrid approach to product development, Shaharuzaman et al. (2020) detailed the conceptual design phase of a natural fibre composite side-door impact beam (TRIZ and biomimetic). Similarly, using the same method, Yusof et al. (2020) conceptually designed an oil palm polymer composite automobile crash box (ACB). As a means of applying biomimicry functions and TRIZ principles, Lim et al. (2018) developed the TRIZ biomimicry contradiction matrix via text mining and latent Dirichlet allocation (LDA).

This research study explores how the Integrated Method of TRIZ, Biomimetics, and Morphological Chart can be used to design a natural fibre composite bicycle helmet liner conceptually. As a multi-criteria decision-making (MCDM) technique, Grey Relational Analysis (GRA) was used to select the optimal design strategy.

## DESIGN BRIEF

Existing research is vital because it provides a roadmap for future studies to adhere to the rules and restrictions outlined for bike helmet liner product design (PDS). PDS also includes geometry limiting rules for cyclist helmet liner profiles. It is essential to know the head circumference of the rider to identify the correct size bike helmet. You can measure the circumference of your head by wrapping a flexible tape measure around the widest part of your head, about an inch above your eyebrows (<https://www.rei.com/learn/expert-advice/bicycle-helmet.html>). Also, the PDS data can be utilised as a guide when choosing eco-friendly and biodegradable materials to manufacture a low-density product.

In addition, Malaysian authorities only permit helmets with the UN ECE R22.05 and SIRIM certifications ([https://www.motoworld.com.my/index.php?route=blog/post&post\\_id=48](https://www.motoworld.com.my/index.php?route=blog/post&post_id=48)); therefore, the helmet's design must adhere to those standards. ECE R22.05 is an international standard adopted by Malaysia (UN R22.05), where the Economic Commission for Europe is shortened to "ECE". In 1958, the UN formally adopted it to describe rules for cars and bicycles. In this case, 22 refers to Regulation No. 22, updated with the 2005 Series of Standard Amendments. Sixty-two countries have confirmed their acceptance of ECE R22.05. In 1972, the United States Department of Transport released Regulation No. 22, which specified minimum standards for helmets in terms of head coverage, field of vision, user hearing, helmet projections, and material durability. The helmets are tested for inflammability, combustibility, cold, heat, moisture resistance, shock absorption, penetration, rigidity, and chain strap strength. The guideline additionally details the required labelling for helmets. It includes the helmet's dimensions, weight, and donning procedures (<https://www.godigit.com/motor-insurance/two-wheeler-insurance/helmets/>

helmet-safety-ratings). Finally, the new design for the bicycle helmet's liner must adhere to the specifications in Table 1: Liner thickness of 9 mm, head circumference of 52–64 cm, and maximum cycling helmet weight of 1 kg.

Table 1  
Product design specification summary for cyclist helmet (Leng et al., 2022; Kassar et al., 2016; Bhudolia et al., 2021; Novak et al., 2019)

Specification	Description
Head circumference (cm)	52–64
Liner thickness (mm)	9
Weight (g)	<1000

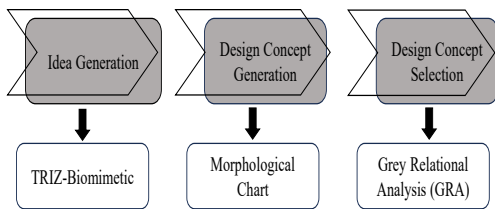


Figure 1. Research methodology based on the TRIZ-Biomimetic and GRA methods

## METHODOLOGY

Figure 1 depicts the framework of the natural fibre composite cycling helmet's conceptual design and development process. The TRIZ-Biomimetic approach was first used to generate ideas. Then, a morphological chart might be used to transform the suggestions made by TRIZ-Biomimetic into concrete plans. Afterwards, a 1:1 scale 3D model was created to aid the design process by displaying all the finer details. Finally, utilising the GRA technique, choose the best design idea for the composite cyclist helmet liner.

### Specifying Geometry and Generating Ideas with TRIZ-Biomimetic Integration

The TRIZ contradiction matrix for the intended scope of the investigation is displayed in Table 2. 40 TRIZ creative concepts were used to find solutions after an exact match was made between 39 engineering parameters from both sources. The most pertinent principal answers from Table 2 served as the foundation for developing a brand-new cycling helmet liner design. This study aims to develop a novel conceptual design for a cyclist's helmet liner made of natural fibre composites with superior energy absorption properties. So, this design needed a more stringent result, which could be achieved by minimising #2, the Weight of the stationary object. Thus, Inventive Principles #3, Local Quality, and #26, Copying were selected for implementation. The plan's finer points are laid forth in Table 3. Concepts for a new bike helmet liner design were generated using an integrated TRIZ and biomimetic approaches (Table 4).

The TRIZ technique yielded a broad answer, which the designers had to flesh out (Cerniglia et al., 2008). A morphological chart might be used to translate the suggestions made by TRIZ. Moreover, a biomimetic strategy focused on "solving #26. Copying" would use nature's wisdom to address engineering issues (Milwich et al., 2006; Lenau et al., 2009). A cyclist's helmet's traditional outer geometry profile was kept, but the interior was

rethought and redesigned to accommodate new ideas and concepts. This iteration of designs uses integrated geometry by considering symmetric body form design alongside the ideas of energy absorption inspired by spider webs (<https://www.sutori.com/en/story/biomimicry-in-architecture-spider-web-concept-and-behavior-in-design--9VBrkakv7qv6i7gmcmlKJnYJ>), pomelo peels (Ortiz et al., 2018), and honeycombs (<https://architizer.com/blog/inspiration/collections/heavenly-honeycomb-buildings/>).



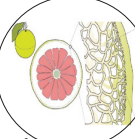
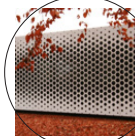
Table 2  
 TRIZ contradiction matrix for natural fibre reinforced composites cyclist helmet design

39 Engineering parameters		TRIZ principle solutions to 40 inventive principles
Improving parameters	Worsening parameters	
#2. Weight of stationary object	#8. Volume of stationary object	#2 Taking out #26 Copying #40 Composite materials #6 Universality
	#27. Reliability	#10 Preliminary action #40 Composite materials #8 Anti-weight #3 Local quality
	#14 Strength	#15 Dynamics #28. Mechanics Substitution #40 Composite materials #8 Anti-weight

Table 3  
 Details of conceptual design strategy

TRIZ solution principle	TRIZ recommended solution	Strategy for design
#40. Composite materials	Change from uniform to composite (multiple) materials	Use composition where pineapple fibre with reinforced thermoplastic polyethene for maximum energy absorption by raising composite toughness
#26. Copying	Use simpler and inexpensive copies (virtual reality, natural); replace an object or process with optical copies	Biomimetic methods provide intellectual concepts inspired by nature (how plants or animals absorb energy during collisions to survive and save their lives)
#3. Local quality	Make each part of an object function in conditions most suitable for its operation; Make each part achieve different and useful functions	Creating optimum cross section using a hybrid method to strengthen structure (combination of two or more biomimetics elements)

Table 4  
 Integrated TRIZ, morphological chart, and biomimetics methods for cyclist helmet concept generation

TRIZ solution principle and design strategy	Design feature	Solutions			
		1	2	3	4
#26. Copying Biomimetic methods provide intellectual concepts inspired by nature (how plants or animals absorb energy during collisions to save and survive their lives)	Liner design	Spider web 	Hedgehog spine 	Pomelo peel 	Honeycomb 
	Absorb energy by	Sudden impact	Raised to stand on end	Outstanding damping	Strong structure
#3. Local quality Creating the most optimum body shape using a hybrid approach to strengthen structure (combination of two or more biomimetics elements)	Body shape	Symmetrical		Asymmetry	
	Body type	Foam	Shell	Combined geometry	Multilayer

**Designing a Concept for a Cyclist's Helmet by Combining TRIZ, Biomimetics, and Morphological Charts**

Figure 2 shows four new ideas for bike helmets using a mix of TRIZ, biomimetics, and morphological charts. The 3D model was made on a 1:1 scale, making it easier to see how the design elements fit together. C1 said that the first design idea copied the spiderweb structure for the liner. The spider orb-web frame silk structure is more potent per unit weight than high-tensile steel, with a very high toughness of 2.5108 J/m<sup>3</sup> or 1.5105 J/kg (Gosline et al., 1986). It means spider web structures can take in much energy when hitting something (Du et al., 2011; Ko et al., 2004). For the second concept idea, C2, the model was made in a honeycomb shape. They used a honeycomb structure, which can be light, intense, last a long time, and save money (Boria & Forasassi, 2008; Gavrilă & Rosu, 2011). The third idea, C3, for the liner of a cyclist's helmet, was based on pomelo peel. In recent years, researchers have become interested in the structure of pomelo peel because it is so good at damping and releasing energy (Ortiz et al., 2018). Model C4 combines spiderweb



and honeycomb structures to make a very dense model. All concept designs had the same standard external geometry profile to keep the shape and function of bike helmets the same (Cristóbal et al., 2018).

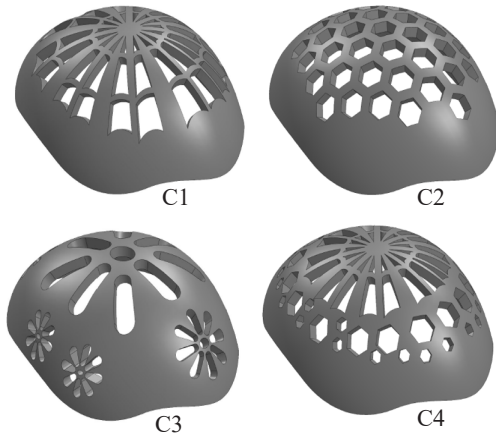


Figure 2. 3D CAD model of new liner cyclist helmet concept designs

### The GRA Method to Choose the Best Concept Design for a Cyclist's Helmet

The grey relational analysis is an impact assessment model measuring how similar the comparability and reference sequences are by comparing their relationship grades. For the study's calculations, Microsoft Excel 2016 was used. There are four main steps to the GRA method. The first step is making a grey relational sequence, which includes making a comparability sequence referring to the attributes of each option. A perfect target or reference sequence similar to the developed grey relational series is well-defined after converting all performance

criteria into a single comparability sequence. After that, the grey relational coefficient is calculated for each attribute. The final step is to determine the grey relational grade (GRG) from the coefficients and weights given to the performance attributes. The GRG found is used to rank the options (Geum et al., 2011; Jayakrishna & Vinodh, 2017; Maidin et al., 2022).

Two main attributes and the sub-attributes that go with them were taken from the PDS main document. The data came from recent and well-known literature, and cyclists' opinions were gathered using an electronic survey questionnaire sent to helmet-wearing cyclists. It was done to determine the most important thing in designing helmet liners. Both aspects of the conceptual design were chosen to help us determine what was most important to them. Cost and weight are the most important considerations when selecting a concept design for a cyclist's helmet, as shown in Figure 3. The better a part can absorb energy, the less weight it has. Also, the design must be cost-effective. It can be done by leaving out optional

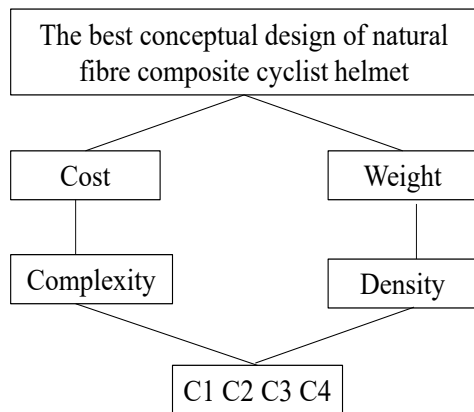


Figure 3. GRA framework for pineapple fibre polymer composites cyclist helmet concept design selection



features when making the part. The goal of the project was set at the top level. The relative importance of each of the compared criteria was based on the numerical values seen in the attributes of the overall cyclist helmet concept designs (Table 5). The part's volume and mass were calculated using the Solidworks 2021 software. These values were calculated numerically based on the properties of the pineapple material found in other research studies. The cost of this project is related to the manufacturing complex. The value is marked with a 0 and 1, representing low to medium scale, respectively.

Table 5

*Overall concept design attributes of pineapple polymer composite cyclist helmet liner (Leng et al., 2022; Maidin et al., 2023; Maidin et al., 2022; Kassar et al., 2016; Geum et al., 2011; Jayakrishna & Vinodh, 2017; Bhudolia et al., 2021; Novak et al., 2019)*

Parameter	C1	C2	C3	C4	GRA Generating
Mass (g)	446.41	459.47	465.21	453.17	SIB
Volume (mm <sup>3</sup> )	372009.7	382895.5	387677.5	377643.5	SIB
Manuf. complex	0	1	1	1	SIB

Poisson's ratio: 0.394, Tensile strength: 898.5 MPa, 1 = medium, 0 = low

The design for concept 1 used less material and was thus considered less expensive than the others that used more material, such as designs for concepts 2, 3, and 4. However, concept 2 was the easiest to draw and design during the design stage, followed by concepts 3 and 4, whereas concept 1 was the most difficult to create due to the complicated elements.

## RESULTS AND DISCUSSION

The best design concept was chosen using Grey Relational Analysis (GRA), a multi-criteria decision-making (MCDM) method. The concept selection process calculation was prepared at the outset of the research. However, it has not been presented here for the sake of brevity. The entire evaluation results using the GRA method are presented in Table 6. Overall, the GRA results of the cyclist helmet concept design selection demonstrated that the C1 design scored the highest grey relational grade (GRG) with a value of 1.0000, followed by the C4 design with a value of 0.4989 at the second rank. The third and fourth ranks are from C2 and C3, with grades of 0.3901 and 0.3333, respectively. Along with the general conclusions, each concept criterion's scores were recorded and transformed into a graph (Figure 4).

The Distinguishing Coefficient ( $\delta$ ) is a crucial component of GRA, a top multi-criteria decision-making (MCDM) model of grey system theory, which was created by Chinese researchers in the 1980s (Deng, 1989). However, generally, researchers assume  $\delta = 0.5$ , whereas the present study adapted this practice (Maidin et al., 2022; 2023). According

Table 6  
Findings of grey relational grade and rank for natural fibre composites of cyclist helmet liner conceptual design

CONCEPT	GRG	RANK
C1	1.0000	1
C4	0.4989	2
C2	0.3901	3
C3	0.3333	4

to other researchers, the values  $\delta$  variance have little impact on the components' GRA rankings. The study, however, showed that differences in  $\delta$  may impact the rank order (Mahmoudi et al., 2020). Primarily, the distinguishing coefficient in this investigation was set at 0.5, and the result is shown in Figure 5. Figure 6 depicts the results for grey relational coefficients when various distinguishing coefficients were utilised. In addition, as part of error analysis, this work assessed the effect on GRA results when the differentiating coefficients were adjusted to 0.1, 0.3, 0.7, and 0.9, correspondingly. The consistency of the results is shown in Table 7.

