

Pertanika Journal of
**SCIENCE &
TECHNOLOGY**

JST

VOL. 29 (2) APR. 2021



A scientific journal published by Universiti Putra Malaysia Press

Pertanika Journal of Science & Technology

About the Journal

Overview

Pertanika Journal of Science & Technology (PJST) is the official journal of Universiti Putra Malaysia published by UPM Press. It is an open-access online scientific journal which is free of charge. It publishes the scientific outputs. It neither accepts nor commissions third party content.

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

PJST is a **quarterly** (January, April, July and October) periodical that considers for publication original articles as per its scope. The journal publishes in **English** and it is open to authors around the world regardless of the nationality.

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 was founded in 1978. A decision was made in 1992 to streamline *Pertanika* into three journals as *Pertanika Journal of Tropical Agricultural Science*, *Pertanika Journal of Science & Technology*, and *Pertanika Journal of Social Sciences & Humanities* to meet the need for specialised journals in areas of study aligned with the interdisciplinary strengths of the university.

After almost 28 years, as an interdisciplinary Journal of Science & Technology, the journal now focuses on research in science and engineering and its related fields.

Goal of *Pertanika*

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 are guaranteed to receive a decision within 14 weeks. The elapsed time from submission to publication for the articles averages 5-6 months.

Abstracting and indexing of *Pertanika*

The journal is indexed in SCOPUS (Elsevier), Clarivate-Emerging Sources Citation Index [ESCI (Web of Science)], BIOSIS, National Agricultural Science (NAL), Google Scholar, MyCite and ISC.

Future vision

We are continuously improving access to our journal archives, content, and research services. We have the drive to realise exciting new horizons that will benefit not only the academic community, but society itself.

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 takes 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

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. All *Pertanika* journals have ISSN as well as an e-ISSN.

Pertanika Journal of Science & Technology: ISSN 0128-7680 (*Print*); ISSN 2231-8526 (*Online*).

Lag time

A decision on acceptance or rejection of a manuscript is reached in 3 to 4 months (average 14 weeks). The elapsed time from submission to publication for the articles averages 5-6 months.

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

Refer to *Pertanika's Instructions to Authors* at the back of this journal.

Editorial process

Authors 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 usually sent to reviewers. Authors are encouraged to suggest names of at least three potential reviewers at the time of submission of their manuscript to *Pertanika*, but the editors will make the final choice. The editors are not, however, bound by these suggestions.

Notification of the editorial decision is usually provided within ten to fourteen weeks 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, three referees independently evaluate the scientific quality of the submitted manuscripts.

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 seven steps to the editorial review process:

1. The Journal's Chief Executive Editor (CEE) and the Editorial Board Members (EBMs) examine the paper to determine whether it is appropriate for the journal and should be reviewed. If not appropriate, the manuscript is rejected outright and the author is informed.
2. The CEE sends the article-identifying information having been removed, to 2 or 3 reviewers who are specialists in the subject matter represented by the article. The CEE requests them to 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 (EiC) examines the review reports and decides whether to accept or reject the manuscript, invites the author(s) to revise and resubmit the manuscript, or seek additional review reports. Final acceptance or rejection rests with the CEE and EiC, who reserve the right to refuse any material for publication. In rare instances, the manuscript is accepted with almost no revision. Almost without exception, reviewers' comments (to the author) are forwarded to the author. If a revision is indicated, the editor provides guidelines to the authors 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 CEE along with specific information describing how they have answered' the concerns of the reviewers and the editor, usually in a tabular form. The author(s) may also submit a rebuttal if there is a need especially when the authors disagree with certain comments provided by reviewer(s).
5. The CEE sends the revised paper 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 EiC examines their comments and decides whether the paper is ready to be published, needs another round of revisions, or should be rejected. If the decision is to accept, the CEE is notified.
7. The CEE 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 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 online.

Pertanika Journal of

**SCIENCE
& TECHNOLOGY**

Vol. 29 (2) Apr. 2021



A scientific journal published by Universiti Putra Malaysia Press



EDITOR-IN-CHIEF

Luqman Chuah Abdullah

Chemical Engineering

CHIEF EXECUTIVE EDITOR

Mohammad Jawaid

Polymer Composites

UNIVERSITY PUBLICATIONS COMMITTEE

CHAIR

Zulkifli Idrus

EDITORIAL STAFF

Journal Officers:

Kanagamalar Silvarajoo, *ScholarOne*

Siti Zuhaila Abd Wahid, *ScholarOne*

Tee Syin-Ying, *ScholarOne*

Ummi Fairuz Hanapi, *ScholarOne*

Editorial Assistants:

Ku Ida Mastura Ku Baharom

Siti Juridah Mat Arip

Zulinaardawati Kamarudin

PRODUCTION STAFF

Pre-press Officers:

Nur Farrah Dila Ismail

Wong Lih Jiun

WEBMASTER

IT Officer:

Munir Hayat Md Bahrain

EDITORIAL OFFICE

JOURNAL DIVISION

Putra Science Park,
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 Serdang, Selangor, Malaysia

General Enquiry

Tel: +603 9769 1622 | 1616

E-mail:

executive_editor.pertanika@upm.edu.my

URL: www.journals-jd.upm.edu.my

PUBLISHER

UPM PRESS

Universiti Putra Malaysia
43400 UPM, Serdang, Selangor, Malaysia.

Tel: +603 9769 8855 | 8854

Fax: +603 9679 6172

E-mail: penerbit@upm.edu.my

URL: <http://penerbit.upm.edu.my>



EDITORIAL BOARD 2020-2022

Abdul Latif Ahmad

Chemical Engineering
Universiti Sains Malaysia, Malaysia

Adem Kilicman

Mathematical Sciences
Universiti Putra Malaysia, Malaysia

Ahmad Zaharin Aris

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

Azlina Harun@Kamaruddin

Enzyme Technology, Fermentation Technology
Universiti Sains Malaysia, Malaysia

Bassim H. Hameed

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

Biswajeet Pradhan

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

Daud Ahmad Israf Ali

Cell Biology, Biochemical, Pharmacology
Universiti Putra Malaysia, Malaysia

Hari M. Srivastava

Mathematics and Statistics
University of Victoria, Canada

Ho Yuh-Shan

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

Hsiu-Po Kuo

Chemical Engineering
National Taiwan University, Taiwan

Ivan D. Rukhlenko

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

Lee Keat Teong

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

Mohamed Othman

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

Mohammad Jawaid

Polymers and Plastics, Colloid and Surface Chemistry, Composite and Hybrid Materials, Chemical Engineering
Universiti Putra Malaysia, Malaysia

Mohd. Ali Hassan

Bioprocess Engineering, Environmental Biotechnology
Universiti Putra Malaysia, Malaysia

Mohd Sapuan Salit

Concurrent Engineering and Composite Materials
Universiti Putra Malaysia, Malaysia

Najafpour Darzi Ghasem

Bioprocess Technology, Chemical Engineering, Water and Wastewater Treatment Technology, Biochemical Engineering and Biotechnology, Bioethanol, Biofuel, Biohydrogen, Enzyme and Fermentation Technology
Babol Noshirvani University of Technology, Iran

Nor Azah Yusof

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

Norbahiah Misran

Communication Engineering
Universiti Kebangsaan Malaysia, Malaysia

Roslan Abd-Shukur

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

Wing Keong Ng

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

INTERNATIONAL ADVISORY BOARD 2018-2021

Adarsh Sandhu

Editorial Consultant for Nature Nanotechnology and Contributing Writer for Nature Photonics, Physics, Magnetoresistive Semiconducting Magnetic Field Sensors, Nano-BioMagnetism, Magnetic Particle Colloids, Point of Care Diagnostics, Medical Physics, Scanning Hall Probe Microscopy, Synthesis and Application of Graphene Electronics-Inspired Interdisciplinary
Research Institute (EIIRIS), Toyohashi University of Technology, Japan

Graham Megson

Computer Science
The University of Westminster, U.K

Kuan-Chong Ting

Agricultural and Biological Engineering
University of Illinois at UrbanaChampaign, USA

Malin Premaratne

Advanced Computing and Simulation
Monash University, Australia

Mohammed Ismail Elnaggar

Electrical Engineering
Ohio State University, USA

Peter J. Heggs

Chemical Engineering
University of Leeds, U.K

Ravi Prakash

Vice Chancellor, JUIT, Mechanical Engineering, Machine Design, Biomedical and Materials Science
Jaypee University of Information Technology, Indian

Said S.E.H. Elnashaie

Environmental and Sustainable Engineering
Penn. State University at Harrisburg, USA

Suhash Chandra Dutta Roy

Electrical Engineering
Indian Institute of Technology (IIT) Delhi, India

Vijay Arora

Quantum and Nano-Engineering Process
Wilkes University, USA

Yi Li

Chemistry, Photochemical Studies, Organic Compounds, Chemical Engineering
Chinese Academy of Sciences, Beijing, China

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 © 2018-19 Universiti Putra Malaysia Press. All Rights Reserved.



Pertanika Journal of Science & Technology
Vol. 29 (2) Apr. 2021

Contents

Foreword	i
<i>Mohammad Jawaid</i>	
Engineering Sciences	
High Level Synthesis Optimizations of Road Lane Detection Development on Zynq-7000	707
<i>Panadda Solod, Nattha Jindapetch, Kiattisak Sengchuai, Apidet Booranawong, Pakpoom Hoyingcharoen, Surachate Chumpol and Masami Ikura</i>	
Farmers' Participation in Irrigation Management in the Punggur Utara Irrigation Area, Indonesia	725
<i>Dyah Indriana Kusumastuti, Vera Chania Putri, Dwi Jokowinarso and Endro Prasetyo Wahono</i>	
Pseudo-colour with K-means Clustering Algorithm for Acute Ischemic Stroke Lesion Segmentation in Brain MRI	743
<i>Abang Mohd Arif Anaqi Abang Isa, Kuryati Kipli, Ahmad Tirmizi Jobli, Muhammad Hamdi Mahmood, Siti Kudnie Sahari, Aditya Tri Hernowo and Sinin Hamdan</i>	
Improved High Dynamic Range for 3D Shape Measurement based on Saturation of the Coloured Fringe	759
<i>Shanyu Chua, Chee Chin Lim, Swee Kheng Eng, Yen Fook Chong and Chiun Tai Loh</i>	
The Built Environment and its Impact on Transit based Transportation Users Walking Activity in Semarang, Indonesia	771
<i>Anita Ratnasari Rakhmatulloh, Diah Intan Kusumo Dewi and Dinar Mutiara Kusumo Nugraheni</i>	
Ethylene Yield from a Large Scale Naphtha Pyrolysis Cracking Utilizing Response Surface Methodology	791
<i>Mohamad Hafizi Zakria, Mohd Ghazali Mohd Nawawi and Mohd Rizal Abdul Rahman</i>	
Smartphone Application Development for Rice Field Management Through Aerial Imagery and Normalised Difference Vegetation Index (NDVI) Analysis	809
<i>Nor Athirah Roslin, Nik Norasma Che'Ya, Rhushalshafira Rosle and Mohd Razi Ismail</i>	

Information, Computer and Communication Technologies

Review

A Review Article on Software Effort Estimation in Agile Methodology 837
Pantjawati Sudarmaningtyas and Rozlina Mohamed

Reducing Cognitive Impairment Among Dementia Users Through Mobile Application 863
Nur Atheera Mohd Hassan, Aslina Baharum, Zaidatul Haslinda Abdullah Sani, Kent Chau and Noorsidi Aizuddin Mat Noor

Photonics Rib Waveguide Dimension Dependent Charge Distribution and Loss Characterization 885
Angie Teo Chen Chen, Mohammad Rakib Uddin and Foo Kui Law

Optimizing Placement of Field Experience Program: An Integration of MOORA and Rule-Based Decision Making 895
Okfalisa Okfalisa, Rizka Hafsari, Gusman Nawanir, Saktioto Toto and Novi Yanti

Assessing Malaysian University English Test (MUET) Essay on Language and Semantic Features Using Intelligent Essay Grader (IEG) 919
Wee Sian Wong and Chih How Bong

Artificial Neural Network Intelligent System on the Early Warning System of Landslide 943
Aghus Sofwan, Sumardi, Najib and Indrah Wendah Atma Bhirawa

Performance Evaluation of Different Membership Function in Fuzzy Logic Based Short-Term Load Forecasting 959
Oladimeji Ibrahim, Waheed Olaide Owonikoko, Abubakar Abdulkarim, Abdulrahman Okino Otuoze, Mubarak Akorede Afolayan, Ibrahim Sani Madugu, Mutiu Shola Bakare and Kayode Elijah Adedayo

Mathematical Sciences

Comparison of Single and MICE Imputation Methods for Missing Values: A Simulation Study 979
Nurul Azifah Mohd Pauzi, Yap Bee Wah, Sayang Mohd Deni, Siti Khatijah Nor Abdul Rahim and Suhartono

Robust Multivariate Correlation Techniques: A Confirmation Analysis using Covid-19 Data Set 999
Friday Zinzendoff Okwonu, Nor Aishah Ahad, Joshua Sarduana Apanapudor and Festus Irismisose Arunaye

Generalized Fibonacci Search Method in One-Dimensional Unconstrained Non-Linear Optimization 1017
Chin Yoon Chong, Soo Kar Leow and Hong Seng Sim

Mathematical Modeling and Availability Analysis of Leaf Spring Manufacturing Plant <i>Sohan Lal Tyagi, Shikha Bansal, Priyanka Agarwal and Ajay Singh Yadav</i>	1041
Penalized LAD-SCAD Estimator Based on Robust Wrapped Correlation Screening Method for High Dimensional Models <i>Ishaq Abdullahi Baba, Habshah Midi, Leong Wah June and Gafurjan Ibragimov</i>	1053
Relative Risk Estimation for Human Leptospirosis Disease in Malaysia Based on Existing Models and Discrete Space-Time Stochastic Sir Model <i>Sufi Hafawati Ideris, Muhammad Rozi Malim and Norshahida Shaadan</i>	1071
Applied Sciences and Technologies	
Factors in Ergonomic Design of 6-to-18-month Baby Carriers for Elderly People <i>Ariya Atthawuttikul and Sorat Khongkharat</i>	1097
Numerical Analysis of the Crack Inspections Using Hybrid Approach for the Application the Circular Cantilever Rods <i>Saddam Hussein Raheemah, Kareem Idan Fadheel, Qais Hussein Hassan, Ashham Mohammed Aneel, Alaa Abdulazeez Turki Al-Taie and Hussein Kadhim Sharaf</i>	1109
Radiation Modified CS/PVA Film with PVP Coating as Cu Adsorbent Material <i>Norhashidah Talip, Amalina Muhammad Afifi, Mohd Yusof Hamzah, Maznah Mahmud and Sarada Idris</i>	1119
An Investigation on Indoor Temperature of Modern Double Storey House with Adapted Common Passive Design Strategies of Malay Traditional House <i>Maryam Qays Oleiwi and Mohd Farid Mohamed</i>	1135
Assessment by Simulation of Different Topological Integration of Solar Photovoltaic Plant in Medium Voltage Distribution Networks <i>Md. Milon Uddin, Mushfiqur Rahman, Md. Tanzid Ridwan Hossain and Md. Habibur Rahman</i>	1159
Earth Sciences	
Volume Transport Variability in the Western Equatorial Pacific and its Relations to Halmahera Throughflow <i>Marlin Chrisye Wattimena, Agus Saleh Atmadipoera, Mulia Purba, I Wayan Nurjaya and Fadli Syamsudin</i>	1171

Solution of Collinearity Problem in Two-Dimensional Electrical Resistivity Tomography using Wenner Array <i>Mustapha Adejo Mohammed Nordiana Mohd Muztaza and Rosli Saad</i>	1199
Potential Groundwater Exploration in Use of 2-D Electrical Resistivity Tomography (ERT) Techniques at the Department of Agriculture Kelantan Research and Developmental Platform Padang Raja Kelantan <i>Wan Fazilah Fazlil Ilahi, Nur Hidayu Abu Hassan, Mohd Razi Ismail, Nik Norasma Che'Ya, Zulkarami Berahim, Mohamad Husni Omar, Nurul Idayu Zakaria and Mohamed Azwan Mohamed Zawawi</i>	1219
Environmental Sciences	
Efficacy of Intelligent Mosquito System (I.M.O.S) with Xmos Mini Aerosol Against <i>Aedes</i> in 17 th College, Universiti Putra Malaysia <i>Latifah Saiful Yazan, Banulata Gopalsamy, Siti Najiha Abu Bakar, Khairul Aiman Manan, Noranis Shahida Shahidan and Lee Yean Wang</i>	1229
Modelling <i>In-situ</i> Factors Affecting Bud's Growth of <i>Rafflesia kerrii</i> Meijer in Lojing Highlands, Kelantan, Peninsular Malaysia <i>Abdul Hamid Mar Iman, Nor Hizami Hassin, Muhamad Azahar Abas and Zulhazman Hamzah</i>	1243
Ecological Roles of Benthic Community in Seafloor Northwards of Pulau Indah, Klang, Malaysia <i>Mohd Sophian Mohd Kasihmuddin and Zaidi Che Cob</i>	1267
Medical and Health Sciences	
Radiographic Measurement of Cochlear in Sudanese Population using High Resolution Computed Tomography (HRCT) <i>Lubna Bushara, Mohamed Yousef, Ikhlas Abdelaziz, Mogahid Zidan, Dalia Bilal and Mohammed El Wathig</i>	1289
Prevalence of Vitreous & Retinal Disorders among Sudanese Diabetic Patients: A B-Scan Ultrasonography Study <i>Mohamed Yousef, Safaa Bashir, Awadalla Wagealla, Mogahid Zidan, Mahmoud Salih Babiker and Mona Mohamed</i>	1299
Material Sciences	
Poly lactide and its Composites on Various Scales of Hardness <i>Abraham Kehinde Aworinde, Eyere Emagbetere, Samson Oluropo Adeosun and Esther Titilayo Akinlabi</i>	1313
Chemical Sciences	
SETS: A Seed-Dense-Expanding Model-Based Topological Structure for the Prediction of Overlapping Protein Complexes <i>Soheir Noori, Nabeel Al-A'araji and Eman Al-Shamery</i>	1323

Foreword

Welcome to the 2nd 2021 issue for the *Pertanika Journal of Science & 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 35 articles; 34 regular articles and a review article. The authors of these articles come from different countries namely Bangladesh, India, Indonesia, Iraq, Malaysia, Nigeria, Saudi Arabia, Sudan and Thailand.

Another article that we wish to highlight is on “Artificial neural network (ANN) intelligent system on the early warning system of landslide” by Aghus Sofwan et al. from Diponegoro University, Indonesia. The ANN intelligent system uses sensors to obtain landslide causative parameters: rainfall, the slope of the ground, soil moisture, and vibration. The ANN system output delivered three-level conditions: the safe, the standby, and the hazardous. Two methods the feed-forward backpropagation (FFBP) and the cascade forward backpropagation (CFBP) were analyzed. The FFBP method performance is better than that of the CFBP method in terms of safe condition level. However, the CFBP method is superior to the FFBP method in standby and hazard conditions. Hence, the researchers deploy the ANN CFBP method into the developed hardware system. Detailed information on this study is presented on page 943.

The next article is on robust multivariate correlation techniques: a confirmation analysis using Covid-19 data set by Friday Zinzendoff Okwonu and co-researchers from Universiti Utara Malaysia and Delta State University. Robust multivariate correlation techniques are proposed to determine the strength of the association between two or more variables of interest since the existing multivariate correlation techniques are susceptible to outliers when the data set contains random outliers. The performances of the proposed techniques were compared with the conventional multivariate correlation techniques. The comparative analysis of the proposed methods and the conventional methods showed that the proposed methods were insensitive to influential observations or random outliers than the conventional methods. The study based on these methods showed moderate to a strong association between confirmed, discharged, and death cases. It implies that someone infected with the virus can only be discharged or die. Detailed information on this study is presented on page 999.

A regular article titled “Factors in ergonomic design of 6-to-18-month baby carriers for elderly people” was written by Ariya Atthawuttikul and Sorat Khongkharat from Thailand. Their aim was to formulate an ergonomic design factor for 6-to-18-month baby carrier from simulated loading point data and physical parameters of an average Thai

elderly person. They have synthesized four ergonomic design factors namely shoulder and hip posture while using the carrier, the seat and backrest of the baby carrier, the load-bearing points on the body of the carrying person, and the type of baby carrier. As a result, the authors obtained valuable information on loading points of the baby carriers and on the body of an average Thai elderly person. Detailed information on this study can be found on page 1097.

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

Dr. Mohammad Jawaid

executive_editor.pertanika@upm.edu.my

High Level Synthesis Optimizations of Road Lane Detection Development on Zynq-7000

Panadda Solod^{1*}, Nattha Jindapetch¹, Kiattisak Sengchuai¹, Apidet Booranawong¹, Pakpoom Hoyingcharoen¹, Surachate Chumpol² and Masami Ikura²

¹Department of Electrical Engineering, Faculty of Engineering, Prince of Songkla University, Hat Yai, Songkhla 90112, Thailand

²Toyota Tsusho Nexty Electronics (Thailand) co, Ltd Bangkok, Thailand

ABSTRACT

In this work, we proposed High-Level Synthesis (HLS) optimization processes to improve the speed and the resource usage of complex algorithms, especially nested-loop. The proposed HLS optimization processes are divided into four steps: array sizing is performed to decrease the resource usage on Programmable Logic (PL) part, loop analysis is performed to determine which loop must be loop unrolling or loop pipelining, array partitioning is performed to resolve the bottleneck of loop unrolling and loop pipelining, and HLS interface is performed to select the best block level and port level interface for array argument of RTL design. A case study road lane detection was analyzed and applied with suitable optimization techniques to implement on the Xilinx Zynq-7000 family (Zybo ZC7010-1) which was a low-cost FPGA. From the experimental results, our proposed method reaches 6.66 times faster than the primitive method at clock frequency 100 MHz or about 6 FPS. Although the proposed methods cannot reach the standard real-time (25 FPS), they can instruct HLS developers for speed increasing and resource decreasing on an FPGA.

ARTICLE INFO

Article history:

Received: 21 September 2020

Accepted: 30 December 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.01>

E-mail addresses:

6010120066@psu.ac.th (Panadda Solod)

nattha.s@psu.ac.th (Nattha Jindapetch)

ak.kiattisak@hotmail.com (Kiattisak Sengchuai)

bapidet@eng.psu.ac.th (Apidet Booranawong)

hpakpoom@eng.psu.ac.th (Pakpoom Hoyingcharoen)

surachate@th.nexty-ele.com (Surachate Chumpol)

ikura@th.nexty-ele.com (Masami Ikura)

* Corresponding author

Keywords: Array partitioning, FPGA, high level synthesis (HLS), HLS interface, loop pipelining, loop unrolling

INTRODUCTION

Advanced Driving Assistant Systems (ADAS) are enhanced self-driving in autonomous driving cars. There are many sub-systems in ADAS, such as Lane Keeping Assistant System (LKAS) and

Lane Departure Warning System (LDWS) (Chen & Boukerche, 2020). Both LKAS and LDWS contain road lane detection, which is one of the complex systems in ADAS. The important thing about road lane detection, which causes the accuracy of detection is the various environment. Therefore, many algorithms are developed to resolve the problem. The perception model for autonomous vehicles in a variety of environments such as dynamic movements of obstacles, parked and moving vehicles, poor quality lines, shape curve, and the strange lane was reviewed (Feniche & Mazri, 2019). The typical model for lane detection consists of five steps: image cleaning, feature detection, model application, tracking integration, and coordinates translation. Most of lane detection algorithm contains straight line detection, which is Hough Transform (HT).

Due to the requirements of road lane detection, which are real-time computation and low-power consumption. A real-time lane detection was implemented using simple filter and Kalman filter on a high-performance device (IMX6Q) (Lee et al., 2017) and Canny-Hough is modified (Hwang & Lee, 2016) to achieve high-speed computation. To achieve both requirements to develop on a hardware device, Lu et al. (2013), Guan et al. (2017), and Marzotto et al. (2010) developed parallel architectures for the road lane detection on hardware devices (Virtex-5, ML505, ALTERA DE2, and ALTERA DE2-115). Although the high-speed computation with low-power consumption can be recompense by using hardware devices, the development time still long and high cost. In Promrit and Suntiamorntut (2017), blob detection and HT were applied for road lane detection on the low-cost hardware device (Zynq-7000 family). To decrease the development time, Khongprasongsiri et al. (2018) and Panda et al. (2018) considered the optimization technique on HLS (loop-pipelining and loop-unrolling) to resolve the bottleneck computation and found out the suitable factor for loop algorithm. However, the HLS optimization techniques on hardware devices analysis with complex algorithms such as lane detection have not been completely analyzed yet.

In this study, we proposed HLS optimization processes and analysis methods on hardware devices (Zynq-ZC7010) for implementing the complicated algorithms, especially the nested-loop for high-speed and low-resource usage on low-cost devices. The contributions are the proposed HLS optimization processes that should be done sequentially. In each process, we proposed an analysis method to consider the suitable optimization techniques should be applied to implement on a device. Road lane detection in Solod et al. (2018) was selected to be a case study, which contained four main steps: pre-processing, edge detection, line detection, and angle calculation. The experimental results illustrated the improvement after applied each HLS optimization process that can be the instructions for speed increasing and resource decreasing on an FPGA.

PRELIMINARIES

This section explains basic knowledge of lane detection (edge detection and Hough transform) and HLS optimization techniques (pipelining, loop unrolling, array partitioning,

HLS interface). Since this paper does not propose either the new architecture or algorithms for the lane detection, the original theories, as well as the HLS optimization techniques are explained. This is a case study to be applied with the proposed HLS optimization processes.

Edge Detection

There are many methods to perform edge detection. Prewitt, Sobel, Robert, and Canny edge detections use the gradient method to detect the edge by looking at the maximum and minimum value in the first derivative of the image. The gradient method of Prewitt, Sobel, and Robert to detect edge has different gradient values of x and y direction as respectively shown in Figures 1, 2 and 3.

Meanwhile, Canny edge detection also uses the gradient method, but it is more complicated steps than the others. In the edge detection process, there are six steps: (i) reducing the noise by using Gaussian filter, (ii) finding edge by taking the gradient of the image (the gradient value is the same as Sobel edge detection), (iii) finding edge strength, (iv) finding the edge direction by the Equation 1, (v) tracing the related edge direction to the direction in an image, then perform non-maximum suppression, and (vi) finally eliminating streaking by hysteresis.

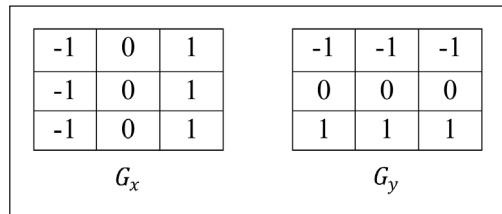


Figure 1. X and Y direction gradient value of Prewitt Edge Detection

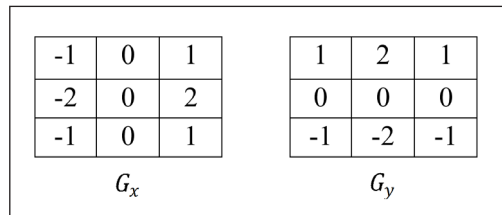


Figure 2. X and Y direction gradient value of Sobel Edge Detection

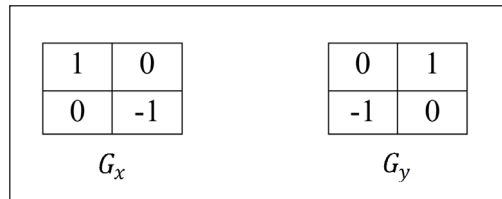


Figure 3. X and Y direction gradient value of Robert Edge Detection

$$\theta = \tan^{-1}(G_x/G_y) \tag{1}$$

Hough Transform

For lines and circles identification, Hough transform is one of a feature extraction to process by using a voting procedure to detect incomplete instances of objects of a particular type of shape. Generally, according to Equation 2, the straight line is capable to represent the linear equation with two important parameters: slope (m) and intercept

(c). Equation 3 represents the polar form of a line. Figure 4 shows the transformation of Hough, where ρ represents the vertical distance of the straight line on x-y plane from the origin position. The angle between ρ and horizontal axes is represented by θ . A position on the blue straight line on x-y plane indicates to each curve on θ - ρ plane. An identical straight line will cause an identical intersection position of the curve on θ - ρ plane. Therefore, the length of the straight line is able to represent the number of intersection curves.

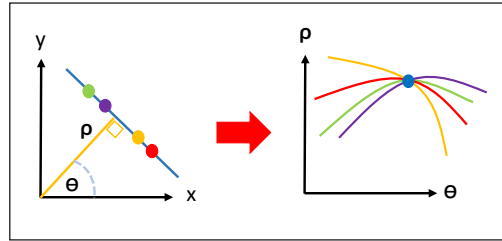


Figure 4. Hough transform concept

$$y = mx + c \quad [2]$$

$$\rho = x \cos\theta + y \sin\theta \quad [3]$$

Pipelining

Pipelining is a method for rapidity optimization, which decreases the latency of the operations by the initiation interval (II) decreasing for a function or loop. Figure 5 is a sample of the pipelining technique applied to a for-loop, which contains 3 operations: Rd, Cmp, and Wr. The left-hand side of Figure 5 is an operating sequence of the process without pipelining. Each iteration needs three clock cycles for processing. Six clock cycles are needed for two iterations processing due to the last operation of the first iteration must be completed before the first operation in the consecutive iteration starts. In contrast, pipelining is applied to the same for-loop with II equal to 1. In the second clock cycle, the first operation in the consecutive iteration can be immediately operated as shown on the

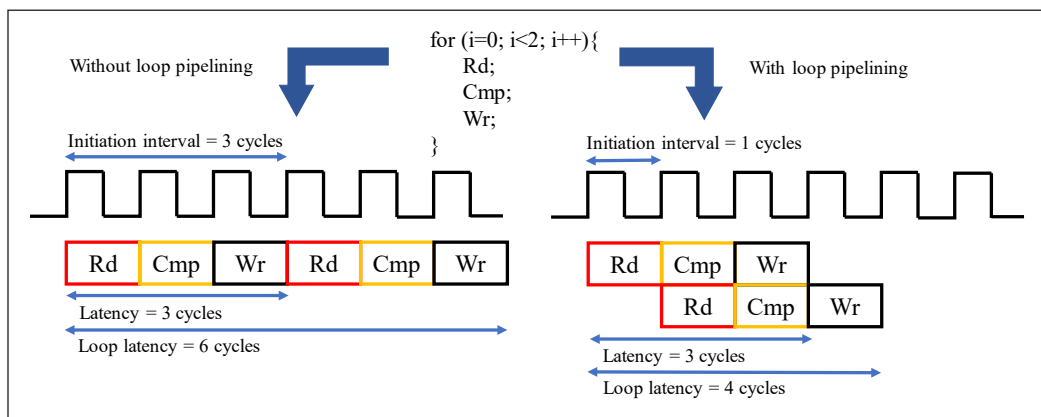


Figure 5. The example of loop pipelining for for-loop

right-hand side of Figure 5. The number of clock cycles for processing is reduced from six to four clock cycles.

Unrolling

Unrolling is a performance optimization technique as well. It allows some or all iterations to be concomitant. Figure 6 is an example of applying loop unrolling. Without unrolling, each element of the array needs 1 clock cycle for processing. Therefore, six iterations are needed to process six elements of array whereas only two iterations are needed in case of using unrolling with factor equals three due to three elements of array are processed in 1 clock cycle.

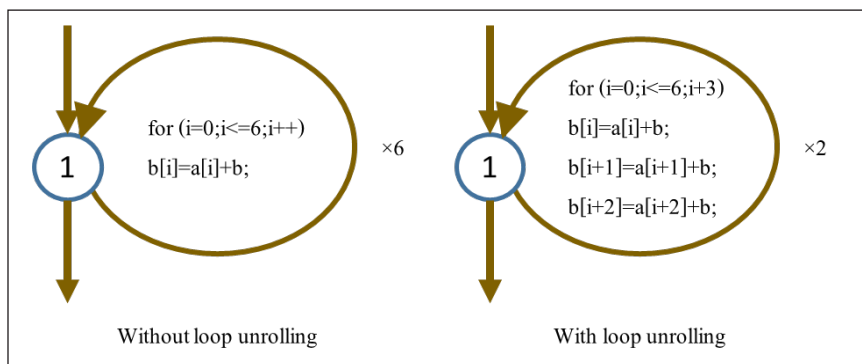


Figure 6. The example of loop unrolling for for-loop

Array Partitioning

Array partitioning is a method for speed and region optimization provided by Xilinx Vivado HLS. An original array is divided into sub-group of smaller arrays and stored into separate banks. Figure 7 demonstrates three types of array partitioning, which are block, cyclic, and complete types. In the case of the block type, a large array is divided into balanced blocks whose array elements are consecutively arranged. In the case of the cyclic type, a large array is divided into a balanced block and interleaved with elements. In the case of the complete type, an array is divided into single elements.

HLS Interface

Xilinx Vivado HLS supports the specification of I/O protocol types. Port interface is created by the synthesis of interface based on efficacy industry-standard interface and manual interface specifications, with the manner of the interface is illustrated in the input source code. Three port types on RTL design, including clock and reset port, block-level interface protocols, and port-level interface protocols are created by Xilinx Vivado. Typically, the

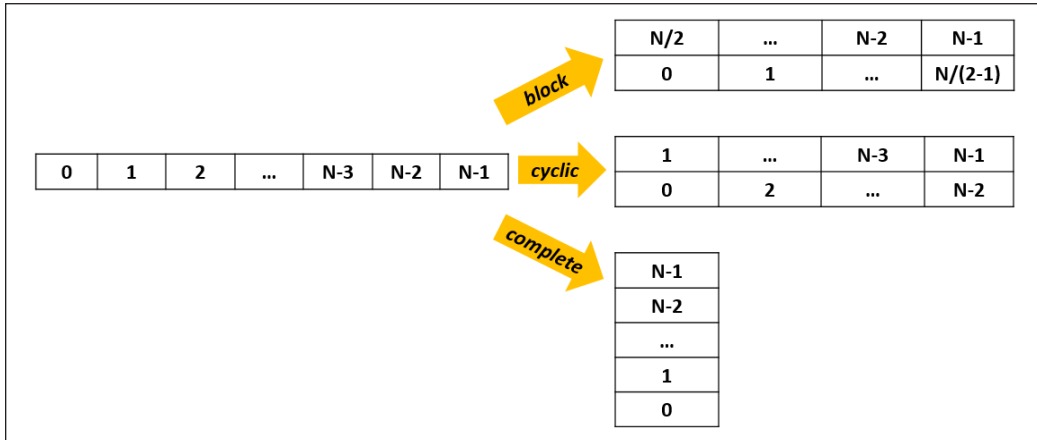


Figure 7. The type of array partitioning

protocol of the block-level interface is gathered in the design. These signals control the block are autonomous to port-level I/O protocols. These ports determine when the block can begin data processing, demonstrate when new inputs can be asserted new inputs, and demonstrate whether the system is idle or the operation completed. After the block-level protocol has been used to start the operation of the block, the port-level I/O protocols are used to sequence data into and out of the block.

METHODS

Road Detection Framework

A framework of lane tracking proposed in this paper is shown in Figure 8. There are four main steps: 1. pre-processing, 2. edge detection, 3. line detection, and 4. angle calculation. Edge detection and line detection are implemented on the PL part to increase the operation speed, and others are implemented on the PS part. Pre-processing expanded in Figure 9 includes image cropping, image dividing into two parts (left side image and right side image) to separate left and right line, an RGB to grayscale image conversion, and grayscale to binary image conversion. According to the result in Solod et al. (2018), Robert edge detection is selected to perform edge detection. Hough line transform is then

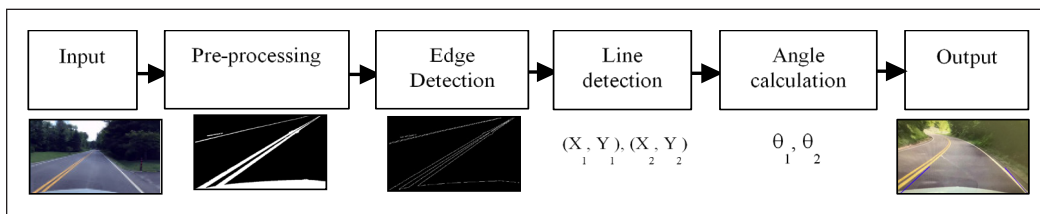


Figure 8. Road lane detection framework

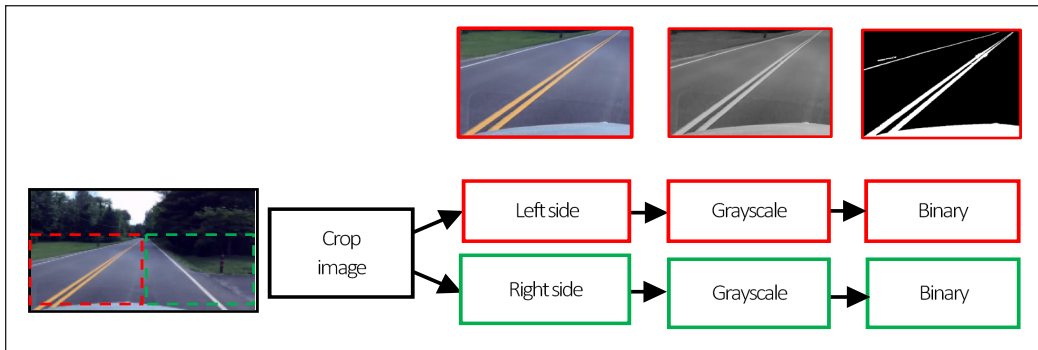


Figure 9. Pre-processing step of road lane detection

applied to the line detection. The positions of (x_1, y_1) and (x_2, y_2) obtained from the line detection step are used to calculate the angles (θ_1 and θ_2).

The angle calculation is divided into three cases to recognize the road path curve and prevent fallibility from the previous step. θ_1 represents the angle of the left side and θ_2 represents the angle of the right side. The first case is straight lane as shown in Figure 10, which θ_1 must be over 90° and θ_2 must be less than 90° . The second case is the left curve as shown in Figure 11, in which θ_1 and θ_2 must be less than 90° . The last case is the right curve as shown in Figure 12, in which θ_1 and θ_2 must be over 90° .

HLS Optimization Processes

According to the operation time of road lane detection spends most time in the process. Edge detection and line detection are implemented on PL using HLS on Zynq-7000, which allows C and C++ languages and authorizes efficient method as loop unrolling, loop pipelining, or array

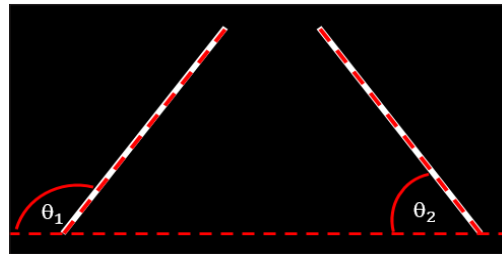


Figure 10. The measurement angle of straight lane

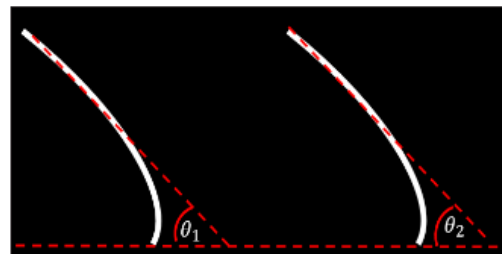


Figure 11. The measurement angle of left curve lane

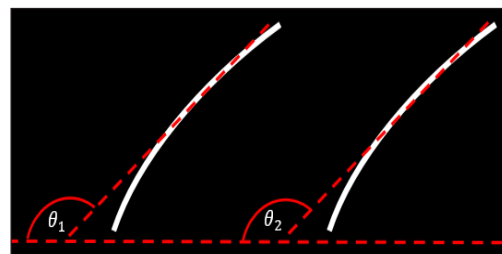


Figure 12. The measurement angle of right curve lane

partitioning, etc. The proposed HLS optimization processes are divided into four steps as follows.