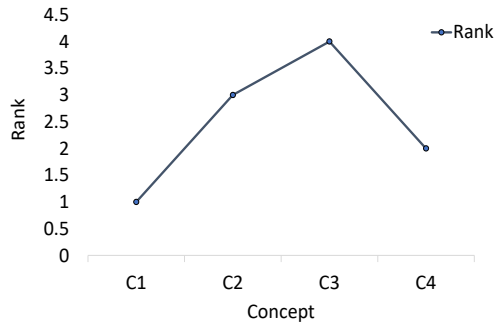


Figure 4. GRA final synthesis results with respect to all criteria

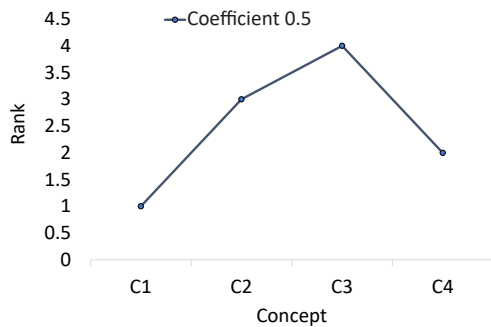


Figure 5. The distinguishing coefficient of 0.5

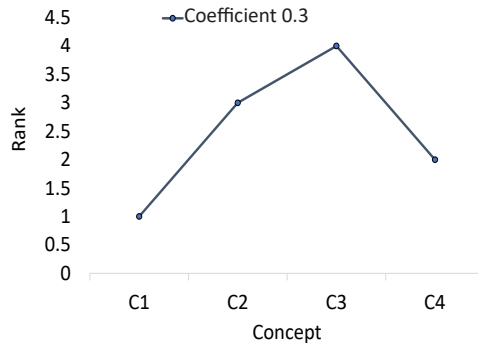
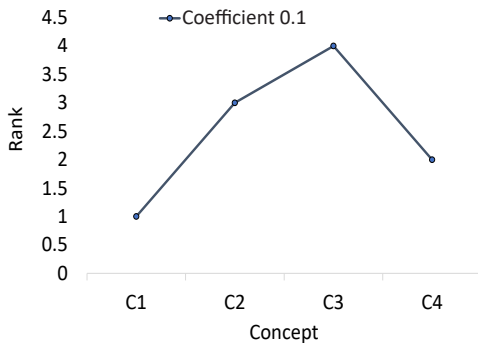


Figure 6. Impact of distinguishing coefficient on the results of GRA to find the best natural fibre composites of cyclist helmet liner conceptual design

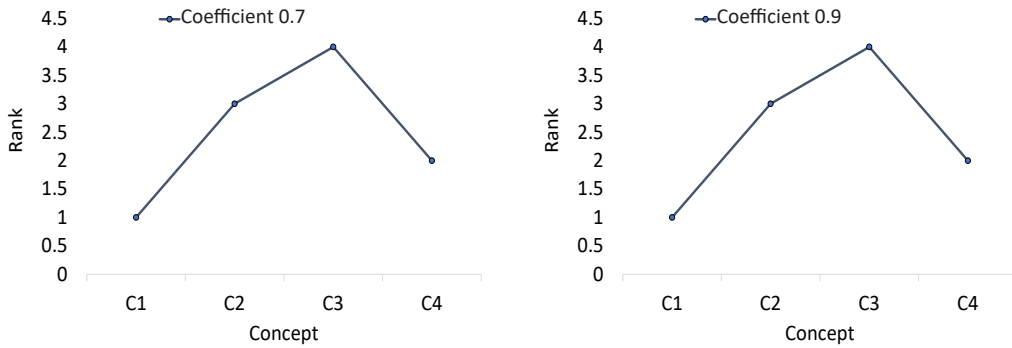


Figure 6. Continue

Table 7  
Error analysis summary based on five circumstances

Rank	Original Results	0.1	0.3	0.5	0.7	0.9
#1	C1	1	1	1	1	1
#2	C4	2	2	2	2	2
#3	C2	3	3	3	3	3
#4	C3	4	4	4	4	4

## CONCLUSION

After the GRA design selection, concept design C1 was selected as the best design concept for the natural fibre composites of cyclist helmet liner conceptual design when the highest grey relational grade (GRG) value and rank were compared. The generated results were acceptable due to the consistency value throughout the evaluation. Besides that, the stability of validation of the analysis was obtained by changing the multi-distinguishing coefficient value for error analysis, still ranked the C1 design first among the three designs. As a result, there was justification for approving the suggested design. Furthermore, the hybrid TRIZ-Biomimetics-Morphological Chart method revealed the ability to be used in conceptual design idea generation, enhancement, and development. Moreover, the GRA method has been shown as a systematic method in the process of conceptual design selection to achieve the goal or design solution, mainly when undertaking the conceptual design of natural fibre composites of cyclist helmet liners.

## ACKNOWLEDGEMENTS

The authors thank Universiti Putra Malaysia (Serdang, Malaysia) and Universiti Teknikal Malaysia Melaka (Durian Tunggal, Malaysia) for the support and the opportunity to perform this study as well as the Ministry of Higher Education Malaysia for awarding the scholarship to the principal author to conduct this research project.

## REFERENCES

- Ahmad, F., Choi, H. S., & Park, M. K. (2015). A review: Natural fiber composites selection in view of mechanical, light weight, and economic properties. *Macromolecular Materials and Engineering*, 300(1), 10-24. <https://doi.org/10.1002/mame.201400089>.
- Al-Ghazali, N. A., Aziz, F. N. A. A., Abdan, K., & Nasir, N. A. M. (2022). Mechanical properties of natural fibre reinforced geopolymer composites: A review. *Pertanika Journal of Science and Technology*, 30(3), 2053-2069. <https://doi.org/10.47836/pjst.30.3.16>
- Baker, S. P., Li, G., & Fowler, C., & Dannenberg A. (1993) *Injuries to Bicyclists: A National Perspective*. National Academies. <https://trid.trb.org/view/405159>
- Bharath, B., Kumar, G. C., Shivanna, G., Hussain, S. S., Chandrashekhar, B., Raj, B. A. S., Kumar, S. A., & Girisha, C. (2018). Fabrication and mechanical characterization of bio-composite helmet. *Materials Today: Proceedings*, 5(1), 2716-2720. <https://doi.org/10.1016/j.matpr.2018.01.053>
- Bharath, K. N., & Basavarajappa, S. (2016). Applications of biocomposite materials based on natural fibers from renewable resources: A review. *Science and Engineering of Composite Materials*, 23(2), 123-133. <https://doi.org/10.1515/secm-2014-0088>
- Bhudolia, S. K., Gohel, G., Subramanyam, E. S. B., Leong, K. F., & Gerard, P. (2021). Enhanced impact energy absorption and failure characteristics of novel fully thermoplastic and hybrid composite bicycle helmet shells. *Materials and Design*, 209, Article 110003. <https://doi.org/10.1016/j.matdes.2021.110003>
- Blanco, D. H., Cernicchi, A., & Galvanetto, U. (2014). Design of an innovative optimized motorcycle helmet. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, 228(2), 95-110. <https://doi.org/10.1177/175433711351874>
- Bliven, E., Rouhier, A., Tsai, S., Willinger, R., Bourdet, N., Deck, C., Madey, S. M., & Bottlang, M. (2019a). A novel strategy for mitigation of oblique impacts in bicycle helmets. *Journal of Forensic Biomechanics*, 10(1), Article 1000141. <https://doi.org/10.4172/2090-2697.1000141>
- Bliven, E., Rouhier, A., Tsai, S., Willinger, R., Bourdet, N., Deck, C., Madey, S. M., & Bottlang, M. (2019b). Evaluation of a novel bicycle helmet concept in oblique impact testing. *Accident Analysis & Prevention*, 124, 58-65. <https://doi.org/10.1016/j.aap.2018.12.017>
- Boria, S., & Forasassi, G. (2008). Honeycomb sandwich material modelling for dynamic simulations of a crash-box for a racing car. *WIT Transactions on The Built Environment*, 98, 167-176. <https://doi.org/10.2495/SU080171>
- Cerniglia, D., Lombardo, E., & Nigrelli, V. (2008). Conceptual design by TRIZ: An application to a rear underrun protective device for industrial vehicle. *AIP Conference Proceedings*, 1060(1), 328-331. <https://doi.org/10.1063/1.3037085>
- Cristóbal, J. R. S., Carral, L., Diaz, E., Fraguela, J. A., & Iglesias, G. (2018). Complexity and project management: A general overview. *Complexity*, 2018, Article 4891286. <https://doi.org/10.1155/2018/4891286>
- Deng, J. L. (1989). Grey information space. *Journal of Grey System*, 1(2), 103-117.
- Du, N., Yang, Z., Liu, X. Y., Li, Y., & Xu, H. Y. (2011). Structural origin of the strain-hardening of spider silk. *Advanced Functional Materials*, 21(4), 772-778. <https://doi.org/10.1002/adfm.201001397>

- Gavrila, L., & Rosu, P. (2011). Tensile properties analysis of honeycomb structures. *Review of the Air Force Academy*, 18(1), 25-28.
- Geum, Y., Cho, Y., & Park, Y. (2011). A systematic approach for diagnosing service failure: Service-specific FMEA and grey relational analysis approach. *Mathematical and Computer Modelling*, 54(11-12), 3126-3142. <https://doi.org/10.1016/j.mcm.2011.07.042>
- Gosline, J. M., Demont, M. E., & Denny, M. W. (1986). The structure and properties of spider silk. *Endeavour*, 10(1), 37-43. [https://doi.org/10.1016/0160-9327\(86\)90049-9](https://doi.org/10.1016/0160-9327(86)90049-9)
- Høyevik, A. (2018). Bicycle helmets—to wear or not to wear? A meta-analysis of the effects of bicycle helmets on injuries. *Accident Analysis & Prevention*, 117, 85-97. <https://doi.org/10.1016/j.aap.2018.03.026>
- Ilevbare, I. M., Probert, D., & Phaal, R. (2013). A review of TRIZ, and its benefits and challenges in practice. *Technovation*, 33(2-3), 30-37. <https://doi.org/10.1016/j.technovation.2012.11.003>.
- Jayakrishna, K., & Vinodh, S. (2017). Application of grey relational analysis for material and end of life strategy selection with multiple criteria. *International Journal of Materials Engineering Innovation*, 8(3-4), 250-272. <https://doi.org/10.1504/IJMATEI.2017.090241>
- Kassar, S., Siblini, S., Wehbi, B., Abro, O., & Shehadeh, M. (2016, November 11-17). *Towards a safer design of helmets: Finite element and experimental assessment*. [Paper presentation]. International Mechanical Engineering Congress and Exposition, Arizona, USA. <https://doi.org/10.1115/IMECE2016-66778>
- Ko, F. K., & Jovicic, J. (2004). Modeling of mechanical properties and structural design of spider web. *Biomacromolecules*, 5(3), 780-785. <https://doi.org/10.1021/bm0345099>
- Lenau, T. A. (2009, August 24-27). *Biomimetics as a design methodology possibilities and challenges*. [Paper presentation] DS 58-5: Proceedings of ICED 09, the 17<sup>th</sup> International Conference on Engineering Design, California, USA.
- Leng, B., Ruan, D., & Tse, K. M. (2022). Recent bicycle helmet designs and directions for future research: A comprehensive review from material and structural mechanics aspects. *International Journal of Impact Engineering*, 168, Article 104317. <https://doi.org/10.1016/j.ijimpeng.2022.104317>
- Li, M., Ming, X., Zheng, M., Xu, Z., & He, L. (2013). A framework of product innovative design process based on TRIZ and patent circumvention. *Journal of Engineering Design*, 24(12), 830-848. <https://doi.org/10.1080/09544828.2013.856388>.
- Lim, C., Yun, D., Park, I., & Yoon, B. (2018). A systematic approach for new technology development by using a biomimicry-based TRIZ contradiction matrix. *Creativity and Innovation Management*, 27(4), 414-430. <https://doi.org/10.1111/caim.12273>.
- Mahmoudi, A., Javed, S., Liu, S., & Deng, X. (2020). Distinguishing coefficient driven sensitivity analysis of GRA model for intelligent decisions: Application in project management. *Technological and Economic Development of Economy*, 26(3), 621-641. <https://doi.org/10.3846/tede.2019.11890>.
- Maidin, N. A., Sapuan, S. M., Taha, M. M., & Yusoff, M. Z. M. (2022). Material selection of natural fibre using a grey relational analysis (GRA) approach. *BioResources*, 17(1), 109-131. <https://doi.org/10.15376/BIORES.17.1.109-131>

- Maidin, N. A., Sapuan, S. M., Mastura, M. T., & Zuhri, M. Y. M. (2023). Materials selection of thermoplastic matrices of natural fibre composites for cyclist helmet using an integration of DMAIC approach in six sigma method together with grey relational analysis approach. *Journal of Renewable Materials*, 11(5), 2381-2397. <https://doi.org/10.32604/jrm.2023.026549>
- Mansor, M. R., Sapuan, S. M., Salim, M. A., Akop, M. Z., & Musthafah, M. T. (2016). Concurrent design of green composites. In D. Verman, S. Jain, X. Zhang & P. C. Gope (Eds.), *Green Approaches to Biocomposite Materials Science and Engineering* (pp.48-75). IGI Global. <https://doi.org/10.4018/978-1-5225-0424-5.ch003>
- McKenzie, B. (2014). *Modes less traveled: Bicycling and walking to work in the united states, 2008-2012* (No. ACS-25). US Department of Commerce, Economics and Statistics Administration. chrome-extension://efaidnbmnnnibpcajpcgclefindmkaj/https://lede-admin.usa.streetsblog.org/wp-content/uploads/sites/46/2014/05/acs-25.pdf
- Milwich, M., Speck, T., Speck, O., Stegmaier, T., & Planck, H. (2006). Biomimetics and technical textiles: Solving engineering problems with the help of nature's wisdom. *American Journal of Botany*, 93(10), 1455-1465. <https://doi.org/10.3732/ajb.93.10.1455>
- Novak, J., Burton, D., & Crouch, T. (2019). Aerodynamic test results of bicycle helmets in different configurations: Towards a responsive design. *Proceedings of the Institution of Mechanical Engineers, Part P: Journal of Sports Engineering and Technology*, 233(2), 268-276. <https://doi.org/10.1177/1754337118822613>
- Ortiz, J., Zhang, G., & McAdams, D. A. (2018). A model for the design of a pomelo peel bioinspired foam. *Journal of Mechanical Design*, 140(11), 1-5. <https://doi.org/10.1115/1.4040911>.
- Pappadis, M. R., Lundine, J. P., Kajankova, M., Hreha, K. P., Doria, N., Cai, X. C., & Flanagan, J. E. (2023). Education on the consequences of traumatic brain injury for children and adolescents with TBI and families/caregivers: A systematic scoping review. *Brain Injury*, 37(1), 1-23. <https://doi.org/10.1080/02699052.2022.2145357>
- Pickering, K. L., Efendy, M. G. A., & Le, T. M. (2016). A review of recent developments in natural fibre composites and their mechanical performance. *Composites Part A: Applied Science and Manufacturing*, 83, 98-112. <https://doi.org/10.1016/j.compositesa.2015.08.038>
- Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine*, 50(Supplement), S106-S125. <https://doi.org/10.1016/j.ypmed.2009.07.028>
- Rivara, F. P., Thompson, D. C., Patterson, M. Q., & Thompson, R. S. (1998). Prevention of bicycle-related injuries: Helmets, education, and legislation. *Annual Review of Public Health*, 19(1), 293-318. <https://doi.org/10.1146/annurev.publhealth.19.1.293>
- Sapuan, S. M. (2017). Conceptual design in concurrent engineering for composites. In *Composite Materials: Concurrent Engineering Approach* (pp.141-207). Oxford. <https://doi.org/10.1016/B978-0-12-802507-9.00005-2>
- Sapuan, S. M., & Mansor, M. R. (2014). Concurrent engineering approach in the development of composite products: A review. *Materials & Design*, 58, 161-167. <https://doi.org/10.1016/j.matdes.2014.01.059>

- Sapuan, S. M., & Mansor, M. R. (2015). Design of natural fiber-reinforced composite structures. In R. D. S. G. Campilho (Ed.), *Natural Fiber Composites* (pp. 255-278). CRC Press.
- Shaharuzaman, M. A., Sapuan, S. M., Mansor, M. R., & Zuhri, M. Y. M. (2020). Conceptual design of natural fiber composites as a side-door impact beam using hybrid approach. *Journal of Renewable Materials*, 8(5), 549-563. <https://doi.org/10.32604/jrm.2020.08769>
- Takhounts, E. G., Craig, M. J., Moorhouse, K., McFadden, J., & Hasija, V. (2013, November 11-13). *Development of Brain Injury Criteria (BrIC) 2013-22-0010*. [Paper presentation]. 57<sup>th</sup> Stapp Car Crash Conference, Orlando, Florida.
- The Worldwatch Institute. (2013). *Vital Signs Volume 20: The Trends That Are Shaping Our Future*. Island Press.
- Ulrich, K. T., & Eppinger, S. D. (2012). Chapter 2: Development process and organizations. In *Product Design and Development* (5th ed., pp.11-32). McGraw-Hill Irwin.
- Vincent, J. F. V. (2002). Survival of the cheapest. *Materials Today*, 5(12), 28-41. [https://doi.org/10.1016/S1369-7021\(02\)01237-3](https://doi.org/10.1016/S1369-7021(02)01237-3)
- Vincent, J. F. V. (2009). Biomimetics - A review. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 223(8), 919-939. <https://doi.org/10.1243/09544119JEIM561>
- Yamashina, H., Ito, T., & Kawada, H. (2002). Innovative product development process by integrating QFD with TRIZ. *International Journal of Production Research*, 40(5), 1031-1050. <https://doi.org/10.1080/00207540110098490>
- Yusof, N. S. B., Sapuan, S. M., Sultan, M. T. H., & Jawaaid, M. (2020). Conceptual design of oil palm fibre reinforced polymer hybrid composite automotive crash box using integrated approach. *Journal of Central South University*, 27(1), 64-75. <https://doi.org/10.1007/s11771-020-4278-1>





## Improved GCC Technique: A Comprehensive Approach to Color Cast Rectification and Image Enhancement

Danny Ngo Lung Yao<sup>1\*</sup>, Abdullah Bade<sup>2</sup>, Iznora Aini Zolkifly<sup>1</sup> and Paridah Daud<sup>1</sup>

<sup>1</sup>*School of Information Technology, Faculty of Business and Technology, UNITAR International University, 47301 Petaling Jaya, Selangor, Malaysia*

<sup>2</sup>*Mathematics Visualization (MathViz) Research Group, Faculty of Science and Natural Resources, Universiti Malaysia Sabah, 88400 Kota Kinabalu, Sabah, Malaysia*

### ABSTRACT

The domain of underwater imaging is riddled with multifarious challenges, such as light attenuation, scattering, and color distortion, which can have a detrimental impact on the quality of images. In order to address these challenges, the Generalized Color Compensation (GCC) technique has been introduced, which utilizes color compensation and color mean adjustment to rectify color cast while integrating contrast enhancement via the Contrast Limited Adaptive Histogram Equalization (CLAHE). Nevertheless, the performance of GCC is limited due to the production of bright and smooth images. To overcome this challenge, we have introduced the improved GCC approach, which employs color compensation and color mean adjustment to rectify color cast. Subsequently, a contrast-enhanced image is generated through CLAHE to improve image contrast, while the detail-enhanced image is produced via a cumulative distribution function. Furthermore, image fusion between the detail-enhanced and contrast-enhanced images yields a superior-quality image. Our experimental results demonstrate the effectiveness of our proposed technique in improving the visual quality of underwater images. Objective metrics such as Underwater Image Quality Measure (UIQM) demonstrate that our technique surpasses GCC in terms of image sharpness, colorfulness, and contrast.

### ARTICLE INFO

#### Article history:

Received: 17 May 2023

Accepted: 15 August 2023

Published: 03 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.24>

#### E-mail addresses:

[danny.ngo@unitar.my](mailto:danny.ngo@unitar.my) (Danny Ngo Lung Yao)

[abb@ums.edu.my](mailto:abb@ums.edu.my) (Abdullah Bade)

[iznora@unitar.my](mailto:iznora@unitar.my) (Iznora Aini Zolkifly)

[paridah69@unitar.my](mailto:paridah69@unitar.my) (Paridah Daud)

\*Corresponding author

*Keywords:* Color cast, GCC, underwater image enhancement

### INTRODUCTION

As the demand for underwater exploration intensifies, the significance of underwater computer vision has correspondingly escalated in acquiring crucial information.

However, the inherent lack of visibility in an underwater environment presents a substantial limitation to the potential of underwater computer vision, given that clear underwater imagery is an indispensable prerequisite. Consequently, mounting interest has been in augmenting the quality of underwater images to address this burgeoning need.

In contrast to images captured in an atmospheric environment, where light scattering effects pose the primary obstacle, underwater image quality is primarily impacted by light absorption induced by the water medium. Light energy exponentially dissipates as it penetrates deeper into the water, with the attenuation rate varying based on the wavelength of the light. Notably, different colors of light exhibit distinct penetration ranges, with red light, having the longest wavelength, attenuating more rapidly than the shorter wavelengths of green and blue light as it descends further into the water.

A digital image can be defined as a matrix of pixels, each assigned a unique color intensity value that represents color information, including RGB colors. However, underwater image quality is susceptible to degradation caused by light absorption, leading to color attenuation, especially red. This color attenuation issue causes the remaining colors to dominate, leading to color casts such as blueish, yellowish, or greenish tones.

In addition to color casts, color attenuation also leads to distortion, wherein irregular color regions appear in the images. Furthermore, underwater images are prone to various forms of degradation, including poor visibility range, loss of image contrast, image noise, and haze effects. The Generalized Color Compensation (GCC) technique (Yao et al., 2022) has been recognized for its impressive ability to remove color cast from images. However, the output images from GCC often exhibit undesired qualities such as excessive brightness and loss of texture. In light of this, the present study endeavors to address this issue by proposing a novel solution to enhance the performance of the GCC technique.

## Literature Review

Underwater image enhancement has a long history dating back to the inception of the Underwater Image Formation Model (UIFM) (Jaffe, 1990). The primary objective of underwater image enhancement techniques is to optimize the visual quality of underwater images. Research in this field has been broadly categorized into image enhancement, restoration, fusion, and learning-based enhancement. The former refers to optimizing the visual quality of images directly, while the latter strives to recover clean images from degraded ones by employing physical image formation models. Image fusion involves integrating image features from multiple sources, and learning-based enhancement employs machine learning algorithms to optimize the underwater image enhancement process.

Spatial domain and transform domain techniques are sub-categorized under image enhancement techniques, offering different approaches to enhance the quality of underwater images. Spatial domain techniques, typically low in complexity and suitable for real-time

implementation, are widely used for enhancing underwater images by adjusting the pixel intensity. Various spatial domain techniques have improved underwater images, such as image equalization, smoothing, and sharpening. However, these techniques often result in undesired effects, such as reddish or abnormal color regions. To address the color cast problem in underwater images, various color balance techniques, such as the Grey World algorithm and Underwater White Balance (UWB) (Ancuti et al., 2018), have been introduced. However, they sometimes introduce other degradation, such as color distortion and oversaturation.

Transform domain techniques differ from spatial domain techniques by initially converting the underwater image into a different domain before implementing further enhancements. These techniques have lower complexity and enable manipulation of the image frequency component, but automation poses challenges and may not improve every image component simultaneously. Common transformation techniques such as Empirical Mode Decomposition (EMD) (Çelebi & Ertürk, 2010) and Wavelet Transform (WT) (Singh et al., 2015) have been utilized in enhancing underwater images.

In image restoration, various techniques are used to correct common issues such as color cast, loss of contrast, and haze effects in images. Among these techniques are polarization-based, prior-based, and model modification techniques. Model modification techniques aim to improve the accuracy of image restoration by creating a more precise physical model. The Atmospheric Scattering Model (ASM) (Narasimhan & Nayar, 2002) is widely used to describe image formation; however, it neglects the absorption effects that significantly affect underwater image formation. Therefore, there is a need to develop a more accurate mathematical model for underwater image formation.

The Sea-Thru model (Akkaynak & Treibitz, 2019) is a model that considers the coefficient governing the increase in backscatter with distance and the signal attenuation coefficient, which depends on object range and reflectance. This model has successfully removed the color cast, restored accurate colors in underwater images, and recovered depth dependency. Besides, Pei and Chen's revised underwater image formation model (2022) has been proposed to remove the underwater effects while tackling the haze issues. However, further research is necessary to precisely define underwater image formation, as the development of model modification techniques has shown.

Underwater images often suffer from loss of contrast, color cast, and haze effects. Many haze removal techniques have been developed to address these issues, including polarization-based and prior-based techniques. Polarization-based techniques (Schechner & Karpel, 2004) involve using polarizers to capture images and exploit the relationship between underwater image degradation and partially polarized scattered airtight to restore image visibility. Prior-based techniques like Dark Channel Prior (DCP) (He et al., 2009) and Intensity Attenuation Difference Prior (IADP) (Carlevaris-Bianco et al., 2010) leverage

statistical observations made on images to restore image quality. However, prior-based techniques have limitations in removing absorption effects due to a lack of a comprehensive mathematical model.

The improvement of underwater images has a rich history in utilizing image fusion techniques. The Multi-scale Fusion (MSF) technique (Ancuti et al., 2012) is one such technique that combines a color-corrected image with a contrast-enhanced image through MSF. While image fusion techniques can integrate desired features from different images, they can also result in distortion. For example, the UWB technique occasionally produces color distortion despite removing the color cast. Unlike traditional image fusion, which involves contrast-enhanced image and color-balanced images, Non-Subsampled Shearlet Transform (NSST) based fusion (Lin et al., 2023) involves the fusion of detail-enhanced image and edge-enhanced images to deal with color distortion and low contrast issues.

Learning-based enhancement techniques have recently gained popularity for optimizing underwater image enhancement processes. These techniques rely on training data extracted from features to select and extract the optimal information for enhancement. Several learning-based enhancement techniques have been developed for underwater image enhancement, including dictionary learning, Convolutional Neural Networks (CNN), and Generative Adversarial Networks (GAN). For example, adaptive color mapping with a learning-based technique (Farhadifard et al., 2015) sharpened images based on sparse representation using learned dictionaries. CNNs have been commonly used for training underwater image enhancement processes, including UIE-net, Encoding-decoding deep CNN (Sun et al., 2017), and Underwater Residual CNN (URCNN) (Hou et al., 2018). Moreover, GAN-based techniques such as Feature-based Conditional GAN (MLFcGAN) (Liu et al., 2019) for multi-scale feature extraction and Cast-GAN (Li & Cavallaro, 2020) for color cast removal have also been developed. Both CNN and GAN have shown effectiveness in feature extraction and training to obtain optimal parameters for underwater image enhancement processes. Nonetheless, these techniques' efficacy heavily depends on the chosen enhancement approach.

## **MATERIALS AND METHODS**

The Generalized Color Compensation (GCC) technique recently stabilized the color compensation process, and no more color distortion appears as an aftereffect. Nevertheless, the GCC technique tends to smooth and brighten the image. The improved GCC technique has been proposed to rectify the color cast and improve the image detail to address these issues of the GCC technique. Figure 1 shows the process of the proposed technique.

The initial step in underwater image enhancement involves obtaining the input underwater image. However, due to the light attenuation in water, the resulting image often has a color cast that needs to be rectified. Color compensation is employed to ensure that

the image accurately represents the natural colors of the scene to tackle this issue. The proposed technique recompenses the color through Equation 1, where  $I'_c$  and  $I_c$  represent the compensated and color intensity, respectively. Moreover,  $I_{max}$  is the highest intensity among the color channels,  $C$ .

$$I'_c = I_c + (\overline{I_{max}} - \overline{I_c}) \left(1 - \frac{c}{255}\right) \left(\frac{I_{max}}{255}\right), c \in \{r, g, b\} \tag{1}$$

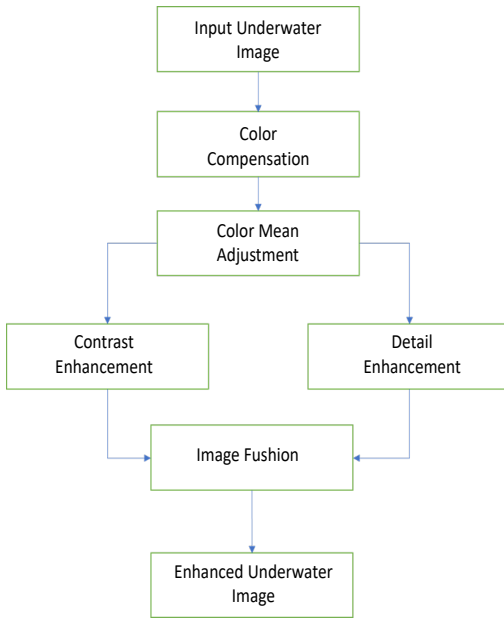


Figure 1. Improved GCC technique

Furthermore, to guarantee a balanced color representation, the image's color mean is adjusted to achieve color consistency throughout the image. It will help to rectify the color cast in underwater images. The proposed technique adjusts the color means through Equation 2, where the  $I_c^{ma}$  and  $I'_{max}$  are the color mean adjusted intensity and the highest intensity among the compensated color channels, respectively.

$$I_c^{ma} = I'_c - \overline{I_c} + \overline{I'_{max}} \tag{2}$$

A contrast-enhanced version,  $I_c^{contrast}$  is produced through CLAHE to enhance the image's contrast. In addition, a detail-enhanced version,  $I_c^{detail}$ , is produced to improve the image detail. The z score for each color channel of the contrast-enhanced

image is first computed, where the z scores will be used to compute the cumulative distribution function,  $pvalue_c$ . The detail-enhanced version image is computed as Equation 3.

$$I_c^{detail} = (I_c^{ma})(pvalue_c) \tag{3}$$

Finally, the contrast-enhanced and detail-enhanced versions are fused into a single enhanced underwater image to produce an improved image with greater visual clarity and detail.