- Step 1: array sizing is performed to decrease the resource usage on PL part.
- Step 2: loop analysis is performed to determine which loop must be loop unrolling or loop pipelining.
- Step 3: array partitioning is performed to resolve the bottleneck of loop unrolling and loop pipelining.
- Step 4: HLS interface is performed to select the best block level and port level interface for array argument of RTL design.

These four HLS optimization processes should be done sequentially. In each process, an analysis method is proposed to consider the suitable optimization techniques should be applied to implement on an FPGA device. If the optimization processes are not done in suggested sequence, the optimization resistant result may be obtained. This behavior is illustrated in the experimental results and discussions section.

Array Sizing. The first point to consider is the array sizing. Analysis and calculation of the appropriate array size for block memory storage capacity can decrease over the allocation of block RAM (BRAM) by considering the amount of BRAM in Equation 4. At array sizing step, 2 parameters are considered. The first parameter is memory depth, which caused by the amount of array element. For example, consider a 16-bits 1-dimension array with 2050 word lines. Due to the requirement of the memory depth, the memory depth must over 2050. The at least memory depth for this array is $4096(2^{12})$, the data width is 16 bits and the size of BRAM is 18K bits (18KBRAM). According to Equation 4, $4(2^2)$ 18K BRAM is allocated to keep the array. Likewise, the word line of the array is reduced to 2048 we need to allocate $4(2^2)$ BRAM 18K to store an array. Likewise, the size of the array is reduced to 2048, the amount of BRAM is decreased to $2(2^1)$.

In this work, we reduced the input image from 1920×1080 pixels to 200×200 pixels. From the Hough line transform theory, the maximum diagonal is reduced from 1445 to 224, which equates to the value of rho in the first dimension size of $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ array. The second parameter is the data width. According to the concept of Hough line transform, the value of $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ will not over the value of rho parameter, which has the maximum value to the diagonal line of the image, so the data width can be reduced from 16 bits to 8 bits. Theta also decreases from 0° - 360° to 30° - 150° . Thus the second dimension is reduced from 360 to 120, which is sufficient for road lane detection.

$$\text{Block RAM} = \frac{\text{memory depth} \times \text{data width}}{\text{block RAM size}} \quad [4]$$

Loop Analysis. From the theory of Robert edge detection and Hough line transform, we summarize into four main nested loops that are agreeable to C language as shown in Figure 13. The first step is started with a 2-layer nested loop of a 16-bit-2-dimension array (hough1[rho][theta] and hough2[rho][theta]) is initialized to 0 for intersection times counting that starts with zero. The iteration of the inner layer equals the number of theta and the outer layer equals the number of rho. The second step is divided into two minor nested loops, which are a 4-layer nested loop for edge detection and a 3-layer nested loop to count intersection times in θ - ρ plane for every pixel. The third step is a 2-layer nested loop for the Hough line transform normalization by considering the value of all elements of hough1[rho][theta] and hough2[rho][theta]. The final step is a 3-layer nested loop for lane tracking by array hough1[rho][theta] and hough2[rho][theta] checking to determine the position of the line in the image.

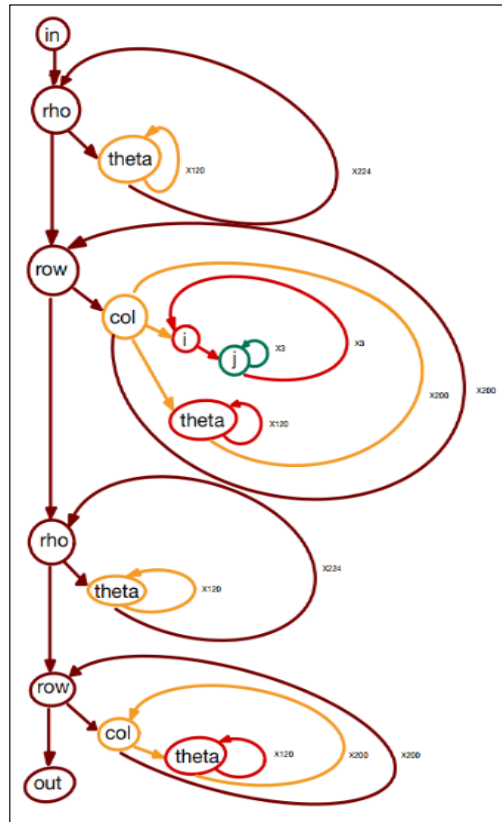


Figure 13. The nested loops of C language for edge and line detection

From all steps of edge detection and line detection in Figure 13, we found that the nested loop is in every step, which caused speed reduction. In this work, we consider each nested loop with an optimization technique that is appropriate with the nested loop type. Both 2-dimension arrays hough1[rho][theta] and hough2[rho][theta] are initialized in the first nested loop, which results in each iteration is independent. All array elements can be initialized simultaneously. Generally, as shown in Figure 14, two iterations in the first nested loop need four clock cycles to operate. In the case of loop unrolling is added, two iterations need only two clock cycles to operate. Edge detection in the first minor nested loop is dependent value in each iteration. According to Robert edge detection theory, both x_weight and y_weight are convolution results, which are cumulative variables. Therefore, loop unrolling is not necessary for the first minor nested loop.

$$hough_{(1,2)}[rho][theta] = hough_{(1,2)}[rho][theta] + 1 \quad [5]$$

The second minor nested loop is counting times of intersection in θ and ρ space according to the Hough line transform. Equation 5 demonstrates that $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ are cumulative variables, which is dependent on other iteration as well. Therefore, loop pipelining is appropriate to add in the second minor nested loop. The schedule of this step will be changed following Figure 15, without loop pipelining the next element of $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ cannot be started reading before the previous element finished processing, whereas adding loop pipelining, the next element of $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ can start reading before previous element finish processing.

$\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ are normalized in third nested loop. This step is similar to the first nested loop, in which the result in each iteration is independent. Loop unrolling will get less latency than loop pipelining, thus loop unrolling should be selected if the resource is sufficient.

The value of every element of $\text{hough1}[\rho][\theta]$ and $\text{hough2}[\rho][\theta]$ is checked to mark the position of the lane so the result in each iteration is not influencing to the other. Loop unrolling should be selected as well as the third nested loop. However, loop

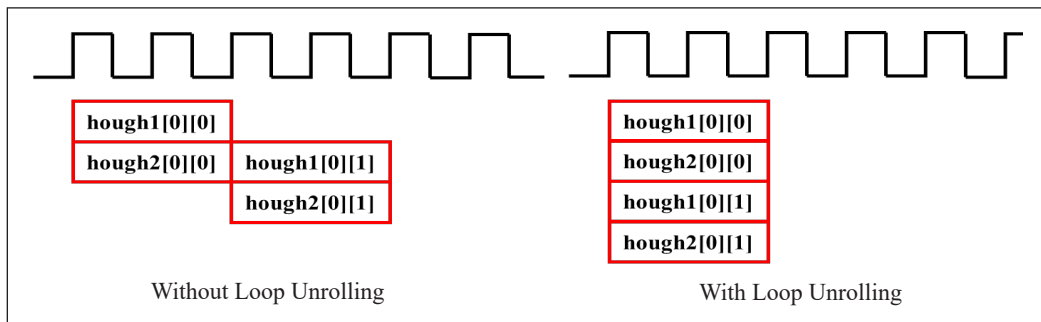


Figure 14. Loop unrolling applying in 1st nested loop

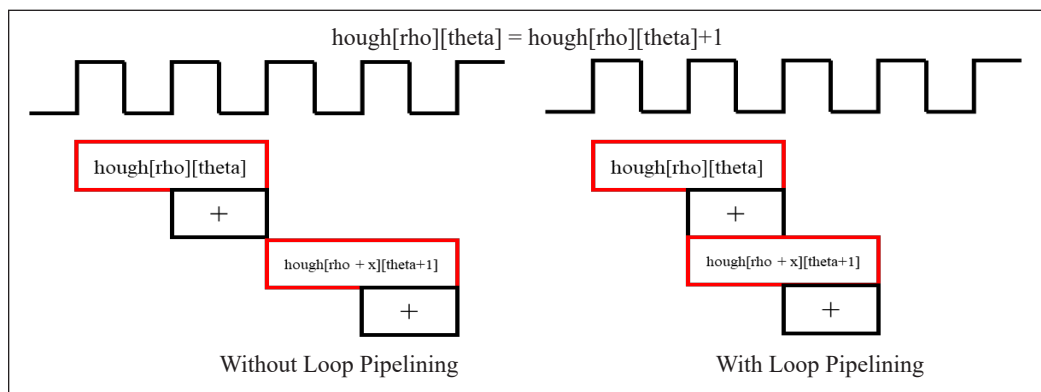


Figure 15. Loop pipelining applying in 2nd minor nested loop

pipelining must be selected instead of loop unrolling because of the limitation of resources on the device.

Array Partitioning for Unrolling and Pipelining. Array partitioning significantly reduces that can decrease the latency. This technique encourages loop pipelining and loop unrolling, such as in the second minor nested loop. Although loop pipelining is added, $hough1[rho][theta]$ and $hough2[rho][theta]$ are still stored in the same bank of memory, the bottleneck will occur because each BRAM allows only one read operation of one element at that time. Loop pipelining will not get the highest efficiency as expected. Thus, the second element ($hough1[rho + x][theta + 1]$ and $hough2[rho + x][theta + 1]$) of array $hough1[rho][theta]$ and $hough2[rho][theta]$ cannot be accessed until the first element ($hough1[rho][theta]$ and $hough2[rho][theta]$) is successfully accessed. Due to the bottleneck issue, loop unrolling in Figure 14 cannot get the highest efficiency as well. We found that array partition can be resolve or ameliorate this problem. In this part, an array partitioning block type, F equal to 6, and D equal to 2 are applied to $hough1[rho][theta]$ and $hough2[rho][theta]$ in Figure 16. The original array on the top is stored in one bank of memory. This array is divided into 6 banks of memory after array partitioning is applied. Memory depth in each bank equals 8,192.

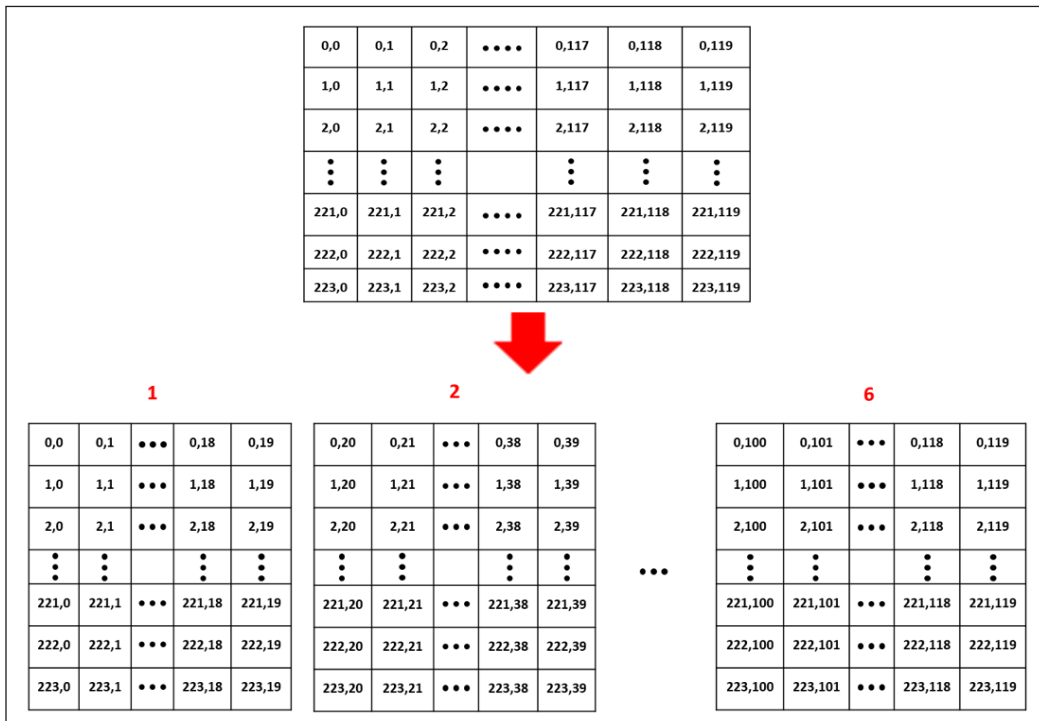


Figure 16. Array partitioning type of $hough1[rho][theta]$ and $hough2[rho][theta]$

HLS interface for Array Type. Register Transfer Level (RTL) description is created by Xilinx HLS, in which an input/output operation must be performed through a port in the design. In this work, an RTL has 1-input port and 1-output port, which is all 1-dimension array. We consider port design that suitable for an array type. There are `ap_ctrl_none`, `ap_ctrl_hs` and `ap_ctrl_chain` for block-level interface. There are `axis`, `s_axilite`, `m_axi`, `ap_hs`, `ap_memory`, `bram`, `ap_fifo`, and `ap_bus` for data transmission.

RESULTS AND DISCUSSIONS

This experiment aims to increase the speed of road lane detection by adding HLS optimization techniques including array sizing, loop unrolling, loop pipelining, and HLS interface management. The comparison of resource usage and operating time are discussing as well.

Process Profiling

In the beginning, the profile of the operation time of each process in road lane detection is extracted on Intel® Core™ i7-7500U CPU @ 2.70 GHz, which input file has size 720x1280 pixels and frame rate 24 fps. The detailed profiling of process activities illustrates the time-consuming processes that must be implemented as hardware accelerators on the PL part. We found that the step of edge detection and line detection step spent most processing time as shown in Table 1.

Table 1
Comparison of operation time in each process

Optimization	Resource (%)				Latency
	DSP	BRAM	CLB	FF	
Default	17	1,710	11	4	15,847,183,592
Array sizing	13	57	11	3	103,709,185

Array Sizing Results

The results as shown in Table 1, we found that edge detection and line detection should select to be hardware accelerator to increase operation speed. However, the resource, which is spent on PL part is over the limitation. Therefore, array sizing should be considered by the method described in the previous section. The input image is cropped and resized from 1920x1080 to 200x200 pixels. Therefore, `hough1[rho][theta]` and `hough2[rho][theta]` is resized from 2203x120 to 224x120 according to HT theory, which in rho is the feasible perpendicular length of the input image. As shown in Table 2 latency is reduced about 152 times faster and BRAM is reduced form 1,710% to 57%.

Table 2
Comparison of default and array sizing

Lane detection process	Processing Time (%)	Processing Time (FPS)	Processing Time (s/frame)
Pre-processing	30.60	0.8928	0.1049
Edge detection	32.30	0.9424	0.1107
Line detection	36.10	1.0533	0.1238
Angle calculation	1.00	0.0292	0.0034

Loop Unrolling and Loop Pipelining Results

The optimization techniques, loop unrolling, and loop pipelining are considered into inner-nested loop of edge and line detection to analyze the suitable usability. In case of considering to least latency, loop unrolling is sufficient to add along with first, third, and last nested loop of Figure 13. Loop pipelining is sufficient to add along with the second nested loop (both of edge detection and line detection nested loop). Although loop unrolling can reduce the latency more than loop pipelining in third and last nested loop, the number of CLBs is more than the limitation. Therefore, loop pipelining is sufficient to add along with third and last nested loop instead of loop unrolling. The proposed method as shown in Table 3 is the combination of sufficient optimization techniques in all nested loop that can reduce latency at clock frequency 100 MHz from 103,709,185 clock cycles to 15,789,018 clock cycles or about 6.57 times is reduced. Both of loop unrolling and loop pipelining is latency reducing optimization by throughput increasing or initiation interval decreasing. In case of $\text{hough1}[\rho][\theta]$, $\text{hough2}[\rho][\theta]$ and two more arrays, $\text{sin}[\theta]$ and $\text{cos}[\theta]$,

Table 3
Comparison of loop unrolling and loop pipelining in each loop

Process	Method	Resource (%)				Latency (clock cycle)
		DSP	BRAM	CLB	FF	
All	Default	12	30	9	3	103,709,185
1 st	Unrolling	12	30	27	3	103,695,297
	Pipelining	12	30	10	3	103,708,738
Edge detection	Unrolling	13	30	9	3	102,739,561
	Pipelining	13	30	10	3	98,012,241
2 nd	Unrolling	135	30	228	72	61,853,749
	Pipelining	13	30	10	3	46,056,556
Edge + Line detection	Combination	15	30	10	3	45,046,530
3 rd	Unrolling	12	30	177	79	98,457,416
	Pipelining	12	30	15	6	98,457,432
4 th	Unrolling	11	30	372	56	72,509,185
	Pipelining	12	30	11	3	74,828,791
Proposed method (1)		15	30	34	6	15,789,018

which are constants stored in the same block of memory, the loop pipelining and loop unrolling cannot get the best efficiency because of the limitation of memory accessing.

Array Partitioning Results

Since hough1[rho][theta], hough2[rho][theta] and two more arrays, sin[theta] and cos[theta] are constants stored in the same block of memory, the loop pipelining and loop unrolling cannot get the best efficiency cause of the limitation of memory accessing Table 4 is the result of adding array partitioning to hough1[rho][theta], hough2[rho][theta], sin[theta] and cos[theta] compare to the first proposed method. Both hough1[rho][theta] and hough2[rho][theta] arrays are sufficient with array partitioning block type with F equal to 6 at equal to is 2. The latency from the experiment is reduced from 15,789,018 clock cycles to 15,777,837 clock cycles, which is only 0.99 times faster. Although Adding array partitioning complete type to these arrays instead of array partitioning block type would be faster, the resources on the device in this work will not enough to implement. In contrast, cos[theta] and sin[theta] are sufficient with array partitioning complete type and the latency is reduced from 15,789,018 clock cycles to 15,697,034 clock cycles compared to the second proposed method.

Table 4
 Comparison of array partitioning

Optimization	Resource (%)				Latency (clock cycle)
	DSP	BRAM	CLB	FF	
Proposed method (1)	15	30	34	6	15,789,018
block type, F 2, D 2	15	30	27	6	15,782,316
block type, F 6, D 2	17	44	24	8	15,777,837
hough1 & hough2					
block type, F 12, D 2	17	44	26	9	15,803,597
cyclic type, F 2, D 2	15	30	28	6	25,774,041
cyclic type, F 6, D 2	17	44	24	7	25,769,578
cyclic type, F 12, D 2	15	30	28	6	25,774,041
sin & cos					
Complete type	17	42	39	8	15,697,034
Block type, F 2	18	42	27	9	15,777,845
Block type, F 6	17	42	27	9	15,777,837
Block type, F 12	17	42	27	9	15,777,846
Proposed method (2)	17	42	39	8	15,697,034

HLS Interface Management Results

In the case of the HLS interface, there are three types of block type interfaces available for arrays, namely ap_ctrl_none, ap_ctrl_hs, and ap_ctrl_chain. There are many types of port-level interfaces available for arrays, namely axis, s_axilite, m_axi, ap_hs, ap_memory, bram,

ap_fifo, and ap_bus. In this experiment, the block-level interface and port-level interface are matched to the input and output of RTL that Xilinx HLS has created. Following Figure 17, the first column represents a block-level interface, the second column represents a port-level interface for the input and the final column is represent a port-level interface for output. The result of the HLS interface management after all nested loop is added by the optimization technique is shown in Table 5, which compares to the default of a block-level interface and port-level interface, generated by Xilinx HLS. The least latency occurs with all of the block-level interfaces while the port-level interface of input should be axis and the port-level interface of output should be ap_memory.

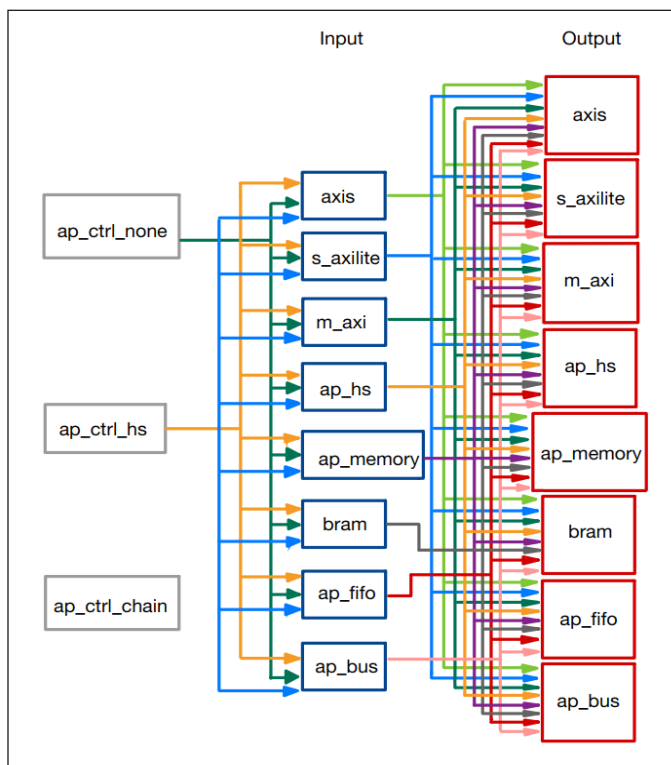


Figure 17. HLS interface matching for array type

Table 5
Comparison of HLS interface

Method	Latency (clock cycles)	Processing Performance (FPS)
Default	15,847,183,592	0.006
Proposed (1)	15,789,018	6.334
Proposed (2)	15,697,034	6.371
Proposed (3)	15,616,232	6.404

CONCLUSION

This paper has presented the optimization for complex road lane detection method, which supports lane curve recognition by angle calculation and loop analysis with optimization technique to increase the rapidity of operation underneath the limitation of resources on Xilinx Zynq-7000 (ZC7010).

The sequence of the optimization contains four steps: array sizing to reduce the resources, loop unrolling and loop pipelining to increase operation speed by parallel operation, array partitioning to prevent the bottleneck that can occur from the loop unrolling and loop pipelining step, and HLS interface management to get the best data transmission.

From the experiment, we found that an array sizing is suitable for low memory devices and adjustable to the method or experiment. Loop unrolling and loop pipelining is latency reducing optimization techniques that suitable to different kinds of loops. In the case of loop unrolling is suitable for dependent loops, which are loops that the result of each iteration in loop does not affect to other iteration. In case of loop pipelining is suitable for dependent loops, which are loops that the result of each iteration in loop effect to other iteration. Array partitioning can be both of area reducing and latency reducing optimization. This technique supports loop unrolling and loop pipelining to achieve efficiency as much as possible. The determination of factor (F) and dimension (D) of array partitioning depends on the sequence of the algorithm. The type of C argument, which is an input and output in RTL will be suitable with a different type of HLS interface.

An array sizing, loop unrolling, loop pipelining, and array partitioning have been applied to edge detection and line detection using Xilinx Vivado HLS. The summary resource and latency of the proposed optimization techniques compared to the default method are shown in Table 6. Following Tables 7 and 8, the resource utilization and speed are compared between the proposed method on Zybo-ZC7010 with the other algorithms on different architectures. Most of the performance in other algorithms on larger FPGA devices is faster according to the clock frequency but spent a lot of resources as well. Although the proposed method cannot be reached in real-time (25 FPS), it could reduce resource consumption and the latency from 103,951,104 clock cycles to 15,616,232 clock cycles, or latency is reduced 6.66 times at clock frequency 100 MHz.

In addition, the proposed method in this work can reach higher performance by implementing on the high-performance devices. The optimizations can further gain the benefit on the devices with sufficient resources. Since our procedure optimization method only considered the inner-layer of nested-loop. Therefore, the speed performance still can be increased by nested-loop layer analysis along with the procedure of the proposed method that can be an instruction for the HLS developers.

Table 6

The summary of the proposed optimization techniques compared the default method

Optimization	Resource				Latency (clock cycle)
	<i>DSP</i>	<i>BRAM</i>	<i>CLB</i>	<i>FF</i>	
Default	17	42	39	8	15,697,034
HLS interface	16	42	39	8	15,616,232
Proposed method (3)	16	42	39	8	15,616,232

Table 7

The resource utilization comparison of the proposed method with other methods on different architectures

Architectures	Resource Utilization				
	<i>Slices</i>	<i>LUTs</i>	<i>On-chip Memory (Kbit)</i>	<i>DSP</i>	Embedded Multiplier
Virtex-5 (ML505) (El Hajjouji et al., 2020)	1,119	1,996	1,625	1	1
Altera DE2 (Marzotto et al., 2010)	14,945	14,945	1,555	-	15
Altera DE2-115 (Lu et al., 2013)	29,431	2,589	3,052	8	8
Altera Stratix (Guan et al., 2017)	41,115	1,459	1,604	48	16
Zybo (ZC7010-1) (proposed)	2,816	6,864	101	13	-

Table 8

The speed comparison of the proposed method with other methods on different architectures

Method	F (MHz)	Resolution	Performance (FPS)
Virtex-5 (ML505) (El Hajjouji et al., 2020)	200	640 × 480	68
Altera DE2 (Marzotto et al., 2010)	115	1920 × 1080	25
Altera DE2-115 (Lu et al., 2013)	200	1,024 × 768	64
Altera Stratix (Guan et al., 2017)	200	1,024 × 768	185
Zybo (ZC7010-1) (Proposed)	100	1920 × 1080	6

ACKNOWLEDGEMENTS

This research is financed by Prince of Songkla University, Hat Yai, Songkhla, Thailand.

REFERENCES

- Chen, Y., & Boukerche, A. (2020). A Novel Lane Departure Warning System for Improving Road Safety. In *ICC 2020-2020 IEEE International Conference on Communications (ICC)* (pp. 1-6). IEEE Conference Publishing. <https://doi.org/10.1109/ICC40277.2020.9149085>

- El Hajjouji, I., Mars, S., Asrih, Z., & El Mourabit, A. (2020). A novel FPGA implementation of hough transform for straight lane detection. *Engineering Science and Technology, an International Journal*, 23(2), 274-280. <https://doi.org/10.1016/j.jestch.2019.05.008>
- Feniche, M., & Mazri, T. (2019). Lane detection and tracking for intelligent vehicles: A survey. In *2019 International Conference of Computer Science and Renewable Energies (ICCSRE)* (pp. 1-4). IEEE Conference Publishing. <https://doi.org/10.1109/ICCSRE.2019.8807727>.
- Guan, J., An, F., Zhang, X., Chen, L., & Mattausch, H. J. (2017). Real-time straight-line detection for XGA-size videos by hough transform with parallelized voting procedures. *Sensors*, 17(2), Article 270. <https://doi.org/10.3390/s17020270>
- Hwang, S., & Lee, Y., (2016). FPGA-based real-time lane detection for advanced driver assistance systems. In *Proceedings of 2016 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)* (pp. 218-219). IEEE Conference Publishing. <https://doi.org/10.1109/APCCAS.2016.7803937>
- Khongprasongsiri, C., Kumhom, P., Suwansantisuk, W., Chotikawanid, T., Chumpol, S., & Ikura, M. (2018). A hardware implementation for real-time lane detection using high-level synthesis. In *2018 International Workshop on Advanced Image Technology (IWAIT)* (pp. 1-4). IEEE Conference Publishing. <https://doi.org/10.1109/IWAIT.2018.8369730>
- Lee, D. K., Shin, J. S., Jung, J. H., Park, S. J., Oh, S. J., & Lee, I. S. (2017). Real-time lane detection and tracking system using simple filter and Kalman filter. In *2017 Ninth International Conference on Ubiquitous and Future Networks (ICUFN)* (pp. 275-277). IEEE Conference Publishing. <https://doi.org/10.1109/ICUFN.2017.7993792>
- Lu, X., Song, L., Shen, S., He, K., Yu, S., & Ling, N. (2013). Parallel hough transform-based straight line detection and its FPGA implementation in embedded vision. *Sensors*, 13(7), 9223-9247. <https://doi.org/10.3390/s130709223>
- Marzotto, R., Zoratti, P., Bagni, D., Colombari, A., & Murino, V. (2010). A real-time versatile roadway path extraction and tracking on an FPGA platform. *Computer Vision and Image Understanding*, 114(11), 1164-1179. <https://doi.org/10.1016/j.cviu.2010.03.015>
- Panda, P. R., Sharma, N., Kurra, S., Bhartia, K. A., & Singh, N. K. (2018). Exploration of loop unroll factors in high level synthesis. In *2018 31st International Conference on VLSI Design and 2018 17th International Conference on Embedded Systems (VLSID)* (pp. 465-466). IEEE Conference Publishing. <https://doi.org/10.1109/VLSID.2018.115>
- Promrit, P., & Suntiamorntut, W. (2017, July). Design and development of lane detection based on FPGA. In *2017 14th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-4). IEEE Conference Publishing. <https://doi.org/10.1109/JCSSE.2017.8025909>
- Solod, P., Sengchuai, K., Booranawong, A., Hoyingcharoen, P., Chumpol, S., Ikura, M., & Jindapetch, N. (2018, October 30 - November 2). Model based design approach for road lane tracking [Paper presentation]. In *Asia Pacific Conference on Robot IoT System Development and Platform 2018 (APRIS2018)*. Prince of Songkla University (PSU) Phuket Campus, Thailand.

Farmers' Participation in Irrigation Management in the Punggur Utara Irrigation Area, Indonesia

Dyah Indriana Kusumastuti*, Vera Chania Putri, Dwi Jokowinarso and Endro Prasetyo Wahono

Department of Civil Engineering, Engineering Faculty, University of Lampung, Jl. Sumantri Brojonegoro No.1, Bandar Lampung 35145, Indonesia

ABSTRACT

Farmers' participation in water resource management in Indonesia has been accommodated through Law No. 7/2004 and Government Regulation No. 20/2006 regarding irrigation. In government regulations, farmers' participation in irrigation water management has been described in detail, with one of the components of irrigation network management being operations and maintenance. Most irrigation system problems are related to irrigation networks. Farmers' participation in the Water Users Association (WUA), a farmer institution responsible for managing and developing irrigation networks at the tertiary level, determines how irrigation network management activities are carried out. The purpose of this study was to determine the effect of the socio-economic characteristics of farmers and of group participation of a WUA in carrying out irrigation management activities in operations and maintenance activities. The result is connected to the physical condition of the irrigation network. The research took place in the village of Sumber Rejo in the

Punggur Utara irrigation area, which is in Lampung Province, Indonesia, and is where WUA Harapan Maju is active. Samples were determined based on a random sampling method, which found that the sample size required for the study was 80 farmers. The results show that of socio-economic factors, the most influential is the number of farmers' dependents, followed by the land area owned by each farmer. It was also found that group participation of WUA Harapan Maju

ARTICLE INFO

Article history:

Received: 07 September 2020

Accepted: 25 November 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.02>

E-mail addresses:

kusumast@gmail.com (Dyah Indriana Kusumastuti)

chania.vera@yahoo.com (Vera Chania Putri)

d.jokowinarso@gmail.com (Dwi Jokowinarso)

epwahono@eng.unila.ac.id (Endro Prasetyo Wahono)

* Corresponding author

was categorised as “barely” adequate. This result was confirmed by the condition of the irrigation system in the Punggur Utara irrigation area, which was also “barely” adequate, and the system’s operations and maintenance needed to be improved.

Keywords: Farmer, group participation, irrigation management, operations and maintenance, WUA

INTRODUCTION

Irrigation management plays an important role in agriculture, as irrigation is a key factor that may affect farmer productivity (Azhar et al., 2014; Carracelas et al., 2019; Dirwai et al., 2019; Hudzari et al., 2016; Mateos et al., 2002; Meena et al., 2019; Valipour, 2014). In developing countries, irrigation management systems have not performed well in the past because their management was entirely (and poorly) controlled by governments (Abernethy, 2010). Due to weak institutional arrangements, maintenance cannot be performed optimally, which causes both physical infrastructure and infrastructure management to deteriorate. These problems have impacted efficient and timely water delivery and farmer productivity in turn.

Since the 1980s, instruments of reform have been implemented in many developing countries (Alaerts, 2020; Baldwin et al., 2016; Mollinga & Bolding, 2004; Sehring, 2007). The most significant paradigm of these reforms is a shift from construction and rehabilitation to operations and maintenance, which can improve infrastructure conditions (Suhardiman & Giordano, 2014). To improve infrastructure conditions in Indonesia, the reforms introduced the formation of Water Users Associations (WUAs) (Gany et al., 2019). These involve farmers in system management and irrigation management transfer, thereby allowing farmers greater control and responsibility (Mukherji et al., 2009).

Irrigation management transfer is a reform package aimed at full or partial devolvement of irrigation management responsibilities from the government to farmers themselves through the establishment of WUAs (FAO, 2007; Mukherji et al., 2009). Participatory irrigation management and irrigation management transfer became the national irrigation strategies for most developing countries and are considered the best solution for irrigation management in developing countries (Fulazzaky, 2017; Poddar et al., 2011).

In irrigation management, it is vital to perform collective actions based on farmers’ collaborative efforts or cooperation in WUAs. Even so, the results are mixed. Some cases of collective action for irrigation management face difficulties (Letsoalo & Van Averbeke, 2004; Moustafa, 2004), whereas other cases are successful (Ghosh et al., 2019; Huamanchumo et al., 2008; Rap, 2006; Vandersypen et al., 2008).

In Indonesia, the farmers’ participatory approach to water resource management has been accommodated by Law No. 7/2004. In compliance with the terms of article 41 of Law No. 7/2004, Government Regulation No. 20/2006 (PP No. 20/2006) on irrigation

has formulated a more detailed description of farmers' participation in irrigation water management. One of the components of the irrigation networks scheme management is operations and maintenance and is related to the daily activities of farmers on their irrigated lands. Such activities target reducing government subsidies, improving water-use efficiency, and improving annual food production (Fulazzaky, 2017).

Farmers are the main actors in all operational activities and irrigation maintenance. In carrying out the series of activities involved in irrigation, individual farmers are influenced by many factors, one being farmers' socio-economic characteristics. These characteristics include age, education level, duration of farming, the position of the farmer in the organisation, family size, land area, and income level (Khalkheili & Zamani, 2009). Apart from these characteristics, another influential element in a series of operations and maintenance activities is the group dynamics of the WUA. Therefore, in this study, the effect of farmers' socio-economic aspects, as well as the dynamics of a WUA are reviewed to understand the effect of operations and maintenance activities, both in physical and non-physical terms. The study is conducted by distributing questionnaire to the group members to gather the information about the socio-economic characteristics of the farmers as well as data collection and field observation to assess the physical condition of the irrigation network.

WUA Group Participation

Within a WUA group, cooperation among farmers is necessary to determine and influence the behaviour of the group and its members to effectively achieve the group's aims. Achieving group aims depends on the energy and creativity of members in carrying out their activities. In other words, the development of farmer groups depends on the internal dynamics of the groups, which are determined by the group members. Assessing group dynamics means assessing the strength contained in a group, as this determines group behaviour and group members' ability to achieve the group's goals. Ghosh et al. (2019) used several parameters to assess the group dynamics of some WUAs in India, such as group participation, decision-making procedures, group atmosphere, empathy, interpersonal trust, and social support. Supriyadi (2008) conducted similar research in the Polokarto District, Sukoharjo Regency, Central Java Province, Indonesia. Supriyadi's study found that group dynamics such as group goals, group structure, task functions, coaching and maintenance, cohesiveness, and pressure within the group were related to operations and maintenance of irrigation infrastructure. In contrast, the effectiveness of the group and its hidden agenda were unrelated to operations and maintenance of irrigation infrastructure. Other research was conducted by Desrimon and Fiatno (2018) on the influence of social and economic aspects on operations and maintenance. This research found that social, economic, and institutional factors simultaneously influence aspects of operations and

maintenance. These findings have motivated the current researchers to investigate irrigation system management in terms of the socio-economic part. In addition to this, the influence of aspects of a WUA in terms of group participation towards operational activities and maintenance of irrigation networks will be investigated.

METHODS

This research was conducted in the Punggur Utara irrigation area, located in the village of Sumber Rejo in the Kota Gajah district of the Central Lampung Regency. The Punggur Utara irrigation area is part of the Sekampung irrigation system area, located at coordinates $105^{\circ}28'E - 105^{\circ}33'E$ and $4^{\circ}54'S - 5^{\circ}01'S$ (Figure 1). Irrigation water for this area is conveyed through the Argoguruh Dam, which uses the Sekampung River. Argoguruh's weir is located at coordinates $5^{\circ}11'57''S - 105^{\circ}10'46''E$.

The Punggur Utara irrigation area consists of 26 units, and this study focused on one of them: Unit 15, otherwise known as BPU 15 (BPU stands for Punggur Utara structures in its Indonesia language). This unit passes through five villages (Sumber Rejo, Purwo Rejo,

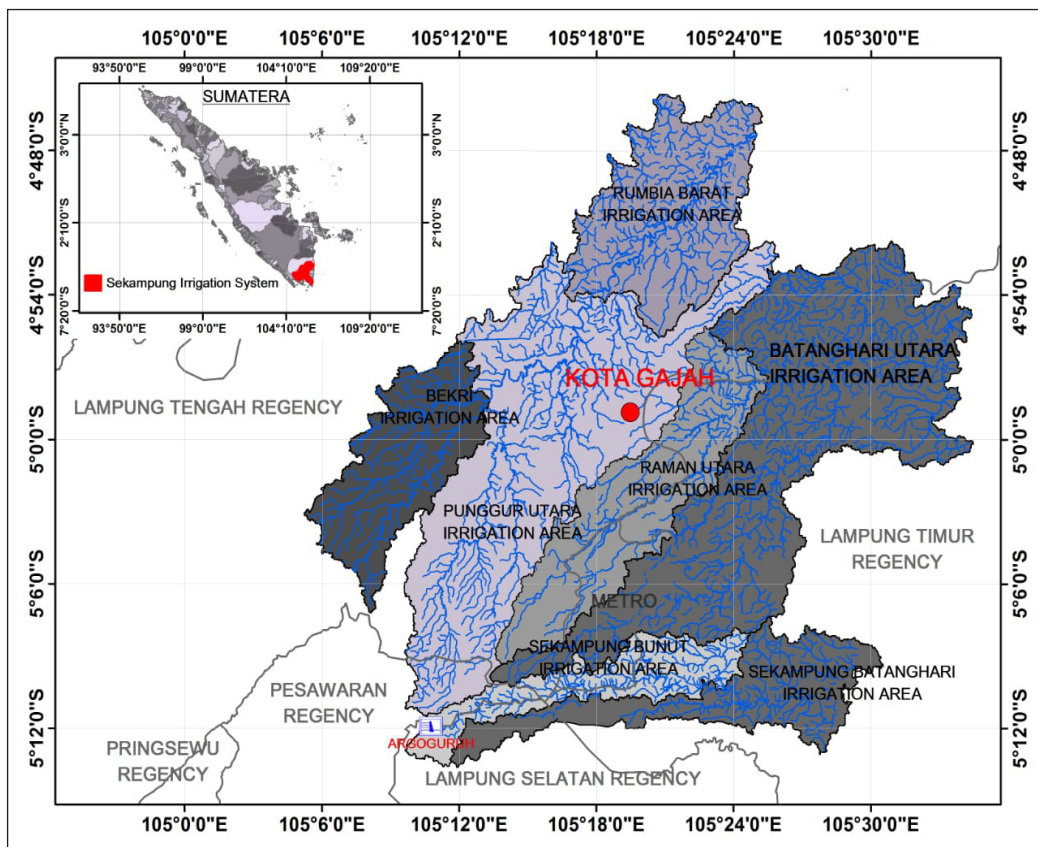


Figure 1. Study area

Bangun Rejo, Buyut Ilir, and Buyut Utara), and each village is managed by a WUA. This research was conducted in the village of Sumber Rejo, where the potential and functional areas of paddy rice fields are 304 ha and 238 ha, respectively.