## RESULTS AND DISCUSSION













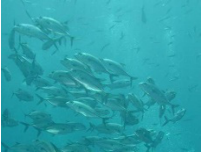


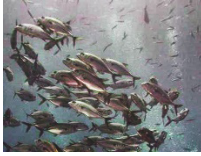




The Underwater Image Enhancement Benchmark Dataset (UIEBD) (Li et al., 2019) is a publicly available dataset of underwater images that has been widely used for testing and benchmarking image enhancement algorithms. The dataset includes 950 images with



varying levels of degradation captured in different underwater environments with various cameras. Therefore, the UIEBD was chosen as the dataset for the performance evaluation.

Table 1 presents the findings of the Visual Comparison Analysis. Our proposed technology demonstrates proficiency in eliminating color casts, including bluish, greenish, and yellowish. In comparison to the GCC technique, the proposed technique yields sharper images. Notably, for input images 1 and 4, the proposed technique outperforms the UWB technique by preserving the image's color vibrancy and delivering a more authentic representation.

Table 1  
Visual comparison analysis

	Input	UWB	GCC	Proposed technique
1				
2				
3				
4				
5				

In order to evaluate the performance of the proposed technique, the Underwater Image Quality Measure (UIQM) (Panetta et al., 2016) was selected to evaluate the image quality in terms of sharpness, colorfulness, and contrast. The sharpness component of UIQM, Underwater Image Sharpness Measure (UISM), measures the level of detail and



focus on the image, while the colorfulness component, Underwater Image Colorfulness Measure UICM), measures the vividness and saturation of colors. The contrast component, Underwater Image Contrast Measure (UIConM), measures the difference between the lightest and darkest parts of the image, indicating the overall visual clarity and definition.

Table 2 presents the results of the UIQM for five different input images processed by three distinct image processing techniques: UWB, GCC, and the proposed technique. Notably, the UIQM value indicates the level of image quality, whereby a higher UIQM value corresponds to superior image quality in terms of three key factors: sharpness, colorfulness, and contrast.

Based on Table 2, the proposed technique consistently yields the highest UIQM scores across most input images, indicating superior image quality compared to the other two techniques. For example, for image 1, the proposed technique yielded a quality metric score of 2.976, which is higher than the scores for UWB and GCC (2.704 and 2.077, respectively). Similar trends were observed for the other input images, underscoring the proposed technique's superior image processing capabilities in sharpness, contrast, and colorfulness.

Table 2  
UIQM

Image	Input	UWB	GCC	Proposed technique
1	-1.397	2.704	2.077	<b>2.976</b>
2	0.121	3.208	3.671	<b>3.788</b>
3	0.845	1.728	2.151	<b>2.237</b>
4	-1.133	<b>3.901</b>	3.119	3.770
5	1.061	3.211	3.272	<b>3.652</b>

Table 3  
UISM

	Input	UWB	GCC	Proposed technique
1	1.379	<b>3.660</b>	2.739	3.531
2	3.011	3.633	<b>4.137</b>	4.074
3	0.897	1.290	<b>1.650</b>	1.535
4	2.029	4.176	3.685	<b>4.204</b>
5	2.052	<b>3.960</b>	3.680	3.742

The outcomes of the UISM are presented in Table 3. Although the proposed approach did not manifest significant advantages over the other two techniques in relation to boosting image sharpness, it demonstrated comparable levels of improvement for the sharpness metric. It strongly implies that the proposed technique holds substantial potential as a feasible alternative for enhancing image sharpness—the proposed technique obtained better results for inputs 2, 3, 4, and 5 than UWB. The proposed technique outperformed inputs 1, 4, and 5 compared to the GCC technique. Therefore, the proposed technique may not outperform the UWB and GCC in every scenario, but its performance in improving the sharpness was considerably balanced.

Table 4 shows the results of UICM that measure image colorfulness. It is also worth noting that the GCC and proposed techniques have varying results across

the different inputs. For instance, the GCC technique performs better than the proposed technique for inputs 2 and 3 but worse for inputs 1, 4, and 5. Therefore, the effectiveness of the proposed technique and GCC in producing colorful images seems to depend on the specific input image. The UWB technique also manifests variable outcomes, displaying superior performance only for inputs 1 and 5. Conversely, the proposed technique illustrates a more stable and predictable performance than the other two.

Table 4  
*UICM*

	Input	UWB	GCC	Proposed technique
1	-116.363	<b>-6.287</b>	-13.020	-12.181
2	-104.545	-10.200	<b>-2.776</b>	-3.696
3	-22.472	-10.195	<b>5.500</b>	1.804
4	-71.359	-22.583	-0.495	<b>0.376</b>
5	-40.806	-30.592	-7.375	<b>-6.060</b>

Table 5  
*UIConM*

	Input	UWB	GCC	Proposed technique
1	0.413	0.504	0.457	<b>0.637</b>
2	0.610	0.678	0.707	<b>0.752</b>
3	0.339	0.457	0.422	<b>0.485</b>
4	0.868	<b>0.919</b>	0.875	0.841
5	0.779	0.812	<b>0.845</b>	0.806

Table 5 presents the results of the UIConM, which assesses the contrast quality of images. The UIConM values are higher for images with better contrast quality. The table shows that the proposed technique consistently outperforms the other techniques for three out of five inputs in terms of contrast quality, with the highest UIConM values for each input. Furthermore, the UWB and GCC techniques exhibit varying outcomes across the inputs. For instance, the UWB technique performs better than the GCC technique for inputs 1, 3, and 4 but worse for inputs 2 and 5. It is because the UWB technique had recovered an imbalanced color, especially for inputs 1 and 4, where the images do not preserve the underwater color well but left the difference between intensities become bigger, which eventually led to the conclusion of obtaining

higher contrast. The proposed technique, on the other hand, preserves the underwater color and, at the same time, obtains better contrast.

## CONCLUSION

Based on the evaluations conducted, it can be concluded that the proposed technique performs better than the other two techniques (UWB and GCC) in terms of overall image quality, as measured by the Underwater Image Quality Measure (UIQM). The proposed technique consistently yields the highest UIQM scores across most input images, indicating superior image quality in sharpness, colorfulness, and contrast. The UISM results also suggest that the proposed technique has comparable levels of improvement for image sharpness, indicating that it holds potential as a feasible alternative for enhancing image sharpness. The UICM and UIConM results show that the proposed technique produces

colorful and high-contrast images that are more stable and predictable than the other two techniques, which exhibit varying outcomes across different input images.

Overall, the proposed technique appears to be a promising image processing technique for eliminating color casts and enhancing overall image quality. Our proposed improved GCC technique has preserved the capability of the GCC technique, which can remove the various color casts like bluish, yellowish, and greenish tones without further distorting the image color. Meanwhile, the proposed method overcomes the smoothing image issue held by the GCC method.

## ACKNOWLEDGEMENTS

This research was supported by the UNITAR International University, Malaysia.

## REFERENCES

- Akkaynak, D., & Treibitz, T. (2019, June 15-20). *Sea-thru: A method for removing water from underwater images*. [Paper presentation]. IEEE/CVF Conference on Computer Vision & Pattern Recognition (CVPR), Long Beach, USA. <https://doi.org/10.1109/CVPR.2019.00178>
- Ancuti, C. O., Ancuti, C., De Vleeschouwer, C., & Bekaert, P. (2018). Color balance & fusion for underwater image enhancement. *IEEE Transactions on Image Processing*, 27(1), 379-393. <https://doi.org/10.1109/TIP.2017.2759252>
- Ancuti, C., Ancuti, C. O., Haber, T., & Bekaert, P. (2012, June 16-21). *Enhancing underwater images & videos by fusion*. [Paper presentation]. IEEE Conference on Computer Vision & Pattern Recognition, Providence, USA. <https://doi.org/10.1109/CVPR.2012.6247661>
- Carlevaris-Bianco, N., Mohan, A., & Eustice, R. M. (2010, September 20-23). *Initial results in underwater single image dehazing*. [Paper presentation]. Oceans 2010 MTS/IEEE Seattle, Seattle, USA. <https://doi.org/10.1109/OCEANS.2010.5664428>
- Çelebi, A. T., & Ertürk, S. (2010, July 7-10). *Empirical mode decomposition based visual enhancement of underwater images*. [Paper presentation]. 2<sup>nd</sup> International Conference on Image Processing Theory, Tools & Applications, Paris, France. <https://doi.org/10.1109/IPTA.2010.5586758>
- Farhadifard, F., Zhou, Z., & von Lukas, U. F. (2015, September 7-9). *Learning-based underwater image enhancement with adaptive color mapping*. [Paper presentation]. 9<sup>th</sup> International Symposium on Image & Signal Processing & Analysis (ISPA), Zagreb, Croatia. <https://doi.org/10.1109/ISPA.2015.7306031>
- He, K., Sun, J., & Tang, X. (2009, June 20-25). *Single image haze removal using dark channel prior*. [Paper presentation]. IEEE Conference on Computer Vision & Pattern Recognition, Miami, Florida. <https://doi.org/10.1109/CVPR.2009.5206515>
- Hou, M., Liu, R., Fan, X., & Luo, Z. (2018, October 7-10). *Joint residual learning for underwater image enhancement*. [Paper presentation] 25<sup>th</sup> IEEE International Conference on Image Processing (ICIP), Athens, Greece. <https://doi.org/10.1109/ICIP.2018.8451209>
- Jaffe, J. S. (1990). Computer modeling & the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, 15(2), 101-111. <https://doi.org/10.1109/48.50695>

- Li, C. Y., & Cavallaro, A. (2020, October 25-28). *Cast-gan: Learning to remove colour cast from underwater images*. [Paper presentation]. IEEE International Conference on Image Processing (ICIP), Abu Dabi, United Arab Emirates. <https://doi.org/10.1109/ICIP40778.2020.9191157>
- Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., & Tao, D. (2019). An underwater image enhancement benchmark dataset & beyond. *IEEE Transactions on Image Processing*, 29, 4376-4389. <https://doi.org/10.1109/TIP.2019.2955241>
- Lin, S., Li, Z., Zheng, F., Zhao, Q., & Li, S. (2023). Underwater image enhancement based on adaptive color correction and improved retinex algorithm. *IEEE Access*, 11, 27620-27630. <https://doi.org/10.1109/ACCESS.2023.3258698>
- Liu, X., Gao, Z., & Chen, B. M. (2019). MLFcGAN: Multilevel feature fusion-based conditional GAN for underwater image color correction. *IEEE Geoscience & Remote Sensing Letters*, 17(9), 1488-1492. <https://doi.org/10.1109/LGRS.2019.2950056>
- Narasimhan, S. G., & Nayar, S. K. (2002). Vision and the atmosphere. *International Journal of Computer Vision*, 48, 233-254. <https://doi.org/10.1023/A:1016328200723>
- Panetta, K., Gao, C., & Agaian, S. (2016). Human-visual-system-inspired underwater image quality measures. *IEEE Journal of Oceanic Engineering*, 41(3), 541-551. <https://doi.org/10.1109/JOE.2015.2469915>
- Pei, S. C., & Chen, C. Y. (2022). Underwater images enhancement by revised underwater images formation model. *IEEE Access*, 10, 108817-108831. <https://doi.org/10.1109/ACCESS.2022.3213340>
- Schechner, Y. Y., & Karpel, N. (2004, June 27-July 2). *Clear underwater vision*. [Paper presentation]. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision & Pattern Recognition, (CVPR), Washington, USA. <https://doi.org/10.1109/CVPR.2004.1315078>
- Singh, G., Jaggi, N., Vasamsetti, S., Sardana, H. K., Kumar, S., & Mittal, N. (2015, February 23-25). *Underwater image/video enhancement using wavelet based color correction (WBCC) method*. IEEE Underwater Technology (UT), Chennai, India. <https://doi.org/10.1109/UT.2015.7108303>
- Sun, X., Liu, L., & Dong, J. (2017, August 4-8). *Underwater image enhancement with encoding-decoding deep CNN networks*. [Paper presentation]. IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People & Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI), San Francisco, USA. <https://doi.org/10.1109/UIC-ATC.2017.8397462>
- Yao, D. N. L., Bade, A., & Waheed, Z. (2022). Recompense the color loss for underwater image using generalized color compensation (GCC) technique. *Journal of Physics: Conference Series*, 2314(1), 012006. <https://dx.doi.org/10.1088/1742-6596/2314/1/012006>

## Transfer Learning for Lung Nodules Classification with CNN and Random Forest

Abdulrazak Yahya Saleh<sup>1\*</sup>, Chee Ka Chin<sup>2</sup> and Ros Ameer Rosdi<sup>1</sup>

<sup>1</sup>Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

<sup>2</sup>Department of Electrical and Electronic, Faculty of Engineering, Universiti Malaysia Sarawak (UNIMAS), 94300 Kota Samarahan, Sarawak, Malaysia

### ABSTRACT

Machine learning and deep neural networks are improving various industries, including healthcare, which improves daily life. Deep neural networks, including Convolutional Neural Networks (CNNs), provide valuable insights and support in improving daily activities. In particular, CNNs enable the recognition and classification of images from CT and MRI scans and other tasks. However, training a CNN requires many datasets to attain optimal accuracy and performance, which is challenging in the medical field due to ethical worries, the lack of descriptive notes from experts and labeled data, and the overall scarcity of disease images. To overcome these challenges, this work proposes a hybrid CNN with transfer learning and a random forest algorithm for classifying lung cancer and non-cancer from CT scan images. This research aims include preprocessing lung nodular data, developing the proposed algorithm, and comparing its effectiveness with other methods. The findings indicate that the proposed hybrid CNN with transfer learning and random forest performs better than standard CNNs without transfer learning. This research demonstrates the potential of using machine learning algorithms in the healthcare industry, especially in disease detection and classification.

*Keywords:* Convolutional Neural Network, CT scan, lung nodules, random forest, transfer learning

### ARTICLE INFO

#### Article history:

Received: 17 May 2023

Accepted: 15 August 2023

Published: 03 November 2023

DOI: <https://doi.org/10.47836/pjst.32.1.25>

#### E-mail addresses:

ysahabdulrazak@unimas.my (Abdulrazak Yahya Saleh)

cheekachin@yahoo.com.my (Chee Ka Chin)

rosameerar@gmail.com (Ros Ameer Rosdi)

\*Corresponding author

### INTRODUCTION

Lung cancer is a common and deadly disease in modern times. Cancer cells initially develop in the lungs but can spread to other organs, including lymph nodes and the brain (Rajadurai et al., 2020). Lung cancer was the most common and deadly cancer worldwide in 2018, making up 11.6% of all cancer cases and deaths (Bray

et al., 2018). It was also Malaysia's second leading cause of death from cancer that year, after breast cancer (Bray et al., 2018), according to the World Health Organization. Men comprise approximately 16.6% of the patients, while women comprise 5.4%. In Malaysia, lung cancer is the third most prevalent type of cancer, affecting men more than women, according to the Malaysia National Cancer Registry Report 2012-2016 (Azizah et al., 2019). Today, Deep Learning is often used in medical image analysis (Alom et al., 2019; Arabahmadi et al., 2022; Salahuddin et al., 2022; Zakaria et al., 2022). Deep Learning is becoming increasingly popular and necessary for reliable and accurate results (Anderson et al., 2018). Deep learning simulates how the human brain processes data and recognizes patterns to make decisions. As technology and algorithms improve, machines can offer more accurate and reliable medical analysis. Identifying cells that are cancerous or malignant is essential for lung cancer therapy.