WUA Harapan Maju at Sumber Rejo Village

WUA Harapan Maju already has a legitimate organisational structure, consisting of a chairman, secretary, treasurer, and an irrigation water distribution regulator who has different duties and responsibilities. The position, status, and role of the caretaker, along with other group members, is described in detail in the group's organisational structure. In this study, the group participation of WUA Harapan Maju was studied. A questionnaire was distributed to group members and included several assessment categories that aimed to represent group participation.

The sample was determined using the probability sampling method (random sample), with a cluster sampling technique. Samples were taken from the distribution of tertiary channels in Sumber Rejo Village. Then, as many samples as were needed were randomly drawn from the cluster. The number of farmers who joined the WUA Harapan Maju, Sumberejo Village is 434. To determine the number of samples to be examined, the Slovin formula was used, and by assuming a margin error of 10%, the sample size obtained was 80 farmers.

Variable

In this research, independent variables consist of some socio-economic characteristics of the farmers, and the dependent variable consists of operational and maintenance irrigation (Table 1).

Table 1
Dependent and independent variables

Independent variable (X)	Dependent variable (Y)
Individual Farmers' Characteristics	Irrigation management (Y)
• Age (X_1)	
• Level of Education (X_2)	
• Farming experience (X_3)	
• Size of Land (X_4)	
• Number of Family Dependents (X_5)	
• Income (X_6)	
Group Characteristics	
• Group Participation (X_7)	

Group Participation Criteria

In this research, group dynamics focus on group members' participation, with indicators of group participation shown in Table 2. These indicators include participation in group goals, participation in group activities, participation in decision making and participation in interaction. Table 3 presents variables of irrigation management, including irrigation channel network, irrigation structure, maintenance, irrigation water, frequency of water distribution, and suitability of the schedule of water distribution.

Before doing further analysis, validity and reliability tests were performed to evaluate the accuracy and precision of the variables, as well as the consistency and stability of the questionnaire. The result of the validity test for variables 1 to 7 (representing farmers' age, farmers' education, farmers' dependents, farmers' experience, farmers' land area, farmers' income, and group participation respectively) shows that t_{computed} values for most variables are larger than t_{table} values, so that the results of the validity test are valid, except for variable 1, farmers' age. Furthermore, the reliability test found that R_{computed} is 0.3425, which is larger than R_{table} , meaning that the questionnaire is consistent and robust.

Each of the seven variables is grouped into three categories and scored 1 to 3 in the ascending order. For example, level of education (X2) is grouped into elementary school, junior high school, and senior high school and higher. For analysis purpose, it is given the scores of 1, 2 and 3 for elementary school, junior high school, and senior high school and higher, respectively. Likert scale is used in the analysis using the scores of 1, 2 and 3 assigned to be poor, adequate, and good, respectively. Correlation analysis between variables is analysed using the Pearson product moment. Then multiple linear regression analysis is applied to quantify the strength of the relationship between the dependent and the independent variables.

Table 2
Variables of group participation

Variable	Indicator	Number of questions
Group Participation	a. Participation in group goals	7
	b. Participation in group activities	11
	c. Participation in decision making	7
	d. Participation in interaction	3

Table 3
Variables of irrigation management

Variable	Indicator
Irrigation Management	a. Irrigation channel network
	b. Irrigation structure
	c. Maintenance
	d. Irrigation water
	e. Frequency of water distribution
	f. Suitability of the schedule of water distribution

RESULTS AND DISCUSSION

Characteristics of Respondents

Based on the calculation of the sample required, 80 respondents participated in this study. The farmers' answers on the questionnaire are classified into three categories for each variable. Age is one aspect of demographic characteristics that can influence participants' activities. It is known from the questionnaire that the youngest respondent is 22 years old, and the oldest is 76 years old. The average age of respondents in the village of Sumberejo is 49.3 years old. Based on the division of farmers' interval categories, the percentage of farmers aged between 21–40, 41–59, and 60–80 years old is 30%, 42.5%, and 27.5% respectively. The survey suggests that the highest percentage being 42.5% for the interval of 41–59 years old.

Farmers dependents are classified into three categories (i.e., 0–2 children, 3–4 children, and 5–6 children). It was found from the questionnaire that the largest percentage is 61.25% for farmers having 3–4 children, and the smallest percentage is 15% for farmers having 0–2 children. It can be said that the middle category, with 3–4 children, makes up most farmers' dependents.

The education levels of the respondents are classified into three categories (i.e., elementary school, junior high school, and senior high school and higher). The survey found that the largest percentage of respondents were formally educated until elementary school (48.75%), while the smallest percentage of respondents were educated until senior high school or higher (21.25%). This means that the farmers participating in WUA Harapan Maju still have a low level of education.

The farming experience of the respondents is divided into three categories (2–21 years, 22–41 years, and 42–60 years), with the highest percentage being 47% for 2–21 years' experience in farming, and the lowest being 25% for farmers having 42–60 years' experience in agriculture. The results suggest that most farmers have the least amount of farming experience (i.e., between 2 to 21 years).

The land area owned by respondents is classified into three categories. The percentages for land areas of 0.125–0.425 ha, 0.426–0.726 ha, and ≥ 0.727 ha are 33.75%, 32.5%, and 33.75% respectively. It is observed that there is no significant difference in the percentage of the land area owned by respondents.

The level of income is divided into three categories: <Rp2,000,000 per month, Rp2,000,000 to Rp4,900,000 per month, and \geq Rp5,000,000 per month. The largest percentage is 61.25% (i.e., farmers with income from farming of <Rp2,000,000, while the smallest percentage is 10% for farmers with farm income \geq Rp5,000,000). It was found that farmers joining WUA Harapan Maju are still in the category of having a low farming income.

The result of farmers' characteristics in WUA Harapan Maju, Sumber Rejo Village, Kota Gajah District, Central Lampung Regency is shown in Table 4.

Table 4
Résumé of farmers' characteristics

Category	Percentage
Age in years (X_1)	
22– 0	30
41–59	42.5
60–76	27.5
Level of education (X_2)	
Elementary School	48.75
Junior High School	30
≥ Senior High School	21.25
Farming experience (years) (X_3)	
2–21	47.5
22–41	27.5
42–60	25
Size of Land (X_4)	
0.125–0.425	33.75
0.426–0.726	32.5
≥ 0.727	33.75
Number of family dependents (X_5)	
0–2	15
3–4	61.25
5 - 6	23.75
Income in rupiahs (X_6)	
< Rp2,000,000	61.25
Rp,2000,000–Rp4,900,000	26.25
≥ Rp. 5,000,000	12.50

Farmers' Participation in WUA Harapan Maju

Based on the 28-question survey, the respondents' highest total score was 67, and the lowest was 41, with an average score of 58. Table 5 shows the scores for each variable in Group Participation. Based on that result, the farmers' participation in WUA Harapan Maju, Sumber Rejo Village is "barely" adequate. Farmers care about the group's goals, and farmers interact well, which is shown by the scores of 2.580 and 2.325 for the variable's

Table 5
Average score for the group participation of WUA Harapan Maju

Group Participation	Av. Total Score	Std Dev.	Average Score
Goal	18	1.925	2.580
Activities	18	3.700	1.636
Decision Making	15	2.597	2.107
Interaction	7	1.302	2.325
			2.162

Remarks: responses are weighted 1–3 (from poor to good)

“goal” and “interaction”, respectively. In decision making, the participation of farmers is sufficient, as evidenced by the score of 2.107. However, only some farmers are actively involved in all operations and maintenance activities, which is shown by the average score of 1.636 for the “activities” variable.

Irrigation Management

Regarding irrigation management, some aspects respondents were asked about include the performances of the irrigation channel network; irrigation structures; maintenance of the irrigation channel and structures; sufficiency of irrigation water for paddy fields; sufficiency of water distribution during one planting period; and suitability of water distribution with the predetermined schedule. Table 6 presents the average scores of the variables in irrigation management. The scores which have scores greater than 2 are “maintenance”, followed by “frequency of water distribution”, while other variables score less than 2. In general, farmers believe that the irrigation management needs to be improved shown by the average score of all variables, 1.875, which is categorised as “barely” adequate.

Table 6
Average score of irrigation management

Operations and Maintenance	Standard Deviation	Average Score
Irrigation channel network	0.600	1.763
Irrigation structures	0.569	1.575
Maintenance	0.656	2.489
Irrigation water	0.613	1.438
Frequency of water distribution	0.792	2.175
Suitability of the schedule	0.828	1.813
		1.875

Remarks: responses are weighted 1–3 (from poor to good)

Irrigation Network in Sumber Rejo Village

The existing condition of primary and secondary irrigation channels in the Punggur Utara irrigation area is mostly made of permanent structures. For tertiary channels, especially in the research area (including BPU 15 Ki 1, BPU 15 Ki 2, BPU 15 Ka, and BF 1 Ka), almost the entire channels are made of mounted stone. Figure 2 shows the top view of the study area, with the irrigation structures of BPU 15 Ki 1, BPU 15 Ki 2, BPU 15 Ka, and BF 1 Ka. Figure 3 presents the conditions of the gates, demonstrating that the surrounding of gate BF 1 Ka was full of trash. Figure 4 shows some cracks in several parts of the channels, as well as weeds growing on the channel banks.

The sufficiency of the irrigation network as a means for delivering irrigation water to the field can be measured through the channel density (CD) and structure density (SD) of irrigation networks. CD is defined as channel length (m) divided by the served area (ha), and it can be considered as sufficient when $CD \geq 50$ m/ha. Structure density, SD, is defined as the number of structures divided by the total area of the paddy field. The structure density is considered to be sufficient when $SD \geq 0.1$.

Information regarding the irrigation network in Sumber Rejo Village, including the channel length, land area, tertiary and quarterly boxes, and culverts, were obtained from the Punggur Utara Irrigation Area Office. Table 7 presents the density for the channels, and Table 8 shows the density of the structures. From the channel and structure densities presented in Tables 8 and 9, it can be said that the tertiary irrigation canal of BPU 15 at Sumber Rejo Village is inadequate both for channel and network densities.

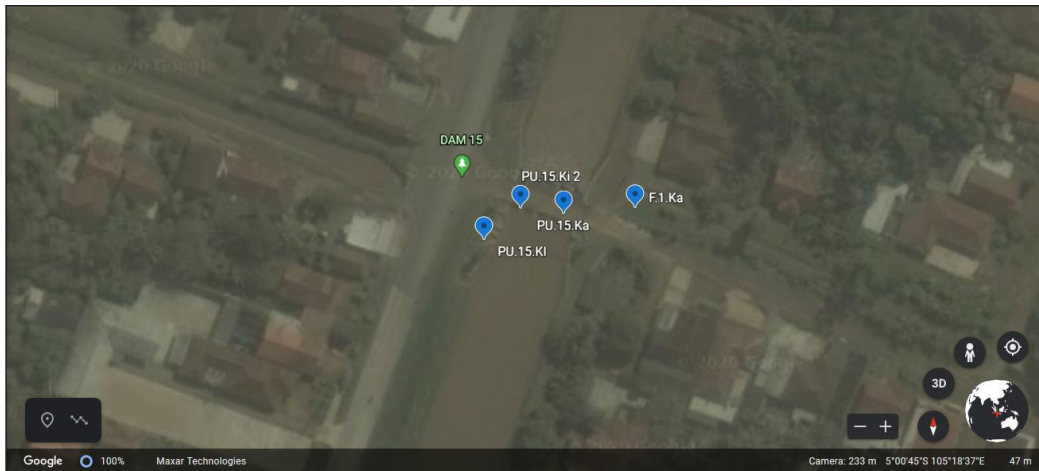


Figure 2. Top view of the study area, with the irrigation structures of BPU 15 Ka, BPU 15 Ki1, BPU 15 Ki2, and BF 1 Ka

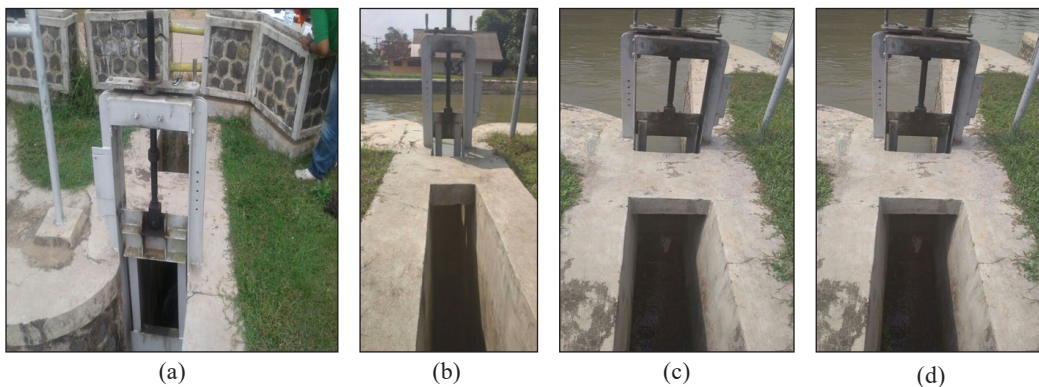


Figure 3. The gate conditions of (a) BPU 15 Ka, (b) BPU 15 Ki1, (c) BPU 15 Ki2, and (d) BF 1 Ka (from left to right respectively) in the Punggur Utara irrigation area at Sumber Rejo Village



Figure 4. Channel conditions in the Punggur Utara irrigation area at Sumber Rejo Village

Table 7
Structures in BPU 15 Tertiary Canal at Sumber Rejo Village

Types of structures	BPU 15 Tertiary Channels			
	BPU 15 Ki 1	BPU 15 Ki 2	BPU 15 Ka	BF 1 Ka
Channel length (m)	400	2100	2650	2015
Potential area (ha)	30	40	107	107
Functional area (ha)	25	45	73	98
Tertiary Box (unit)	1	2	4	2
Quarterly Box (unit)	-	1	4	1
Culverts (unit)	-	-	-	-

Table 8
The density of Tertiary Canal BPU 15 at Sumber Rejo Village

No	Channels	A (ha)	Channel length (m)	CD (m/ha)	Remarks
1.	BPU 15 Ki 1	25	400	16	Inadequate
2.	BPU 15 Ki 2	45	2100	46.67	Inadequate
3.	BPU 15 Ka	73	2650	36.30	Inadequate
4.	BF 1 Ka	98	2015	20.56	Inadequate
Average				29.88	Inadequate

Table 9
The density of structures in BPU 15 At Sumber Rejo

No	Channels	A (ha)	Total boxes and culverts	SD (unit/ha)	Remarks
1.	BPU 15 Ki 1	25	1	0.04	Inadequate
2.	BPU 15 Ki 2	45	3	0.07	Inadequate
3.	BPU 15 Ka	73	8	0.07	Inadequate
4.	BF 1 Ka	98	3	0.03	Inadequate
Average				0.05	Inadequate

Statistical Analysis

Correlation between variables was conducted using the Pearson product moment, and the regression statistics found that the R² determination coefficient was 0.78, which is greater than 0.22, the Strongswan R table for a 5% error level for the number of samples being 80. This shows that there is a positive and significant relationship between the independent variables and dependent variables. It can be said that there is a significant influence between variables of age (X1), education (X2), total dependents (X3), farming experience (X4), land size (X5), farming income (X6), and group participation (X7) on water management (Y), amounting to 78%, with the influence of variations outside the model being 22%.

As there is a correlation between variables and a significant impact of the independent variables to the dependent variable, the next step was to perform multiple linear regression analysis to find the appropriate equation. Table 10 shows the result of this analysis. Based on the Table 10, we know that the P-value of variable X1 is 0.5411 larger than the critical confidence limit, which is 0.05. This means that age, variable X1, has no effect on the variable of operations and maintenance (Y). Omitting variable X1, but still using the result of multiple linear regression analysis, the Equation 1 becomes as follows:

$$Y = 1.7377 + 0.72 X2 + 1.19 X3 + 0.76 X4 + 0.76 X5 + 0.74 X6 + 0.92 X7 \quad (1)$$

Table 10
Multiple linear regression

Intercept	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
	1.7377	0.7529	2,3081	0,0239	0,2369	3,2386
X1	-0.1175	0.1914	-0.6140	0.5411	-0.4992	0.2641
X2	0.7096	0.1572	4.5138	0.0000	0.3962	1.0229
X3	1.1907	0.1873	6.3565	0.0000	0.8173	1.5641
X4	0.7637	0.2102	3.6336	0.0005	0.3447	1.1827
X5	0.7620	0.1617	4.7125	0.0000	0.4396	1.0843
X6	0.7420	0.1732	4.2849	0.0001	0.3968	1.0872
X7	0.9282	0.1466	6.3307	0.0000	0.5369	1.2205

Interpretation of the Results

Based on the results of the study, all variables have a direct relationship to the irrigation management, which in this case, mostly related to the farmers' irrigation networks. However, of all these factors, age does not have a significant relationship or has a very low relationship. This is supported by the results of the validity test that found the age variable was invalid and the result of regression analysis, showing that the age variable has no effect on operations and maintenance activities. This means that regardless of age,

farmers, both young and old, will be involved in carrying out operations and maintenance activities in the irrigation network.

The most influential variable on operations and maintenance activities is farmers' dependents (X3), being 1.19. This indicates that the greater the number of farmers' dependents, the more actively involved the farmers will be in operational and maintenance activities. Farmers who have more dependents are aware that the operation and maintenance of their irrigation network will have an impact on their farming and thus their economic situations. Thus, farmers with more dependents are willing to get more involved in operations and maintenance to meet their economic needs.

The next significant variable of the individual farmer on the activities in operations and maintenance is the land area variable (X5). This shows that farmers' land area determines farmers' involvement in operations and maintenance activities in the irrigation network. Farmers who have a larger land area will be more active in carrying out operations activities and maintenance of irrigation networks for the good of their farming. They will pay more attention to how the operating system works so that their land can be irrigated well, and farmers will often maintain their irrigation network, such as through cleaning trash and weeds from canals and repairing damaged channels.

Farming income (X6) presented in this study is the whole incomes of the farmers both from farming and non-farming. From all respondents, 61.25% have a low income, 26.25% have medium income and only 12.5% have a high income. Some studies show that participation in non-farming activities has a positive impact on household welfare by raising income levels (Ali et al., 2015; Zereyesus et al., 2017). However, there is a concern if non-farming activities generate much income, farmers may pay less attention to the operation and maintenance activities and even leave their farm. A study showed that there is social economic motivation for farmer to leave farming sector in Central Java, Indonesia (Farhani, 2009).

The constant of 1.7377 indicates that the level of the operations and maintenance performed by farmer members of WUA Harapan Maju is not considered good. This is based on the constant value obtained from the regression analysis, which is less than 2. This value is based on the scoring of the answers in the questionnaire distributed to the respondents. The results of the evaluation of operations and maintenance activities based on respondents' answers are as follows:

- Existing irrigation networks are still not sufficient for the farmers. This is also supported by the results of the channel density, which has an average value of 29.88. This value is below the standard of channel density ranging from 50 to 100 m/ha.
- Well-functioning irrigation structures are less than 50%. This is also supported by irrigation structure density, which has an average value of 0.05. This value is below the sufficient values of structure density ranging from 0.11 to 0.40 units/ha.

- The irrigation water discharge is relatively small. By irrigating the land four to five times during one planting period, farmers feel that they have sufficient water. However, water distribution sometimes does not suit the predetermined schedule.

This study has shown that the operation and maintenance activities performed by the WUA farmer members in Harapan Maju was not good. Meanwhile, the results of the analysis show that both channel and irrigation structure densities are not met the common standards. The adequacy of the irrigation infrastructure seems to have effect on farmer participation in operation and maintenance. Therefore, if the government conducts such irrigation works in the future, so that the density of the structures and channels is sufficient (increased more in comparing to current status) farmers are expected to carry out the operation and maintenance activities better.

It is also suspected that poor operation and maintenance activities done by the farmers are also influenced by their low income. In the Government Regulation of the Republic of Indonesia No. 20/2006 (related to irrigation) it states that the participation of the farmer community should be in the form of contributions, orientation, time, energy, materials, and funds. As mentioned earlier, the majority of the farmer members (i.e. 61.25%) have low income, so that they may not be able to contribute sufficiently in terms of funding for the operation and maintenance activities.

It becomes one of the recommendations of this study, that the government improves irrigation infrastructure especially the densities of channels and structures. In addition, to enhance the participation of farmers in irrigation scheme management, especially in operations and maintenance of hydraulic works, the government may provide incentives scheme to WUAs which capable of doing the operation and maintenance well. The incentives may be awarded periodically based on selection system at city/regency, province, or national levels.

CONCLUSION

Based on the characteristics of farmers joining in WUA Harapan Maju at Sumber Rejo Village, it can be concluded that the major motivations for being more active in irrigation management are the number of farmers' dependents and the land area owned by farmers. Interestingly, farmers' age does not have a significant relationship to irrigation management. Overall, the group participation of WUA Harapan Maju is "barely" adequate, although farmers' low activity levels regarding operations and maintenance can be shown by trash found near the irrigation gate and some unrepaired channels. The performance of irrigation management by WUA Harapan Maju at Sumber Rejo can be categorised as "barely" adequate and needs to be addressed. Irrigation management can be improved by getting farmers to participate more in operation and maintenance, as well as by improving the irrigation structures and networks.

ACKNOWLEDGEMENT

The authors thank the Ministry of Research and Higher Education Indonesia for the funding through applied research scheme.

REFERENCES

- Abernethy, C. L. (2010). Governance of irrigation systems: does history offer lessons for today? *Irrigation and Drainage: The journal of the International Commission on Irrigation and Drainage*, 59(1), 31-39. <https://doi.org/10.1002/ird.552>
- Alaerts, G. J. (2020). Adaptive policy implementation: Process and impact of Indonesia's national irrigation reform 1999-2018. *World Development*, 129, 1-14. <https://doi.org/10.1016/j.worlddev.2020.104880>
- Ali, A., Erenstein, O., & Rahut, D. B. (2015). Impact of gender participation in non-farming activities on household income and poverty levels in Pakistan. *Work*, 52(2), 345-351. <https://doi.org/10.3233/WOR-152103>
- Azhar, A. W. M., Nashriyah, M., Hudzairi, H. R. M., Moneruzzaman, K. M., Amir, H. S., Rohaizad, M. R. M., & Ali, A. M. (2014). Effects of irrigation frequencies on aerial agro-morphological parameters of *Dioscorea hispida* Dennst. (Dioscoreaceae). *Journal of Applied Sciences Research*, 8(9), 27-37. <http://dx.doi.org/10.2139/ssrn.2812557>
- Baldwin, E., Washington-Ottombre, C., Dell'Angelo, J., Cole, D., & Evans, T. (2016). Polycentric governance and irrigation reform in Kenya. *Governance: An International Journal of Policy, Administration, and Institution*, 29(2), 207-255. <https://doi.org/10.1111/gove.12160>
- Carracelas, G., Hornbuckle, J., Rosas, J., & Roel, A. (2019). Irrigation management strategies to increase water productivity in *Oriza sativa* (rice) in Uruguay. *Agricultural Water Management*, 222, 161-172. <https://doi.org/10.1016/j.agwat.2019.05.049>
- Desrimon, A., & Fiatno, A. (2018). Analysis of the impact of social, economic and group aspects of Water User Association on operation and maintenance of irrigation network. *Integrated Industrial Engineering Journal*, 1, 46-54.
- Dirwai, T. L., Senzanje, A., & Mudhara, M. (2019). Water governance impacts on water adequacy in smallholder irrigation schemes in KwaZulu-Natal province, South Africa. *Water Policy*, 21(1), 127-146. <https://doi.org/10.2166/wp.2018.149>
- Farhani, A. (2009). *Socio economic motivation of farmers changing their job from farming sector to furniture industrial sector in Serenan village Juwiring district Klaten regency* [Thesis]. Sebelas Maret University, Surakarta, Indonesia.
- FAO. (2007). *Irrigation management transfer: Worldwide efforts and results*. Food and Agricultural Organization of the United Nations.
- Fulazzaky, M. A. (2017). Participation of farmers in irrigation water management in Indonesia: A review. *Irrigation and Drainage*, 66, 182-191. <https://doi.org/10.1002/ird.2085>

- Gany, A. H. A., Sharma, P., & Singh, S. (2019). Global review of institutional reforms in the irrigation sector for sustainable agricultural water management, including water users' associations. *Irrigation and Drainage*, 68(1), 84-89. <https://doi.org/10.1002/ird.2305>
- Ghosh, S., Kolady, D. E., Das, U., Gorain, S., Srivastava, S. K., & Mondal, B. (2019). Spatio-temporal variations in effects of participatory irrigation management (PIM) reform in India: A panel data analysis. *Agricultural Water Management* 222, 48-61. <https://doi.org/10.1016/j.agwat.2019.05.042>
- Huamanchumo, J., Peñay, S. L., & Hendriks, J. (2008). Developing capacity in water users' organisations: The case of Peru. *Irrigation and Drainage* 57, 300-310. <https://doi.org/10.1002/ird.432>
- Hudzari, R. M., Syazili, R., Ssomad, M. A. H. A., Shokeri, A. F. M., & Husin, N. A. (2016). Design and development of innovative highland water filtration system. *World Journal of Engineering and Technology*, 4, 383-390. <https://doi.org/10.4236/wjet.2016.43037>
- Khalkheili, T. A., & Zamani, G. H. (2009). Farmer participation in irrigation management: The case of Doroodzan dam irrigation network, Iran. *Agricultural Water Management*, 96(5), 859-865. <https://doi.org/10.1016/j.agwat.2008.11.008>
- Letsoalo, S., & Van Averbeke, W. (2004, November 8-11). When water is not enough: Institutions, organisations, and conflicts surrounding the sharing of irrigation water at a smallholder scheme in South Africa. In *Proceedings of the Workshop on 'Water Management for Local Development'*, WRM (pp. 362-373). Loskopdam, Pretoria.
- Mateos, L., López-Cortijo, I., & Sagardoy, J. A. (2002). SIMIS: the FAO decision support system for irrigation scheme management. *Agricultural Water Management*, 56(3), 193-206. [https://doi.org/10.1016/S0378-3774\(02\)00035-5](https://doi.org/10.1016/S0378-3774(02)00035-5)
- Meena, R. P., Karnam, V., Tripathi, S. C., Jha, A., Sharma, R. K., & Singh, G. P. (2019). Irrigation management strategies in wheat for efficient water use in the regions of depleting water resources. *Agricultural Water Management*, 214, 38-46. <https://doi.org/10.1016/j.agwat.2019.01.001>
- Mollinga, P., & Bolding, A. (2004). *The politics of irrigation reform*. Ashgate Publishing.
- Moustafa, M. M. (2004). Can farmers in Egypt shoulder the burden of irrigation management? *Irrigation and Drainage Systems*, 18(2), 109-125. <https://doi.org/10.1023/B:IRRI.0000040249.12754.c8>
- Mukherji, A., Fuleki, B., Shah, T., Suhardiman, D., Giordano, M., & Weligamage, P. (2009). *Irrigation reform in Asia: A review of 108 cases of irrigation management transfer*. International Water Management Institute.
- Poddar, R., Qureshi, M. E., & Syme, G. (2011). Comparing irrigation management reforms in Australia and India - A special reference to participatory irrigation management. *Irrigation and Drainage*, 60(2), 139-150. <https://doi.org/10.1002/ird.551>
- Rap, E. (2006). The success of a policy model: Irrigation management transfer in Mexico. *Journal of Development Studies*, 42(8), 1301-1324. <https://doi.org/10.1080/00220380600930606>
- Sehring, J. (2007). Irrigation reform in Kyrgyzstan and Tajikistan. *Irrigation and Drainage Systems*, 21, 277-290. <https://doi.org/10.1007/s10795-007-9036-0>

- Suhardiman, D., & Giordano, M. (2014). Is there an alternative for irrigation reform? *World Development* 57, 91-100. <https://doi.org/10.1016/j.worlddev.2013.11.016>
- Supriyadi, E. (2008). *The relationship between group dynamics of Water User Association and maintenance of irrigation infrastructure in Polokarto District Sukoarjo Regency*. [Unpublished dissertation]. Universitas Sebelas Maret, Surakarta, Indonesia.
- Valipour, M. (2014). A comprehensive study on irrigation management in Asia and Oceania. *Archives of Agronomy and Soil Science*, 61(9), 1247-1271. <https://doi.org/10.1080/03650340.2014.986471>
- Vandersypen, K., Keita, A. C. T., Lidon, B., Raes, D., & Jamin, J. Y. (2008). Didactic tools for supporting participatory water management in collective irrigation schemes. *Irrigation and Drainage Systems*, 22(1), 103-113. <https://doi.org/10.1007/s10795-007-9042-2>
- Zereyesus, Y. A., Embaye, W. T., Tsiboe, F., & Amanor-Boadu, V. (2017). Implications of non-farm work to vulnerability to food poverty-recent evidence from Northern Ghana. *World Development*, 91, 113-124. <https://doi.org/10.1016/j.worlddev.2016.10.015>



Pseudo-colour with K-means Clustering Algorithm for Acute Ischemic Stroke Lesion Segmentation in Brain MRI

Abang Mohd Arif Anaqi Abang Isa^{1*}, Kuryati Kipli¹, Ahmad Tirmizi Jobli², Muhammad Hamdi Mahmood², Siti Kudnie Sahari¹, Aditya Tri Hernowo³ and Sinin Hamdan¹

¹Faculty of Engineering, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

²Faculty of Medicine and Health Science, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

³Faculty of Medicine, Universiti Malaya, 50603 Kuala Lumpur, Malaysia

ABSTRACT

Segmentation of an acute ischemic stroke from a single modality of a greyscale magnetic resonance imaging (MRI) is an essential and challenging task. Recently, there are several numbers of related works on the automatic segmentation of infarct lesion from the input image and give a high accuracy in extraction of infarct lesion. Still, limited works have been reported in isolating the penumbra tissues and infarct core separately. The segmentation of the penumbra tissues is necessary because that region has the potential to recover. This paper presented an automated segmentation algorithm on diffusion-weighted magnetic resonance imaging (DW-MRI) image utilizing pseudo-colour conversion and K-means clustering techniques. A greyscale image contains only intensity information and often misdiagnosed due to overlap intensity of an image. Colourization is the method of adding colours to greyscale images which allocate luminance or intensity for red, green, and blue channels. The greyscale image is converted to pseudo-colour is to intensify the visual perception and deliver more information. Then, the algorithm segments the region of interest (ROI)

using K-means clustering. The result shows the potential of automated segmentation to differentiate between the healthy and lesion tissues with 90.08% in accuracy and 0.89 in dice coefficient. The development of an automated segmentation algorithm was successfully achieved by entirely depending on the computer with minimal interaction.

Keywords: Acute ischemic stroke, clustering, MRI, pseudo-colour, segmentation

ARTICLE INFO

Article history:

Received: 14 September 2020

Accepted: 30 December 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.03>

E-mail addresses:

abganaqi2@gmail.com (Abang Mohd Arif Anaqi Abang Isa)

kkuryati@unimas.my (Kuryati Kipli)

jatirmizi@unimas.my (Ahmad Tirmizi Jobli)

mmhamdi@unimas.my (Muhammad Hamdi Mahmood)

sskudnie@unimas.my (Siti Kudnie Sahari)

hernowo@um.edu.my (Aditya Tri Hernowo)

hsinin@unimas.my (Sinin Hamdan)

* Corresponding author

INTRODUCTION

Stroke is part of the cardiovascular diseases and can be subtyped into ischemic stroke and hemorrhage stroke. Magnetic Resonance Imaging (MRI) technology is utilized extensively in the identification of brain pathologies (Gonzalez & Schawmm 2016). When the patient positioned in a strong magnetic field, magnetization occurs in the tissue. It produces Radio Frequency (RF) signals that are released by the tissue hence creating the MR image. MRI is better than x-ray, Computed Tomography (CT), and ultrasound in differentiating a diseased and healthy areas of the brain.

Currently, radiologists are required to manually delineate the image slices consecutively to detect the diseases of the human brain in MR images, and this considered as time-consuming. Each slice of the medical images is a valuable data source and may determine the outcome of the treatment for the patient. Incorrect segmentation would lead to inaccurate diagnosis and delaying treatment of the disease. Improper segmentation could result in a poor outcome for the patient. Thus, the automated processing of the vast image data source is necessary to overcome these issues.

Challenges exist in developing automated acute ischemic stroke lesion segmentation as the detection of the stroke area on MRI is influenced by the different signal intensities of MRI. Signal strength changes over a period in acute infarction and differs across various magnetic strength of the machine (Merino & Warach 2010, Purushotham et al., 2015). Another issue in developing an automated segmentation from greyscale images from brain MRI is the similarities of the images of diseased areas to the healthy brain. The greyscale images of infected areas range in the grey intensity, thus making it difficult to distinguish the healthy brain tissues with the diseased areas.

Pseudo-colour image processing is a technique that allocates colours to grey values of a grayscale image based on specific indicators. This procedure is computationally simple and fast. The pseudo-colour translation is a useful enhancement technique to transform a grayscale image, therefore, improving the capability of distinguishing the image information which has broadly applied in medical, engineering, and military fields (Jinlong et al., 2012). Meanwhile, K-means clustering segmentation classified under pixel-based segmentation methods, whereby the application of this method is more practical for biomedical images. The number of clusters in the biomedical image generally recognized for images of sections of the human anatomy.

Recently, there are numerous researches on brain MRI segmentation of the acute ischemic stroke. Muda et al. (2017) discussed an automated segmentation approach of brain lesions from DWI by employing the Fuzzy C-means (FCM) segmentation algorithm. The clustering algorithm achieved by merging the spatial neighborhood data points with the standard FCM algorithm. Nag et al. (2017) proposed a segmentation technique based on the Gaussian mixture model (GMM) for the localization of stroke lesions in DWI

images. The segmentation accomplished when several components of GMM generate a multimodal density, and clusters were assigned to maximize the maximum likelihood function. Vupputuri et al. (2017) presented a hierarchical region splitting (HRS) approach to delineate acute ischemic stroke lesions area using diffusion- and perfusion-weighted image (DWI-PWI) maps from adult patients. They rejected outliers of ROIs from several image maps by cross-validation and sensibly fused to generate final ROIs by utilizing the statistical technique.

Meanwhile, Gautam and Raman (2019) introduced the histogram-based thresholding (HT) and random forest-based (RF) brain classification to segment the stroke lesion. A clustering segmentation on greyscale MRI image has been done by Dhankar et al. (2010) and Nitta et al. (2020). In work by Dhankar et al. (2010), the quality of the segmentation was increased with the increasing value of the number of centroids, k and many k were formed, and it was difficult to recognize the segmented regions. Besides, there were 16 probabilities calculated for initial centroids in work by Nitta et al. (2020). These high number of centroids calculated hence increase the computational cost in the k-means clustering segmentation approach.

In this research, we will propose an image segmentation algorithm that consists of several parts. The core part is the conversion of grayscale MRI into the pseudo-colour combined with a clustering algorithm. The proposed approach will help medical experts to differentiate acute ischemic stroke lesion in structural MRI brain image. The MRI modality used in this research is the Diffusion Weighted Image (DWI), and we focus on hyperintense lesions. The critical contribution of this method is that the stroke lesion can be distinguished from healthy brain tissue by enhancing a grayscale image to colourize image. In summary, the contributions of this paper include 1) development of the automatic segmentation algorithm for acute ischemic stroke lesion in structural MRI brain image. 2) validation and evaluation of the performance of the segmentation algorithm against the ground truth.

This paper is organized as follows. In section II will describes the methods and algorithms of this segmentation. Meanwhile, section III presents the experimental results and discusses them. Lastly, the conclusions are given in Section IV.

METHODS

In this section, we present a pseudo-colour conversion with k-means clustering algorithm method for acute ischemic stroke lesion segmentation on the diffusion-weighted image. The flowchart of the proposed methodology of the segmentation algorithm is displayed in Figure 1. We divided the segmentation flow into three processes, namely pre-processing, pseudo-colour translation, and segmentation. Each approach gives a significant contribution to the segmentation outcome.

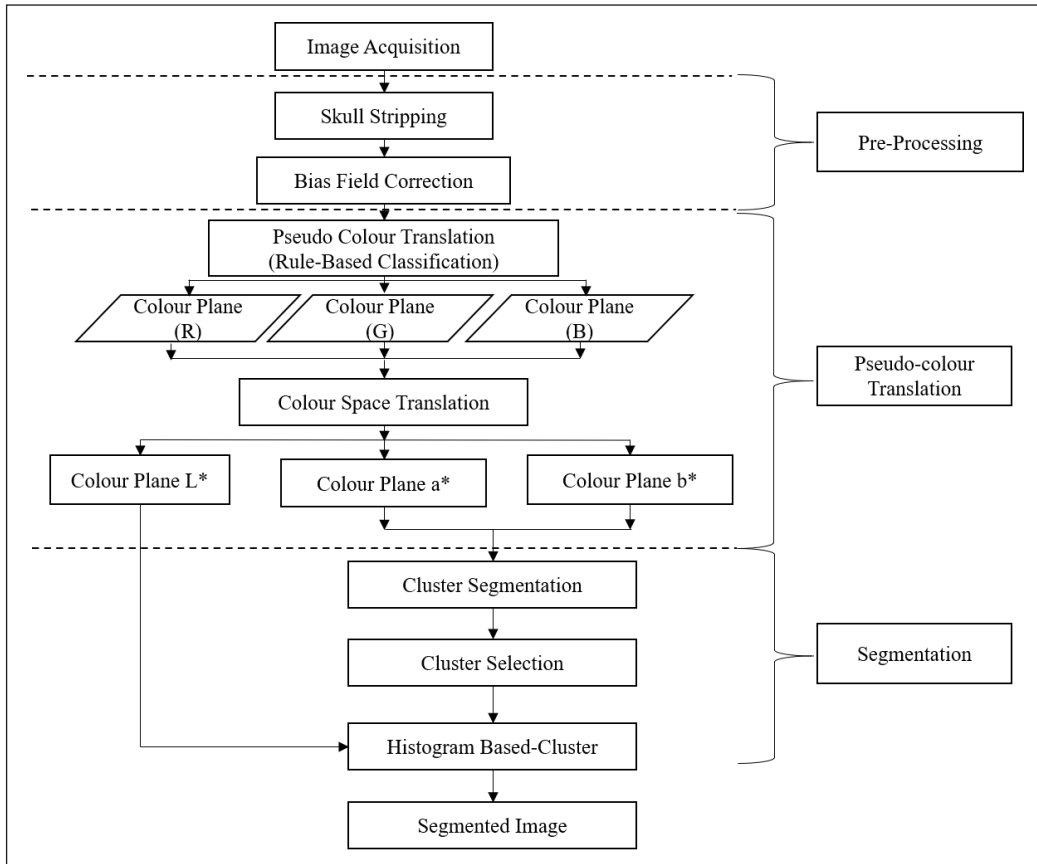


Figure 1. Process flow of the proposed method

Image pre-processing is one of the primary steps in most of the segmentation of the medical image. This step is highly essential to ensure the high accuracy of the segmentation. MR brain images generally consist of some visual artefacts. These artefacts in MRI may be caused by the equipment or the patient itself (Graves & Mitchell, 2013). Interference of the field homogeneity, because of the magnetic material inside or outside of the scanner equipment often occurred while taking images from the patients, and patient-specific artefacts include brain skull and motion artefacts caused by breathing and movement (Krupa & Bekiensinska-Figatowska, 2015). These artefacts need to be eliminated by the pre-processing process before continuing to the subsequent steps.

A pseudo-colour translation is a method to augment the colour contrast of the input grey image, which can improve the accuracy of the brain segmentation performance (Li et al., 2017). The translation process is to obtain practical features and intensify the visual by mapping a single channel of the grey level pixel into Red, Green, and Blue channels (Afruz et al., 2010).

Image segmentation is a method of separating an image into smaller regions with identical attributes. The attributes can be texture, greyscale value, brightness, shape, colour or texture (Tyan et al., 2014; Kipli & Kouzani, 2015). The practice of image segmentation in medical is to provide meaningful information to the medical experts. K-means clustering is one of the popular methods under the clustering algorithm. This method divides or groups similar attributes of data points into one cluster and different attribute into different clusters.

In the subsequent subsections, we will present the proposed pseudo-colour conversion along with the clustering algorithm to segment acute ischemic stroke lesion in detail.

The Proposed Segmentation Algorithm

Initially, brain DWI images are gathered. Then, pre-processing operations are performed to intensify the quality of the images. Removal of non-brain tissues such as skull and bias field correction is proposed for this image. Subsequently, the colour space translation of an RGB colour map is applied, which contains Red, Green, and Blue values for each grey value item. The transformation converts the RGB image to the Commission on Illumination (CIE) Lab (L^*a^*b) colour model to rectify important features before undergoing the clustering process. The colour plane L^*a^*b calculates the tri-stimulus coefficient in terms of luminosity layer L^* , a chromaticity-layer a^* , and chromaticity-layer b^* . This colour space specifies where the colour positioned along the red-green axis and blue-yellow axis. Both a^* and b^* channels include all essential colour information; hence, the colour information can be used as a classifier for clustering accommodate a region of interest. The histogram-based is utilized by the brightness attribute of L^* to exclude the pixels that are not related to the selected cluster. The difference between two layers a^* and b^* are calculated using the Euclidean distance metric. K-means clustering algorithm will find partitions, so that pixel values within each cluster are adjacent to each other based on the Euclidean distance metric obtained.

Image Acquisition

Images from one stroke patient (male, 60 years old) had been obtained in this research. DWI images were obtained on a 1.5T Siemens scanner from Sarawak General Hospital. With the following characteristics: repetition time (TR) = 4.40 s; echo time (TE) = 80 ms; inversion time = one acquisition; flip angle = 90°; field of view (FOV) = 229 mm; 25 slices; and in-plane matrix = 130 × 130; slice thickness, 5 mm; and gap between each slice, 6 mm. Image resolution = 260 x 260. The sample had medical records that had been confirmed by a clinical neuroradiologist. For this study, three slices were giving meaningful information required.

In addition to that, Anatomical Tracings of Lesions After Stroke Release 1.1 or known as ATLAS R1.1 is an open-source dataset comprising of 304 images of T1-weighted MRIs

neuroimaging was being utilized in this research for experiment testing. This dataset includes manually segmented stroke lesions as a gold standard. Image data were gathered on a 1.5 Tesla (T) scanner with a resolution of 0.9mm x 0.9mm x 3.0 mm.

Image Pre-processing

Before segmentation, image pre-processing is a process of removal of non-brain tissues and improves the features of images acquired from the scanners. Several pre-processing operations implemented for this approach comprise of skull stripping, and bias field correction. The raw DWI image had an 8-bit unsigned integer. Skull stripping or non-brain tissue removal is the procedure to eradicate the skull tissues from the brain images and acquired brain tissue before progressing to the segmentation. This procedure is a crucial process due to the forefront to obtain effective segmentation outcomes for precise diagnosis (Kalavathi & Prasath, 2016). A mathematical morphology procedure is proposed to remove the non-brain tissue. Fundamentally, morphology operations that apply to this procedure are dilation and erosion. The image was converted to greyscale before progressing into the binary image while the structuring element used to erode the binary image was disk-shaped with a radius of 55.

Intensity inhomogeneity (IIH), also known as the bias field signal is a smooth and low-frequency signal that distorts the quality of MR images during the acquisition process. The distorted signal produces a low-quality image, thus reduced the high-frequency information of the image, such as the intensity gradient of image pixels. The proposed method to correct the IIH was by employing the energy minimization optimization method. This method has proven to remove the bias field signal from the MRI images accurately (Li et al., 2014).

Pseudo-colour Translation

An 8-bit grayscale image of DWI would be translated into a pseudo-colour image with RGB value. The pseudo-coloured image was a single channel grayscale image that had colour assigned to it. In other words, the gray values of the image would match with the red, green, and blue values so that the shade of gray were displayed as colourized pixels. The translated image showed a pixel with a specified value of each colour. Therefore, the differences in colour value in the translated image reflected the changes in the intensity of the image. A formula to convert the grayscale image to RGB is shown below.

$$\text{gray} = 0.299\text{Red} + 0.587\text{green} + 0.114\text{blue} \text{ (Khalil \& Ali, 2013)}$$

Hence, the translated colour image then be transformed into a CIELab colour model to recover significant attributes. The luminance layer or intensity value of L of a greyscale image is represented with range 0 - 255 while a colour image utilized RGB coordinate

system with Red-Green-Blue corresponded components. Thus, colour plane L^*a^*b will be computed into tristimulus components with the following Equation 1:

$$\begin{aligned} W &= 0.4303R + 0.3416G + 0.1784B, \\ Y &= 0.2219R + 0.7068G + 0.0713B, \\ Z &= 0.0202R + 0.1296G + 0.9393B, \end{aligned} \quad [1]$$

The CIELAB colour model is computed as Equation 2

$$\begin{aligned} L^* &= 116 (h(Y / Y_s)) - 16, \\ a^* &= 500 (h(W / W_s) - h(Y / Y_s)), \\ b^* &= 200 (h(Y / Y_s) - h(Z / Z_s)), \\ h(q) &= \begin{cases} \sqrt[3]{q}, & q > 0.008856 \\ 7.787q + 16 / 116, & q \leq 0.008856 \end{cases} \end{aligned} \quad [2]$$

Where Y_s , W_s , and Z_s are the reference of the white point. Both a^* and b^* layers comprise all required colour information in red-green axis and blue-yellow axis, respectively. Hence, the Euclidean distance will be determined in the clustering process.

K-means Clustering Algorithm

In the proposed method, we applied the clustering algorithm to segment acute ischemic stroke lesion core. The process started with a few pre-processing steps to eliminate non-brain tissues and bias field correction to enhance the image. The grey-level image of raw brain MRI was inadequate to display fine features. Thus, the pseudo-colour transformation was applied to acquire more useful information and intensify the visual.

K-means clustering is a commonly utilized method of the clustering algorithm to separate data into a specific number of groups. The process was started by assigning data points to a group by computing the Euclidean distance among the data points, and the cluster centroid was at the lowest value. The homogeneous (similar) data points were achieved by obtaining less variation within the clusters. Euclidean distance can be calculated as in Equation 3,

$$d = \| x_i - c_j \| \quad [3]$$

where $X = \{x_i \mid i = 1, 2, \dots, l\}$ and $C = \{c_j \mid j = 1, 2, \dots, k\}$ while l is the clustered data and k is the cluster centroids. The resulting step is to recalculate the mean of the cluster centroids

based on their cluster members and regroup the attribute corresponding to the new centroids. The criterion discontinued when all cluster centroids started to merge.

Experimental Setup

There are two experiments conducted to evaluate and validate the performance of the proposed segmentation algorithm. Experiment I was done by using diffusion weighted images (DWI) from the SGH datasets. These images were suggested by the clinical radiologist based on the location of the disease. The segmentation outcome was compared to the ground truth provided by a clinical radiologist. Experiment II is done by selecting 80 T1-weighted brains MRI from the Anatomical Tracings of Lesions After Stroke Release 1.1 dataset (Liew et al., 2018). The criteria for selecting these images were based on the location of acute stroke (left or right hemisphere of the brain) and the size of the acute stroke. The segmentation outcome was compared to the ground truth provided in the dataset.

The segmentation flow process is following the proposed process flow that has been discussed previously. A series of the segmentation process was from the image acquisition, pre-processing, segmentation, and performance evaluation would be carried out. Analysis and interpretation of the results will be conducted.

Evaluation Metrics

In this paper, the accuracy, specificity, sensitivity, and dice coefficient (DC) metrics to calculate the segmentation performance. The accuracy, specificity, sensitivity, and DC metrics are defined as Equation 4, 5, 6, and 7.

$$\text{Accuracy} = \frac{TP+TN}{P+N} = \frac{TP+TN}{TP+FP+TN+FN} \times 100 \quad [4]$$

$$\text{DC} = \frac{2TP}{2TP+FP+FN} \quad [5]$$

$$\text{Sensitivity} = \frac{TP}{TP+FN} \quad [6]$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad [7]$$

Where TP is a true positive, TN is a true negative, FP is a false positive, and FN is a false negative. The TP and TN are the numbers of correctly classified positive and negative classes by the segmentation algorithm. Whereas FP and FN are those images that are incorrectly classified.

The segmentation of the acute ischemic lesions is essential in clinical diagnosis. Therefore, it is necessary to use lesion-wise metrics to assess segmentation performance. Accuracy is a system to measure the degree of closeness of measurements of a quantity compared to the ground truth. On the other hand, DC is a statistical validation system widely used in the MR image segmentation to measure the spatial overlap of the target regions (Zou et al., 2004). In this case, the automated segmentation image was compared with the ground truth provided by the clinical radiologist. Sensitivity refers to the outcome of the test to correctly detect the presence of the target lesions, while specificity relates to the rightly reject the existence of lesions.

RESULTS AND DISCUSSION

Pre-processing

In the proposed method, mathematical morphological operation and energy minimization optimization processes were applied to remove non-brain tissue and bias field correction, respectively. The outer layer of the skull, as observed in Figure 2, was successfully removed from the image (Figure 3). The removal of skull tissues is significant to improve the robustness of the segmentation process and influence the segmentation outcome. The bias field of DWI image had been effectively excluded from the raw image as exhibited in Figure 4. The different intensity value existed in an image might alter the segmentation process afterwards.

Pseudo-colour Translation

Figure 5 represents the image after a pseudo-colour translation. The greyscale image of an MRI was inadequate to support fine features; thus, pseudo-colour translation was applied to the image. The transformation mapped the 8-bit image with grey level values ranges 0 – 255 into Red, Green, and Blue colour.

Segmentation

After the establishment of the pixel values range, the converted image was separated into RGB channels. Then, the RGB channels were translated into CIELab space to calculate the mean of a^* and b^* value. The clustering algorithm occurred to cluster the homogenous pixel values based on the Euclidean distance obtained. Figure 6 shows the segmented result after applying K-means clustering. The image obtained was following cluster selection in the K-means clustering algorithm, and region of interest was extracted successfully. Figure 7 displays the overlaid of the ROI against the original brain image.

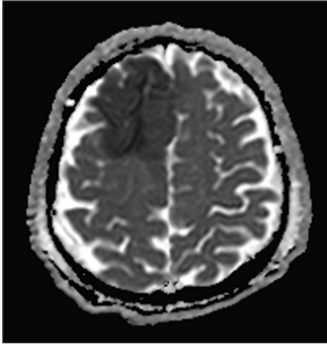


Figure 2. Raw image of DWI

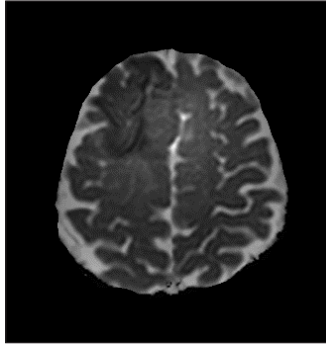


Figure 3. Skull stripped brain image

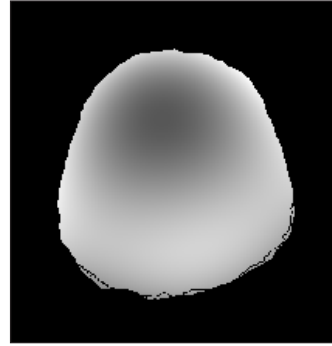


Figure 4. Bias field of an image

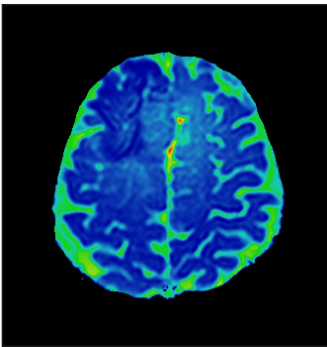


Figure 5. Image after pseudo-colour translation



Figure 6. Segmented region of interest

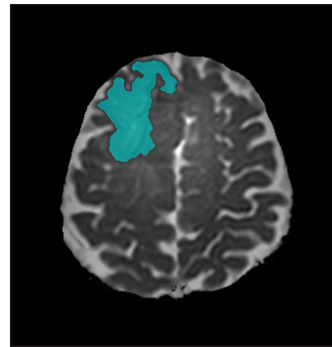


Figure 7. Overlaid of segmented ROI and brain image

Performance Evaluation

The segmentation algorithm was employed for Experiment I which were on DWI images. The performance evaluations metrics for this segmentation algorithm included accuracy, DC, specificity, and sensitivity.

Table 1 shows the achievement of the clustering algorithm for the three DWI images for different time-space. The results show accuracy, DC, specificity, and sensitivity. The maximum accuracy achieved was 90.080% indicating the proposed method was able to segment the acute ischemic stroke, while the average accuracy for three DWI images was 88.790%.

From Table 1, Image 1 provides the lowest accuracy amongst the other two images by 86.820%. The most insufficient accuracy achieved is due to the existence of high-intensity pixel incorporate into the region of interest and caused the over-segmentation and delivers high values in false positive with 0.194 compared with the remaining images.

In Experiment II, the dataset used was ATLAS dataset R1.1. The purpose of this experiment was to obtain the performances of the proposed segmentation algorithm tested on the several different images in the same dataset with the different sizes of the lesion,

Table 1
Performance evaluation segmentation of the region of interest

Index	Accuracy (%)	Dice Coefficient (DC)	Specificity	Sensitivity
Image 1	86.820	0.848	0.194	0.030
Image 2	90.080	0.890	0.071	0.133
Image 3	89.460	0.882	0.140	0.058
Average	88.790	0.873	0.135	0.074

location of the lesion on the brain, and varying intensity of the lesion. MRI modality that was used in this simulation was T1-weighted (T1*w), whereas the location of the lesion was varied (left of right hemispheres) in each of the 80 selected brain images. The selection of these brain images was considered the location of the lesion on the brain image.

Accuracy, DC, sensitivity, and specificity evaluations were applied to compare the performance of the algorithm. Figure 8 and 9 give the boxplots of the segmentation algorithm performance in terms of accuracy, dice coefficient, sensitivity, and specificity. The boxplot represents the distribution of the performance of the 80 images from the ATLAS dataset. The boxplot exhibits the full range of variations (from min to max), and the middle line is the median. Boxplot is a systematic way of displaying the distribution of data based on the maximum, third quartile, median, first quartile, and minimum.

From Figure 8, the maximum accuracy achieved was 89.600%. In comparison, the average accuracy was $74.090\% \pm 1.805\%$ with a 95% confidence level. The lowest accuracy attained was 53.17% due to the undistinguished intensity of the compressed image. The conversion of the pseudo-colour translation may affect the intensity of the image.

From Figure 9, the DC, sensitivity, and specificity of the segmentation algorithm were presented - DC of the algorithm ranging from 0.120 to 0.880 with a median of 0.630. The 95% confidence interval calculated was ± 0.038 . The lowest DC attained was 0.12 due to the distribution of contrast was inconsistent on the image. This condition caused the spatial overlap between the ground truth, and the segmented image was disturbed.

The maximum value sensitivity and specificity of this segmentation algorithm achieved were 1, respectively and indicated that the actual positive values and negative values labelled from these images were accurately categorized. The 95% confidence interval revealed that the sensitivity and specificity mean values were 0.690 ± 0.042 , and 0.780 ± 0.023 , respectively. The sensitivity evaluation is to measure the ratio of positive voxels in the ground truth that was identified as positive by the segmentation. In contrast, evaluation of the specificity measures the ratio of negative voxels in the ground truth that is also identified as unfavorable by the segmentation (Taha & Hanbury, 2015).

The proposed methodology requires uncomplicated pre-processing and colour space translation procedures before undergoing a segmentation task. The segmentation algorithm in this work is easily understood. The overall implementation procedure for the

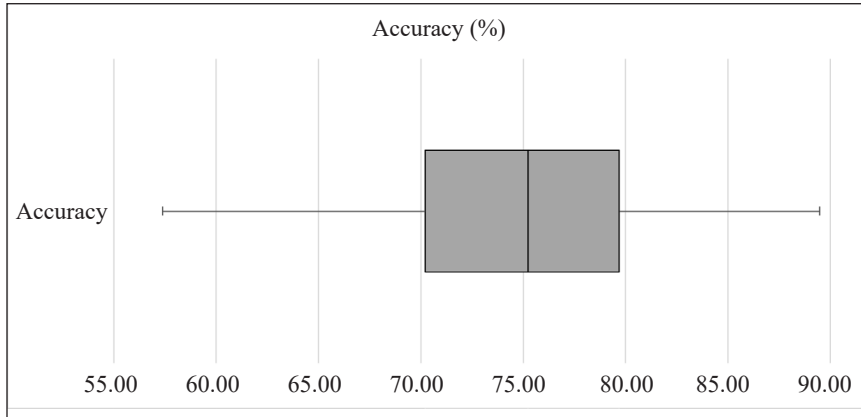


Figure 8. Boxplot of the average segmentation accuracy results

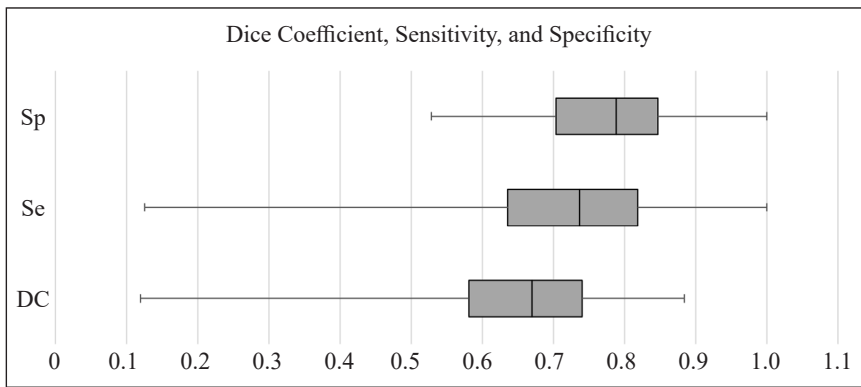


Figure 9. Boxplot of dice coefficient, sensitivity and specificity of the segmentation results

segmentation of the acute stroke lesion region is simple and straight forward. Our work yield the maximum accuracy of 90.080% on diffusion weighted image and 89.600% on T1-weighted image respectively. In addition, maximum dice coefficient achieved from our work were 0.890 and 0.880 for DWI and T1-weighted, respectively. Experimentally, our method exhibited that the proposed segmentation techniques deliver good accomplishment, and the results at par with the existing works (Muda et al. 2017, Nag et al. 2017, Vupputuri et al 2017, and Gautam & Raman, 2019).

The implementation of pseudo-colour segmentation on the grayscale image significantly affect the quality of an image understanding system (Chen & Wang, 2004). The images were segmented into homogeneous areas with similar spectrum information. In our findings, the colour homogeneity was based on the mean value calculated from a* and b* channels from CIELAB colour model. This shows that the spectral information of the image is crucial in imagery data, and the performance of the clustering segmentation significantly increases. While the L* layer separate the dark and light coloured object. The L* layer comprises

the brightness values of each colour and the brightness value of the image was extracted in the clustering algorithm. The extraction of the brightness value reduces the number of iterations in the clustering algorithm. Hence, colour based image segmentation integrate in the clustering reduce the computational cost avoiding pixel features calculation in the image.

This paper proposed a segmentation algorithm based on structural brain MRI that plausibly aid the assessment of the acute ischemic stroke lesion. This proposed system is a ground-breaking approach that able to delineate stroke lesion precisely and delivers the second opinion to the specialist. The automated segmentation of the lesion is a challenging task because of the presence of artefacts, bias field, and brain skull during the acquisition process. In contrast, manual segmentation is a subjective, labour-intensive and varied in results depending on the specialists' expertise, whereas the automated segmentation.

An automated segmentation algorithm for brain MRI is developed by employing pseudo-colour conversion and clustering algorithm. The proposed segmentation algorithm involves a series of pre-processing, pseudo-colour translation and clustering processes. Each step in this image segmentation carries significant contributions towards the outcome. For instance, pre-processing steps was to reduce the miscalculation and error during the estimation of the region of interest. The occurrence of non-uniform grey-level intensities can reduce the segmentation accuracy. At the same time, the pseudo-colour translation is to enhance the visualization of the different brain region, namely, lesion, white matter, grey matter, and cerebrospinal fluid. Finally, the details of the pixel in RGB colour channels are clustered according to the resemblance by employing the K-means clustering algorithm.

CONCLUSION

In this research, we employed the pseudo-colour method on the greyscale DW image to intensify the visual perception. The proposed segmentation algorithm had shown a promising result when applied to the brain DW image focusing on the hyperintense lesions. Pre-processing steps were conducted for skull stripping and bias field correction. The proposed segmentation algorithm was assessed and compared against the ground truth provided by the clinical neuroradiologist. The result showed that the maximum accuracy achieved was 90.08% and capable of segmenting the acute ischemic stroke from brain DWI images successfully. This proposed segmentation algorithm hopefully will help medical experts to differentiate the lesion with the healthy brain tissues. Based on the obtained results, this research potential should be validated using the larger dataset to draw a more concrete conclusion. To achieve these improvement, cooperative advancement across several institutions to form larger dataset of acute ischemic stroke based on brain MR images for efficient development of detection algorithm. Besides, depending on a single modality of MRI is not sufficient to establish a complete acute ischemic stroke thrombolysis. Hence, combining the information from multiple MRI sequences may be helpful.

ACKNOWLEDGEMENTS

We want to thank the Sarawak General Hospital (SGH) for making the brain MRI data available. The first author would like to acknowledge Yayasan Sarawak and Universiti Malaysia Sarawak (UNIMAS) for financial support through the Yayasan Tun Taib scholarship.

REFERENCES

- Afruz, J., Wilson, V., & Umbaugh, S. E. (2010). Frequency domain pseudo-colour to enhance ultrasound images. *Computer and Information Science*, 3(4), 24-34.
- Chen, H. C., & Wang, S. J. (2004). The use of visible color difference in the quantitative evaluation of color image segmentation. In *2004 IEEE International Conference on Acoustics, Speech and Signal Processing* (Vol. 3, pp. 3-593). IEEE Conference Publication. <https://doi.org/10.1109/ICASSP.2004.1326614>
- Dhankar, S., Tyagi, S., & Prasad, T. V. (2010). Brain MRI segmentation using K-means algorithm. In *National Conference on Advances in Knowledge Management, NCAKM 2010* (pp. 1-5). Lingaya's University. <https://doi.org/10.13140/RG.2.1.4979.0567>
- Gautam, A., & Raman, B. (2019). Segmentation of ischemic stroke lesion from 3D MR images using random forest. *Multimedia Tools and Applications*, 78(6), 6559 - 6579. <https://doi.org/10.1007/s11042-018-6418-2>
- Gonzalez, R., & Schwamm, L. (2016). Imaging acute stroke. In *Handbook of clinical neurology*, (pp 293-315). Elsevier's ScienceDirect. <https://doi.org/10.1016/B978-0-444-53485-9.00016-7>
- Graves, M. J., & Mitchell, D. G. (2013). Body MRI artifacts in clinical practice: A physicist's and radiologists' perspective. *Journal Magnetic Resonance Imaging*, 38(2), 269-287. <https://doi.org/10.1002/jmri.24288>
- Jinlong, H. U., Xianrong, P., & Zhiyong, X. U. (2012). Study of grey image pseudo-colour processing algorithms. In *International Symposium on Advanced Optical Manufacturing and Testing Technologies: Large Mirrors and Telescopes*. (Vol. 8415, p. 841519). International Society for Optics and Photonics. <https://doi.org/10.1117/12.977197>
- Kalavathi, P., & Prasath, V. B. S. (2016). Methods on skull stripping of MRI head scan images - A review. *Journal of Digital Imaging*, 29(3), 365-379. <https://doi.org/10.1007/s10278-015-9847-8>
- Khalil, Y. A., & Ali, P. J. M. (2013). A proposed method for colorizing grayscale images. *International Journal of Computer Science and Engineering*, 2(2), 109-114.
- Kipli, K., & Kouzani, A. Z. (2015). Degree of contribution (DoC) feature selection for structural brain MRI volumetric features in depression detection. *International Journal of Computer Assisted Radiology and Surgery*, 10(7), 1003-1016. <https://doi.org/10.1007/s11548-014-1130-9>
- Krupa, K., & Bekiesińska-Figatowska, M. (2015). Artifacts in magnetic resonance imaging. *Polish Journal of Radiology*, 80, 93-106. <https://doi.org/10.12659/PJR.892628>
- Li, C., Gore, J. C., & Davatzikos, C. (2014). Multiplicative intrinsic component optimization (MICO) for MRI bias field estimation and tissue segmentation. *Magnetic Resonance Imaging*, 32(7), 913-923. <https://doi.org/10.1016/j.mri.2014.03.010>

- Li, H., Chen, C., Feng, S., & Zhao, S. (2017). Brain MR image segmentation using NAMS in pseudo-colour. *Computer Assisted Surgery*, 22(0), 170-175. <https://doi.org/10.1080/24699322.2017.1389395>
- Liew, S. L., Anglin, J. M., Banks, N. W., Sondag, M., Ito, K. L., Kim, H., & Stroud, A. (2018). A large, open-source dataset of stroke anatomical brain images and manual lesion segmentations. *Scientific data*, 5(1), 1-11. <https://doi.org/10.1038/sdata.2018.11>
- Merino, J. G., & Warach, S. (2010). Imaging of acute stroke. *Nature Reviews Neurology*, 6(10), 560-571. <https://doi.org/10.1038/nrneurol.2010.129>
- Muda, A. F., Saad, N. M., Abu Bakar, S. A. R., Muda, S., & Abdullah, A. R. (2017, March 15-17). Automated stroke lesion detection and diagnosis system. In *Proceedings of the International MultiConference of Engineers and Computer Scientists*. Hong Kong.
- Nag, M. K., Koley, S., China, D., Sadhu, A. K., Balaji, R., Ghosh, S., & Chakraborty, C. (2017). Computer-assisted delineation of cerebral infarct from diffusion-weighted MRI using Gaussian mixture model. *International Journal of Computer Assisted Radiology and Surgery*, 12(4), 539-552. <https://doi.org/10.1007/s11548-017-1520-x>
- Nitta, G. R., Sravani, T., Nitta, S., & Muthu, B. (2020). Dominant gray level based K-means algorithm for MRI images. *Health and Technology*, 10, 281-287. <https://doi.org/10.1007/s12553-018-00293-1>
- Purushotham, A., Campbell B. C. V., Straka, M., Mlynash, M., Olivo, J., Bammer, R., Kemp, S. M., Albers, G. W., & Lansberg M. G. (2015). Apparent diffusion coefficient threshold for delineation of ischemic core. *International Journal of Stroke*, 10(3), 348-353. <https://doi.org/10.1111/ij.s.12068>
- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3d medical image segmentation: Analysis, selection, and tool. *BMC Medical Imaging*, 15(29), 1-28. <https://doi.org/10.1186/s12880-015-0068-x>
- Tyan, Y. S., Wu, M. C., Chin, C. L., Kuo, Y. L., Lee, M. S., & Chang, H. Y. (2014). Ischemic stroke detection system with a computer-aided diagnostic ability using an unsupervised feature perception enhancement method. *International Journal of Biomedical Imaging*, 2014, 1- 24. <https://doi.org/10.1155/2014/947539>
- Vupputuri, A., Ashwal, S., Tsao, B., Haddad, E., & Ghosh, N. (2017). MRI based objective ischemic core-penumbra quantification in adult clinical stroke. In *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (pp. 3012-3015). IEEE Conference Publication. <https://doi.org/10.1109/EMBC.2017.8037491>
- Zou, K. H., Warfield, S. K., Bharatha, A., Tempany, C. M., Kaus, M. R., Haker, S. J., Wells, W. M., Jolesz, F. A., & Kikinis, R. (2004). Statistical validation of image segmentation quality based on a spatial overlap index. *Academic Radiology*, 11(2), 178-189. [https://doi.org/10.1016/S1076-6332\(03\)00671-8](https://doi.org/10.1016/S1076-6332(03)00671-8)



Improved High Dynamic Range for 3D Shape Measurement based on Saturation Of the Coloured Fringe

Shanyu Chua^{1*}, Chee Chin Lim¹, Swee Kheng Eng¹, Yen Fook Chong¹ and Chiun Tai Loh²

¹*School of Mechatronic Engineering, Universiti Malaysia Perlis, Pauh Putra Campus, 02600 Arau, Perlis, Malaysia*

²*Research and Development Department, Estek Automation Sdn Bhd, Bayan Lepas, 11900 Pulau Penang, Malaysia*

ABSTRACT

Phase-shifting fringe projection methods have been developed for three-dimensional scanning (Zuo et al., 2018). However, the 3-Dimensional (3D) scanning of objects with a high dynamic reflectivity range based on structured light is a challenging task to achieve (Feng et al., 2018). The incorrect intensities captured will cause phase and measurement errors. Thus, this paper proposes a method that improves the current High Dynamic Range (HDR) (Jiang et al., 2016) method to increase the dynamic range. The camera and projector have 3 channels, red, green, and blue, which can absorb and project these lights independently. This paper proposes a method that makes use of this by controlling the intensity of each projected for the camera. Each image can be split into 3 channels and provide 3 images which contain different intensities, then it will be used to compute the 3D information. In general, this is done by controlling the projection of red, green and blue (RGB) channel and apply the Jiang's algorithm (Jiang et al., 2016). The results are compared and analysed with current HDR (Jiang's method) and the regular three-step phase-shifting methods. From the experimental results, it has shown that our proposed method outperforms

the current HDR and the regular three-step phase-shifting methods. Specifically, the proposed method manages to increase the dynamic range of the reflective property of objects. Additionally, our proposed method has also significantly reduced the times of 3D object measurements.

Keywords: 3D scanning, fringe profilometry, phase-shifting profilometry

ARTICLE INFO

Article history:

Received: 18 October 2020

Accepted: 23 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.04>

E-mail addresses:

shanyuchua.ed@gmail.com (Shanyu Chua)

cclim@unimap.edu.my (Chee Chin Lim)

kheng@unimap.edu.my (Swee Kheng Eng)

yenfook@unimap.edu.my (Yen Fook Chong)

* Corresponding author

INTRODUCTION

Machine vision is a popular technology used in industry for automatic inspection and analysis. This method is widely applied in various fields, such as automatic inspection in manufacturing, process control, robot guidance, medical field and entertainment. The main reason for the wide application of this technology in various industries is due to its non-contact property that allows inspection or measurement on material that are soft or easily deformable objects which are usually not suitable for contact inspection (Zuo et al., 2018). It improves the productivity and the quality of the product and provides a competitive advantage to industries (Malamas et al., 2003). The rapidly evolving of computers and image capturing device has allowed this technology to be used in real-time optical inspection system with high accuracy (Huang & Pan, 2015).

Currently, due to the rapid evolution of technology, non-contact 3D shape measurement can acquire fast and precise data. There are many approaches for the 3D shape measurement methods, such as optical interferometry (Pavlicek & Mikeska, 2018), time-of-flight technique (Achar et al., 2017), stereo vision (Zhang, 2018), laser triangulation (Faes et al., 2016), and structured light (Aboali et al., 2017). One of the popular methods is the structured light method and it has been extensively researched and applied in the industries due to its easy hardware setup, high accuracy of measurement, high-density level, high speed and inexpensive (Zhang & Huang, 2006b; Zuo et al., 2018). Fringe projection profilometry (FPP) is the method that projects light with known patterns and applies the triangulation method to obtain the 3D information of the object (Liu et al., 2010).

However, measuring the 3D information of an object's surface with a wide variety of reflective properties and contrast can prove to be a difficult task to perform, especially to achieve consistent high-quality measurement across the entire object (Jiang et al., 2016). For the darker surfaces, lower exposure or low light intensity is used, and higher exposure or high light intensity is used for the brighter surfaces. The specular reflection which is caused by the object with a shiny surface will be acquired as white pixels because of the high reflection of light (Feng et al., 2018). This problem can be solved by spraying a thin layer of diffused powder but this method is not suitable for all objects and some measurements require high precision where the thin layer of spray cannot be ignored (Feng et al., 2018). Therefore, there is a great need for HDR techniques which can overcome this shortcoming.

HDR techniques refer to methods that are used to calculate the surface with a wide variety of surface reflectivity properties (Jiang et al., 2016). Currently, a few HDR techniques have been studied and developed to overcome this problem, such as applying a polarizing filter (Salahieh et al., 2014; Wolff, 1989), capturing images at multiple exposures (Babaie et al., 2015; Song et al., 2017; Zhang & Yau, 2009), capturing images at different light projection intensities (Waddington & Kofman, 2014), and applying algorithm (Chen et al., 2018; Jiang et al., 2016; Wang et al., 2017). The method proposed in Babaie et al.

(2015), Song et al. (2017), Waddington & Kofman (2014), and Zhang & Yau (2009) requires many images which are time-consuming. In the year 2016, Jiang et al. (2016) developed an algorithm that did not rely on the change of camera exposure or the intensity of the projector. By using the three-step phase-shifting fringe (3PSP) and inverse three-step phase fringe, the inverted fringe can be used to complement the saturated pixels. However, it is unable to reconstruct object with surface reflectivity that has a higher dynamic range.

Inspired by Jiang et al. (2016), this paper proposes to enhance the HDR method by projecting controlled red, green and blue (RGB) channel intensity to increase the dynamic range of surface measurement. This is done by projecting the fringe at different RGB channel intensity. The fringe is then split to their respective channel and is used for the phase calculation. Therefore, by controlling the RGB intensity that project brightness that is suitable for the surface reflectivity, Jiang's HDR method can be enhanced.

METHOD

Three-Step Phase-Shifting

The fringe is first generated using the computer. In this experiment, 3PSP is used. 3PSP needs 3 images of fringe pattern with phase-shifts at -120° , 0° and 120° to obtain the 3-D data (Zhang, 2016; Zhang & Huang, 2006a). The fringe pattern is generated using Equation 1, 2, and 3:

$$I_1(x, y) = I'(x, y) + I''(x, y) \cos[\varnothing(x, y) - \alpha] \quad [1]$$

$$I_2(x, y) = I'(x, y) + I''(x, y) \cos[\varnothing(x, y)] \quad [2]$$

$$I_3(x, y) = I'(x, y) + I''(x, y) \cos[\varnothing(x, y) + \alpha] \quad [3]$$

where, I_1 , I_2 , and I_3 are images with phase-shift at -120° , 0° , and 120° , I' is the mean intensity, I'' is the peak-to-valley intensity modulation, $\varnothing(x, y)$ is the wave front phase, and α is phase shift at 120° . The phase, $\varnothing(x, y)$ as shown in Equation 4, can be obtained by using Equation 1, 2, and 3.

$$\varnothing(x, y) = \tan^{-1} \frac{\sqrt{3}[I_1(x, y) - I_3(x, y)]}{2I_2(x, y) - I_1(x, y) - I_3(x, y)} \quad [4]$$

This method is not suitable for scanning objects with a high surface reflectivity because the dark region of the sample captured will be under-exposed while the highly reflective surface is over-exposed. Hence, this cause difficulty in capturing the desired result when

both kinds of surfaces are present (Feng et al., 2018). To overcome this problem, Jiang proposed a HDR method (Jiang et al., 2016) that used the inverted 3PSP to complement the regular 3PSP shown in Current HDR Method.

Current HDR Method

Another 3 inverse fringes are generated, which serve to increase the contrast of the 3D surface to be measured, using Equation 5, 6, and 7.

$$I_1^{inv}(x, y) = I'(x, y) - I''(x, y) \cos[\emptyset(x, y) - \alpha] \quad [5]$$

$$I_2^{inv}(x, y) = I'(x, y) - I''(x, y) \cos[\emptyset(x, y)] \quad [6]$$

$$I_3^{inv}(x, y) = I'(x, y) - I''(x, y) \cos[\emptyset(x, y) + \alpha] \quad [7]$$

Two sets of fringes are generated, where one of it contain fringe with 6-pixel width and 120-pixel width fringe (Zhang, 2016; Zhang & Huang, 2006a). Each set of fringes contain 6 images.

The phase will be computed using Jiang’s method (Jiang et al., 2016). When $I_1(x, y)$ is saturated, it is replaced with $I_1^{inv}(x, y)$ and using Equation 8. When $I_2(x, y)$ is saturated, it is replaced with $I_2^{inv}(x, y)$ and using Equation 9. When $I_3(x, y)$ is saturated, it is replaced with $I_3^{inv}(x, y)$ and using Equation 10. When $I_1(x, y)$ and $I_2(x, y)$ are saturated, it is replaced with $I_1^{inv}(x, y)$ and $I_2^{inv}(x, y)$, and using Equation 11. When $I_1(x, y)$ and $I_3(x, y)$ are saturated, it is replaced with $I_1^{inv}(x, y)$ and $I_3^{inv}(x, y)$ and using Equation 12. When $I_2(x, y)$ and $I_3(x, y)$ are saturated, it is replaced with $I_2^{inv}(x, y)$ and $I_3^{inv}(x, y)$, and using Equation 13. If $I_1(x, y)$, $I_2(x, y)$ and $I_3(x, y)$, then Equation 14.

$$\emptyset(x, y) = \tan^{-1} \frac{-3I_1^{inv}(x, y) + 2I_2(x, y) + I_3(x, y)}{\sqrt{3}[I_1^{inv}(x, y) - I_3(x, y)]} \quad [8]$$

$$\emptyset(x, y) = \tan^{-1} \frac{I_1(x, y) - I_3(x, y)}{\sqrt{3}[I_1(x, y) + I_3(x, y) - 2I_2^{inv}(x, y)]} \quad [9]$$

$$\emptyset(x, y) = \tan^{-1} \frac{-I_1(x, y) - 2I_2(x, y) + 3I_3^{inv}(x, y)}{\sqrt{3}[I_3^{inv}(x, y) - I_1(x, y)]} \quad [10]$$

$$\phi(x, y) = \tan^{-1} \frac{I_1^{inv}(x, y) + I_2^{inv}(x, y) - 3I_3(x, y)}{\sqrt{3}[I_1^{inv}(x, y) - I_3(x, y)]} \quad [11]$$

$$\phi(x, y) = \tan^{-1} \frac{I_3^{inv}(x, y) - I_1^{inv}(x, y)}{\sqrt{3}[2I_2(x, y) - I_1^{inv}(x, y) - I_3^{inv}(x, y)]} \quad [12]$$

$$\phi(x, y) = \tan^{-1} \frac{3I_1(x, y) - I_2^{inv}(x, y) - I_3^{inv}(x, y)}{\sqrt{3}[I_3^{inv}(x, y) - I_1(x, y)]} \quad [13]$$

$$\phi(x, y) = \tan^{-1} \frac{2[B_1(x, y) - B_3(x, y)]}{\sqrt{3}[B_2(x, y) - B_1(x, y) - B_3(x, y)]} \quad [14]$$

where $B_1(x, y) = I_1(x, y) - I_1^{inv}(x, y)$, $B_2(x, y) = I_2(x, y) - I_2^{inv}(x, y)$, and $B_3(x, y) = I_3(x, y) - I_3^{inv}(x, y)$.