### **Literature Review**

Deep learning techniques can analyze CT scan images and identify cancer cells at an early stage to prevent them from becoming fatal (Primakov et al., 2022; Thai et al., 2021). Deep learning image analysis applications on Computed Tomography (CT) scan images to aid in the detection of malignant cells prior to their development and lethality (Primakov et al., 2022; Thai et al., 2021). Deep Learning is ideally suited for image processing tasks, especially object detection and localization (Singh & Gupta, 2019). Deep learning, especially CNN, can achieve high accuracy with abundant data (Zhao et al., 2018). Convolutional Neural Networks (CNNs) need large and precise labeled training data, such as ImageNet, to operate effectively. Unfortunately, large datasets are often unavailable for medical images because of the high cost of expert explanations, ethical concerns, and the lack of images of diseases (Zhao et al., 2018). In addition, models with a high parameter count tend to overfit and cannot learn patterns when working with lesser datasets (Li et al., 2020). Most traditional CNN architectures begin with a high parameter count, making their performance highly dependent on the data size. Therefore, datasets that consist of only hundreds or thousands of instances are incompatible with standard CNN models trained on large-scale datasets like ImageNet (Keshari et al., 2018). It is an issue that researchers need to address if they want to improve the performance of the model when handling a huge volume of annotated data. A standard CNN may prove insufficient in medical imaging with small datasets, where datasets usually comprise a few hundred to thousands of data. Numerous studies have investigated various types of CNN for lung nodule detection, false positive reduction, and classification to address this issue (Halder et al., 2020; Forte et al., 2022; Nakrani et al., 2021; Sharif et al., 2020). However, Table 1 in this document focuses specifically on CNNs employing innovative strategies instead of conventional techniques or conventional CNNs.

Table 1

*A comprehensive review of various types of CNN for pulmonary nodule detection*

	Ref.	Model	Data sets	Key Points
Hybrid CNN	Zhao et al. (2018)	2D LeNet + 2D AlexNet	LIDC/IDRI	A malignancy prediction model was created by combining the settings of the LeNet layers with the parameter settings of AlexNet.
Transfer Learning Based System	Da Nóbrega et al. (2018)	2D CNN, SVM, MLP, KNN, RF, Naïve Bayes	LIDC/IDRI	Eleven 2D CNN models are utilized for feature extraction. Models are ResNet50, DenseNet169, VGG16, Xception, VGG19, Inception-ResNet-V2, DenseNet201, MobileNet, InceptionV3, NASNetMobile and NASNetLarge .SVM, MLP, KNN, RF, and Naïve Bayes are trained separately with collected features.
Transfer Learning	Yamashita et al. (2018)	VGG16	LIDC/IDRI Private Data set	Comparison of CNN models for feature extraction, models based on feature engineering, and models based on transfer learning
Multiscale Feature with Transfer Learning	Tang et al. (2020)	3D U-Net CNN, Transfer learning	TIANCHI17 and LUNA16	In the study, 3D U-Net architecture was used to extract features from input images. Transfer learning was also utilized, and fine-tuning the model helped feature extraction. The experiment showed that layer-wise transfer training improved image recognition accuracy in situations with a limited sample size.
Advanced Off-The-Shelf CNNs	Ali et al. (2018)	3D CNN	LUNA16	Reinforcement learning for medical image analysis and lung nodule detection in CT scan images is a groundbreaking advance.
	Qin et al. (2018)	3D U-Net and 3D DenseNet	LUNA 16	Automatic lung nodule detection combines multitasking residual learning and online mining techniques for hard negative examples, including a 2D U-Net, a 3D DenseNet, and a Region Proposal Network (RPN). The 3D U-Net generates candidate nodes, while the 3D DenseNet is mainly used to reduce false positives.
	Tang et al. (2018)	3D Faster R-CNN and 3D DCNN	Tianchi AI competition	A two-phased framework is used to identify nodules and reduce false positives. The framework utilizes a 3D Faster RCNN to generate nodule samples and a 3D DCNN model to recognize nodule candidates.



Numerous studies have aimed to enhance the precision of early lung cancer detection through Deep Learning. However, there is still a discrepancy between identifying these algorithms and their integration into real medical applications. A hybrid deep learning model has been developed based on Convolutional Neural Networks (CNNs) and Random Forest (RF) for malicious node classification. The reason for using this hybrid approach is that transfer learning is advantageous when dealing with limited data and can leverage knowledge from pre-trained models, while RF offers robustness, feature importance analysis, versatility, and competitive performance across different types of tasks. Our model uses transfer learning from pre-trained CNN models to extract features from node images. The model's performance was evaluated compared to a baseline CNN model without transfer learning.

### MATERIALS AND METHODS

The study is divided into four phases: Dataset Preparation, Research Design, Application and Implementation, and Performance Analysis. The research methodology is presented in Figure 1, while Figure 2 shows the overall framework for the suggested model.

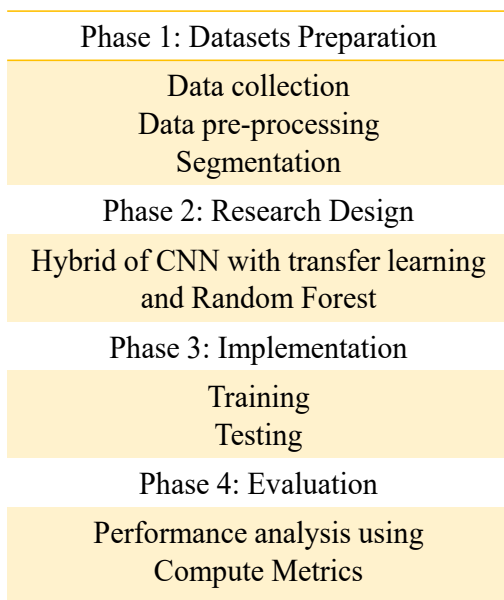


Figure 1. The research methodology employed in the study's workflow

LIDC-IDRI dataset, with different factors used to filter the heterogeneous scans. Due to the minuscule size of pulmonary nodules, only a 2.5 mm slice thickness CT scan or less was considered. CT scans that demonstrated non-uniform slice spacing or missing slices

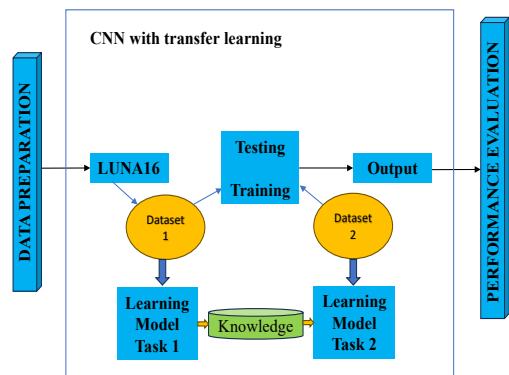


Figure 2. The overarching structure for the proposed methods

This research uses a subset of the LIDC-IDRI dataset known as LUNA16 (Camp, 2022). We use the LUNA16 dataset, a widely used lung cancer CT image dataset, and a subset of the LIDC-IDRI dataset. The LUNA16 dataset is a subset of the

were excluded from the analysis. As a result, the dataset was reduced to 888 CT scans annotated by radiologists, resulting in 36,378 annotations. Only annotations classified as nodules 3 mm were deemed relevant in lung cancer screening protocols. When multiple readers discovered nodules that were close together, their radii were combined, and their positions and diameters averaged.

Lung segmentation data are also provided, consisting of lung segmentation images calculated with automatic algorithms (Peirelinck et al., 2022). However, these segmentation images should not be utilized as the gold standard for any segmentation study. It is worth noting that the primary dataset via LIDC-IDRI was saved in DICOM (.dcm) format, an alternative format for CT scan data. The model was trained with 1528 training datasets, while the remaining CT scans of patients were used for testing and validation. In total, there were 478 test datasets and 382 validation datasets. Figure 3 shows disparate sections from a LUNA16 CT scan.

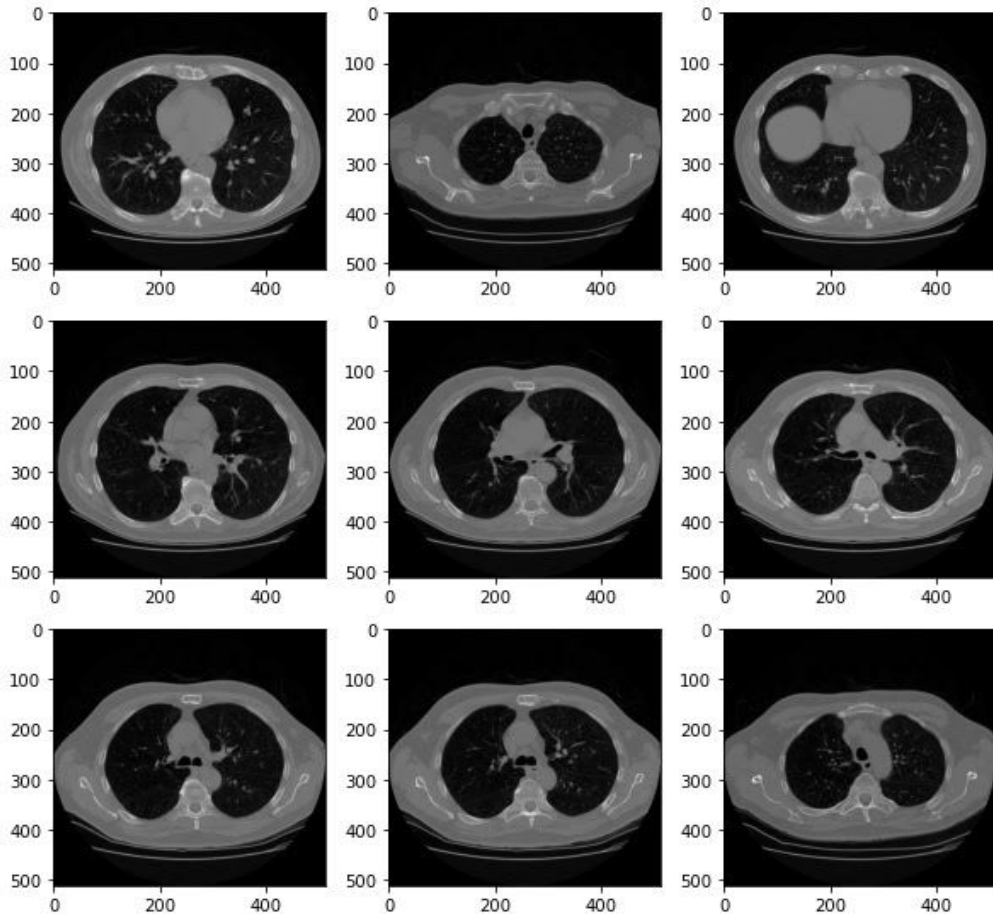


Figure 3. Primary CT scan DICOM slices extraction from LUNA16 dataset

This research employed a pre-trained VGG16 model, trained on 1.2 million natural images from the ImageNet dataset. VGG16 is initially pretrained on the large-scale ImageNet dataset, which contains millions of labeled images from various categories. During pretraining, the model learns to extract useful features from the images and classify them into one of the ImageNet classes. This pretraining step helps VGG16 learn general visual representations that can be applied to various tasks. The overall prediction of VGG16,  $VGG16(x)$ , can be represented as:

$$VGG16(x) = FC(h\{M - 1\}) \quad (1)$$

Where  $FC$  represents the last fully connected layer,  $x$  is the input image,  $M$  is the total number of fully connected layers, and  $h\{M - 1\}$  is the input to the last fully connected layer. A comprehensive framework was developed for the proposed methods in this study, with a CNN algorithm utilized to leverage the LUNA16 dataset for lung nodule classification.

The proposed algorithm utilizes Convolutional Neural Networks (CNN) with transfer learning and is fed the LUNA16 dataset as input. The proposed algorithm used transfer learning, a popular method in which pre-trained models established for one task are used as the basis for another model for a different task (Goodfellow et al., 2016). In addition, transfer learning is also a machine learning technique that involves leveraging knowledge gained from one task or domain and applying it to a different but related task or domain. It enables the reuse of pre-trained models and learned representations, which can significantly benefit new tasks, especially when the new task has limited training data. The underlying idea of transfer learning is that knowledge gained from solving one task can be transferred and utilized to improve the learning or performance of another related task. Instead of starting the learning process from scratch, transfer learning allows models to initialize with already learned features or parameters, speeding up convergence and improving generalization. Transfer learning  $\theta_t$  typically involves fine-tuning the pre-trained model on the target task. It can be represented as:

$$\theta_t = \operatorname{argmin}(\theta_t)L(\theta_t, D) + \alpha L(\theta_p + D_p) \quad (2)$$

Where  $D$  is the target task training set,  $L(\theta_t, D)$  is the loss function for the target task,  $\alpha$  is a trade-off parameter,  $\theta_p$  is the weights and biases of the pre-trained model,  $D_p$  is a separate pretraining dataset (or a subset of the target task data), and  $L(\theta_p + D_p)$  represents the loss function for pretraining.

Two prevalent transfer learning techniques exist the developed model approach and the pre-trained model approach (Peirelinck et al., 2022; Zhuang et al., 2020). Previous research has demonstrated that transfer learning can enhance the discriminative capabilities of a generic dataset and improve the model's ability to generalize to other tasks (Da Nóbrega et al., 2018).

Various scholars have also shown interest in hybridization techniques due to their recent outstanding performance in addressing cancer detection problems by enhancing algorithm performance (Chin et al., 2021a, 2021b; Saleh et al., 2021). Therefore, the proposed approach for classifying lung nodules involves utilizing a hybrid CNN with transfer learning and Random Forest (RF). Random Forest (RF) is a popular machine-learning algorithm for classification and regression tasks. An ensemble learning method combines multiple decision trees to make predictions. The prediction of the random forest  $RF(x)$  can be represented as:

$$RF(x) = \operatorname{argmax}(c) \sum t(x) R \quad (3)$$

Where  $t(x)$  is the prediction of the  $t$ -th decision tree for input  $x$ , and  $c$  represents the possible classes or outputs.