Enhanced HDR Method

Even though the current method is faster and does not require multiple exposures of images, however, it tends to fail during measurement of wide ranges of reflectivity due to only a single exposure (Feng et al., 2018). Thus, our enhanced HDR method is proposed.

With the aim of increasing the dynamic range of the surface from the current HDR method, the image is captured by controlling the RGB intensity. The projector will project the red channel that is set to the lowest intensity, the blue channel that is set to the highest intensity and green channel that is set to the mid-intensity. With this, 3 intensities can be obtained from a single set of images. Each set of images from these 3 images can be used independently to compute the 3D coordinates.

The overall flow chart for this method is as shown in Figure 1. It will start from the channel with the highest intensity, which is the blue channel, and check if all three non-inverse fringes are over-exposed. If it is not over-exposed, it will be used for 3D coordinates computation; and if it is over-exposed, it will proceed to check the second highest intensity channel which is the green channel. Green channel will be checked if all three non-inverse fringes are over-exposed. If it is not over-exposed, it will be used for 3D coordinates computation; and if it is over-exposed, it will proceed to the lowest intensity channel, which is the red channel. If all three non-inverse fringes are still over-exposed, computation will be done using Equation 14.

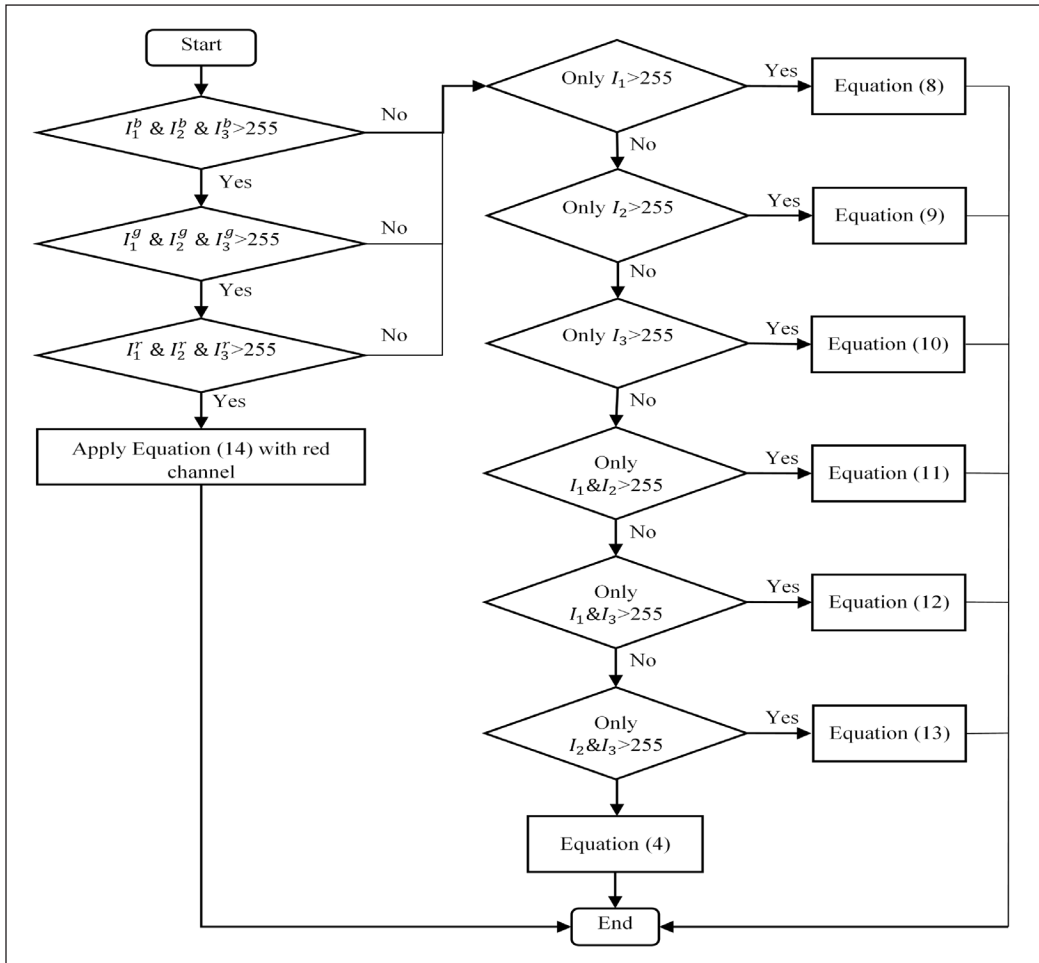


Figure 1. The flowchart of the proposed method

Experiment Setup

The camera used was Basler axA2040-55uc USB 3.0 with Sony IMX265 Complementary Metal-Oxide-Semiconductor (CMOS) sensor. The lens used was TCSM024: 3D bi-telecentric lens with Scheimpflug adjustment. The projector used was a digital light projection (DLP). LightCrafter 4500 was a software used to flash the projector’s memory and control the projector. pylonViewer was used to view and capture the image from the camera. The captured image was then processed using the proposed method. Figure 2 demonstrates the setup of the experiment.

RESULTS AND DISCUSSION

The experiment was done using a sample with high dynamic reflective surfaces. The sample used was a circuit board as shown in Figure 3(a). From Figure 3(a), there are four main

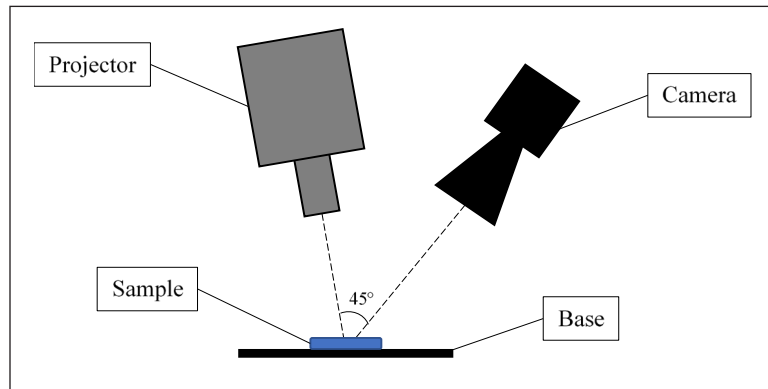


Figure 2. Diagram of the experimental setup

classes of components, as labelled using 1, 2, 3 and 4. Component class 1 and component class 2 consist of a middle Integrated Circuit (IC) and the biggest Integrated Circuit (IC), respectively. Component class 3 consists of three smallest Integrated Circuits (ICs), and the component class 4 consists of ten resistors. For all the Integrated Circuits (ICs), the low reflective property is at the semiconductor wafers (black colour) part, and the high reflective property is at the soldered part. Figure 3(d) shows the current HDR method's result which is better compared to Figure 3(b), and Figure 3(c). These experimental results show that the proposed HDR method has a great improvement in the accuracy to measure objects with high-contrast surfaces.

Figure 4 shows the result of the experiment using a different light setting and luminance. The experiment done by capturing the data in five different intensities, where 1 is the lowest level for light intensity and 5 is the highest light intensity. The data is converted into a heat map for easier visualization. The different light intensities are controlled by using the projector.

The dark blue region in Figure 4 indicates that it is either over-exposed or under-exposed and they are set to 0 mm. At the light intensity 1, the three-step phase-shifting method (Figure 4(a)) shows a better result compared to Jiang's method (current HDR method) and proposed HDR method, Figure 4(b) and Figure 4(c). This is because the current HDR method is developed for the shiny surface which may not be able to include a broad range of reflectivity (Feng et al., 2018). As the light intensity increases, the component class 4 with a white rectangle line surrounding it, as shown in Figure 4(a), starts to fail and become over-exposed when undergoing the regular 3PSP method, as shown in the indicator at light intensity 2 under 3PSP method. This is because the white rectangular line surrounding it has a high reflective property and caused it to be over-exposed. The pinhole of the circuit board from the three-step phase-shifting method is producing error as in Figure 4(d). The shiny surface of the sample cannot be seen when the reflectivity

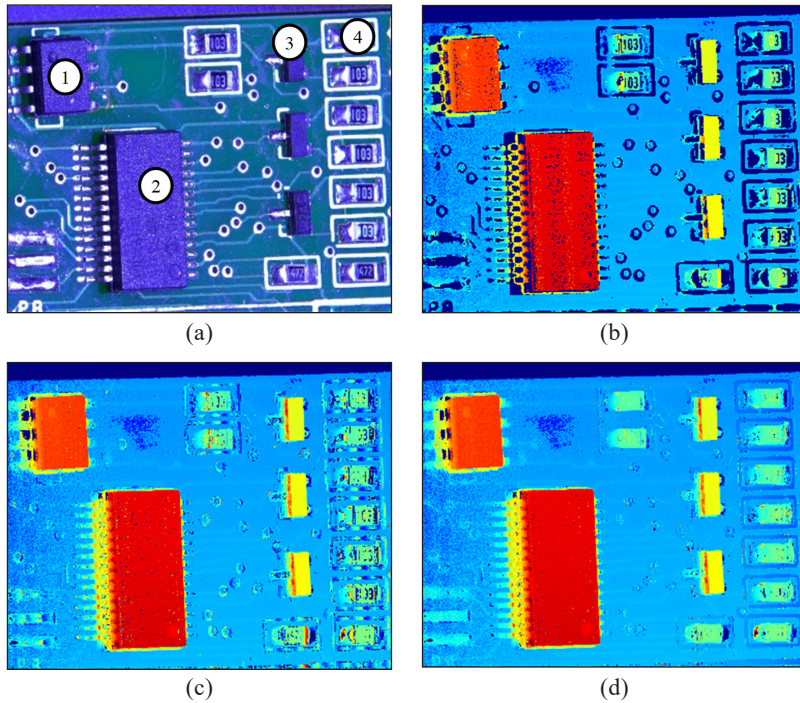


Figure 3. (a) Original image (b) Three-step phase-shifting (c) Current HDR method (d) Proposed HDR Method of the circuit board image

increases as in Figure 4(d). The proposed HDR method, shown in Figure 4(f), and current HDR method, shown in Figure 4(e) are clearly much better comparing with the regular three-step phase-shifting method at light intensity 2. On the other hand, the current HDR method shows improvement as the intensity increases but error starts to appear on the high reflective surface on component class 4 at light intensity 3 under current HDR method (Figure 4(h)), whereas, the proposed method is still able to give good results (Figure 4(i)).

At intensity 4, from the three-step phase-shifting (Figure 4(j)), the black IC is starting to become over-exposed and it is starting to produce error as shown in the indicator. The pinhole of the circuit board from the current method is also starting to get worse as the light intensity is increased (Figure 4(k)). However, the proposed method is still able to produce a satisfying result (Figure 4(l)). At intensity 5 of the three-step phase-shifting, the white part and the shiny part of the circuit board is over-exposed (Figure 4(m)). The result is as expected because the three-step phase-shifting algorithm cannot compute samples with surfaces that have high dynamic range of reflectivity. The current HDR method, however, shows better results, but the white part and the shiny part of the circuit board shows noise and some cannot be computed (Figure 4(n)). The proposed algorithm can compute the 3D data even at a high light intensity on the surface of the sample with a high dynamic range of reflectivity (Figure 4(o)).

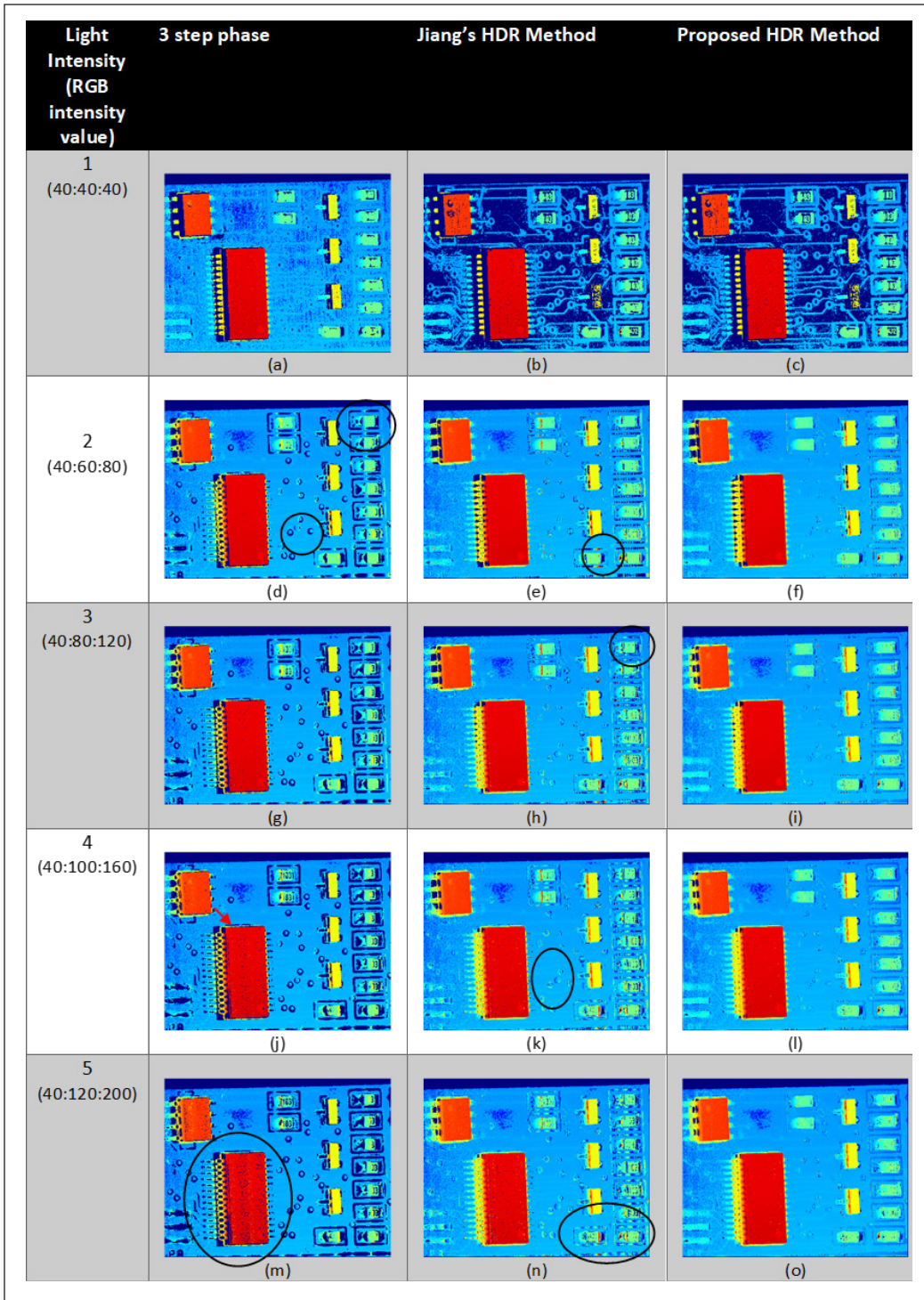


Figure 4. Comparison between the 3 step phase-shifting method, current HDR method and proposed HDR method at different light intensities

Figure 6 displays the reconstruction of the solder ball grid from Figure 5. Solder ball grid shown in Figure 5 has both the highly reflective surface and low reflective surface. The result in Figure 6(a) indicates that the three-step phase shifting is not able to reconstruct the background of the image perfectly as it is brighter and more reflective. Not only that, but the tip of the solder balls is also not well reconstructed. Next, Jiang’s HDR method in Figure 6(b) shows better result in the reconstruction of the solder ball tip, however, the background acquired is worse compared to the three-step phase-shifting method. The proposed method in Figure 6(c) shows a better reconstruction of the overall image compared to the two method. The background can be reconstructed even when it is highly exposed. Through the action of controlling the intensity of each RGB channel, we can reconstruct both the over-exposed region and under-exposed region at a greater range which shows the robustness of the proposed HDR method.

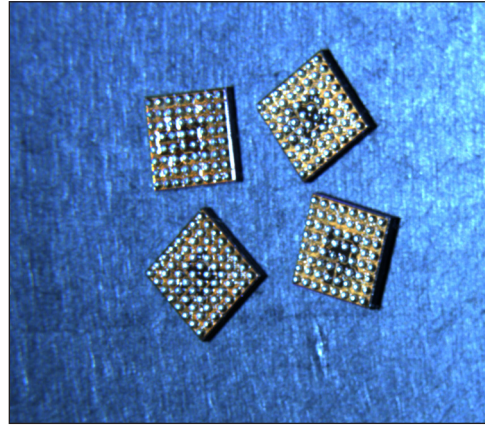


Figure 5. Solder ball grid

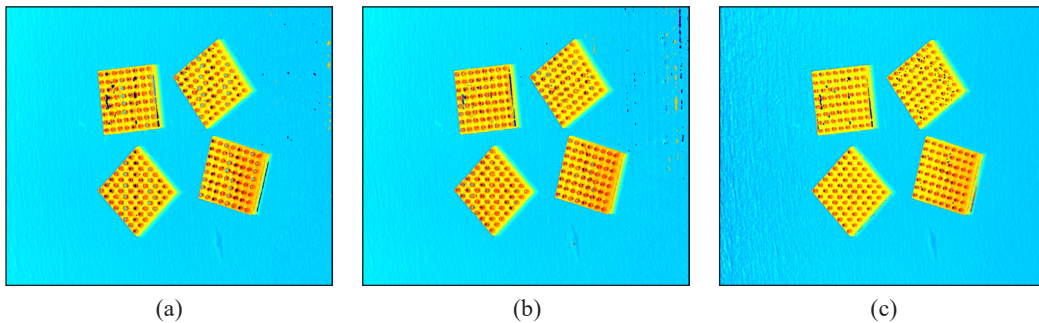


Figure 6. 3D reconstruction of solder ball grid (a) Three-step phase shifting (b) Jiang’s HDR method (c) Proposed HDR method

CONCLUSION

This paper has shown an enhancement of the current HDR method. The proposed method can increase the dynamic range of the 3D shape measurement by controlling different channels of colour projected and isolating each channel.

To increase the dynamic range of the surface reflectivity, the RGB channels are used to control the intensity projected. Not only that, each RGB channels can compute the 3D data by itself. The experiment demonstrates that the 3D shape measurement can be successfully measured even at higher intensity while the current method had already overexposed.

The current development of the HDR method that uses multiple camera exposures or multiple projector intensity requires many sets of images to be taken. This causes the process to be slow. Our proposed method can capture only one set of images and compute the HDR image. Consequently, our proposed method significantly reduces the 3D objects measurement times. However, our proposed method has the limitations of being unable to compute the 3D data at low exposure and there is no quality checking of the fringe captured.

ACKNOWLEDGEMENT

The equipment was provided by Estek Automation Sdn Bhd. This study would not be made possible without the help from Estek Automation Sdn Bhd.

REFERENCES

- Aboali, M., Manap, N. A., Darsono, A. M., & Yusof, Z. M. (2017). Review on three dimensional (3-D) acquisition and range imaging techniques. *International Journal of Applied Engineering Research*, 12(10), 2409-2421.
- Achar, S., Bartels, J. R., Whittaker, W. L., Kutulakos, K. N., & Narasimhan, S. G. (2017). Epipolar time-of-flight imaging. *ACM Transactions on Graphics*, 36(4), 1-8. <https://doi.org/10.1145/3072959.3073686>
- Babaie, G., Abolbashari, M., & Farahi, F. (2015). Dynamics range enhancement in digital fringe projection technique. *Precision Engineering*, 39, 243-251. <https://doi.org/10.1016/j.precisioneng.2014.06.007>
- Chen, C., Gao, N., Wang, X., & Zhang, Z. (2018). Adaptive projection intensity adjustment for avoiding saturation in three-dimensional shape measurement. *Optics Communications*, 410(October 2017), 694-701. <https://doi.org/10.1016/j.optcom.2017.11.009>
- Faes, M., Abbeloos, W., Vogeler, F., Valkenaers, H., Coppens, K., Goedemé, T., & Ferraris, E. (2016). Process Monitoring of Extrusion Based 3D Printing via Laser Scanning. *Computer Vision and Pattern Recognition*, 2016, 1-5. <https://doi.org/10.13140/2.1.5175.0081>
- Feng, S., Zhang, L., Zuo, C., Tao, T., Chen, Q., & Gu, G. (2018). High dynamic range 3-D measurements with fringe projection profilometry: A review. *Measurement Science and Technology*, 29(12), Article 122001. <https://doi.org/10.1016/j.vetmic.2007.04.014>
- Huang, S. H., & Pan, Y. C. (2015). Automated visual inspection in the semiconductor industry: A survey. *Computers in Industry*, 66, 1-10. <https://doi.org/10.1016/j.compind.2014.10.006>
- Jiang, C., Bell, T., & Zhang, S. (2016). High dynamic range real-time 3D shape measurement. *Optics Express*, 24(7), Article 7337. <https://doi.org/10.1364/OE.24.007337>
- Liu, Y., Xi, J., Yu, Y., & Chicharo, J. (2010). Phase error correction based on Inverse Function Shift Estimation in Phase Shifting Profilometry using a digital video projector. In *Optical Metrology and Inspection for Industrial Applications* (Vol. 7855, p. 78550W). International Society for Optics and Photonics. <https://doi.org/10.1117/12.870369>

- Malamas, E. N., Petrakis, E. G., Zervakis, M., Petit, L., & Legat, J. D. (2003). A survey on industrial vision systems, applications and tools. *Image and Vision Computing*, 21(2), 171-188. [https://doi.org/10.1016/S0262-8856\(02\)00152-X](https://doi.org/10.1016/S0262-8856(02)00152-X)
- Pavlicek, P., & Mikeska, E. (2018). White-light interferometry without depth scan. In P. Zemánek (Ed.), *21st Czech-Polish-Slovak Optical Conference on Wave and Quantum Aspects of Contemporary Optics* (Vol. 10976, p. 9). SPIE. <https://doi.org/10.1117/12.2517877>
- Salahieh, B., Chen, Z., Rodriguez, J. J., & Liang, R. (2014). Multi-polarization fringe projection imaging for high dynamic range objects. *Optics Express*, 22(8), Article 10064. <https://doi.org/10.1364/OE.22.010064>
- Song, Z., Jiang, H., Lin, H., & Tang, S. (2017). A high dynamic range structured light means for the 3D measurement of specular surface. *Optics and Lasers in Engineering*, 95(September 2016), 8-16. <https://doi.org/10.1016/j.optlaseng.2013.12.013>
- Waddington, C., & Kofman, J. (2014). Modified sinusoidal fringe-pattern projection for variable illuminance in phase-shifting three-dimensional surface-shape metrology. *Optical Engineering*, 53(8), Article 084109. <https://doi.org/10.1117/1.OE.53.8.084109>
- Wang, M., Du, G., Zhou, C., Zhang, C., Si, S., Li, H., Lei, Z., & Li, Y. (2017). Enhanced high dynamic range 3D shape measurement based on generalized phase-shifting algorithm. *Optics Communications*, 385, 43-53. <https://doi.org/10.1016/j.optcom.2016.10.023>
- Wolff, L. B. (1989). Using polarization to separate reflection components. In *Proceedings CVPR'89: IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 363-369). IEEE Conference Publication. <https://doi.org/10.1109/CVPR.1989.37873>
- Zhang, S. (2016). *High-speed 3D imaging with digital fringe projection techniques*. CRC Press. <https://doi.org/10.1017/CBO9781107415324.004>
- Zhang, S. (2018). High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering*, 106(February), 119-131. <https://doi.org/10.1016/j.optlaseng.2018.02.017>
- Zhang, S., & Huang, P. S. (2006a). High-resolution, real-time three-dimensional shape measurement. *Optical Engineering*, 45(12), Article 123601. <https://doi.org/10.1117/1.2402128>
- Zhang, S., & Huang, P. S. (2006b). Novel method for structured light system calibration. *Optical Engineering*, 45(8), Article 083601. <https://doi.org/10.1117/1.2336196>
- Zhang, S., & Yau, S. T. (2009). High dynamic range scanning technique. *Optical Engineering*, 48(3), Article 033604. <https://doi.org/10.1117/1.3099720>
- Zuo, C., Feng, S., Huang, L., Tao, T., Yin, W., & Chen, Q. (2018). Phase shifting algorithms for fringe projection profilometry: A review. *Optics and Lasers in Engineering*, 109(200), 23-59. <https://doi.org/10.1016/j.optlaseng.2018.04.019>

The Built Environment and its Impact on Transit based Transportation Users Walking Activity in Semarang, Indonesia

Anita Ratnasari Rakhmatulloh^{1*}, Diah Intan Kusumo Dewi¹ and Dinar Mutiara Kusumo Nugraheni²

¹Department of Urban and Regional Planning, Faculty of Engineering, Diponegoro University, Semarang, Prof. Soedarto, Tembalang, Semarang 50275, Indonesia

²Department of Computer Science, Faculty of Science and Mathematics, Diponegoro University, Semarang, Prof. Soedarto, Tembalang, Semarang 50275, Indonesia

ABSTRACT

The increase in mobility rate due to the current rise in population has become a challenge in urban development. This led to the development of pedestrian walkways, which are integrated with transit-based transportation, to minimize problems due to the high level of mobility of the urban population. According to some experts, environmental conditions are one of the main factors capable of affecting pedestrians' frequency. Therefore, this study explores the effect of the built environment, such as density, diversity, and design, on the pedestrian frequency of the Bus Trans Semarang passengers. Data were collected from 9 corridors, consisting of 447 bus stops, directly connected to the pedestrian walkways, as one of the feeders for transit-based transportation facilities. The analysis method used for each variable was the formulation technique, with data simulated using various applications, such as ArcGIS. The linear regression partial t-test model was also analyzed using SPSS, with the effect of pedestrian frequency on dependent variables determined, using the built environment elements (independent variable). The results showed a positive

significance between the diversity variable and pedestrians' frequency along the Bus Trans Semarang corridors. Furthermore, the land that varies with the dominance of commercial and educational area has the highest attraction features that affects a person's frequency to walk.

ARTICLE INFO

Article history:

Received: 16 November 2020

Accepted: 10 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.05>

E-mail addresses:

anita.ratnasari.r@gmail.com (Anita Ratnasari Rakhmatulloh)

diah.intan@pwk.undip.ac.id (Diah Intan Kusumo Dewi)

dinar.mutiara@live.undip.ac.id (Dinar Mutiara Kusumo

Nugraheni)

* Corresponding author

Keywords: Built environment, pedestrian walkways, transit-based transportation

INTRODUCTION

The development of a transit-based transportation system is intensively carried out as an alternative means of private transportation to overcoming population growth in various countries. According to Makarova et al. (2017), the Asian region is generally known for its unhealthy transportation system, which causes various problems, such as congestion, air pollution, and sprawl-related issues (Makarova et al., 2017). Therefore, to overcome this, the World Health Organization (WHO) encourages countries to implement sustainable transportation systems that focus on safe, efficient, accessible, affordable, inclusive, green, and healthy transportational mobility.

A Transit-Oriented Development (TOD) is a form of sustainable transportation that emphasizes integrating urban spatial design to unite people, activities, buildings, and public spaces through connectivity, accessible by walking (ITDP, 2017). Increasing walking intensity (walkability) is considered part of the effective solutions in overcoming unhealthy transportation because it minimizes congestion, increases environmental sustainability, encourages physical activity, improves public health, and enhances urban settlements' appropriateness (Blanco & Alberti, 2009). The pedestrian walkways, commonly referred to as a sidewalk in a transit-based transportation system, is a facility feeder that connects a transit centre with other activity functions (Hu et al., 2013). However, the lack of optimal pedestrian walkways is a challenge for establishing a sustainable transportation system in several countries.

Various studies suggested that the optimization of pedestrian walkways, such as the completeness of design components, an attractive, safe, and comfortable city shape increases residents' tendency to walk to a destination (Özbil et al., 2015). Furthermore, the developments of a land-use planning, and urban design are crucial in increasing pedestrian activity in cities (Timmermans, 2009). However, research on the relationship between the conditions of the built environment, and walking tendency in the transit-based transportation system is rare. Various parties have been inclined to focus on increasing the transportation system's capacity, without observing how the developments affect the environment. Therefore, the development of a new transportation system is linear with the uneven distribution of building density and land-use management, thereby leading to inadequate functioning of several pedestrian walkways.

The built-environment is a transit-based transportation system commonly found in European countries. The initial milestone associated with this technique was first carried out by Cervero & Kockelman (1997) in the European countries to welcome new urbanism concepts. According to them, the built environment means physical features of the urban landscape (i.e., alterations to the natural landscape). These authors formulate three main components to assess the built environment from the same source, namely density, diversity, and design.

The density is defined as the component produced by observing population frequency per household and building quality per area (Cervero & Kockelman, 1997; Monteiro & Campos, 2012; Zhang et al., 2016). Furthermore, this research and the pedestrian activity in previous studies observed that the areas with a high-density level are associated with walking (Oktaviani et al., 2020). A study carried out by Duduta (2013) developed indicators for assessing density components in constructing measurements, which are properly related to building height or the Floor Area Ratio (FAR) analysis method. The study takes a minimum distance of 200m from the stop point, which was recorded, and divided into three categories of level assessment, namely low (1-2 floors), medium (3-8 floors), and high (>8 floors) densities.

The diversity component is represented as a land-use balancing index used to compare the ratio of residential and non-residential (commercial) areas. According to Sung et al. (2013), the level of balancing land-use in an area rises, the activities of pedestrians relatively increased. Furthermore, by simplifying Cervero and Kockelman (1997) research, the diversity component emphasizes the calculation model of residential and non-residential land area ratios, which in this case, is measured by the balancing index formula (Sung et al., 2013). The diversity variable analysis results show the patterns of land use and its changes due to population mobility. Several studies, such as those carried out by Cervero, found that the effects of diversity have long been ignored.

The design component is part of the features referring to the assessment of pedestrian walkways and its integration with the road network system (Dixon, 1996; Munshi, 2016). Similar to diversity, the effects of design also have long been ignored in built environment analysis. However, the design component's measurement is more complicated than density and diversity, with a system of assessing pedestrian walkway completeness elements adjusted to environmental conditions (Zhang et al., 2016). Therefore, this study summarized the design indicators needed to identify the appropriate pedestrian walkways at the research location. This includes the lane width, route count, road intersection number, the presence of ramps, bridges, zebra crossings, and the completeness of street furniture.

The country context plays an essential role in ascertaining the built environment assessment of pedestrian activities. The biggest question that emerged from previous research is whether the context of the European region's built environment is similar to the Asian zone. Furthermore, a research in downtown Montreal, Canada showed that the components of the built environment, density (population, number of jobs, schools, and bus stops), diversity (commercial land use), and design (percentage of main arteries, average length of roads), affected pedestrian activity (Miranda-Morenoa et al., 2011). This city's position in Canada is a provincial capital not included in the national central area. It is observed to be a part of the cities located in developed countries. In this location, the problem encountered was a lack of pedestrians' security on the walkways. However,

this location had a complete pedestrian walkways element, surrounding skyscrapers. As a four-season country, the provision of vegetation along the pedestrian walkways in the city is designed to suit seasonal conditions. This is because the designs that adapt to climate change and periods are being intensified to create sustainable urban life.

The discussion of the built environment, on pedestrian walkways in developing countries, for example, the transit area in Bangkok, Thailand, was developed from an interplay between people's tendency to walk and the diversity of land-use around the transit area (Townsend & Zacharias, 2010). The walkways condition in Bangkok, Thailand, is associated with air pollution and lack of sidewalks (Townsend & Zacharias, 2010; Chalermpong, 2007). Therefore, from this study, it is obtained that the development of commercial, retail, and residential land is relatively high around the transit centre, which shows that the influence of diversity is high on pedestrian movement patterns.

The research carried out in other Asian regions, such as Beijing, China, obtained different results. According to Zhao et al. (2018), the findings on the built environment at the environmental-scale transit centres in Beijing, show that the density component plays a substantial role in influencing pedestrian activities. The study showed that the pedestrian walkway components in Indonesia was complete compared to those in Bangkok. Also, the influence of the vertical building construction policy in the central area of Beijing City is the basis for forming a compact area, with relatively close distances between buildings. Based on that statement, the density of the built environment in this area influences a person's inclination to walk. Therefore, the geographic characteristics with the research's scope carried out in Bangkok and Beijing relatively similar. This is because they are often conducted in locations with relatively flat homogeneous topography, using city-scale research coverage.

Semarang is one of the major cities in Indonesia, located in Central Java Province. The city intensively plans to develop a Transit Oriented Development (TOD) system, accommodate the community's mobility needs, and create a sustainable transportation channel. The requirement needed for the development of the TOD system in this city is the Bus Trans Semarang. The relationship between the built environment and pedestrian activities in the central transit area, along the Bus Trans Semarang corridor, seems interesting because it covers the city centre area with a diverse topography in its surroundings. The problem of pedestrian walkways provision in Semarang City is the same as that of Bangkok, which is associated with pollution related-issues, and lack of sidewalks. Based on further data, it was observed that approximately 88% of Bus Trans Semarang passengers walked to or from the bus stops (Purwanto & Manullang, 2018). This result is certainly in contrast to the lack of pedestrian walkways provision in the city of Semarang, which then causes a gap between availability and demand.

Policies regarding land-use, regulated by City Territory Section (CTS), indicated that Semarang City's use is very diverse. Furthermore, the absence of a firm policy on building construction efforts has led to a high sprawl level in this area. Therefore, this study aims to ascertain the best-built environment model, which affects a person's tendency to walk along the Bus Trans Semarang corridor, from the city centre to the suburbs. It also complements previous research that only focused on density conditions in the built environment analysis. This research is further accompanied with a question, such as, 'Do the built environment components affect the pedestrian activity of transit-based transportation users in Semarang City?'

METHODS

Semarang Study Area

Semarang City is one of the metropolitan cities in Indonesia, located in Central Java Province, with a population of 1.79 million people. According to the data obtained from the Semarang City Statistics Centre in 2019, the population of this area keeps increasing by 50% yearly, accompanied by the provision of a non-optimal public transportation system. This location is one of the pilot project for Bus Rapid Transit (BRT) development in Indonesia, with the provision of Bus Trans Semarang in 2009, to help improve the quality of transportation system (SuaraMerdeka, 2015).

Therefore, this research covers all Bus Trans Semarang routes, consisting of, Corridors I, II, IIIA, IIIB, IV, V, VI, VII, and VIII, as shown in Figure 1. According to Kim et al. (2018), two methods are used to measure the built environment, based on the area's scope, namely, the street level of 50m and the neighborhood position at a maximum radius of 400m, from the observation point. The data showed that the maximum distance for walking in Semarang City was only 200m-250m between destination locations (Purwanto & Manullang, 2018). Therefore, sampling of the observation area, using the neighbourhood level method, with a radius of 200m from the right and left of the road was carried out in this research.

In this study, the entire Bus Trans Semarang corridor has 596 bus-stops, scattered from the city centre to the suburbs, while also connecting various important land-uses, such as terminals, government areas, education fields, and many more, as shown in Table 1. Data from the Semarang City Transportation Office in 2018 showed a significant increase in Bus Trans Semarang passengers from 2010-2018. The total number of passengers between 2010 and 2018 were 369,326 and 10,210,296, respectively. Furthermore, the largest contributor for pedestrians in this study was Corridor I, which had massive bus-stops and fleets compared to the other eight corridors. Based on the data collected from the field observations, pedestrian activities along the Bus Trans Semarang corridor were dominated by school, work, and recreational activities. City centre areas such as Corridors I, II, and IV, had more massive pedestrian activities than other zones.

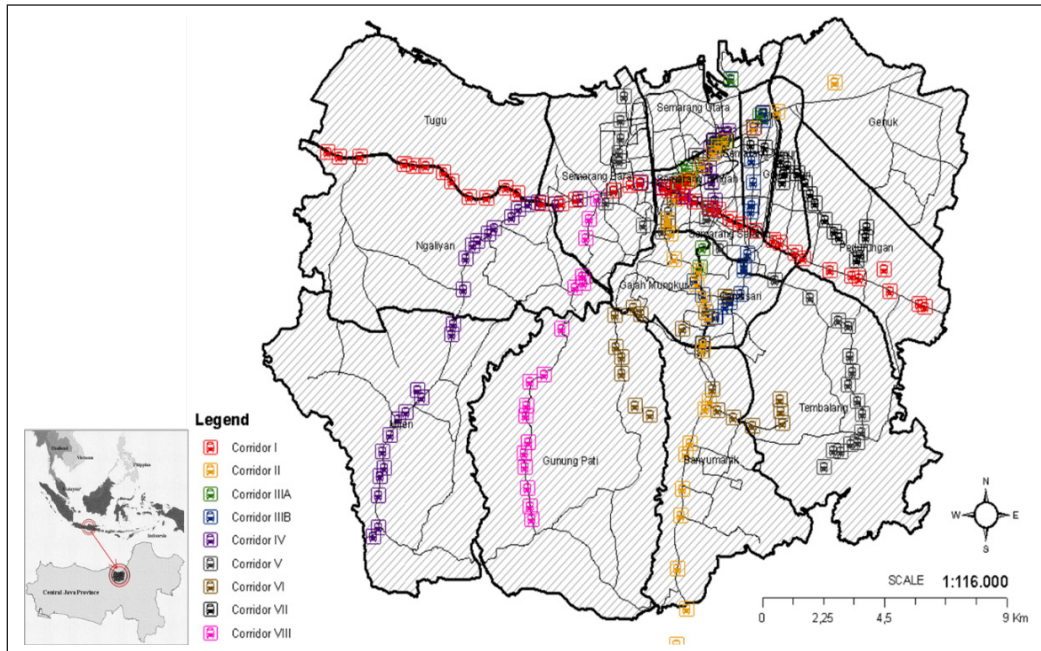


Figure 1. Bus Trans Semarang stops in the study area

Table 1
Bus Trans Semarang corridor profile

Corridor	Route	Number of			Corridor Length (Km)	Operating Years
		Pedestrian (People)	Bus (Unit)	Stops (Unit)		
I*	Mangkang Terminal→Penggaron Terminal.	2,713,852	25	81	60.0	2009
II	Terboyo Terminal→Sisemut Terminal.	838,081	34	74	26.5	2012
IIIA	Tanjung Emas Port→KagokTanjung Emas Port.	166,934	16	45	10.5	2014
IIIB	Tanjung Emas Port→Elizabeth Hospital→Tanjung Emas Port.	111,289		40	10.0	2014
IV	Cangkiran Terminal →Semarang Tawang Station.	520,436	22	87	22.3	2013
V	Meteseh Residential AreaAirport→Marina Beach.	366,319	16	75	25.9	2017
VI	Diponegoro University→Semarang State University.	257,831	16	63	13.4	2017
VII	Terboyo Terminal→Semarang City Hall.	87,886	13	63	6.9	2018
VIII	Cangkiran Terminal→Simpang Lima (City Center).	56,799	20	68	22.4	2019

Note. *corridors with the most pedestrians

Quantitative Analysis

This research utilized the quantitative research technique, which focused on analyzing the built environment in the form of density, diversity, and design. Data was collected through naturalistic direct field observation methods. Furthermore, the density analysis assessed the population frequency by comparing the total populace, using the interpolation method (ArcGIS 10.3). This analysis produced a spatial picture of population density in each Corridor. Also, the density component analysis used the numerical calculation method, such as the Floor Area Ratio (FAR), to compare the number of buildings per floor area. The calculations in accordance with Duduta's (2013) research, were divided into three categories, namely low (Equation 1), medium (Equation 2), and high (Equation 3) densities, as shown in Equation 1-3.

$$\text{Low density} = \frac{\text{number of buildings 1-2 floors}}{\text{the total number of buildings within a 200 meter radius of the stop point}} \quad [1]$$

$$\text{Medium density} = \frac{\text{number of buildings 3-8 floors}}{\text{the total number of buildings within a 200 meter radius of the stop point}} \quad [2]$$

$$\text{High density} = \frac{\text{number of buildings >8 floors}}{\text{the total number of buildings within a 200 meter radius of the stop point}} \quad [3]$$

Diversity components were analyzed, using the land-use ratio analysis, while comparing the residential, commercial, office space, public services, and recreational areas, to the total space around the Bus Trans Semarang bus-stop, in each Corridor. According to Sung et al. (2013), this diversity of land-use is used to compare the residential and non-residential areas, using the balancing index (RNR) formula, as shown in Equation 4.

$$\text{RNR}_i = 1 - \left[\frac{\text{Res-Non res}}{\text{Res+Non res}} \right] \rightarrow \text{The closer to 1, the more balance} \quad [4]$$

The design component was carried out by analyzing the road structure and pedestrian walkways' completeness features, based on movement needs. In this case, the assessment indicators were completeness elements, which include crossings, ramps, intersections, lane width, number of routes, and sidewalk furniture. The analysis of this component was presented to visualize the existing conditions, which were compared with the required pedestrian feature demand data, in Semarang. Further, the overall results of the density, diversity, and design analysis were processed, using the significance of linear regression (partial T-test) and the SPSS Statistics 24 application, to determine the effect of the built environment on walking activities in each Corridor. The model formula used in determining the influence between the variables is shown in Equation 5.

$$Y = a + b(X) \quad [5]$$

Where;

Y = frequency of pedestrians (person)

a = constant

b = coefficient

X = built environment variables

Linear regression had been widely used in previous studies to produce a relationship model between the built environment variables and pedestrian movement patterns in an area. The requirements of linear regression analysis (partial T-test) is that when the results of the coefficients (t) are greater than the t-table (2.35), with the regression data's significant value less than 0.05 (<0.05), then both variables (dependent & independent), influences each other. The built environment variable consists of density (X1), diversity (X2), and design (X3) as independent variables, which correlates with the pedestrian intensity in each Corridor (Y), as shown in Figure 2.

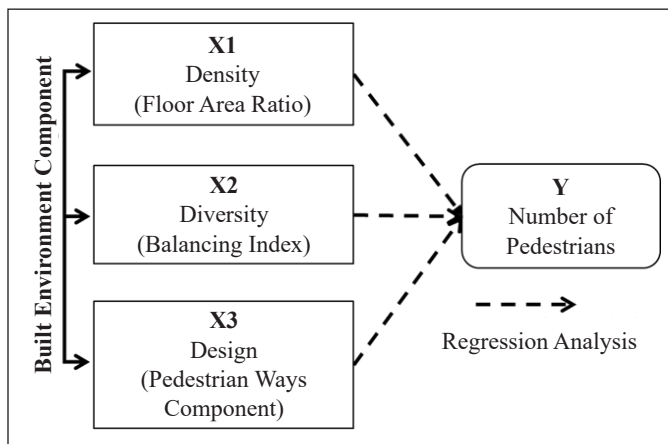


Figure 2. Linear regression analysis diagram

RESULT AND DISCUSSION

Density Analysis

Based on the comparative calculation between total population and land area in Semarang City, the highest population density was observed to occur in the downtown vicinity (Figure 3). This is similar to the theory of urban growth in the world, stating “the closer to the city centre, the higher the population density” (Fee & Hartley, 2011). The high population density increases the demand for public transportation, including BRT, therefore, confirming the need to integrate services in areas with high density (Patankar et al., 2007). Figure 3

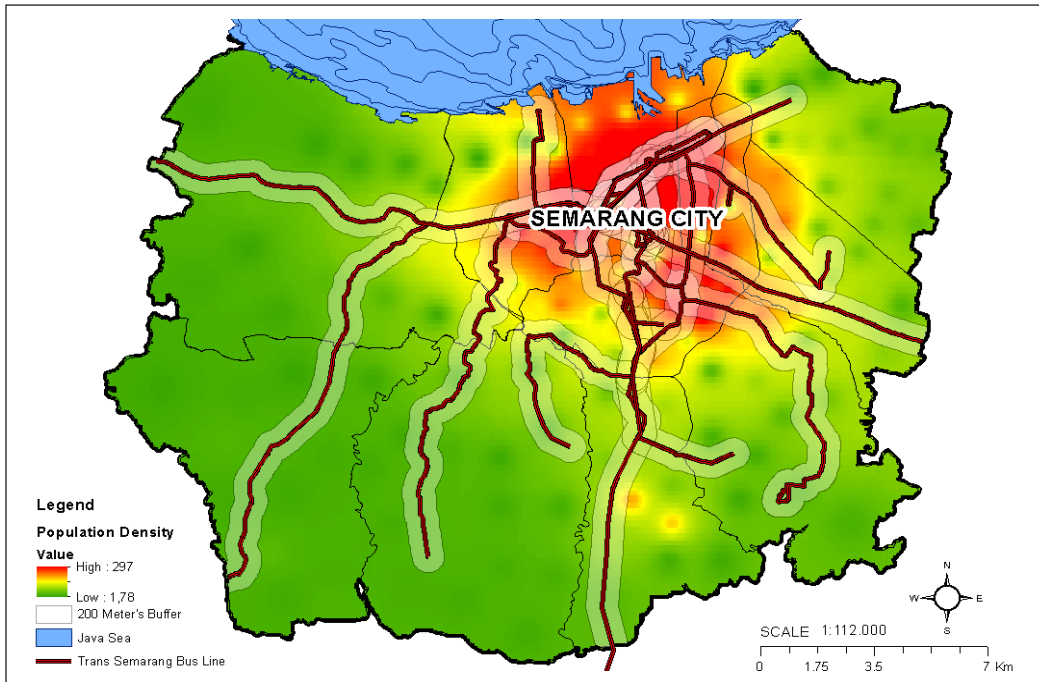


Figure 3: The population density in buffer area, 200 meters from bus corridor

shows the implementation of the Bus Trans Semarang system has paid attention to this theory, as observed in all routes (I-VIII) passing through the city centre and numerous bus-stops provided.

In addition to calculating population density, the component was measured by observing the quality of building. Table 2 shows the calculation of density values, with IIIB having the highest total, compared to other routes. The density in this corridor was dominated by 1-2 floor buildings, indicated by the low frequency of 0.983. Meanwhile, data from the observations showed the entire Bus Trans Semarang corridor were dominated by developments in a horizontal direction and following the road network. According to Table 2, with regard to the element density conditions for each corridor, only three out of nine routes had a relatively high frequency value (33%).

Furthermore, the research on transportation hubs in Japan discovered a high Floor Area Ratio (FAR) with optimized vertical building construction, improves land-use quality, while creating effective movement for pedestrians (Yang & Yao, 2019). Thus, high density areas surrounded by vertical buildings, minimized the use of motorized vehicle and increased pedestrian intensity (Cervero & Kockelman, 1997; Monteiro & Campos, 2012; Zhang et al., 2016). Also, the density analysis along the Bus Trans Semarang corridor, produced results contrary to the study by Zhao et al. (2018), stating the dominance of density effect on the built environment and pedestrian activity in Beijing was high. This was in line with

Table 2
The condition of the element density of all corridor radius 200 meters (Total)

Corridor	Building			Total (unit)	Lenght of corridor (Km)	Large of corridor (Ha)	Total Density (unit/ Ha)	Density		
	1-2 floors (unit)	3-8 floors (unit)	>8 floors (unit)					Low	Medium	High
(1)	(2)	(3)	(4)	(5)	(6)	200*(6) = (7)	(5)/(7)	(2)/(5)	(3)/(5)	(4)/(5)
I	4,358	290	8	4,656	60.0	1,200	4	0.897	0.096	0.007
II	3,833	130	3	3,966	26.5	530	7	0.960	0.043	0.001
IIIA	3,258	84	6	3,348	10.5	210	16	0.958	0.039	0.003
IIIB*	6,249	85	6	6,340	10.0	200	32	0.983	0.016	0.001
IV	714	167	21	902	22.3	446	2	0.758	0.172	0.028
V	968	152	18	1,138	25.9	518	2	0.817	0.164	0.039
VI	3,587	121	0	3,708	13.4	268	14	0.936	0.064	0.000
VII	654	110	2	766	6.9	138	6	0.824	0.149	0.002
VIII	171	96	7	274	22.4	448	1	0.632	0.341	0.027

Note. *corridor with the highest density

the Chinese government's regulation of intensifying vertical development, to implement a sustainable city development, and is different from the existing regulation condition in Semarang, with no emphasis on vertical building. However, increased vertical building growth is required, especially in downtown areas, in terms of land-use optimization, with the creation of a pedestrian-friendly and sustainable transit-based transportation system.

Diversity Analysis

Table 3 shows the equilibrium calculation results, with the total balance index of all the corridors close to 1, thus, categorized as areas with mixed land-use, possessing high parameters (Sung et al., 2013). Furthermore, the highest balancing index of 1.110, was obtained in the Corridor II area, with trade and service centre as the main land-use. The area's function as a centre for trade and services, has led to increasingly diverse land-use, and the proportion of residential (60 hectares) to non-residential (59 hectares) land in the area is very stable within a 200m radius, from the transit centre. As one of the longest routes, Corridor I had the smallest balancing index of 0.395 (getting away from 1). In this corridor, the proportion of residential to non-residential land-use area is unbalanced. The residential land amounts to 98.70ha, while non-residential areas dominated by trade and services, offices, and industry, was only 40.22ha.

In addition to the total calculation results, the table above also explains the stop points with the highest balancing index value on each route. This aims to serve as the main sample, showing the spatial balancing index conditions through land use, per building. In general, Figure 4 explains the position of the balancing index sample point for each

Table 3
Condition element diversity for all corridor radius 200 meters (Total)

Corridor	Res (Ha)	Non-res (Ha)	Balancing Index (1)/(2) = (3)	Landuse (4)
I**	98.70	40.22	0.395	Settlements, industry, trade and services, warehousing, worship, defense and security, offices, education, health, recreation, public spaces, sports.
II*	60.59	58.72	1.110	Offices, trade and services, police education and sports, industry, and housing.
IIIA	64.58	89.44	0.831	Defense and security, education, health, offices, public spaces, recreation, settlements, trade and services, and worship.
IIIB	56.55	23.04	1.184	
IV	261.32	339.46	0.870	Offices, worship, health, trade and services, education, housing, recreation, defense and security, and transportation.
V	215.86	311.21	0.820	Education, health, industry, offices, defense and security, housing, sports, recreation, trade and services, and worship.
VI	57.68	31.10	1.314	Defense and security, education, health, offices, settlements, recreation, trade and services, and worship.
VII	129.01	247.47	0.685	Education, health, industry, offices, defense and security, housing, sports, recreation, trade and services, and worship.
VIII	158.80	167.19	0.938	Trade and services, worship, health, education, offices, recreation, defense and security, and settlements.

Note. Res (Residential area), Non-res (Non-residential area): *corridor with the highest balancing index; **corridor with the lowest balancing index

corridor on the distribution of land use in Semarang City. Meanwhile, Figure 5, section H shows the location of the highest balancing index for Corridor I is the same with VIII, at the Amarta bus-stop point, with a value of 1.10, and dominated by non-residential land-use, including trade and services. Meanwhile, Figure 5, section A show the location with the highest balancing index for Corridor II area is the City Hall bus-stop, with a value of 0.96, and dominated by non-residential land-use, in the form of offices as well as educational buildings. Figure 5, section B shows the location with the highest balancing index on Corridor IIIA is the Raden Patah bus-stop point, with a value of 1.06, and dominated by non-residential land-use in the form of trade as well as services, with school buildings. However, Figure 5, section C shows the highest balancing index for Corridor IIIB was located at the Don Bosco stopping point, with a value of 1.06, and dominated by residential land use, including housing and school buildings.

Furthermore, Figure 5, section D shows the location with the highest balancing index in the Corridor IV area is Nyaliyan Square bus-stop, with a value of 0.99, and dominated by non-residential land use, in the form of trade as well as services, with educational buildings. Meanwhile, Figure 5, section E shows the highest balancing index in Corridor

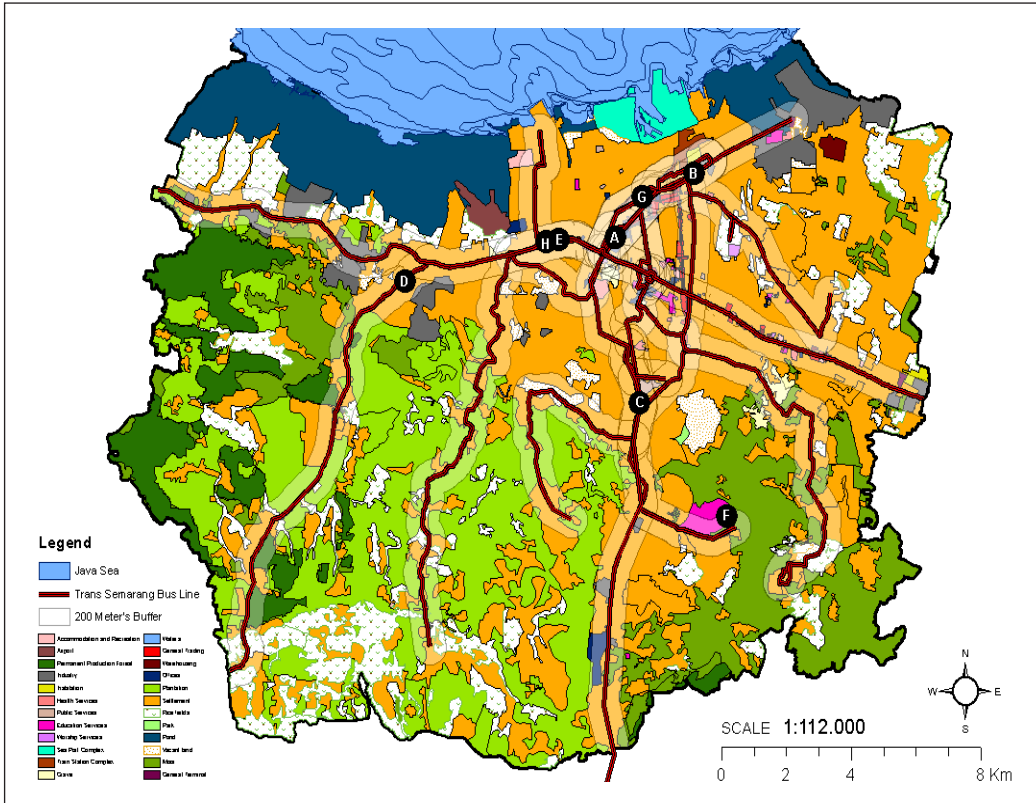


Figure 4. The distribution of land use and location of balancing index sample points in Semarang City

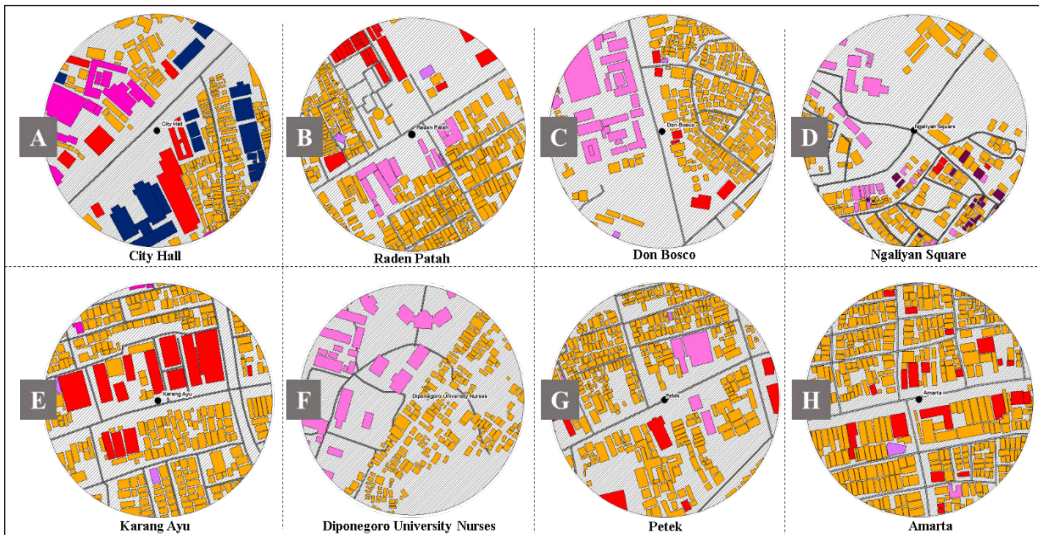


Figure 5. The land use corridor and highest balancing index value (200 meters) ; (A) Corridor II, (B) Corridor IIIA, (C) Corridor IIIB, (D) Corridor IV, (E) Corridor V, (F) Corridor VI, (G) Corridor VII, (H) Corridor I & VIII.

V area, is located at the Karang Ayu bus-stop, with a value of 1.00 (very balanced), and dominated by settlement land-use. Figure 5, section F shows the highest balancing index value in Corridor VI, is located at the Diponegoro University Nursing bus-stop (educational area), with a value of 1.04, and dominated by non-residential land-use, in the form of educational buildings. Meanwhile in Figure 5 section G, shows the the highest balancing index in Corridor VII is at the Petek bus-stop point, with a value of 0.99, and dominated by non-residential land use, in the form of trade as well as services.

The diversity analysis results for all Bus Trans Semarang corridors had the same pattern of balance index. In addition, locations with the highest balancing index, were dominated by land-use for education, trade, and services. Several studies stated commercial land-uses (educational with trade and services) were able to increase people's interest in walking activity, within the area (Miranda-Morenoa et al., 2011; Townsend & Zacharias, 2010; Chalermpong, 2007). Based on field observations, activities in the area of education, trade and services are conducted only at morning as well as evening, especially on weekdays (Monday-Friday), and pedestrian walkways are rarely used during holidays. In addition, an increase in attractions is required to bring in and improve activities, thus, allowing public facilities, including pedestrian walkway, to function optimally. According to Untermann and Lynn (1984), areas with mixed land-use characteristics on foot are more attractive, because the resulting movements are more diverse. Therefore, various types of land-use are required to improve population activities' pattern, consequently, optimizing the usage of space and creating a sustainable pedestrian walkways system.

Design Analysis

The pedestrian walkways design analysis showed Corridors IV, I, and II are routes with the highest number of inclines, crossings, and irregularities in comparison with the others as indicated in Table 4. Moreover, a linear pattern is produced towards the pedestrian walkway facilities when compared with the data presented in Table 1 and the corridors with the highest number of pedestrians were observed to have the tendency of having more complete facilities. It was also discovered that the closeness to the city center where there is a denser population activity level leads to a better and more complete walkway facilities design. Meanwhile, some locations observed to be far from the city center such as those traversed by Corridor VI have pedestrian walkways with a width of 1m-2m and are only equipped with vegetations as shown in Table 4 and Figure 6.

The sidewalks along the Bus Trans Semarang corridor needs improvement due to the differences observed between the current condition and user needs. The region's undulating topography and climatic conditions makes the average walking distance in Semarang City to be only 200m-250m. Therefore, the completeness of the design components and city environmental conditions for the pedestrian walkways has the ability to increase people's

Table 4
Complete pedestrian walkways design for the Bus Trans Semarang corridor radius 200 meters (Total)

Corridor	Pedestrian walkways Large (m)	Lanes of the road	Number of			Street Furniture
			Inter sections	Ramp	Crossings	
I*	1-5	2 – 6	172	267	42	Shade vegetation, bollards, lights, trash cans, traffic signs, seats, special lanes for the disabled.
II*	1-5	2	165	138	38	Traffic signs (no stopping sign, no parking), street lights, ways with disabilities, and potted vegetation.
IIIA	1-5	1 - 2	95	3	13	Lights, bollards, vegetation, traffic signs, trash cans, fire hydrants.
IIIB	1-1.5	1 - 2	246	0	22	Shade vegetation, and traffic signs.
IV*	1-3	2 - 4	110	396	30	Shade vegetation, bollards, lights, traffic signs, and special lanes for the disabled.
V	0.5-2.5	1 - 2	170	27	31	Shade vegetation, lights, traffic signs, benches, and special lanes for people with disabilities.
VI	1-3	2	132	40	18	Traffic lights and signs.
VII	1-2.5	1 - 2	127	21	12	Shade vegetation, lights, traffic signs, and ways for the disabled.
VIII	1-3	2 - 6	24	184	16	Shade vegetation, special lanes for people with disabilities, and traffic signs.

Note. corridor with the most complete design elements

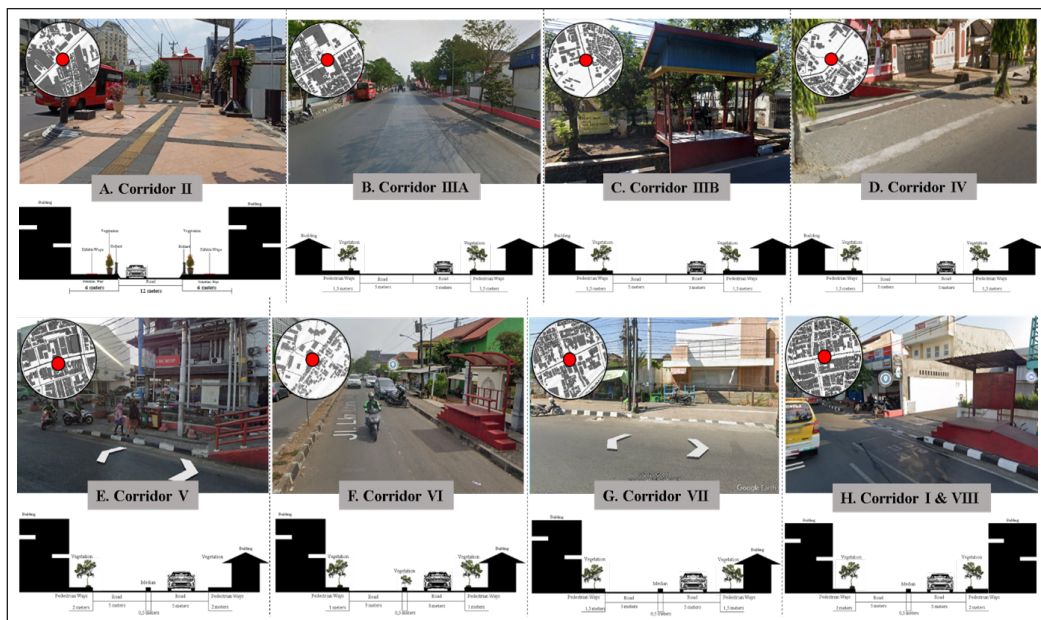


Figure 6. Pedestrian walkways conditions throughout the Bus Trans Semarang corridor. Sample of pedestrian walkways for corridor II, IIIA, IIIB, IV, V, VI, VII, and VIII.

willingness to walk to and from a destination (Özbil et al., 2015). It has also been reported that completed facilities such as ramps, sidewalks, vegetation, and several others affect walking activities in the area (Dixon, 1996; Munshi, 2016). However, the country has a tropical climate and year-round sunshine which causes some level of discomfort in engaging in outdoor activities such as walking. This means there is a need for adjustment based on the climate by using vegetation as shade which has been reported to be applicable in four-seasoned countries as observed in Montreal, Canada (Miranda-Morenoa et al., 2011). This vegetation also has the ability to produce sufficient oxygen and reduce carbon dioxide from motorized pollution at the same time. This, therefore, means it is necessary to have a pedestrian walkway with complete facilities and focus on both environmental and climatic conditions in order to achieve a sustainable transportation system.

The Built Environment Model Affecting Pedestrians

The linear regression test conducted using the partial coefficient T-test method was used to determine the significance of the relationship between two or more variables. This involves the application of two things which are the coefficient T and the significant results as the benchmarks. The number of pedestrian passengers at the Bus Trans Semarang per corridor (Y) was used as the dependent variable while the built environment which included the total density (X1), balancing index diversity (X2), and the number of walkway design elements (X3) in a radius of 200 from corridors were used as the independent variables. The results presented in Table 5 showed the T-value of density (X1) is 0.454 with a significance value of 0.669 (>0.05) and this means this variable did not significantly influence the pedestrian intensity (Y) due to the fact that the calculated T-value was lesser than the T-table value of 2.353 while the significance is greater than 0.05.

The table also shows diversity (X2) had a coefficient T-value of 7.028 (> 2.353) with a significance calculated value of 0.001 (<0.05) and this means there is a significant relationship between the pedestrian intensity (Y) and diversity (X2). This shows that

Table 5
The results of data processing linear regression (partial T-test) SPSS

Code	Independent Variable	Unstandardized B	T*	Sig**
Coefficient		2.630		
X1	Density (Floor Area Ratio calculation)	0.011	0.454	0.669
X2	Diversity (The ratio of residential and non-residential land or Balancing Index)	2.172	7.028	0.001
X3	Design (The completeness of the pedestrian walkway elements)	0.000	1.624	0.165

Note. The F table value is 2.35: *If T count> T the data table has a close relationship; **Sig <0.05 has a significant relationship meaning

diversity has a positive influence on the intensity of pedestrians at the research location and this is reflected in the more balanced use of land in some areas which led to a higher frequency of people taking a walk. Moreover, design (X3) was found not to have a significant relationship with the pedestrian intensity (Y) as observed in the smaller calculated T-value of 1.624 in comparison with the 2.353 recorded on the T-table with the significance calculated also indicated to be greater than the standard of tolerance (> 0.05).

The linear regression model in Equation 6 showed the diversity variable had a positive effect on the frequency of pedestrians as indicated by the higher level of balancing index in an area which led to an increase in the number of pedestrians. However, when the probability of diversity or balancing index increased by 2.172, the frequency of pedestrians was observed to have risen by 2.630million in a year as indicated in Equation 6.

$$Y = 2.630 + 2.172 \times X \quad [6]$$

These results are similar to the findings of the research conducted in the City of Montreal, Canada where land diversity was reported to be influencing the tendency of a person to walk around the transit center (Miranda-Moreno et al., 2011). The land designed for trade and services with school buildings were found to have the highest pedestrian movements. Moreover, a study in Bangkok also showed commercial land tends to encourage people to walk around transit centers (Townsend & Zacharias, 2010; Chalermpong, 2007). The environmental conditions along the transit-based transportation routes in Bangkok City have been observed to have the closest characteristics to the area studied in this present research. This involved the relative spreading out of the building density and land-use from the city center to the suburbs as well as the non-optimal facilities provided for the pedestrian walkways. Meanwhile, the differences observed are focused on the research area coverage, topographical conditions, and urban development policies. For example, the density of Montreal and Beijing tends to influence people to walk frequently due to their flat topography, vertical development policies, and the relatively close distance between buildings and transit centers. The pedestrian design elements in Montreal are also relatively complete when compared to Semarang and Bangkok City of Thailand. The model produced from the analysis is, however, presented in the following Figure 7.

CONCLUSION

The analysis showed the diversity component is an important feature of the built environment which significantly influences how frequent people walk along the Bus Trans Semarang. This was indicated by a high balancing index between residential and non-residential land-uses with the commercial lands dominated by activities in the form of trade, education, and service areas. Furthermore, the high mixed use of land areas tends

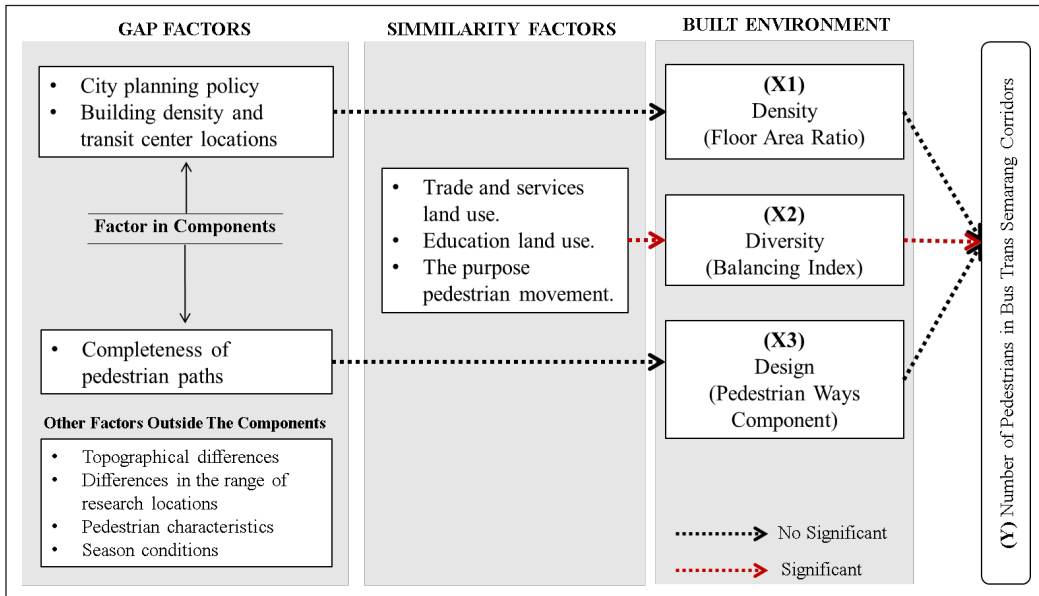


Figure 7. Model for the influence of built environment on pedestrian activity

to affect people’s frequency to walk due to the diversity in their patterns of activity. The model also showed the existence of a positive relationship between the diversity variable and the frequency of pedestrians and this further indicates the ability of a higher level of balancing index to increase the frequency of pedestrians in the study area.

The two other components including density and design were, however, observed not to have any effect on how frequent people walk. Moreover, the distance between the buildings tends to be far apart due to the very minimal optimization of vertical development. The lack of pedestrian facilities along the Bus Trans Semarang corridor, especially in the city center area, is discovered to be part of the reasons the components did not have influence on the frequency to walk in the city of Semarang. There is also the fear to avoid the effect of this issue on people’s desire to use transit-based transportation such as the Bus Trans Semarang. Therefore, a follow-up measure is expected to be made available to adjust pedestrians’ demands to the built environment in order to create a sustainable transportation system.

ACKNOWLEDGEMENTS

The authors are grateful to the Directorate of Research and Community Service, Deputy of Strengthening Research and Development of the Ministry of Technology/National Research, and the Innovation Agency of the Republic of Indonesia, (2020 fiscal year) for funding this research.

REFERENCES

- Blanco, H., & Alberti, M. (2009). Hot, congested, crowded and diverse: emerging research agendas in planning. *Progress in Planning*, 71(4), 153-205. <https://doi.org/10.1016/j.progress.2009.03.001>
- Cervero, R., & Kockelman, K. (1997). Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research part D: Transport and Environment*, 2(3), 199-219. [https://doi.org/10.1016/S1361-9209\(97\)00009-6](https://doi.org/10.1016/S1361-9209(97)00009-6)
- Chalermpong, S. (2007). Rail transit and residential land use in developing countries: Hedonic study of residential property prices in Bangkok, Thailand. *Transportation Research Record*, 2038(1), 111-119. <https://doi.org/10.3141/2038-15>
- Dixon, L. B. (1996). Bicycle and pedestrian level-of-service performance measures and standards for congestion management systems. *Transportation Research Record*, 1538(1), 1-9. <https://doi.org/10.1177/0361198196153800101>
- Duduta, N. (2013). Direct ridership models of bus rapid transit and metro systems in Mexico City. *Transportation Research Board*, 2394(1), 93-99. <https://doi.org/10.3141/2394-12>
- Fee, K., & Hartley, D. (2011). *Urban growth and decline: The role of population density at the city core*. Federal Reserve Bank of Cleveland.
- Hu, J. H., Zhan, C. Z., Cheng, Z. F., & Wang, B. (2013). A research of pedestrian evacuation simulation for BRT station based on fine grid method. *Procedia Engineering*, 52, 137-144. <https://doi.org/10.1016/j.proeng.2013.02.118>
- ITDP. (2017). *The BRT standard 2016*. Institute for Transportation & Development Policy Indonesia.
- Kim, C., Parent, O., & Hofe, R. V. (2018). The role of peer effects and the built environment on individual travel behavior. *Environment and Planning B: Urban Analytics and City Science*, 45(3), 452-469. <https://doi.org/10.1177/2399808317740354>
- Makarova, I., Pashkevich, A., Shubenkova, K., & Mukhametdinov, E. (2017). Ways to increase population mobility through the transition to sustainable transport. *Procedia Engineering*, 187, 756-762. <https://doi.org/10.1016/j.proeng.2017.04.434>
- Miranda-Moreno, L. F., Morency, P., & El-Geneidy, A. M. (2011). The link between built environment, pedestrian activity and pedestrian-vehicle collision occurrence at signalized intersections. *Accident Analysis and Prevention*, 43(5), 1624-1634. <https://doi.org/10.1016/j.aap.2011.02.005>
- Monteiro, F. B., & Campos, V. B. G. (2012). A proposal of indicators for evaluation of the urban space for pedestrians and cyclists in access to mass transit station. *Procedia - Social and Behavioral Sciences*, 54, 637-645. <https://doi.org/10.1016/j.sbspro.2012.09.781>
- Munshi, T. (2016). Built environment and mode choice relationship for commute travel in the city of Rajkot, India. *Transportation Research Part D*, 44, 239-253. <https://doi.org/10.1016/j.trd.2015.12.005>
- Oktaviani, I., Dewi, D. I. K., & Rakhmatulloh, A. R. (2020). Can building density influence the amount of BRT trans Semarang ridership? In *IOP Conference Series: Earth and Environmental Science* (Vol. 409, No. 1, p. 012038). IOP Publishing. <https://doi.org/10.1088/1755-1315/409/1/012038>

- Özbil, A., Yeşiltepe, D., & Argin, G. (2015). Modeling walkability: The effects of street design, street-network configuration and land-use on pedestrian movement. *Journal of the Faculty of Architecture*, 12(3), 189-207.
- Patankar, V. M., Kumar, R., & Tiwar, G. (2007). Impacts of bus rapid transit lanes on traffic and commuter mobility. *Journal of Urban Planning and Development*, 133(2), 99-106. [https://doi.org/10.1061/\(ASCE\)0733-9488\(2007\)133:2\(99\)](https://doi.org/10.1061/(ASCE)0733-9488(2007)133:2(99))
- Purwanto, E., & Manullang, O. R. (2018). Evaluasi trotoar sebagai feeder non motorized untuk mendukung bus rapid transit (BRT) di Kota Semarang [Evaluation of sidewalk as a non motorized feeder to support bus rapid transit (BRT) in Semarang City]. *Jurnal Pembangunan Wilayah dan Kota*, 14(1), 17-27. <https://doi.org/10.14710/pwk.v14i1.17336>
- SuaraMerdeka. (2015, October 09). Trans Semarang jadi percontohan [Trans Semarang Becomes Pilot]. *Suara Merdeka*. Retrieved October 20, 2015, from <https://www.suaramerdeka.com/>
- Sung, H., Go, D., & Choi, C. G. (2013). Evidence of Jacobs's street life in the great Seoul city: Identifying the association of physical environment with walking activity on streets. *Cities*, 35, 164-173. <https://doi.org/10.1016/j.cities.2013.07.010>
- Timmermans, H. (Ed.). (2009). *Pedestrian behaviour: Models, data collection and application*. Emerald Group Publishing Limited.
- Townsend, C., & Zacharias, J. (2010). Built environment and pedestrian behavior at rail rapid transit stations in Bangkok. *Transportation*, 37(2), 317-330. <https://doi.org/10.1007/s11116-009-9226-8>
- Untermann, R. K., & Lynn, L. (1984). *Accommodating the pedestrian: Adapting towns and neighborhoods for walking and bicycling*. Van Nostrand Reinhold.
- Yang, C., & Yao, M. (2019). Ultra-high intensity redevelopment of the core area of Japanese rail transit hub station. *International Journal of Sustainability and Planning*, 14(3), 245-259. <https://doi.org/10.2495/SDP-V14-N3-245-259>
- Zhang, Y., Ding, C., & Huang, L. (2016). The built environment and the frequency of cycling trips by Urban elderly: Insights from Zhongshan, China. *Journal of Asian Architecture and Building Engineering*, 15(3), 511-518. <http://doi.org/10.3130/jaabe.15.511>
- Zhao, C., Nielsen, T. A. S., Olafsson, A. S., Carstensen, T. A., & Meng, X. (2018). Urban form, demographic and socio-economic correlates of walking, cycling, and e-biking: Evidence from eight neighborhoods in Beijing. *Transport Policy*, 64, 102-112. <https://doi.org/10.1016/j.tranpol.2018.01.018>



Ethylene Yield from a Large Scale Naphtha Pyrolysis Cracking Utilizing Response Surface Methodology

Mohamad Hafizi Zakria^{1,2*}, Mohd Ghazali Mohd Nawawi¹ and Mohd Rizal Abdul Rahman^{2,3}

¹*School of Chemical and Energy Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor, Malaysia*

²*Manufacturing and Engineering Division, Pengerang Refining Company Sdn Bhd, PICMO B2, Pengerang Integrated Complex, 81600 Johor, Malaysia*

³*Group Technical Solutions, Petroliaam Nasional Berhad (PETRONAS), The Intermark Tower, 55000 Kuala Lumpur, Malaysia*

ABSTRACT

Statistical software is a robust application that has proven reliable worldwide. However, it is not normally used in the actual large scale olefin plant as it relies on the simulation software by Olefin Licensor should any issue rises. The study was conducted in a newly commissioned large scale olefin plant to see the impact of various operating variables on the ethylene yield from Short Residence Time (SRT) VII Furnace. The analysis was conducted utilizing statistical analysis, Response Surface Methodology (RSM) in Minitab Software Version 18 to develop a reliable statistical model with a 95% confidence level. The historical data was taken from the Process Information Management System

(PIMS) Software, PI Process Book Version 2015, and underwent both residuals and outliers removal prior to RSM analysis. 10 variables were shortlisted from the initial 15 identified variables in the studied SRT VII via Regression analysis due to RSM limitation to conduct the larger analysis in Minitab Software Version 18. The Response Optimizer tool showed that the ethylene yield from naphtha pyrolysis cracking in the studied plant could be maximized at 34.1% with control setting at 600.39 kg/hr of Integral Burner Flow, 6.81% of Arch O₂, 113.42 Barg of Steam Drum Pressure,

ARTICLE INFO

Article history:

Received: 16 November 2020

Accepted: 22 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.06>

E-mail addresses:

fizietranung@gmail.com; mohamadhafizi.zakri@prefchem.com

(Mohamad Hafizi Zakria)

ghazalinawawi@utm.my (Mohd Ghazali Mohd Nawawi)

mohdrizal@petronas.com.my (Mohd Rizal Abdul Rahman)

* Corresponding author

496.96°C of Super High Pressure (SHP) Temperature, 109.11 t/hr of SHP Boiler Feed Water (BFW) Flow, 92.78 t/hr of SHP Flow, 63.50 t/hr of Naphtha Feed Flow, and -13.38 mmHg of Draft Pressure.