The Python programming language was used to train and test the algorithm, followed by a comprehensive performance metrics evaluation. After dataset preparation and research design, the proposed algorithm underwent training and testing, and the resulting models were evaluated based on performance. The hybrid method is implemented using Python programming on a computer system with an Intel Core i5 10th Gen processor, 16 GB RAM, and NVIDIA GeForce RTX 2060 GPU support with 6 GB RAM. The optimal hyperparameters that need to be set for this research include the number of trees, which is set to three, and the random state, which is set to 42, in the function of the random forest. The dataset is divided into two groups, namely the Train set and Test set, with most of the dataset allocated to training and the remaining portion reserved for model testing. The performance measure metrics are also analyzed based on testing accuracy, sensitivity, specificity, and F1-score, which can be represented as:

$$\text{Testing Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (6)$$

Where  $TP$ ,  $TN$ ,  $FP$ , and  $FN$  are True Positive, True Negative, False Positive, and False Negative, respectively.

## RESULTS AND DISCUSSION

The library pydicom is used to demonstrate the image field and the metadata information in the CT scan images to visualize the dataset. As the images have been deposited in

DICOM format, a standard browser cannot be used to access them. An example of the metadata stored in the file is shown in Figure 4. The watershed algorithm has been utilized to segment the dataset. The watershed algorithm has been considered a traditional image segmentation and separation algorithm. In image processing, the grayscale image can be subjected to a procedure similar to a geological watershed or drainage split, which divides adjacent drainage basins. This procedure, aptly named "watershed," employs the image as though it had been a topographic map, with each point's brightness denoting its height. As a result, it becomes possible to identify lines that follow the crests of ridges. In 1997, M. Couprie and G. Bertrand coined the term "topological watershed." The watershed algorithm begins with user-defined markers and treats pixel values as a local topography (elevation) (Bertrand, 2005).

```

-----
(0008, 0005) Specific Character Set          CS: 'ISO_IR 100'
(0008, 0016) SOP Class UID                  UI: CT Image Storage
(0008, 0018) SOP Instance UID              UI: 1.2.840.113654.2.55.247817952625791837963403492891187883824
(0008, 0060) Modality                      CS: 'CT'
(0008, 103e) Series Description             LO: 'Axial'
(0010, 0010) Patient's Name                PN: '00cba091fa4ad62cc3200a657aeb957e'
(0010, 0020) Patient ID                    LO: '00cba091fa4ad62cc3200a657aeb957e'
(0010, 0030) Patient's Birth Date          DA: '19000101'
(0018, 0060) KVP                           DS: None
(0020, 000d) Study Instance UID            UI: 2.25.86208730140539712382771890501772734277950692397709007305473
(0020, 000e) Series Instance UID          UI: 2.25.11575877329635228925808596800269974740893519451784626046614
(0020, 0011) Series Number                 IS: "3"
(0020, 0012) Acquisition Number           IS: "1"
(0020, 0013) Instance Number              IS: "134"
(0020, 0020) Patient Orientation           CS: ''
(0020, 0032) Image Position (Patient)      DS: [-145.500000, -158.199997, -356.200012]
(0020, 0037) Image Orientation (Patient)   DS: [1.000000, 0.000000, 0.000000, 0.000000, 1.000000, 0.000000]
(0020, 0052) Frame of Reference UID        UI: 2.25.83033509634441686385652073462983801840121916678417719669650
(0020, 1040) Position Reference Indicator  LO: 'SN'
(0020, 1041) Slice Location                DS: "-356.200012"
(0028, 0002) Samples per Pixel             US: 1
(0028, 0004) Photometric Interpretation    CS: 'MONOCHROME2'
(0028, 0010) Rows                          US: 512
(0028, 0011) Columns                       US: 512
(0028, 0030) Pixel Spacing                 DS: [0.597656, 0.597656]
(0028, 0100) Bits Allocated                US: 16
(0028, 0101) Bits Stored                   US: 16
(0028, 0102) High Bit                      US: 15
(0028, 0103) Pixel Representation          US: 1
(0028, 0120) Pixel Padding Value           US: 63536
(0028, 1050) Window Center                 DS: "40.0"
(0028, 1051) Window Width                  DS: "400.0"
(0028, 1052) Rescale Intercept             DS: "-1024.0"
(0028, 1053) Rescale Slope                 DS: "1.0"
(7fe0, 0010) Pixel Data                    OW: Array of 524288 elements
-----

```

Figure 4. The metadata is stored within a solitary DICOM file

The Watershed algorithm floods the basins of various markers until they intersect on watershed lines. Often, these markers have been selected as local minima in an image, and the basins are filled accordingly. This semantic segmentation approach helps emphasize the lung area and generate binary masks. To begin, external and internal markers from the images of CT scans were obtained using binary dilations and then added to a full dark image using watershed methods. The watershed marker eliminates image noise and detects cancerous cells in the lungs. The removal of external noise from the images is illustrated

in Figure 5, where a binary mask is applied to the image, with black pixels indicating cancer cells. An integrated Sobel filter and watershed algorithm were used to improve segmentation to remove the outer lung layers. An internal marker is then used to create a lung filter using Numpy's bitwise operations, which extracts the heart from the CT scan images. To ensure more accurate segmentation than previous methods, morphological operations, and gradients are used to complete the lung filter. Figure 6 displays the segmented lung after applying the Sobel filter application. The process created 2388 labeled images comprising almost 12 patients with CT scan data, with a comparable number of cancerous and non-cancerous patients.

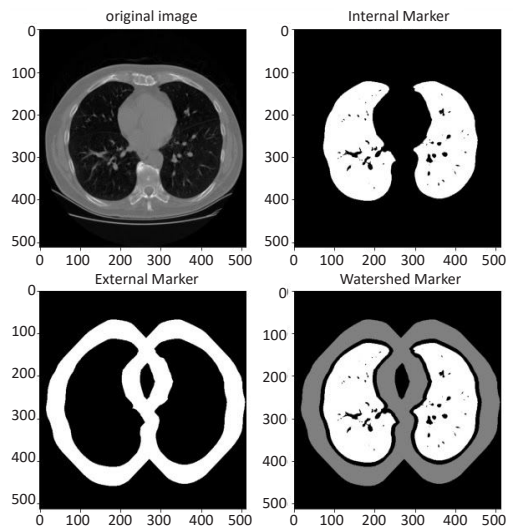


Figure 5. Distinct markers are derived from the watershed algorithm applied to the CT scan

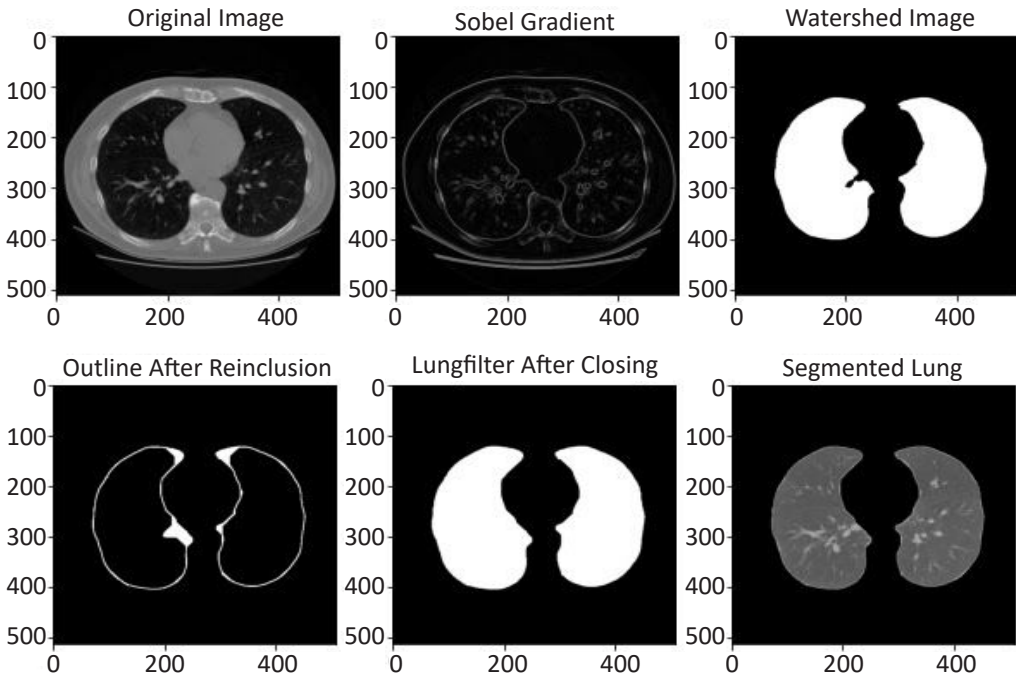


Figure 6. Visualization of lung segmentation



The model that was suggested is a combination of two powerful techniques: CNN with transfer learning and Random Forest. In order to contrast the performance of both models, a standard CNN was built as the first model, which uses convolution layers, flattened fully connected layers, max pooling, and dropout in the middle layers. Figure 7 illustrates this basic approach without the advantage of transfer learning. The second hybrid model applies transfer learning on a pre-trained VGG-16 model and Random Forest as a classifier. The last three layers were modified to incorporate the Random Forest classifier. The base estimator used in this hybrid model is 3, with a random state 42. Figure 8 displays the illustration of the suggested model. Subsequently, the dual models, one with Transfer Learning and the other with Random Forest, were implemented and trained on segmented lungs. A batch size of 64 was used for the image data generator, and 20 images were used in each epoch for 30 epochs, except for the first model, which had 20 images in each epoch. The shape of the training images for the first model was (32,32,3) while the proposed model had a shape of (128,128,3). The application of augmentation techniques facilitated the training of models on various augmentations, including shear range, zoom range, horizontal flip, rotation range, and center shift, thereby achieving superior outcomes. In the final layer, a solitary binary classification node distinguishes between cancerous and non-cancerous lungs. Additionally, TensorFlow Keras callbacks were utilized to preserve the most accurate model and execute a comprehensive 30-epoch training session to chart the comparative graphs.

Table 2 compares the proposed hybrid model's testing accuracy and Area Under the Curve (AUC) and the standard CNN without transfer learning. The findings display that the proposed hybrid model beats the standard model and achieves better test accuracy and AUC. The proposed model is considered preferable over the standard model despite the fact

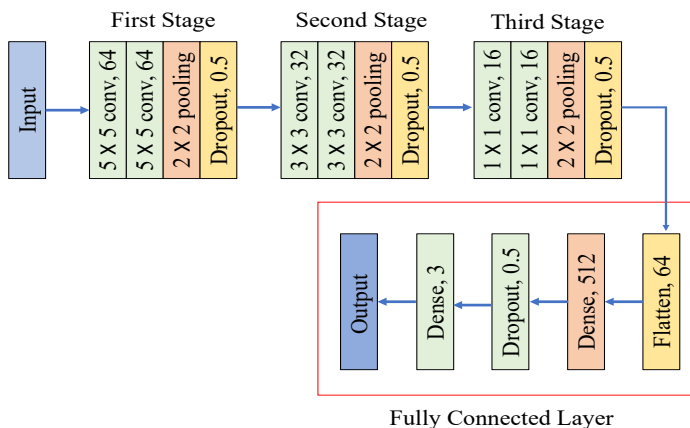


Figure 7. Standard CNN without transfer learning



that the standard model may still be suitable for solving classification tasks. Furthermore, Figure 9(a) shows that the proposed model attained an AUC of 0.985, better than the original model's AUC of 0.983, as shown in Figure 9(b).

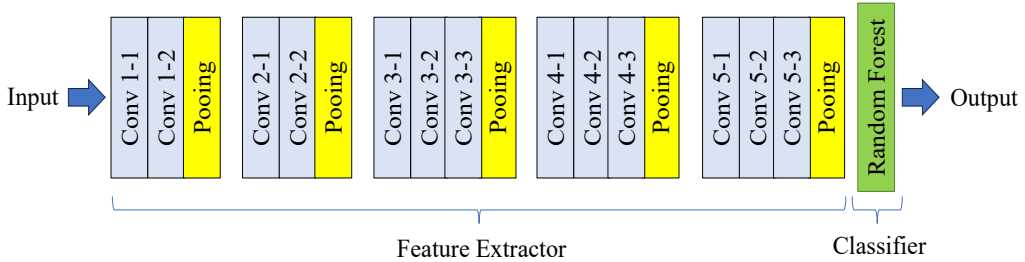


Figure 8. Hybrid CNN with transfer learning and Random Forest

Table 2  
Overview of the accuracy performance of both models

Models	Testing Accuracy	Area Under the Curve (AUC)
Standard Model	93.72%	0.983
Proposed Model (Hybrid CNN with Transfer Learning and Random Forest)	98.53%	0.985

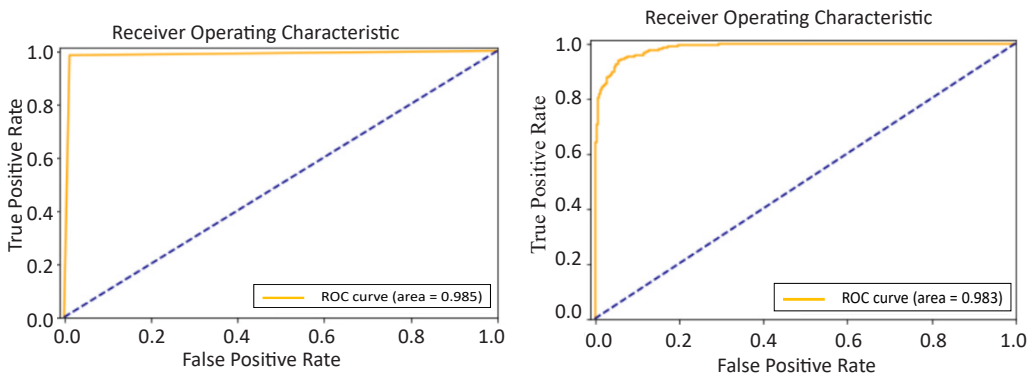


Figure 9. Receiver Operating Characteristic (ROC) Curve Performance (a) Proposed Hybrid Model (b) Standard Model

In terms of the confusion matrix, the hybrid model proposed in this study outperformed the typical CNN model without transfer learning in terms of accuracy. All of the CT scan images are classified correctly for the proposed hybrid model, as shown in Figure 10(a), while some of the CT scan images are still classified incorrectly for the standard CNN model, as shown in Figure 10(b). Based on Table 3, the performance metrics of CT image classification through the confusion matrix, as presented in Figure 10, proved that

the proposed hybrid model is better than the standard model. This confusion matrix and performance measures show that both models can classify most CT scan images. It can be proven in Figures 11(a) and 11(b), which are some CT scan images predicted correctly based on the standard and suggested models.

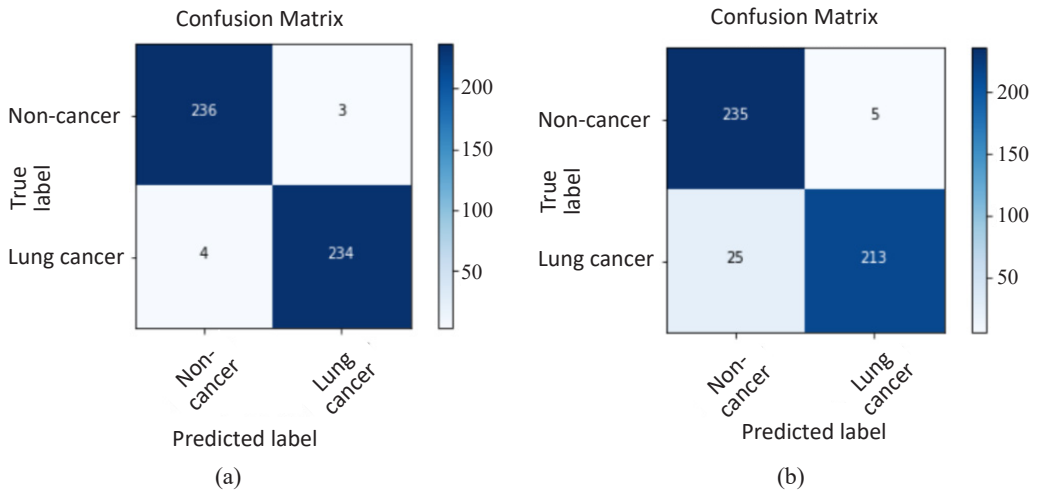


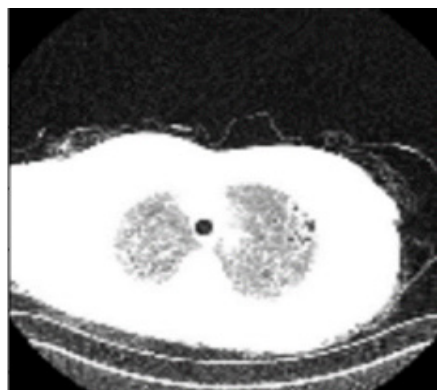
Figure 10. Confusion Matrix Performance (a) Proposed Hybrid Model (b) Standard Model

Table 3  
Performance measure

Models	Sensitivity	Specificity	F1-Score
Standard Model	97.92%	89.54%	94.00%
Proposed Model (Hybrid CNN with Transfer Learning and Random Forest)	98.74%	98.32%	98.54%



(a)



(b)

Figure 11. Prediction and actual label are (a) non-cancer, (b) lung cancer

A comparative study in Table 4 shows that the proposed CNN with transfer learning and Random Forest model performs better than the existing techniques. Only work from (Saleh et al., 2021) performed better than the proposed hybrid work in specificity and AUC. During study development, several challenges arose that significantly impacted the results and outcomes. One of the biggest challenges was the dataset format used, which was in dmc format as opposed to ordinary formats of image processing such as jpg. The DMC format saves information about patients and a 3D CT scan image, so it was necessary to convert the images to jpg format before feeding them into the models for training and testing. It made the research more labor-intensive and time-consuming.

Table 4  
*Comparative study*

Models	Testing Accuracy	Sensitivity	Specificity	F1-Score	AUC
3D U-Net CNN (Tang et al., 2020)	96.80%	92.40%	94.60%	-	0.941
3D CNN (Ali et al., 2018)	64.40%	58.90%	55.30%	-	-
3D U-Net and 3D DenseNet (Qin et al., 2018)	-	96.70%	-	-	-
Deformable CNN (Haiying et al., 2021)	-	95.80%	-	-	-
CNN (Sheng et al., 2021)	90.00%	-	-	-	-
AlexNet CNN (Agarwal et al., 2021)	96.00%	-	-	-	-
Deep Neural Networks Ensemble (Ardimento et al., 2021)	96.49%	98.73%	-	-	-
CNN-SVM (Saleh et al., 2021)	97.91%	97.90%	<b>99.32%</b>	-	<b>1.000</b>
Proposed Hybrid Model	<b>98.53%</b>	<b>98.74%</b>	98.32%	<b>98.54%</b>	0.985

## CONCLUSION

Deep learning is crucial in acquiring profound knowledge and aiding medical professionals in comprehending a patient's condition, significantly enhancing their quality of life. The medical field has seen an increasing adoption of machine learning and deep neural networks due to their ability to improve detection and classification, ultimately benefiting patients. Convolutional Neural Networks (CNNs) have gained widespread popularity

in medical imaging applications, especially for CT and MRI scan analysis, recognition, and classification tasks. However, achieving high accuracy and performance with CNN algorithms requires extensive training using large datasets.

A hybrid CNN employing transfer learning and random forest techniques has been developed to improve lung nodule classification. This hybridization has shown promise in enhancing image classification. In the future, the hybrid algorithm will be applied to various medical imaging fields. This technique can benefit from a wide range of imaging modalities for improving image classification, detection, and segmentation.

## ACKNOWLEDGEMENTS

The authors acknowledged the financial support from the Ministry of Higher Education Malaysia through the Fundamental Research Grant Scheme (FRGS) FRGS/1/2021/SS0/UNIMAS/02/4. Moreover, we thank the Faculty of Cognitive Sciences and Human Development, Universiti Malaysia Sarawak (UNIMAS), for their support and funding of the publication.

## REFERENCES

- Agarwal, A., Patni, K., & Rajeswari, D. (2021, July 8-10). *Lung cancer detection and classification based on Alexnet CNN*. [Paper presentation]. International Conference on Communication and Electronics System (ICCES 2021), Coimbatre, India. <https://doi.org/10.1109/ICCES51350.2021.9489033>
- Ali, I., Hart, G. R., Gunabushanam, G., Liang, Y., Muhammad, W., Nartowt, B., Kane, M., Ma, X., & Deng, J. (2018). Lung nodule detection via deep reinforcement learning. *Frontiers in Oncology*, 8, 108. <https://doi.org/10.3389/fonc.2018.00108>
- Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan, M., Essen, B. C. V., Awwal, A. A. S., & Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 292. <https://doi.org/10.3390/electronics8030292>
- Anderson, J., Rainie, L., & Luchsinger, A. (2018, December 10). *Artificial intelligence and the future of humans*. Pew Research Center. <chrome-extension://efaidnbmnnnibpajpglecfindmkaj/http://tony-silva.com/eslefl/miscstudent/downloadpagearticles/AIhumanfuture-pew.pdf>
- Arabahmadi, M., Farahbakhsh, R., & Rezazadeh, J. (2022). Deep learning for smart healthcare—A survey on brain tumor detection from medical imaging. *Sensors*, 22(5), 1960. <https://doi.org/10.3390/s22051960>
- Ardimento, P., Aversano, L., Bernardi, M. L., & Cimitile, M. (2021, July 18-22). *Deep neural networks ensemble for lung nodule detection on chest CT scans*. [Paper presentation]. International Joint Conference on Neural Networks (IJCNN 2021), Shenzhen, China. <https://doi.org/10.1109/IJCNN52387.2021.9534176>
- Azizah, A., Hashimah, B., Nirmal, K., Siti Zubaidah, A., Puteri, N., Nabihah, A., Sukumaran, R., Balqis, B., Nadia, S. M. R., Sharifah, S. S. S., Rahayu, O., Nur Alham, O., & Azlina, A. A. (2019). *Malaysia national cancer registry report (Report 2012-2016)*. National Cancer Registry. [chrome-extension://efaidnbmnnnibpajpglecfindmkaj/https://www.moh.gov.my/moh/resources/Penerbitan/LaporanUmum/2012-2016%20\(MNCRR\)/MNCRR\\_2012-2016\\_FINAL\\_\(PUBLISHED\\_2019\).pdf](chrome-extension://efaidnbmnnnibpajpglecfindmkaj/https://www.moh.gov.my/moh/resources/Penerbitan/LaporanUmum/2012-2016%20(MNCRR)/MNCRR_2012-2016_FINAL_(PUBLISHED_2019).pdf)

- Bertrand, G. (2005). On topological watersheds. *Journal of Mathematical Imaging and Vision*, 22(2), 217-230. <https://doi.org/10.1007/s10851-005-4891-5>
- Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians*, 68(6), 394-424. <https://doi.org/10.3322/caac.21492>
- Camp, B. V. B. (2022). *Data from The Lung Image Database Consortium (LIDC) and Image Database Resource Initiative (IDRI): A completed reference database of lung nodules on CT scans (LIDC-IDRI)*. The Cancer Imaging Archive (TCIA) Public Access. <https://pubmed.ncbi.nlm.nih.gov/21452728/>
- Chin, C. K., Mat, D. A. A., & Saleh, A. Y. (2021a). Hybrid of convolutional neural network algorithm and autoregressive integrated moving average model for skin cancer classification among Malaysian. *IAES International Journal of Artificial Intelligence*, 10(3), 707-716. <https://doi.org/10.11591/ijai.v10.i3.pp707-716>
- Chin, C. K., Mat, D. A. A., & Saleh, A. Y. (2021b, April 9-11). *Skin cancer classification using convolutional neural network with autoregressive integrated moving average*. [Paper presentation]. International Conference on Robot Systems and Applications (ICRSA 2021), Chengdu, China. <https://doi.org/10.1145/3467691.3467693>
- Da Nóbrega, R. V. M., Peixoto, S. A., da Silva, S. P. P., & Rebouças Filho, P. P. (2018, June 18-21). *Lung nodule classification via deep transfer learning in CT lung images*. [Paper presentation]. International Symposium on Computer-based Medical Systems (CBMS 2018), Karlstad, Sweden. <https://doi.org/10.1109/CBMS.2018.00050>
- Forte, G. C., Altmayer, S., Silva, R. F., Stefani, M. T., Libermann, L. L., Cavion, C. C., Youssef, A., Forghani, R., King, J., Mohamed, T. L., Andrade, R. G. F., & Hochegger, B. (2022). Deep learning algorithms for diagnosis of lung cancer: A systematic review and meta-analysis. *Cancers*, 14(16), 3856. <https://doi.org/10.3390/cancers14163856>
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Haiying, Y., Zhongwei, F., Ding, D., & Zengyang, S. (2021, May 25-27). *False-positive reduction of pulmonary nodule detection based on deformable convolutional neural networks*. [Paper presentation]. International Conference on Bioinformatics and Computational Biology (ICBCB 2021), Taiyuan, China. <https://doi.org/10.1109/ICBCB52223.2021.9459209>
- Halder, A., Dey, D., & Sadhu, A. K. (2020). Lung nodule detection from feature engineering to deep learning in thoracic CT images: A comprehensive review. *Journal of Digital Imaging*, 33(3), 655-677. <https://doi.org/10.1007/s10278-020-00320-6>
- Keshari, R., Vatsa, M., Singh, R., & Noore, A. (2018, June 18-23). *Learning structure and strength of CNN filters for small sample size training*. [Paper presentation]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2018), Utah, USA.
- Li, Z., Yao, H., & Ma, F. (2020, February 3-7). *Learning with small data*. [Paper presentation]. International Conference on Web Search and Data Mining (WSDM 2020), New York, USA.
- Nakrani, M. G., Sable, G. S., & Shinde, U. B. (2021). A comprehensive review on deep learning based lung nodule detection in computed tomography images. In S. C. Satapathy, V. Bhateja, B. & Janakiramaiah, Y. W. Chen (Eds.) *Intelligent system design* (pp.107-116). Springer Link. [https://doi.org/10.1007/978-981-15-5400-1\\_12](https://doi.org/10.1007/978-981-15-5400-1_12)

- Peirelinck, T., Kazmi, H., Mbuwir, B. V., Hermans, C., Spiessens, F., Suykens, J., & Deconinck, G. (2022). Transfer learning in demand response: A review of algorithms for data-efficient modelling and control. *Energy and AI*, 7, 100126. <https://doi.org/10.1016/j.egyai.2021.100126>
- Primakov, S. P., Ibrahim, A., van Timmeren, J. E., Wu, G., Keek, S. A., Beuque, M., Granzier, R. W. Y., Lavrova, E., Scrivener, M., Sanduleanu, S., Kayan, E., Halilaj, I., Lenaers, A., Wu, J., Monshouwer, R., Geets, X., Gietema, H. A., Hendriks, L. E. L., Morin, O., ... & Lambin, P. (2022). Automated detection and segmentation of non-small cell lung cancer computed tomography images. *Nature Communications*, 13(1), 3423. <https://doi.org/10.1038/s41467-022-30841-3>
- Qin, Y., Zheng, H., Zhu, Y. M., & Yang, J. (2018, April 15-20). *Simultaneous accurate detection of pulmonary nodules and false positive reduction using 3D CNNs*. [Paper presentation]. International Conference on Acoustics, Speech and Signal Processing (ICASSP 2018), Alberta, Canada. <https://doi.org/10.1109/ICASSP.2018.8462546>
- Rajadurai, P., How, S. H., Liam, C. K., Sachithanandan, A., Soon, S. Y., & Tho, L. M. (2020). Lung cancer in Malaysia. *Journal of Thoracic Oncology*, 15(3), 317–323. <https://doi.org/10.1016/j.jtho.2019.10.021>
- Salahuddin, Z., Woodruff, H. C., Chatterjee, A., & Lambin, P. (2022). Transparency of deep neural networks for medical image analysis: A review of interpretability methods. *Computers in Biology and Medicine*, 140, 105111. <https://doi.org/10.1016/j.combiomed.2021.105111>
- Saleh, A. Y., Chin, C. K., Penshie, V., & Al-Absi, H. R. H. (2021). Lung cancer medical images classification using hybrid CNN-SVM. *International Journal Advanced in Intelligence Information*, 7(2), 151-162. <https://doi.org/10.26555/ijain.v7i2.317>
- Sharif, M. I., Li, J. P., Naz, J., & Rashid, I. (2020). A comprehensive review on multi-organs tumor detection based on machine learning. *Pattern Recognition Letters*, 131, 30-37. <https://doi.org/10.1016/j.patrec.2019.12.006>
- Sheng, J., Li, Y., Cao, G., & Hou, K. (2021, July 18-22). *Modeling nodule growth via spatial transformation for follow-up prediction and diagnosis*. [Paper presentation]. International Joint Conference on Neural Networks (IJCNN 2021), Shenzhen, China. <https://doi.org/10.1109/IJCNN52387.2021.9534163>
- Singh, G. A .P., & Gupta, P. K. (2019). Performance analysis of various machine learning-based approaches for detection and classification of lung cancer in humans. *Neural Computing and Applications*, 31(10), 6863–6877. <https://doi.org/10.1007/s00521-018-3518-x>
- Tang, H., Kim, D. R., & Xie, X. (2018, April 4-7). *Automated pulmonary nodule detection using 3D deep convolutional neural networks*. [Paper presentation]. International Symposium on Biomedical Imaging (ISBI 2018), Washington, USA. <https://doi.org/10.1109/ISBI.2018.8363630>
- Tang, S., Yang, M., & Bai, J. (2020). Detection of pulmonary nodules based on a multiscale feature 3D U-Net convolutional neural network of transfer learning. *PLoS One*, 15(8), e0235672. <https://doi.org/10.1371/journal.pone.0235672>
- Thai, A. A., Solomon, B. J., Sequist, L. V., Gainor, J. F., & Heist, R. S. (2021). Lung cancer. *The Lancet*, 398(10299), 535–554. [https://doi.org/10.1016/S0140-6736\(21\)00312-3](https://doi.org/10.1016/S0140-6736(21)00312-3)
- Yamashita, R., Nishio, M., Do, R. K. G., & Togashi, K. (2018). Convolutional neural networks: An overview and application in radiology. *Insights Into Imaging*, 9(4), 611–629. <https://doi.org/10.1007/s13244-018-0639-9>