Keywords: Ethylene yield, naphtha cracking, olefin plant, pyrolysis cracking, response surface methodology, statistical analysis

INTRODUCTION

Thermal and catalytic cracking are the mature technologies in producing polymer grade olefins from gaseous and liquid feedstock such as ethane, propane, naphtha, and gasoil (Fakhroleslam & Sadrameli, 2019, 2020; Sadrameli, 2015; Sadrameli, 2016). These feedstocks are available from various sources worldwide and may be applied to various technologies in producing olefins (Fakhroleslam & Sadrameli, 2019).

Apart from well-known thermal and catalytic cracking technologies, there are also other technologies being studied for olefins production such as propane dehydrogenation (Akporiaye et al., 2001; Darvishi et al., 2016), Fluidized Catalytic Cracking (FCC) (Sadrameli, 2016), olefin metathesis (Astruc, 2005), and reaction optimization in FCC process (Akah & Al-Ghrami, 2015). However, they are not widely applied to industrial applications to produce olefins due to the higher production cost (Fakhroleslam & Sadrameli, 2019).

Thermal cracking or also known as pyrolysis cracking in the olefin industry causes hydrocarbon bonds to break and resulting in the formation of the desired small and unsaturated molecule (Van de Vijver et al., 2015; Vangaever et al., 2020) such as olefins. Ethylene is an example of an olefin product that is produced from high-temperature pyrolysis cracking reaction (Shi et al., 2015; Song et al., 2019; Van de Vijver et al., 2015; Vangaever et al., 2020).

The pyrolysis cracking process in the steam cracker furnace is divided into three main sections normally referred to as radiation, convection, and flue gas stack (Fakhroleslam & Sadrameli, 2019; Fakhroleslam & Sadrameli, 2020; Karimzadeh et al., 2009; Sadrameli, 2015; Sadrameli, 2016). Over the years, numerous improvements have been established to improve the process efficiencies to these three sections in the steam cracker furnace to achieve the higher ethylene yield (Belinskaya et al., 2019; Han et al., 2016; Peller et al., 2018; Podrojková et al., 2018).

This study was conducted in the actual large scale plant with naphtha liquid as a feedstock to the steam cracker furnace utilizing pyrolysis cracking technology. The studied steam cracker furnace was Short Residence Time (SRT) VII from ABB Lummus, designed to produce 1,100 KTA capacity of polymer grade ethylene product. SRT VII is among the most promising technology available in the market proven reliable and efficient in producing the olefins (Caballero et al., 2015; Kuritsyn et al., 2008).

The study in the actual large scale plant is challenging as it is often affected by various process fluctuation (Feli et al., 2017; Zakria et al., 2016) due to frequent process variation in the feedstock composition, upstream plant performance, downstream readiness, and utility availability. Figure 1 shows the configuration of the SRT VII and its auxiliaries in the studied plant.

The plant was designed to process the continuous naphtha feed from the upstream plant. The feed enters the furnace and mixes with Dilution Steam (DS) in the convection section. The mixing is initiated at the start of the process to improve olefin selectivity by reducing the partial pressure of naphtha feed (Karimi et al., 2017; Masoumi et al., 2006) and favors the ethylene yield from the reversible reaction following Le' Chatelier's Principles (Cai et al., 2017; Epstein, 1978).

The mixed feed from the convection section flows into the radiation section with the operating Tube Metal Temperature (TMT) of 1,050 °C - 1,180 °C. This extremely high temperature is supplied to the SRT VII to maximize the ethylene yield generation. The cracked gas from the radiation section then will flow into the Transfer Line Exchanger (TLE) for rapid cooling of effluent and generation of Super High Pressure (SHP) Steam. The SHP steam generated from TLE will be used at the downstream process to drive the Charge Gas Compressor (CGC).

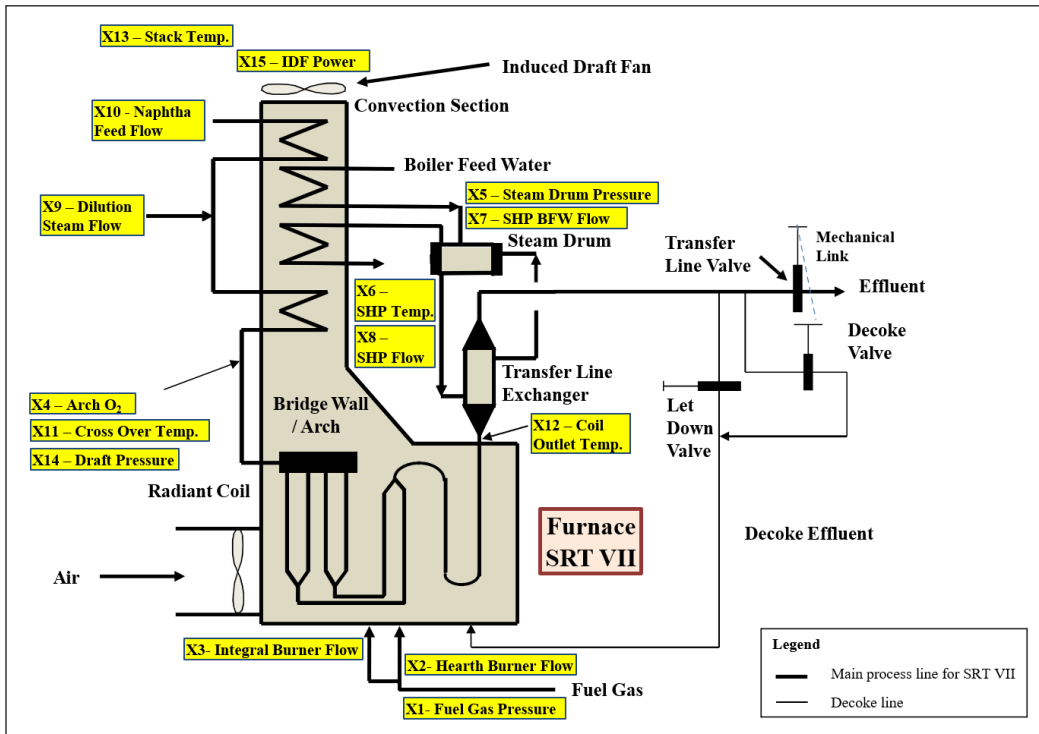


Figure 1. General arrangement for SRT VII at the studied plant inclusive of tags and descriptions for the selected variables

Various studies have been conducted from simulation, lab experiments and pilot scale plants that looking at various process parameters to optimize the olefins yield such as ethylene from the furnace. The specific control variables such as Coil Outlet Temperature (COT) (Song & Tang, 2018; Yu et al., 2018), Dilution Steam ratio (Berreni & Wang, 2011; Geng et al., 2016), naphtha feedstock (Geng et al., 2016; Nian et al., 2015) and fuel gas consumption (Yu et al., 2018) were normally studied to optimize the olefins yield. However, there is also an opportunity to explore other variables in the steam cracker furnace which are not widely discussed such as Arch O₂, Steam Drum, and Draft Pressure to see their relations towards the olefins yield.

Response Surface Methodology (RSM) is the mathematical and statistical techniques used to model and analyze problems where the responses are influenced by a few factors for optimization (Montgomery, 2017). The method has proven successful in a number of recent studies conducted for various subjects (Amirov & Vakhshouri, 2020; Reddy et al., 2020; Sun et al., 2016a; Sun et al., 2016b).

However, RSM application in large scale olefin plants is not common. The industrial olefin plant usually relies on the process simulation provided by the specific Olefin Licensor in the market should any issue rises such as ABB Lummus, KTI-Technip, Linde AG (Pyrocrack), Stone and Webster, and M.W. Kellogg (Sadrameli, 2015). Examples of currently available software are SPYRO (Dente et al., 1979; Van Goethem et al., 2001), SHAHAB (Toufighi et al., 2004), CRACKER (Joo et al., 2000; Joo & Park, 2001), CHEMKIN (Kee et al., 2006; Reyniers et al., 2017; Van Cauwenberge et al., 2017) and CRACKSIM (Hillewaert et al., 1988; Van Geem et al., 2008; Willems & Froment, 1988).

Normally, the simulation software comes with its own limitation such as robust simulation only applicable for specific furnace section such as radiation or convection section only. Some Olefin Licensor also provides the simulation software with a limited number of licenses for the olefin plant with a high price, and normally does not come with full access to safeguard the proprietary olefin plant design. This often makes the process monitoring and its troubleshooting for the olefins plant worldwide difficult. The optimization of ethylene yield via a systematic approach is challenging if the worldwide large scale olefin plant fully relies on the limited access and complex software by Olefin Licensor alone.

This study is significant in closing the gap in providing a practical solution (economical, process width, and easy application) for process monitoring and troubleshooting at the actual large scale olefin plant. It provides a guideline to the Panel Operators and Technologists in the studied plant to develop the maximized ethylene yield model based on a set of actual historical data from the large scale olefin plant using RSM in Minitab Software. Besides, this study also closes the gap in showing the impact of the other parameters that are not widely discussed at the large scale steam cracker furnaces such as Arch O₂, Steam Drum Pressure, and Draft Pressure.

METHODS AND EQUIPMENT

Equipment/Tools

The study utilized the newly commissioned SRT VII furnace with its auxiliaries comprised of main components such as TLEs, Burners, Steam Drum, and Induced Draft Fan (IDF). The steam cracker furnace was designed with 93 t/hr of processing capacity for naphtha liquid by Lummus Technology Heat Transfer (LTHT), United States, and constructed by Toyo Engineering, Japan.

The data used for the analysis was extracted based on actual plant data during the Start of Run (SOR) condition from the Process Information Management System (PIMS) Software, PI Process Book Version 2015 while the analysis was conducted using RSM in Minitab Software Version 18.

Methods

Table 1 shows 15 variables associated with the studied SRT VII furnace with the dedicated tag namely X1 –X15. The specific location of each tag was shown in Figure 1. The tags were chosen without segregation of “controlled” and “output” variables to see the relation of all variables towards Ethylene Yield represented by Y1.

Although these 15 variables covering almost all parameters in the studied steam cracker furnace, the environmental impacts such as SO_x and NO_x emissions at the furnace stack were excluded as the environmental aspect was not prioritized for this study.

The date for analysis was selected on January 24, 2020, 1700 hrs to February 2, 2020, 1200 hrs (211 hrs). The data was extracted as an hourly (average, time weighted) from the PI Process Book with a total of 3,376 data (represented by 1 output and 15 input variables). The data used for analysis focused on the Ethylene Yield at the normal operating range of the studied plant which was 28.70% to 30.67%. This small range was selected to safeguard the plant safety and equipment integrity during the study as advised by the olefin plant licensor.

The Paraffins, Olefins, Naphthenes, and Aromatics (PONA) composition in Naphtha feedstock were 60.92 % vol, 1.02 % vol, 25.97 % vol, and 12.09 % vol respectively. The Initial Boiling Point (IBP) and Final Boiling Point (FBP) for the naphtha feedstock were 34.1 °C and 166.1 °C respectively.

Box Plot, Run Chart, and Individual-Moving Range (I-MR) Chart were used for the data stability verification while the Normality Test and Graphical Summary were used to check on the data normality for these 3,376 data. The P-Value in both stability and normality verification for these tools was ensured at above 0.05 to proceed as normal data in the Regression and RSM analysis. The analysis may utilize the Box-Cox data transformation methodology if the P-Value generated from stability and normality verification was lower than 0.05.

Regression analysis was initiated to all 15 identified variables before the RSM analysis to shortlist only 10 most significant variables. This step was taken due to the limitation in Minitab Software Version 18 in handling the bigger analysis for RSM. The Regression was conducted a few times, via one by one variable elimination in each Regression to remove insignificant variables with $VIF > 10$ and $P\text{-Value} > 0.05$.

The variable elimination followed the sequence which started from the highest VIF value until all variables achieved $VIF < 10$. After all of the remaining variables were determined at $VIF < 10$, the elimination continued for variables starting from the highest P-Value until only 10 variables left in the intermediate Regression for further analysis using RSM.

RSM was conducted to the remaining 10 variables via the Historical Design of Experiment (DOE) method. The intervals were set at 95% in the Minitab (Analyze Response Surface Design interface setting) to ensure the analysis was conducted at the 95% of confidence level. The elimination of the insignificant variables was also conducted in RSM one by one, started from the Two-ways Interaction, Square, and finally to the Linear relation. This sequence was essential to maintain the hierarchical model in RSM.

The elimination in RSM also started from the highest VIF and P-Value until all $VIF < 10$ and $P\text{-Value} < 0.05$ achieved. VIF measures the severity of multicollinearity that exists in the ordinary least squares Regression analysis. $VIF > 10$ is not recommended (Hair et al., 1995) due to its multicollinearity relation which may affect the P-value and finally contribute to the unreliable model. However, the variable with a higher VIF and P-Value in Linear relation may be considered into the final model should the variable exist in the model as Square or Two-Way Interaction to maintain the hierarchical RSM model. This must also be supported by a good R-Square at 75% or higher.

The validation for the final RSM equation model was conducted using the Probability Plot for the residuals. The residuals in the Probability Plot represented the difference between the predicted data from the final equation model versus the actual plant data in the studied plant. This plot was important in measuring the degree of regression line vertically might miss a data point. The P-Value of higher than 0.05 in the Probability Plot was a good validation criterion for the reliable final equation model for the Historical DOE method in RSM.

Surface Plot was also utilized for four variables with the highest impact which was determined from the highest coefficient factor in the final equation model. Surface Plot is important to surface the relation and contribution of two continuous variables towards the fitted response value of ethylene yield in form of a three-dimensional view in Minitab Software.

Finally, the Response Optimizer tool was applied to the final model to predict the maximum value of ethylene yield with the process settings for the identified significant

variables in RSM analysis. The setting in Response Optimizer was useful as a guideline for the studied olefin plant to maximize the predicted response of the ethylene yield.

RESULTS AND DISCUSSION

Table 1 shows the 1st Regression (initial) and 6th Regression (intermediate) that were conducted to select only 10 variables for RSM analysis.

The 1st Regression was started to establish the overall relationship between each variable. The 2nd Regression was continued with the elimination of variables with VIF >10. From the analysis, the VIF for all variables successfully reduced to <10 after the 2nd Regression. Besides, some variables with VIF >10 in 1st Regression such as X10 was successfully reduced to 9.68 in the 2nd Regression and therefore these variables remained in the intermediate regression. This condition occurred due to the multicollinearity factors for X10 was reduced as the first variable, X13 removed from the analysis.

The 3rd – 6th Regressions were conducted with the removal of variables with P-Value >0.05. The 6th Regression was determined as the intermediate Regression as 10 variables were successfully shortlisted. The sequence of elimination for 2nd – 6th Regression conducted were X13 (VIF: 17.54), X9 (P-Value: 0.759), X1 (P-Value: 0.720), X15 (P-Value: 0.542), and X11 (P-Value: 0.564). This one by one variable elimination was taken to ensure only

Table 1
Identified 10 variables for the RSM analysis

Tag with Description	Initial (1 st Regression)			Intermediate (6 th Regression)			New RSM Tag
	Coef.	P-Value	VIF	Coef.	P-Value	VIF	
Constant	-1.3	0.976		4.4	0.883		
X1 Fuel Gas Pressure	0.049	0.850	3.02				
X2 Hearth Burner Flow	-0.000852	0.000	8.50	-0.000917	0.000	7.20	A
X3 Integral Burner Flow	0.02804	0.000	2.59	0.02937	0.000	2.41	B
X4 Arch O ₂	-0.281	0.094	3.72	-0.234	0.111	2.89	C
X5 Steam Drum Pressure	-0.420	0.004	4.41	-0.443	0.002	4.14	D
X6 SHP Temperature	-0.0254	0.237	3.49	-0.0211	0.290	3.05	E
X7 SHP BFW Flow	0.1537	0.000	5.17	0.1617	0.000	4.06	F
X8 SHP Flow	-0.0696	0.013	3.05	-0.0768	0.003	2.66	G
X9 Dilution Steam Flow	-0.000031	0.801	1.10				
X10 Naphtha Feed Flow	0.1400	0.000	10.96	0.1112	0.000	6.83	H
X11 Crossover Temperature	-0.0007	0.991	9.93				
X12 Coil Outlet Temp.	0.0766	0.156	3.21	0.0797	0.049	1.83	J
X13 Stack Temperature	0.0888	0.362	17.54				
X14 Draft Pressure	0.126	0.305	1.40	0.131	0.243	1.20	K
X15 ID Fan Power	-0.0401	0.391	2.14				

insignificant variables were removed as the VIF and P-Value will significantly change after each variable eliminated from the analysis.

RSM analysis continued with the new tags for the identified 10 variables in the intermediate Regression. The new tags were named A - K as shown in Table 1. These changes were necessary to avoid confusion in the RSM analysis as 10 variables in RSM analysis resulted in a total of 45 Two-Way Interactions, 10 Squares, and 10 Linear relations.

The R-Square from the final model was good at 87.22 % regardless it was conducted in the actual large scale plant where process fluctuation often occurred. This value translated to 87.22% of the variability in data was accounted for in the model. It is adequate the meet the data variability which is advised at more than 75% (Haaland, 1989; Omar et al., 2009). Table 2 shows the details of the Analysis of Variance (ANOVA) for the RSM analysis.

The remaining 15 degree of freedoms (DFs) in Table 2 indicated that a total of 41 Two-Way Interactions, 7 Squares, and 2 Linear relations were removed during one by one variable elimination. Some variables with P-Value >0.05 were retained in the final model which were E (SHP Temperature), and K (Induced Draft Fan Power). These variables were accepted to maintain the hierarchical RSM model as their relation also existed in the Square

Table 2
ANOVA for the final RSM model

Source	DF	Adj SS	Adj MS	F-Value	P-Value	VIF
Model	15	26.2656	1.75104	70.54	0.000	
Linear	8	7.7088	0.96360	38.82	0.000	
B	1	0.4423	0.44228	17.82	0.000	2.32
C	1	0.3016	0.30155	12.15	0.001	7.26
D	1	2.2811	2.28106	91.89	0.000	4.68
E	1	0.0415	0.04148	1.67	0.198	2.32
F	1	0.1801	0.18006	7.25	0.008	5.12
G	1	0.2121	0.21208	8.54	0.004	3.45
H	1	0.4239	0.42394	17.08	0.000	8.04
K	1	0.0845	0.08452	3.40	0.067	1.30
Square	3	1.5394	0.51313	20.67	0.000	
C*C	1	0.7332	0.73324	29.54	0.000	1.40
D*D	1	0.4887	0.48868	19.69	0.000	2.01
E*E	1	0.1521	0.15212	6.13	0.014	1.33
2-Way	4	1.6858	0.42144	16.98	0.000	
C*E	1	0.4283	0.42826	17.25	0.000	4.08
C*G	1	1.5186	1.51865	61.18	0.000	2.60
C*K	1	0.1418	0.14176	5.71	0.018	2.48
G*H	1	0.6591	0.65908	26.55	0.000	1.44
Error	155	3.8476	0.02482			
Total	170	30.1132				

and Two-Way Interaction. Besides, they were also supported by a high R-Square of 87.22%. The final equation model established from the RSM analysis is shown in Equation 1.

$$\begin{aligned}
 Y1 = & 26519 + 0.02231 (B) + 188.4 (C) - 280.8 (D) - 41.2 (E) + 0.0714 (F) - \\
 & - 10.54 (G) - 6.69 (H) + 6.62 (K) + 1.924 (C^2) + 1.226 (D^2) \\
 & + 0.0452 (E^2) - 0.618 (CE) + 0.919 (CG) - 1.013 (CK) \\
 & + 0.0750 (GH)
 \end{aligned} \tag{1}$$

The Equation 1 shows that 8 out of 10 variables were considered in the final model. Although E (SHP Temperature), and K (Induced Draft Fan Power) were accepted with P-Value >0.05, they were accepted to maintain the hierarchical order in the final RSM model. Hierarchical order was prioritized over P-Value in the RSM analysis to obtain a reliable final equation model (Gaitonde et al., 2017; Noordin et al., 2004; Tanaka et al., 2007). The hierarchical order prioritization in this analysis was supported by the high R-square at 87.22% compared to the advised R-square of >75% (Haaland, 1989; Omar et al., 2009). Besides, the confidence level was also already set at 95% for all intervals throughout the RSM analysis in the Analyze Response Surface Design interface in Minitab.

The coefficient factors in the final equation model will be a good basis in understanding the operating behavior and the criticality for the significant variables towards the higher Y1 (Ethylene Yield). Figure 2 shows the Probability Plot for residuals conducted for data validation to the final equation model.

The P-Value from the Probability Plot for the residuals was observed at higher than 0.05 which was 0.571. This high P-Value represented a good data distribution and model prediction in explaining the validity of the final equation model. The residuals in this plot compared the predicted data from the equation model and the actual data in the studied plant used for the RSM analysis. The P-Value of higher than 0.05 validated the acceptance for the final equation models using the Historical DOE method in the RSM analysis.

Figure 3 shows the actual data versus the predicted data distribution using the final model generated from the RSM analysis. The data were uniformly distributed at the normal operating range in producing Y1 (Ethylene Yield) at 28.8% – 29.5%. However, as this study was conducted in the actual large scale plant, the process condition was observed fluctuating for the Y1 (Ethylene Yield) at 30.1 – 30.5%. However, this data was accepted into the establishment of the RSM model due the data was observed normal and stable from the normality and stability verification conducted earlier. These data had also proven reliable from the high R-square value of 87.2% in the final model regardless the study was conducted with the fluctuating process condition at the studied olefin plant.

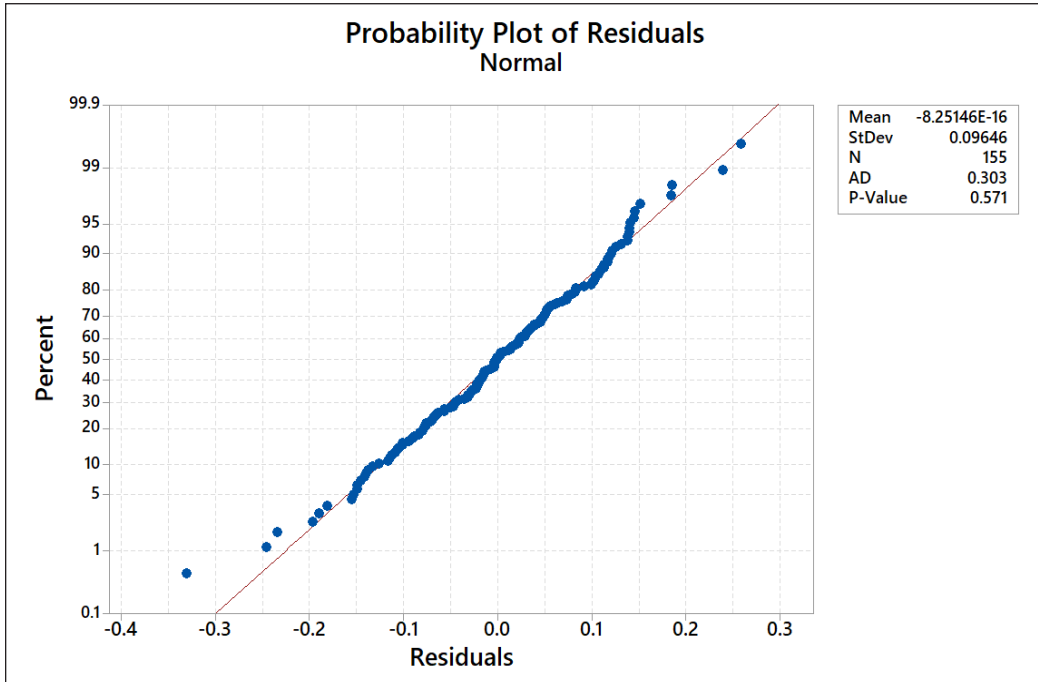


Figure 2. Validation using Probability Plot for the final equation model residuals

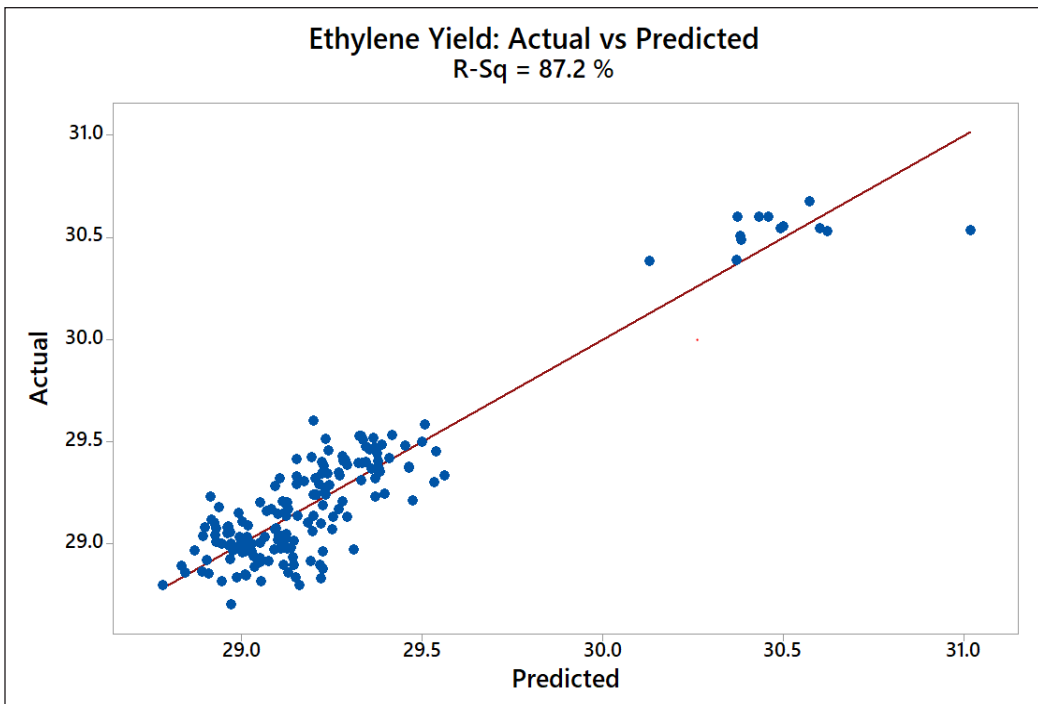


Figure 3. Actual vs predicted value for Y1 (Ethylene yield)

Figure 4 shows the Surface Plot of variables with 4 highest coefficient factors in the equation model in achieving Y1 (Ethylene Yield) which were 188.4 of C (Arch O₂), -280.8 of D (Steam Drum Pressure), -41.2 of E (SHP Temperature), and -10.54 of G (SHP Flow).

The value for non-tested variables was held at mean value, \bar{x} ; 6.38% of C (Arch O₂), 114.05 Barg of D (Steam Drum Pressure), 499.29 °C of E (SHP Temperature), and 90.50

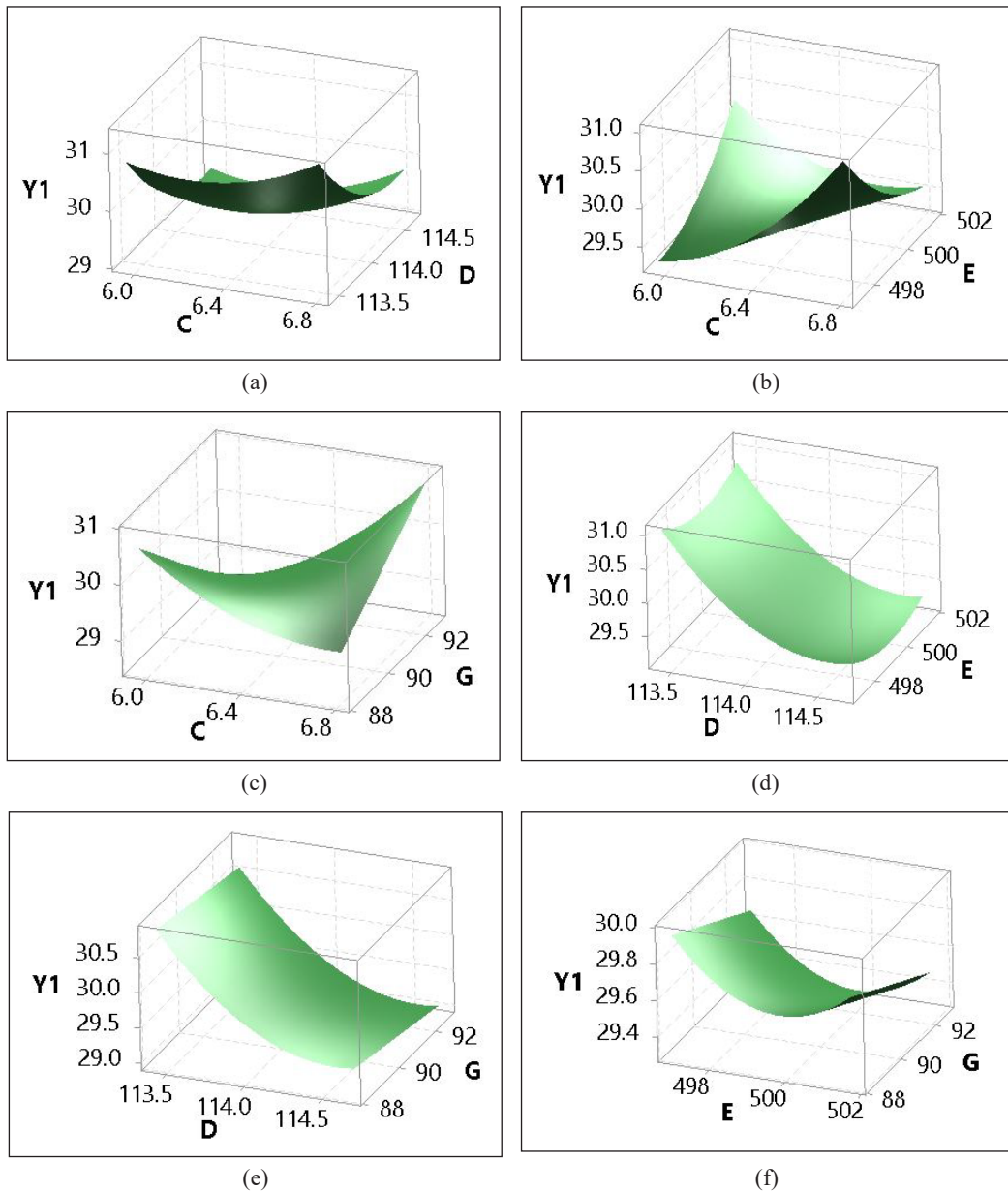


Figure 4. Surface Plot for most impactful variables towards Y1; (a) C vs D, (b) C vs E, (c) C vs G, (d) D vs E, (e) D vs G, and (f) E vs G

t/h of G (SHP Flow). The light on the Surface Plot was also set to indicate Y1 (Ethylene Yield) value at maximum.

Figure 4 (a) and 4 (b) indicate the higher value of C (Arch O₂) combined with the lower value of D (Steam Drum Pressure) or E (SHP Temperature) will result in higher Y1 (Ethylene Yield). However, operating the studied SRT VII with higher E (SHP Temperature) also will result in higher Y1 (Ethylene Yield) due to its significant quadratic relation that exists in the model compared to E (SHP Temperature). It was also good to note that operating the plant with the combination of lower C (Arch O₂) and lower E (SHP Temperature) should be avoided to prevent the poor Y1 (Ethylene Yield) as shown in the Surface Plot.

Figure 4 (c) illustrates the existence of both quadratic and cubic relations between C (Arch O₂) and G (SHP Flow). Operating the plant with lower C (Arch O₂) combined with the lower G (SHP Flow) will favor the higher Y1 (Ethylene Yield). However, the higher Y1 (Ethylene Yield) may also be obtained via operating the higher C (Arch O₂) and G (SHP Flow).

Figure 4 (d), and 4 (e) show the same relation between D (Steam Drum Pressure) vs E (SHP Temperature) and D (Steam Drum Pressure) vs G (SHP Flow). Maximum Y1 (Ethylene Yield) can be achieved by manipulating the lower D (Steam Drum Pressure) for both variables. In these plots, both E (SHP Temperature) and G (SHP Flow) did not show significant impact when combined with D (Steam Drum Pressure) in achieving the targeted Y1 (Ethylene Yield).

Figure 4 (f) demonstrates that manipulating lower E (SHP Temperature) with lower G (SHP Flow) may also result in the higher Y1 (Ethylene Yield). However, Y1 (Ethylene Yield) will also reduce after reaching a certain operating condition of E (SHP Temperature) due to its quadratic relation as shown in the surface plot.

Table 3 displays the result of Multiple Response Prediction utilizing the RSM analysis in Minitab Software while Figure 5 shows the Response Optimizer evaluation for the process condition taken from the identified 8 significant variables to achieve the maximized Y1 (Ethylene Yield).

In general, the Response Optimizer result showed that Y1 (Ethylene Yield) could be maximized at 34.1% with the controlled operating range as illustrated in Figure 5. The recommended process setting for the highest identified coefficient in Equation 1, D (Steam

Table 3
Multiple Response Prediction for final equation model

Response	Fit	Standard Error Fit	Confidence (95%)	
			Confidence Interval	Prediction Interval
Y1	34.096	0.420	(33.267, 34.925)	(33.210, 34.982)

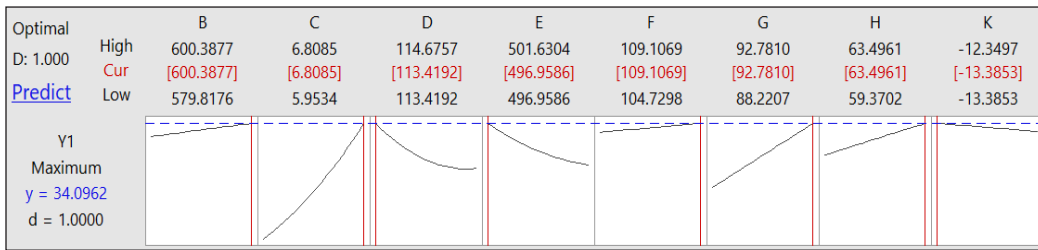


Figure 5. Process condition to achieve maximized Y1 from Response Optimizer

Drum Pressure) was 113.42 Barg. This value favored towards the lowest reading of D (Steam Drum Pressure) which was supported by a huge negative coefficient factor of -288 in the equation model. The setting for the 2nd biggest coefficient factor, C (Arch O₂) was identified at the highest side of 6.81%. This setting explained the big positive coefficient factor of 188.4 for C (Arch O₂).

Figure 5 is helpful as a guide to control the process range condition at studied SRT VII for a better Y1 (Ethylene Yield) with a 95% confidence level. The High and Low range setting in Figure 5 also can be referred to by Panel Operators and Operations Engineers in the studied plant to maximize the Y1 (Ethylene Yield).

Operating the furnace at a higher Coil Outlet Temperature (COT) was normally practiced by olefin plants worldwide to improve the ethylene yield. It was also recommended from various studies and reviews (Fakhroleslam & Sadrameli, 2019; Fakhroleslam & Sadrameli, 2020; Nabavi et al., 2009; Nabavi et al., 2011; Sadrameli, 2015; Song & Tang, 2018). However, it is interesting that COT was not identified as the significant variable in this study. Besides COT, another interesting finding was Hearth Burner Flow that was widely studied also did not show significant relation to ethylene yield in this study (Han et al., 2016; Yu et al., 2018; Zhang et al., 2017).

Previous studies conducted widely using simulation software with a focus on only controlled variables such as COT and fuel gas consumption at burners without full consideration on the inter-related variables in the furnace. For example, Steam Drum Pressure which shown the highest coefficient factor in the final equation model was not discussed widely in the literature compared to the COT and other controlled variables. This due to Steam Drum Pressure is not a primary control variable and therefore it was less attractive for the optimization improvement studies.

RSM was proven a robust statistical tool in finding the significant relations for each identified variable in the studied SRT VII. Although COT and Hearth Burner Flow were not listed as one of the significant variables, the final equation model showed that the highest coefficient was contributed by Arch O₂, Steam Drum Pressure, SHP Temperature, and SHP Flow. These 4 variables were highly dependent on the operating temperature inside the steam cracker furnace which is normally controlled by the COT and Hearth Burner Flow.

However, as this study was conducted with the removal of VIF >10 to reduce multicollinearity in the final equation model, the highly inter-related variables were removed and the more significant variables were taken for the establishment of the reliable final equation model. If the analysis was conducted without removal of the variable with VIF >10, the COT and Hearth Burner Flow may exist in the final model as significant variables. However, it may not yield a reliable final equation model due to the huge multicollinearity factor that existed.

The approach in the study for the establishment of a reliable ethylene yield model was proven successful from the high R-square obtained at 87.22% and high P-Value of 0.571 for Probability Plot of the residuals. These were observed regardless of the study was conducted in the actual large scale olefin plant that came with process fluctuation. This study also showed that less attractive variables in the steam cracker furnace should not be neglected in the optimization study in achieving a higher ethylene yield.

CONCLUSION

The RSM analysis was successfully conducted at the studied plant utilizing Minitab Software Version 18. The Ethylene Yield can be maximized at 34.10% with the optimized value of 600.39 kg/hr of Integral Burner Flow, 6.81% of Arch O₂, 113.42 Barg of Steam Drum Pressure, 496.96 °C of SHP Temperature, 109.11 t/hr of SHP BFW Flow, 92.78 t/hr of SHP Flow, 63.50 t/hr of Naphtha Feed Flow, and -13.38 mmHg of Draft Pressure.

ACKNOWLEDGMENTS

The authors would like to acknowledge Pengerang Refining Company Sdn Bhd for providing robust tools and reliable equipment for data collection, analysis, and model validation. Besides, a special thank is also dedicated to Universiti Teknologi Malaysia for the academic support and advice in completing this study.