- Zakaria, R., Abdelmajid, H., & Zitouni, D. (2022). Deep learning in medical imaging: A review. In J. K. Mandal, S. Misra, J. S. Banerjee & S. Nayak (Eds.) *Application of machine intelligence in engineering* (pp.131-144). CRC Press. <https://doi.org/10.1201/9781003269793-15>
- Zhao, X., Liu, L., Qi, S., Teng, Y., Li, J., & Qian, W. (2018). Agile convolutional neural network for pulmonary nodule classification using CT images. *International Journal of Computer Assisted Radiology and Surgery*, 13(4), 585-595. <https://doi.org/10.1007/s11548-017-1696-0>
- Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., ... & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, 109(1), 43-76. <https://doi.org/10.1109/JPROC.2020.3004555>





# REFEREES FOR THE PERTANIKA JOURNAL OF SCIENCE & TECHNOLOGY

Vol. 32 (1) Jan. 2024

The Editorial Board of the Pertanika Journal of Science and Technology wishes to thank the following:

Ab Al-Hadi Ab Rahman  
*(UTM, Malaysia)*

Abdul Hakeem Memon  
*(University of Sindh, Pakistan)*

Abdullah Hisam Omar  
*(UTM, Malaysia)*

Ahmad Farid Abu Bakar  
*(UM, Malaysia)*

Ali H. Jawad  
*(UiTM, Malaysia)*

Amirul Ridzuan Abu Bakar  
*(UNIMAP, Malaysia)*

Ananthu Retnam  
*(Department of Chemistry, Malaysia)*

Anazida Zainal  
*(UTM, Malaysia)*

Ateeq Ur Rehman  
*(FUI, Pakistan)*

Aytuğ Onan  
*(IKCU, Turkey)*

Azlin Ahmad  
*(UiTM, Malaysia)*

Azni Zain Ahmed  
*(UiTM, Malaysia)*

Bentajer Ahmed  
*(Cadi Ayyad University, Morocco)*

Chen Shaohua  
*(SCAU, China)*

Chew Xinying  
*(USM, Malaysia)*

Chin Kim On  
*(UMS, Malaysia)*

Chua Han Bing  
*(Curtin University, Malaysia)*

Emienour Muzalina Mustafa  
*(UMT, Malaysia)*

Erzam Marlisah  
*(UPM, Malaysia)*

Hazlee Azil Illias  
*(UM, Malaysia)*

Hazlina Hamdan  
*(UPM, Malaysia)*

Jamal Khatib  
*(Beirut Arab University, Lebanon)*

Jusoh Ismail  
*(UNIMAS, Malaysia)*

Kwong Qi Jie  
*(UiTM, Malaysia)*

Mailin Mission  
*(UMS, Malaysia)*

Mansoorian Hossein Jafari  
*(UMSHA, Iran)*

Mimi Hani Abu Bakar  
*(UKM, Malaysia)*

Mohamed Ansari Nainar  
*(UNITEN, Malaysia)*

Mohammad Alibakhshikenari  
*(UC3M, Spain)*

Mohammad Ahmed Alomari  
*(UNITEN, Malaysia)*

Mohammed Fouad El Basuni  
*(Tanta University, Egypt)*

Mohd Dilshad Ansari  
*(CMRCET, India)*

Mohd Firrdhaus Mohd Sahabuddin  
*(UM, Malaysia)*

Mohd Rosdzimin Abdul Rahman  
*(UPNM, Malaysia)*

Muhammad Ahsan  
(Institut Teknologi Sepuluh Nopember, Indonesia)

Muhd Ridzuan Mansor  
(UTeM, Malaysia)

Nor Hanuni Ramli  
(UMP, Malaysia)

Normala Halimoon  
(UPM, Malaysia)

Nur Novilina Arifianingsih  
(ITB, Indonesia)

Pavol Bozek  
(STU, Slovakia)

Renato Procopio  
(University of Genoa, Italy)

Retantyo Wardoyo  
(Universitas Gadjah Mada, Indonesia)

Rohana Hassan  
(UiTM, Malaysia)

Roya Abedi  
(University of Tabriz, Iran)

Saka Cafer  
(Siirt Universitesi, Turkey)

Sapto Purnomo Putro  
(Diponegoro University, Indonesia)

Shreshivadasan Chelliapan  
(UTM, Malaysia)

Siti Hajar Othman  
(UTM, Malaysia)

Sofianita Mutalib  
(UiTM, Malaysia)

Sri Hartati  
(Universitas Gadjah Mada, Indonesia)

Syazwani Idrus  
(UPM, Malaysia)

Syerina Azlin Md Nasir  
(UiTM, Malaysia)

Uswah Khairuddin  
(UTM, Malaysia)

Wan Fairos Wan Yaacob  
(UiTM, Malaysia)

Wan Nor Shuhadah Wan Nik  
(UniSZA, Malaysia)

Wan Zuraida Wan Mohd Zain  
(UiTM, Malaysia)

Yudong Zhang  
(University of Leicester, UK)

Zainah Md. Zaim  
(UMP, Malaysia)

---

CMRCET – CMR College of Engineering and Technology  
FUI – Foundation University Islamabad  
IKCU – İzmir Katip Çelebi University  
ITB – Bandung Institute of Technology  
SCAU – South China Agricultural University  
STU – Slovak University of Technology  
UC3M – Universidad Carlos III de Madrid  
UiTM – Universiti Teknologi MARA  
UKM – Universiti Kebangsaan Malaysia  
UM – Universiti Malaya  
UMP – Universiti Malaysia Pahang  
UMS – Universiti Malaysia Sabah

UMSHA – Hamedan University of Medical Sciences  
UMT – Universiti Malaysia Terengganu  
UNIMAP – Universiti Malaysia Perlis  
UNIMAS – Universiti Malaysia Sarawak  
UniSZA – Universiti Sultan Zainal Abidin  
UNITEN – Universiti Tenaga Nasional  
UPM – Universiti Putra Malaysia  
UPNM – Universiti Pertahanan Nasional Malaysia  
USM – Universiti Sains Malaysia  
UTeM – Universiti Teknikal Malaysia Melaka  
UTM – Universiti Teknologi Malaysia

---

While every effort has been made to include a complete list of referees for the period stated above, however if any name(s) have been omitted unintentionally or spelt incorrectly, please notify the Chief Executive Editor, *Pertanika* Journals at [executive\\_editor.pertanika@upm.edu.my](mailto:executive_editor.pertanika@upm.edu.my)

Any inclusion or exclusion of name(s) on this page does not commit the *Pertanika* Editorial Office, nor the UPM Press or the university to provide any liability for whatsoever reason.

**Selected papers from the International Conference on Data Science  
and Emerging Technologies (2022)**

**Guest Editors:** Yap Bee Wah and Simon Fong

- Improved GCC Technique: A Comprehensive Approach to Color  
Cast Rectification and Image Enhancement 453  
*Danny Ngo Lung Yao, Abdullah Bade, Iznora Aini Zolkifly and  
Paridah Daud*
- Transfer Learning for Lung Nodules Classification with CNN and  
Random Forest 463  
*Abdulrazak Yahya Saleh, Chee Ka Chin and Ros Ameera Rosdi*

Soft Set-based Parameter Reduction Algorithm Through a Discernibility Matrix and the Hybrid Approach for the Risk-Factor Prediction of Cardiovascular Diseases by Various Machine Learning Techniques <i>Menaga Anbumani and Kannan Kaniyaiah</i>	265
<i>Review Article</i>	
Understanding the Degradation of Carbofuran in Agricultural Area: A Review of Fate, Metabolites, and Toxicity <i>Nurul Syuhada Baharudin, Harlina Ahmad and Md Sohrab Hossain</i>	285
Phenolics-Enhancing <i>Piper sarmentosum</i> (Roxburgh) Extracts Pre-Treated with Supercritical Carbon Dioxide and its Correlation with Cytotoxicity and $\alpha$ -Glucosidase Inhibitory Activities <i>Mohd Shahrul Ridzuan Hamil, Fauziahanim Zakaria, Lee Chong Yew, Norzilawati Pauzi, Khamsah Suryati Mohd and Noor Hafizoh Saidan</i>	323
Estimation of Leachate Volume and Treatment Cost Avoidance Through Waste Segregation Programme in Malaysia <i>Josfirin Uding Rangga, Sharifah Norkhadijah Syed Ismail, Irniza Rasdi and Karmegam Karuppiah</i>	339
Comparison Using Intelligent Systems for Data Prediction and Near Miss Detection Techniques <i>Lek Ming Lim, Saratha Sathasivam, Mohd. Tahir Ismail, Ahmad Sufriil Azlan Mohamed, Olayemi Joshua Ibidoja and Majid Khan Majahar Ali</i>	365
Development of GIS-based Ground Flash Density and its Statistical Analysis for Lightning Performance Evaluation of Transmission Lines in Peninsular Malaysia <i>Nurzanariah Roslan, Ungku Anisa Ungku Amirulddin, Mohd Zainal Abidin Ab. Kadir and Noradlina Abdullah</i>	395
Preparation of Activated Carbon from Sugarcane Bagasse Using Microwave-assisted $ZnCl_2$ Chemical Activation: Optimization and Characterization Study <i>Atiqa Rahmawati, Fadzkurisma Robbika and Yuafni</i>	419
Conceptual Design and Selection of Natural Fibre Reinforced Composite Cyclist Helmet Liner Using an Integrated Approach <i>Nurul Ain Maidin, Mohd Sapuan Salit, Mastura Mohammad Taha and Mohd Zuhri Mohamed Yusoff</i>	437

<i>Review Article</i>	
Chromatography and Spectroscopy Methods for the Analysis of Nicotine and Other Chemical Ingredients in E-Liquid Formulation: A Review <i>Mohd Rashidi Abdull Manap, Noor Hazfalinda Hamzah, Qhurratul Aina Kholili, Fatin Abu Hasan and Azhana Alhumaira</i>	127
The Mechanical Performance of Polymer Concrete Incorporating Waste Tin Fibres <i>Hosseinkhah Ashkan, Shuhairy Norhisham, Mohd Supian Abu Bakar, Agusril Syamsir, Mohammed Jalal Abdullah, Muhammad Rizal Muhammad Asyraf, Maiyozi Chairi, Mutiara Yetrina, Jihan Melasari and Muhammad Farid</i>	143
The Mechanical Performance of Polymer Concrete Incorporating Waste Tin Fibres <i>Hosseinkhah Ashkan, Shuhairy Norhisham, Mohd Supian Abu Bakar, Agusril Syamsir, Mohammed Jalal Abdullah, Muhammad Rizal Muhammad Asyraf, Maiyozi Chairi, Mutiara Yetrina, Jihan Melasari and Muhammad Farid</i>	161
Proportions of Green Area and Tree Health on University Campus: The Impact of Pavement Presence <i>Nur Ain Aiman Abd Rahim, Rahmad Zakaria, Asyraf Mansor, Mohd Ashraf Mohamad Ismail and Nik Fadzly N Rosely</i>	185
The Effect of Nutrients in Anodic Chamber to the Performance of Microbial Fuel Cell (MFC) <i>Nashley Ursula Mundi Ujai, Siti Kudnie Sahari, Marini Sawawi, Kuryati Kipli, Asmahani Awang, Mohamad Rusop Mahmood, Lilik Hasanah, Abdul Rahman Kram and Zainab Ngaini</i>	205
Investigation of Blended Seaweed Waste Recycling Using Black Soldier Fly Larvae <i>Iva Yeniseptiariva, I Wayan Koko Suryawan, Wisnu Prayogo, Sapta Suhardono and Ariyanti Sarwono</i>	217
An Attribute-based Data Privacy Classification Through the Bayesian Theorem to Raise Awareness in Public Data Sharing Activity <i>Nur Aziana Azwani Abdul Aziz, Masnida Hussin and Nur Raidah Salim</i>	235
High-performance THz Metallic Axial Mode Helix Antenna with Optimised Truncated Hollow Cone Ground Plane for 6G Wireless Communication System <i>Zahraa Raad Mayoof Hajiyat, Alyani Ismail, Aduwati Sali and Mohd. Nizar Hamidon</i>	249

# Pertanika Journal of Science & Technology

## Vol. 32 (1) Jan. 2024

### Content

Foreword <i>Mohd Sapuan Salit</i>	i
Physicochemical Impacts on Bacterial Communities in Putrajaya Lake, Malaysia <i>Nurul Najwa Fariah Mat Lazim, Afiqah Mohamed, Zana Ruhaizat Zana Rudin, Fatimah Md Yusoff, Ikhsan Natrah and Shahrizim Zulkifly</i>	1
Towards Maximising Hardware Resources and Design Efficiency via High-Speed Implementation of HMAC based on SHA-256 Design <i>Shamsiah Suhaili, Norhuzaimin Julai, Rohana Sapawi and Nordiana Rajae</i>	31
Effect of Natural Ventilation on Thermal Performance of Different Residential Building Forms in the Hot-dry Climate of Jordan <i>Esraa Shehadeh Abbaas, Mazran Ismail, Ala'eddin Ahmad Saif and Muhamad Azhar Ghazali</i>	45
Hydrocyanic Acid, Protein Concentration, and Phytochemical Compounds of Pulut and White Varieties in Young and Matured Cassava ( <i>Manihot Esculenta</i> , Crantz) <i>Sarah Idris, Rosnah Shamsudin, Mohd Zuhair Mohd Nor, Mohd Noriznan Mokhtar and Siti Salwa Abd Gani</i>	67
Attention-based Spatialized Word Embedding Bi-LSTM Model for Sentiment Analysis <i>Kun Zhu and Nur Hana Samsudin</i>	79
Analysis of Pure-pursuit Algorithm Parameters for Nonholonomic Mobile Robot Navigation in Unstructured and Confined Space <i>Izzati Saleh, Azwati Azmin, Azan Yunus and Wan Rahiman</i>	99
Use of Enhanced Greedy Algorithm for Load Balancing in Cloud Computing <i>Hanaa Osman, Asma'a Yassin Hammo and Abdulnasir Younus Ahmad</i>	113



Pertanika Editorial Office, Journal Division,  
Putra Science Park,  
1st Floor, IDEA Tower II,  
UPM-MTDC Center,  
Universiti Putra Malaysia,  
43400 UPM Serdang,  
Selangor Darul Ehsan  
Malaysia

<http://www.pertanika.upm.edu.my>  
Email: [executive\\_editor.pertanika@upm.edu.my](mailto:executive_editor.pertanika@upm.edu.my)  
Tel. No.: +603- 9769 1622

**PENERBIT**  
**UPM**  
UNIVERSITI PUTRA MALAYSIA  
**P R E S S**

<http://www.penerbit.upm.edu.my>  
Email: [penerbit@upm.edu.my](mailto:penerbit@upm.edu.my)  
Tel. No.: +603- 9769 8851