REFERENCES

- Akah, A., & Al-Ghrami, M. (2015). Maximizing propylene production via FCC technology. *Applied Petrochemical Research*, 5(4), 377-392. <https://doi.org/10.1007/s13203-015-0104-3>
- Akporiaye, D., Jensen, S., Olsbye, U., Rohr, F., Rytter, E., Rønnekleiv, M., & Spjelkavik, A. I. (2001). A Novel, Highly Efficient Catalyst for Propane Dehydrogenation. *Industrial & Engineering Chemistry Research*, 40(22), 4741-4748. <https://doi.org/10.1021/ie010299+>
- Amirov, N., & Vakhshouri, A. R. (2020). Numerical modeling and optimization of product selectivity and catalyst activity in Fischer-Tropsch synthesis via response surface methodology: Cobalt carbide particle size and H₂/CO ratio effects. *International Journal of Hydrogen Energy*, 45(56), 31913-31925. <https://doi.org/10.1016/j.ijhydene.2020.08.219>

- Astruc, D. (2005). The Metathesis Reactions: From a Historical Perspective to Recent Developments. *New Journal of Chemistry*, 29(1), 42-56. <https://doi.org/10.1039/b412198h>
- Belinskaya, N., Altynov, A., Bogdanov, I., Popok, E., Kirgina, M., & Simakov, D. S. A. (2019). Production of Gasoline Using Stable Gas Condensate and Zeoforming Process Products as Blending Components. *Energy & Fuels*, 33(5), 4202-4210. <https://doi.org/10.1021/acs.energyfuels.9b00591>
- Berreni, M., & Wang, M. (2011). Modelling and dynamic optimization of thermal cracking of propane for ethylene manufacturing. *Computers & Chemical Engineering*, 35(12), 2876-2885. <https://doi.org/10.1016/j.compchemeng.2011.05.010>
- Caballero, D. Y., Biegler, L. T., & Guirardello, R. (2015). Simulation and optimization of the ethane cracking process to produce ethylene. *Computer Aided Chemical Engineering*, 37, 917-922. <https://doi.org/10.1016/B978-0-444-63578-5.50148-1>
- Cai, Y., Yang, S., Fu, S., Zhang, D., & Zhang, Q. (2017). Investigation of portevin–Le chatelier band strain and elastic shrinkage in al-based alloys associated with Mg contents. *Journal of Materials Science & Technology*, 33(6), 580-586. <https://doi.org/10.1016/j.jmst.2016.05.012>
- Darvishi, A., Davand, R., Khorasheh, F., & Fattahi, M. (2016). Modeling-based optimization of a fixed-bed industrial reactor for oxidative dehydrogenation of propane. *Chinese Journal of Chemical Engineering*, 24(5), 612-622. <https://doi.org/10.1016/j.cjche.2015.12.018>
- Dente, M., Ranzi, E., & Goossens, A. G. (1979). Detailed prediction of olefin yields from hydrocarbon pyrolysis through a fundamental simulation model (SPYRO). *Computers & Chemical Engineering*, 3(1), 61-75. [https://doi.org/10.1016/0098-1354\(79\)80013-7](https://doi.org/10.1016/0098-1354(79)80013-7)
- Epstein, L. G. (1978). The Le Chatelier Principle in optimal control problems. *Journal of Economic Theory*, 19(1), 103-122. [https://doi.org/10.1016/0022-0531\(78\)90058-3](https://doi.org/10.1016/0022-0531(78)90058-3)
- Fakhroeslam, M., & Sadrameli, S. M. (2019). Thermal/catalytic cracking of hydrocarbons for the production of olefins; a state-of-the-art review III: Process modeling and simulation. *Fuel*, 252, 553-566. <https://doi.org/10.1016/j.fuel.2019.04.127>
- Fakhroeslam, M., & Sadrameli, S. M. (2020). Thermal cracking of hydrocarbons for the production of light olefins: A review on optimal process design, operation, and control. *Industrial & Engineering Chemistry Research*, 59(27), 12288-12303. <https://doi.org/10.1021/acs.iecr.0c00923>
- Feli, Z., Darvishi, A., Bakhtyari, A., Rahimpour, M. R., & Raeissi, S. (2017). Investigation of propane addition to the feed stream of a commercial ethane thermal cracker as supplementary feedstock. *Journal of the Taiwan Institute of Chemical Engineers*, 81, 1-13. <https://doi.org/10.1016/j.jtice.2017.10.025>
- Gaitonde, V. N., Manjaiah, M., Maradi, S., Karnik, S. R., Petkar, P. M., & Paulo Davim, J. (2017). Multiresponse optimization in wire electric discharge machining (WEDM) of HCHCr steel by integrating response surface methodology (RSM) with differential evolution (DE). In J. P. Davim (Ed.), *Computational Methods and Production Engineering* (pp. 199-221). Woodhead Publishing.
- Geng, Z., Wang, Z., Zhu, Q., & Han, Y. (2016). Multi-objective operation optimization of ethylene cracking furnace based on AMOPSO algorithm. *Chemical Engineering Science*, 153, 21-33. <https://doi.org/10.1016/j.ces.2016.07.009>

- Haaland, P. D. (1989). *Experimental design in biotechnology*. Marcel Dekker.
- Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis with readings* (4th Ed.). Prentice-Hall.
- Han, Y., Geng, Z., Wang, Z., & Mu, P. (2016). Performance analysis and optimal temperature selection of ethylene cracking furnaces: A data envelopment analysis cross-model integrated analytic hierarchy process. *Journal of Analytical and Applied Pyrolysis*, 122, 35-44. <https://doi.org/10.1016/j.jaap.2016.10.025>
- Hillewaert, L. P., Dierickx, J. L., & Froment, G. F. (1988). Computer generation of reaction schemes and rate equations for thermal cracking. *AIChE Journal*, 34(1), 17-24. <https://doi.org/10.1002/aic.690340104>
- Joo, E., Lee, K., Lee, M., & Park, S. (2000). CRACKER — a PC based simulator for industrial cracking furnaces. *Computers & Chemical Engineering*, 24(2), 1523-1528. [https://doi.org/10.1016/S0098-1354\(00\)00558-5](https://doi.org/10.1016/S0098-1354(00)00558-5)
- Joo, E., & Park, S. (2001). Pyrolysis Reaction Mechanism for Industrial Naphtha Cracking Furnaces. *Industrial & Engineering Chemistry Research*, 40(11), 2409-2415. <https://doi.org/10.1021/ie000774o>
- Karimi, H., Cowperthwaite, E., Olayiwola, B., Farag, H., & McAuley, K. (2017). Modelling of heat transfer and pyrolysis reactions in an industrial ethylene cracking furnace. *Canadian Journal of Chemical Engineering*, 96(1), 33-48. <https://doi.org/10.1002/cjce.22844>
- Karimzadeh, R., Godini, H. R., & Ghashghaee, M. (2009). Flowsheeting of steam cracking furnaces. *Chemical Engineering Research and Design*, 87(1), 36-46. <https://doi.org/10.1016/j.cherd.2008.07.009>
- Kee, R. J., Rupley, F. M., Miller, J. A., Coltrin, M. E., Grcar, J. F., Meeks, E., Moffat, H. K., Lutz, A. E., Dixon-Lewis, G., Smooke, M. D., Warnatz, J., Evans, G. H., Larson, R. S., Mitchell, R. E., Petzold, L. R., Reynolds, W. C., Caracotsios, M., Stewart, W. E., Glarborg P., ... & Puduppakkam, K. V. (2006). *CHEMKIN Release 4.1*. SReaction Design.
- Kuritsyn, V., Arapov, D., Ekimova, A., & Yakupov, A. (2008). Modeling of pyrolysis of straight-run naphtha in a large-capacity type SRT-VI furnace. *Chemistry and Technology of Fuels and Oils*, 44(3), 180-189. <https://doi.org/10.1007/s10553-008-0038-x>
- Masoumi, M., Sadrameli, S. M., Towfighi, J., & Niaei, A. (2006). Simulation, optimization and control of a thermal cracking furnace. *Energy*, 31(4), 516-527. <https://doi.org/10.1016/j.energy.2005.04.005>
- Montgomery, D. C. (2017). *Design and analysis of experiments* (8th Ed.). John Wiley & Sons.
- Nabavi, S., Rangaiah, G., Niaei, A., & Salari, D. (2009). Multiobjective optimization of an industrial LPG thermal cracker using a first principles model. *Industrial & Engineering Chemistry Research*, 48(21), 9523-9533. <https://doi.org/10.1021/ie801409m>
- Nabavi, S., Rangaiah, G., Niaei, A., & Salari, D. (2011). Design optimization of an LPG thermal cracker for multiple objectives. *International Journal of Chemical Reactor Engineering*, 9(1), 1-34. <https://doi.org/10.1515/1542-6580.2507>
- Nian, X., Wang, Z., & Qian, F. (2015). Strategy of changing cracking furnace feedstock based on improved group search optimization. *Chinese Journal of Chemical Engineering*, 23(1), 181-191. <https://doi.org/10.1016/j.cjche.2014.09.027>

- Noordin, M. Y., Venkatesh, V. C., Sharif, S., Elting, S., & Abdullah, A. (2004). Application of response surface methodology in describing the performance of coated carbide tools when turning AISI 1045 steel. *Journal of Materials Processing Technology*, 145(1), 46-58. [https://doi.org/10.1016/S0924-0136\(03\)00861-6](https://doi.org/10.1016/S0924-0136(03)00861-6)
- Omar, W. N. N. W., Nordin, N., Mohamed, M., & Amin, N. A. S. (2009). A two-step biodiesel production from waste cooking oil: Optimization of pre-treatment step. *Journal of Applied Sciences*, 9(17), 3098-3103. <https://doi.org/10.3923/jas.2009.3098.3103>
- Peller, A., Jambor, B., Hájeková, E., Hudec, P., Hadvinová, M., & Horňáček, M. (2018). Hydrogenation of liquid hydrocarbon fractions for production of non-conventional diesel fuel. *Petroleum and Coal*, 60(6), 1112-1119.
- Podrojková, N., Orinák, A., Oriňáková, R., Procházková, L., Čuba, V., Patera, J., & Smith, R. (2018). Effect of different crystalline phase of ZnO/Cu nanocatalysts on cellulose pyrolysis conversion to specific chemical compounds. *Cellulose*, 25(10), 5623–5642. <https://doi.org/10.1007/s10570-018-1997-7>
- Reddy, P. V., Reddy, B., & Ramulu, P. (2020). Mathematical modelling for prediction of tube hydroforming process using RSM and ANN. *International Journal of Industrial and Systems Engineering*, 35(1), 13-27. <https://doi.org/10.1504/IJISE.2020.106848>
- Reyniers, P. A., Schietekat, C. M., Kong, B., Passalacqua, A., Van Geem, K. M., & Marin, G. B. (2017). CFD simulations of industrial steam cracking reactors: Turbulence–chemistry interaction and dynamic zoning. *Industrial & Engineering Chemistry Research*, 56(51), 14959-14971. <https://doi.org/10.1021/acs.iecr.7b02492>
- Sadrameli, S. M. (2015). Thermal/catalytic cracking of hydrocarbons for the production of olefins: A state-of-the-art review I: Thermal cracking review. *Fuel*, 140, 102-115. <https://doi.org/10.1016/j.fuel.2014.09.034>
- Sadrameli, S. M. (2016). Thermal/catalytic cracking of liquid hydrocarbons for the production of olefins: A state-of-the-art review II: Catalytic cracking review. *Fuel*, 173, 285-297. <https://doi.org/10.1016/j.fuel.2016.01.047>
- Shi, H., Su, C., Cao, J., Li, P., Liang, J., & Zhong, G. (2015). Nonlinear adaptive predictive functional control based on the Takagi–Sugeno model for average cracking outlet temperature of the ethylene cracking furnace. *Industrial & Engineering Chemistry Research*, 54(6), 1849-1860. <https://doi.org/10.1021/ie503531z>
- Song, G., & Tang, L. (2018). Optimization model for the transfer line exchanger system. *Computer Aided Chemical Engineering*, 44, 1015-1020. <https://doi.org/10.1016/B978-0-444-64241-7.50164-6>
- Song, H., Su, C. L., Shi, H., Li, P., & Cao, J. T. (2019). Improved predictive functional control for ethylene cracking furnace. *Measurement and Control*, 52(5-6), 526-539. <https://doi.org/10.1177/0020294019842602>
- Sun, Y., Yang, G., Li, K., Zhang, L., & Zhang, L. (2016a). CO₂ mineralization using basic oxygen furnace slag: process optimization by response surface methodology. *Environmental Earth Sciences*, 75(19), 1-10. <https://doi.org/10.1007/s12665-016-6147-7>
- Sun, Y., Zhang, J., & Zhang, L. (2016b). NH₄Cl selective leaching of basic oxygen furnace slag: Optimization study using response surface methodology. *Environmental Progress & Sustainable Energy*, 35(5), 1387-1394. <https://doi.org/10.1002/ep.12365>

- Tanaka, T., Sakai, K., Yamashita, Y., Sakamoto, N., & Koyamada, K. (2007). Hierarchical response surface methodology for parameter optimization: Efficiency of a hierarchical RSM with a hessian matrix. In K. Koyamada, S. Tamura, & O. Ono (Eds.), *Systems Modeling and Simulation* (pp. 213-217). Tokyo, Japan: Springer Japan.
- Toufighi, J., Karimzadeh, R., Saedi, G., Hosseini, S., Morafahi, M., Mokhtarani, B., Niaee, A., & Sadr, A. M. (2004). SHAHAB-A PC-based software for simulation of steam cracking furnaces (Ethane and Naphtha). *Iranian Journal of Chemical Engineering*, 1(2), 55-70.
- Van Cauwenberge, D. J., Vandewalle, L. A., Reyniers, P. A., Van Geem, K. M., Marin, G. B., & Floré, J. (2017). Periodic reactive flow simulation: Proof of concept for steam cracking coils. *AIChE Journal*, 63(5), 1715-1726. <https://doi.org/10.1002/aic.15530>
- Van de Vijver, R., Vandewiele, N. M., Bhoorasingh, P. L., Slakman, B. L., Khanshan, F. S., Carstensen, H. H., Reyniers, M. F., Marin, G. B., West, R. H., & Van Geem, K. M. (2015). Automatic mechanism and kinetic model generation for gas- and solution-phase processes: A perspective on best practices, recent advances, and future challenges. *International Journal of Chemical Kinetics*, 47(4), 199-231. <https://doi.org/10.1002/kin.20902>
- Van Geem, K., Reyniers, M., & Marin, G. (2008). Challenges of modeling steam cracking of heavy feedstocks. *Oil & Gas Science and Technology*, 63(1), 79-94. <https://doi.org/10.2516/ogst:2007084>
- Van Goethem, M. W. M., Kleinendorst, F. I., van Leeuwen, C., & van Velzen, N. (2001). Equation-based SPYRO® model and solver for the simulation of the steam cracking process. *Computers & Chemical Engineering*, 25(4), 905-911. [https://doi.org/10.1016/S0098-1354\(01\)00655-X](https://doi.org/10.1016/S0098-1354(01)00655-X)
- Vangaever, S., Reyniers, P. A., Symoens, S. H., Ristic, N. D., Djokic, M. R., Marin, G. B., & Van Geem, K. M. (2020). Pyrometer-based control of a steam cracking furnace. *Chemical Engineering Research and Design*, 153, 380-390. <https://doi.org/10.1016/j.cherd.2019.10.023>
- Willems, P. A., & Froment, G. F. (1988). Kinetic modeling of the thermal cracking of hydrocarbons. 1. Calculation of frequency factors. *Industrial & Engineering Chemistry Research*, 27(11), 1959-1966. <https://doi.org/10.1021/ie00083a001>
- Yu, K., While, L., Reynolds, M., Wang, X., Liang, J. J., Zhao, L., & Wang, Z. (2018). Multiobjective optimization of ethylene cracking furnace system using self-adaptive multiobjective teaching-learning-based optimization. *Energy*, 148, 469-481. <https://doi.org/10.1016/j.energy.2018.01.159>
- Zakria, M. H., Omar, A. A., & Bustam, M. A. (2016). Mercury removal of fluctuating ethane feedstock in a large scale production by sulphur impregnated activated carbon. *Procedia Engineering*, 148, 561-567. <https://doi.org/10.1016/j.proeng.2016.06.511>
- Zhang, Y., Reyniers, P. A., Du, W., Qian, F., Van Geem, K. M., & Marin, G. B. (2017). Incident radiative heat flux based method for the coupled run length simulation of steam cracking furnaces. *Industrial & Engineering Chemistry Research*, 56(14), 4156-4172. <https://doi.org/10.1021/acs.iecr.6b05013>

Smartphone Application Development for Rice Field Management Through Aerial Imagery and Normalised Difference Vegetation Index (NDVI) Analysis

Nor Athirah Roslin¹, Nik Norasma Che'Ya^{1*}, Rhushalshafira Rosle¹ and Mohd Razi Ismail²

¹Department of Agriculture Technology, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

²Department of Crop Science, Faculty of Agriculture, University Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

In the current practices, farmers typically rely on the traditional method paper-based for farming data records, which leads to human error. However, the paper-based system can be improved by the mobile app technology to ease the farmers acquiring farm data as all of the farm information will be stored in digital form. This study aimed to develop a smartphone agricultural management app known as Padi2U and implement User Acceptance Test (UAT) for end-users. Padi2U was developed using Master App Builder software and integration with the multispectral imagery. Padi2U provides recommendations

based on the Department of Agriculture's (DOA), such as rice check, pest and disease control, and weed management. Through the Padi2U, farmers can access the field data to understand the crop health status online using the Normalised Difference Vegetation Index (NDVI) map derived from the multispectral images. The NDVI is correlated to the Soil Plant Analysis Development (SPAD) value, corresponding to $R^2 = 0.4012$. UAT results showed a 100 percent satisfaction score with suggestions

ARTICLE INFO

Article history:

Received: 2 November 2020

Accepted: 10 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.07>

E-mail addresses:

norathirahroslin@gmail.com (Nor Athirah Roslin)

niknorasma@upm.edu.my (Nik Norasma Che'Ya)

rhushalshafira.rosle@gmail.com (Rhushalshafira Rosle)

razi@upm.edu.my (Mohd Razi Ismail)

*Corresponding author

were given to enhance the Padi2U performance. It shows that Padi2U can be improved to help farmers in the field monitoring virtually by integrating multispectral imagery and information from the field.

Keywords: Chlorophyll content analysis; multispectral imagery; smart farming; smartphone application

INTRODUCTION

Agriculture is one of Malaysia's most significant sectors as it contributes to the country's economy. Smart farming is the idea of enhancing farm management by implementing technology as part of routine farming activities (Muangprathub et al., 2019). For instance, the application of the Internet of Things (IoT) (Jaiganesh et al., 2017) in smart agriculture would provide information on farm management in terms of plant and pest control based on the data collected from the site (Khanna & Kaur, 2019). A significant aspect of agricultural management is the monitoring of crop growth and output during the developmental stages. It helps the farmer to introduce timely measures at the end of the season that ensure maximum yields. In addition to ensuring successful farm management through the smart farming system, the inclusion of technology in agriculture has been shown to save costs by agronomists (Sushanth & Sujatha, 2018). In the agricultural sector, agricultural innovations have been used to assist farmers in their farm management, such as image analysis (Razali et al., 2009; Hudzari et al., 2010; Ishak & Hudzari et al., 2010), Unmanned Aerial Vehicle (UAV) technology (Liu et al., 2018), sensor technology (Nam et al., 2017), Global Navigation Satellite System (GNSS) (Guo et al., 2018), GIS (Zhang & Cao, 2019), and smartphone application technology (Bueno-Delgado et al., 2016). These innovations have been developed to help farmers track and control their crops at home or in the workplace. UAV, for example, is useful for monitoring the entire plantation whereby the farmer may individually observe and examine the crop (Abdullah et al., 2019). In addition, the smartphone app is also useful in introducing farmers to digital agriculture in which farmers can obtain information about crops and manage their field according to the crop status.

The Normalized Difference Vegetation Index (NDVI) created by multispectral UAV is a plant health indicator that is used to quantify the condition of vegetation in order to assess plant germination and its productivity (Rosle et al., 2019). NDVI values vary from -1 to 1, whereby -1 means non-living things, while 1 applies to living things (Christensen, 2019). Furthermore, it has been used to determine crop health status via the detection of chlorophyll content in the crop (Rosle et al., 2019). Smartphone apps have been widely used for different functions in agricultural management. For paddy, the existing system uses the Internet of Things (IoT) and smartphone technology for field management (Kularbphetong et al., 2019). Online at Sawah application shows high acceptance by the users as it is

efficient, easy to learn and satisfactory to use (Simorangkir et al., 2018). Another mobile app, Aptani (Siahaan & Wijaya, 2018) was tailor-made for farmers to provide planting techniques of paddy and methods of controlling pests and diseases in paddy. An automated irrigation system for paddy was also developed in a smartphone app (Barkunan et al., 2019). I-PEDIA is an app developed to identify paddy diseases (Majid et al., 2013). SmartSPAD is another app used in agriculture and it estimates the chlorophyll content (Vesali et al., 2015). Insect Shot is an app that can monitor the brown planthoppers in the paddy field (Watcharabutsarakham et al., 2016). Table 1 shows a summary of the mobile application for agriculture. Based on our review and knowledge, most of the agriculture apps developed limited to one crop management element compared to Padi2U app developed consists of all elements in crop management.

On the other hand, crop monitoring and management will be very useful from the multispectral imagery UAV that can be generated NDVI analysis for health status in the field. Ideally, trends in the future are crop monitoring using IoT and smartphone is one of the common platforms for everyone that can help farmers monitor the crop in an effective way. Thus, this study's objectives were to develop a smartphone app known as Padi2U for agriculture and analyse the relationship between Normalised Difference Vegetation Index (NDVI) generated from UAV and chlorophyll content measured by Soil Plant Analysis Development (SPAD) meter at the ground.

Paddy (*Oryza sativa*) is one of the main crops in Malaysia's agriculture sector (Hassan et al., 2019) and grows rice in this country for staple food. Depending on environmental factors and paddy variety, the complete paddy growth cycle is 120-150 days (Yuzugullu et al.). The paddy growth begins with the flooding of paddy fields for a few weeks before the paddy is transplanted onto the field (Yuzugullu et al., 2017). The first stage of paddy growth is vegetation, whereby germination occurs to initiate the panicle formation (Yuzugullu et al., 2017). The paddy production enters the second stage, the reproductive phase when the panicle successfully forms from the tiller base. At this phase, flowering begins and it enters the stage of maturation (Yuzugullu et al., 2017). Grain filing is done at the ripening stage and excess water in the field is drained to facilitate the harvesting process (Yuzugullu et al., 2017). With total plantation areas of 686,050 hectares, paddy is the most planted crop in Malaysia. Nevertheless, its production is low (Hassan et al., 2017), likely due to the limited amount of fertiliser used (Fatah et al., 2017), excessive pesticide usage (Fatah et al., 2017), climate change (Alam et al., 2019), ineffective irrigation system (Ibrahim & Alam, 2016), Inappropriate technology practice (Hassan et al., 2019), poor farm management (Norasma et al., 2019), unsatisfactory soil conditions (Yusof et al., 2019), disease and pest outbreaks (Yusof et al., 2019) and the output of farmers (Yusof et al., 2019; Alam et al., 2019). Such circumstances cause Malaysia's rice sector to be at a critical point. As a result, to satisfy local demand, Malaysia still needs to import rice from other countries, such as Thailand

and Vietnam. Such circumstances cause Malaysia's rice sector to be at a critical point. As a result, to satisfy local demand, Malaysia still needs to import rice from other countries, such as Thailand and Vietnam.

It is difficult for many rice farmers to conduct crop monitoring and sampling, soil testing, and maintenance of equipment in the paddy sector. In current practice, farmers usually rely on paper-based records to keep all the field data, which is inconvenient as it can be time-consuming for the farmers, subsequently lead to poor management. The optimum planting time is also difficult to predict, apart from tackling any pests and diseases in the plantation (Fountas et al., 2015).

Table 1

Mobile application function and type of crop

Mobile Application	Crop	References
Diagnose disease	Cassava	Mwebaze and Owomugisha (2016)
	Coffee	Manso et al. (2019)
	Tobacco	Valdez-Morones et al. (2018)
	Corn	Simorangkir et al. (2018)
	Paddy	Nasir et al. (2018)
Pest monitoring	Paddy	Watcharabutsarakham et al. (2016)
Field management	Paddy	Kularbphettong et al. (2019)
Planting technique	Paddy	Siahaan and Wijaya (2018)
Irrigation system	Paddy	Barkunan et al. (2019)
Chlorophyll content measurement	Corn	Vesali et al. (2015)

MATERIAL AND METHODS

The study area located at Ladang Merdeka, Kampung Lundang Paku, Ketereh, Kelantan, Malaysia with total area 20 acres (Figure 1). Rice variety planted at this site known as PadiU Putra developed by Universiti Putra Malaysia (UPM) researcher. This paddy variety resistance to leaf blight disease was planted by broadcasting on 30th January 2019.

Data Acquisition

Image Acquisition. Aerial imagery was collected using a DJI Phantom (Shenzhen DJI Sciences and Technologies Ltd., Guangzhou, China) and a multicopter UAV, XR q350 pro mounted with multispectral sensor Parrot Sequoia (Parrot, France, Paris) (Figure 2).



Figure 1. Paddy plot study at Ladang Merdeka, Keteroh, Kelantan.



Figure 2. A multirotor UAV, XR q350 pro.

The flight altitude was 70 m with a 2.13 cm/pix of spatial resolution. Aerial images were collected for different paddy growth stages starting at 12 Day After Planting (DAP), 23 DAP, 36 DAP, 55 DAP, 69 DAP, 75 DAP, 94 DAP, 104 DAP, and 118 DAP (Figure 3).



Figure 3. A pilot controlled the UAV with remote control from ground.

Ground Control Point (GCP) was placed at the corner of the paddy field for the georeferencing process (Figure 4). Before fly the UAV, the sensor was calibrated using white panel references provided by the Parrot Sequoia. Data collection was conducted in the morning from 8:30 am to 11:00 am.



Figure 4. The GCP placed at the corner of the paddy field.

Chlorophyll Concentration Measurement

The leaf chlorophyll concentration measured on fully expand leaf by using a hand-held device known as SPAD-502 meter (Konica Minolta Sensing, Inc., Osaka, Japan) based on the leaf transmittance wavelength measurement for red (650 nm) and infrared (940 nm). This device used to measure the concentration of the chlorophyll in leaf in the safest way and without causing destruction to the leaf sample. The SPAD meter reading was taken at the midpoint of the leaf. Ten SPAD meter readings at each points were averaged to represent the mean SPAD readings. The ground data collection began on 22nd February 2019 (23 DAP), 7th March 2019 (36 DAP) and 26th March 2019 (55 DAP) (Figure 5).



Figure 5. Ground data collection process.

There were eight polyvinyl chloride (PVC) poles about 1 meter height were placed randomly at eight different location around the paddy field area (Figure 6).

Eight points were marked to collect the ground data and at each points ten readings were measured and recorded by using the SPAD-502 meter (Figure 7). The ten leaves sample were taken within 1 meter radius from the PVC poles. The SPAD data collection point were made based on NDVI classification representing the different ranges in its value. The points were used to validate the NDVI value, which had represented the features in the study area. However, due to time constraints and limited human resources and equipment, the study only managed to collect a small number of SPAD data. Therefore, for future research, it is advisable to increase the number of SPAD data to increase accuracy.

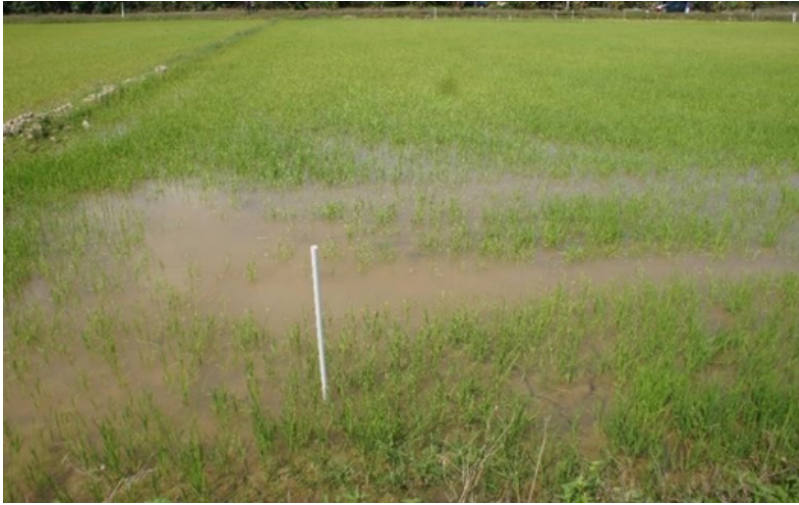


Figure 6. The PVC pole use as marker for ground data collection.

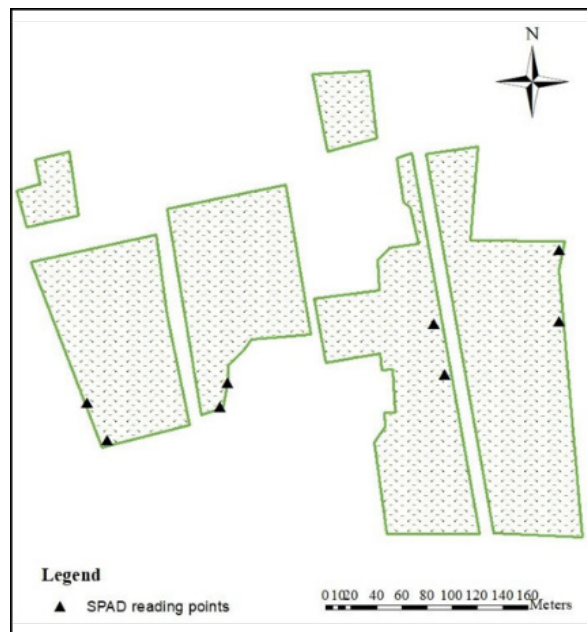


Figure 7. Eight points marked on the map for chlorophyll content data collection.

The ground images captured with high-resolution Canon camera PowerShot SX260 HS (Canon, Japan, Tokyo) (Figure 8).



Figure 8. Data collection for ground data images using Canon camera PowerShot SX260 HS (Canon, Japan, Tokyo).

Image Processing

Image processing process began with import all raw images into Agisoft Photoscan Professional software (Agisoft LLC, St. Peterburg, Russia) for mosaic and align images to produce an orthomosaic image by using Structure from Motion (SfM) algorithm (Figure 9). The orthomosaic image was then processed in ArcGIS 10.1 software, Arc Map to calculate the vegetation index value and produce the map.

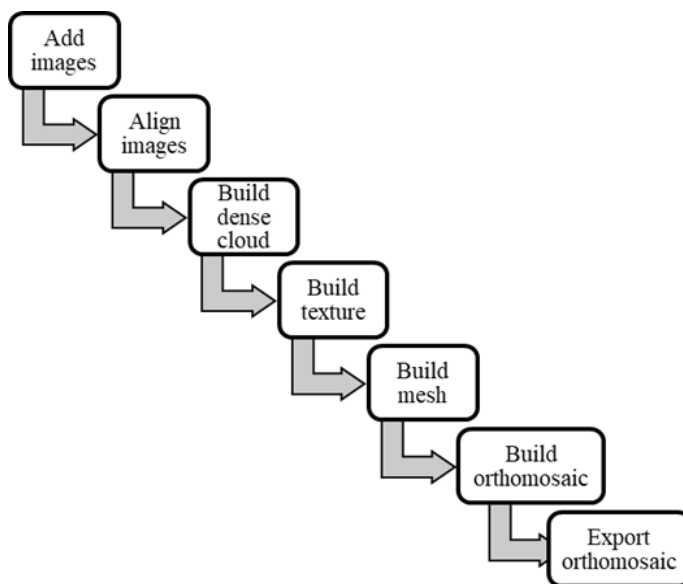


Figure 9. Image processing process in Agisoft Photoscan Professional.

Data Analysis

Vegetation Analysis for Different Vegetation Indices. NDVI is vegetation index was chosen to measure the crop health by using its algorithm formula (Equation 1). There are two bands used to calculate the NDVI value, Near Infrared (NIR) band (770-900 nm) and red band (630-690 nm) based on the absorption in the red caused by the pigmented of chlorophyll and reflection in the infrared due to the structure of the leaf cellular. This NDVI calculation process conducted by using Arc Map software (Esri, California, USA).

$$NDVI = (NIR - Red) / (NIR + Red) \quad (1)$$

Where

NIR = reflection value of near-infrared band

RED = reflection value of red range of the band

Statistical Analysis between Vegetative Indices (NDVI/ NDRE/ SAVI/ OSAVI) with SPAD Readings

The statistical analysis conducted using Microsoft Excel between vegetation index NDVI values from multispectral images analysis in ArcMap correlated with leaf chlorophyll concentration result obtained from the SPAD meter reading. The Pearson correlation analysis used to get the correlation analysis and regression value between the NDVI and chlorophyll concentration. The root mean square error (RMSE) was calculated to predict how dispersed the NDVI values from SPAD reading using Equation 2.

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (Predicted_i - Actual_i)^2}{N}} \quad (2)$$

Where

Predicted_{*i*} = The predicted value for the *i*th observation

Actual_{*i*} = The observed (actual) value for the *i*th observation

N = Total number of observations

Padi2U Mobile App Development

The development of Padi2U involved four stages, namely early design phase, research design phase, research development phase, and research implementation phase.

Early Design Phase. The early design phase aimed to identify the problems faced by the farmers. They were found to have difficulties in finding information about their crop and farms. The information was scattered in many different platforms.

Research Design Phase. The research design phase designed the content of Padi2U. Based on the problems identified in the first phase, the content to be included was information of

agriculture agency, variety of paddy, planting schedule, weather, site information, drone images including multispectral images and red green blue (RGB) images, pest and disease information, supplier contact information, and farmer report form.

Architecture of Padi2U Application. The architecture of Padi2U consisted of three layers: Presentation Layer, Business Layer, and Data Access Layer (Figure 10). In the Presentation Layer, users can view the Graphic User Interface (GUI) of Padi2U and make their decision based on listed examples of Paddy Disease on the menu. The Business Layer is where all the workflows of the app and the actual business of the app take place. Lastly, the Data Access Layer consists of data utilities and access and functions as the back end of the app. For example, if a user requests for disease information, the Business Layer will connect the user to the disease GUI. The database of the disease menu is stored in the Mobile Builder Apps located in Kuala Lumpur, Malaysia.

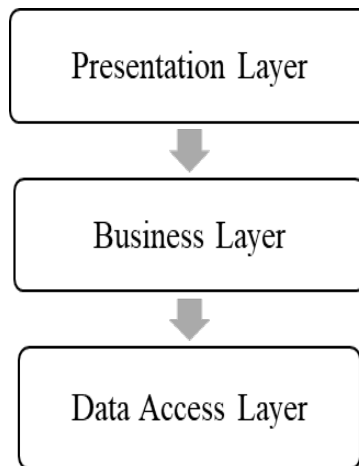


Figure 10. Architecture of Padi2U mobile app.

Padi2U App Menu. The menu and content in the Padi2U app (Figure 11).

Research Development Phase. This phase started with data collection followed by the installation process and system configuration. All the information were analysed and arranged in the database using the Master App Builder software (Malaysia, Kuala Lumpur). The menus were first created before the information was uploaded onto the menus according to the design. The information was in the form of text and images. All the imagery and analysis based on the multispectral and NDVI map were stored in the Padi2U database using Master App Builder software.

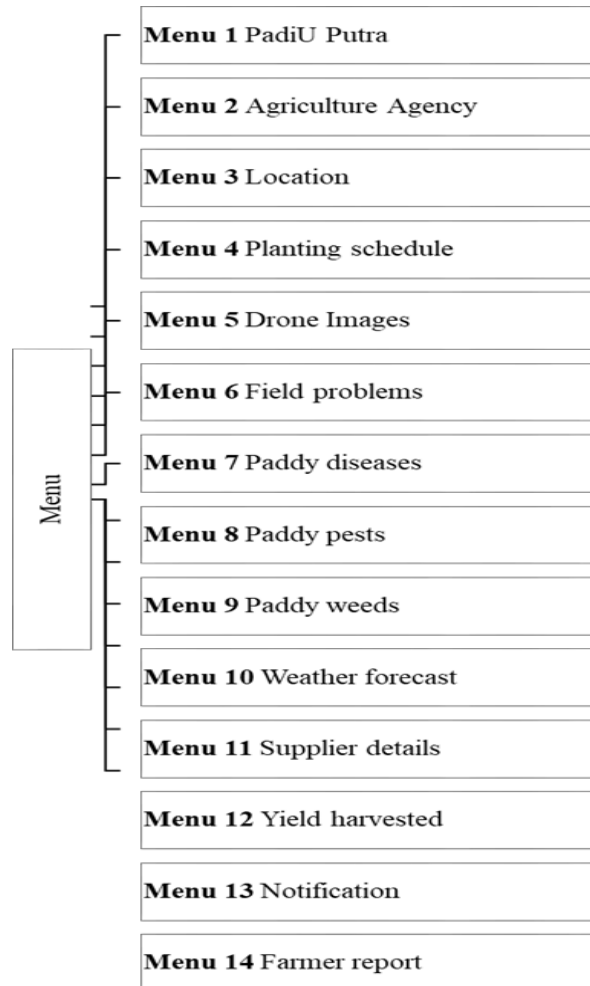


Figure 11. Padi2U app main menu.

Figure 12 shows the software of Master App Builder (Apps Master Builder, Malaysia, Kuala Lumpur) that contains the design of the mobile app, colours, and features. The developer chose one of the design that was suitable and easy to use. The features for the menu originated from a custom page where the text and images were uploaded according to the template. The features on the custom page could also be rearranged based on the content of the menu. For example, under the Drone Images menu and Field Problem menu, the feature chosen was Gallery because this menu could store a lot of pictures and it also enables the users to have a better view by zooming into the images. As for the weather menu, its features were designed specifically for the forecasting of weather. Next, the section on Supplier Information applied the contact information feature while Farmer Report used

a form-based feature that allowed the users to fill in the form before submitting to the back end users. During the development process, the developers could view the app being developed by installing the previewer from the Google Play Store on their smartphone. By filling up the login information, the preview page would be able to show the developers how the app operated on the smartphone before being published.

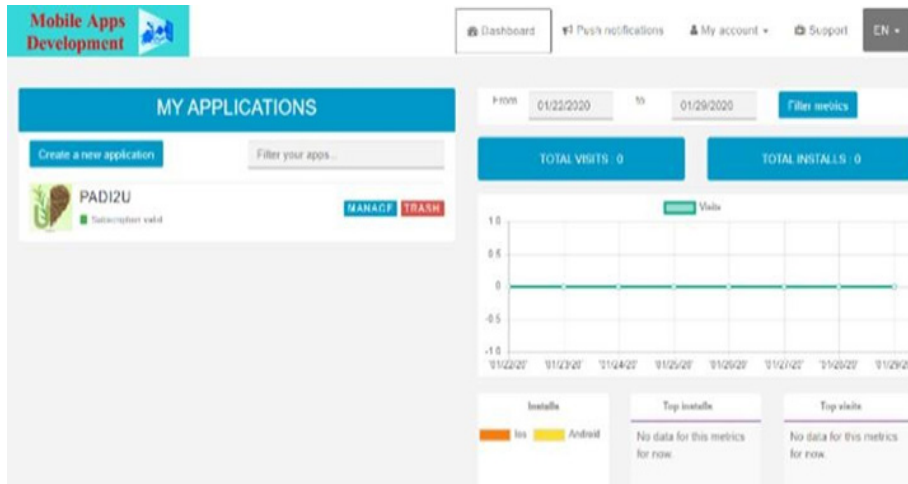


Figure 12. Mobile app development software, Master App Builder (MAB) (Apps Master Builder, Malaysia, Kuala Lumpur).

Research Implementation Phase. The final stage involved the research implementation in which all the functions in Padi2U were developed and tested on the users. Padi2U used Malay, the local language to enable a better understanding of the programme, especially among farmers.

User Acceptance Test (UAT). The UAT was conducted among 13 respondents at Pertubuhan Peladang Kawasan (PPK) Sungai Ketereh, Kelantan (Malaysia) (Figure 13). The respondents selected and decided by the agriculture officer of PPK Sungai Ketereh. Descriptive statistics was used to describe the socio demographic profiles of the respondents. The respondents consisted of farmers, agriculture officers, clerks, and students, and their age ranged from 19 years old to 40 years old. All of them were smartphone users with different operating systems, iOS, and Android. Most the respondents used the mobile app from their smartphones. They tried out the Padi2U app before answering the questionnaire. The questionnaire is divided into two parts, Part A and Part B. In Part A, the question more on respondents' details such as occupation, age, smartphone user, type of operating system, internet user, time spent using the smartphone, and preferable to use

the website or mobile app from their smartphone. In Part B, the questions about the mobile app interface, the information in the app, and its performance. They were also encouraged to provide any recommendations to improve the app.



Figure 13. The respondents testing the Padi2U during UAT.

RESULT AND DISCUSSION

Chlorophyll Content and Vegetation Index Analysis

The average of SPAD readings and NDVI values shown in Table 2 shows data obtained from data collection on 22nd February 2018 (23 DAP), 7 Mac 2018 (36 DAP) and 27 Mac 2018 (55 DAP). The SPAD values are the measurement of chlorophyll concentration in the leaf measured by a hand-held and non-destructive device, SPAD 502 meter. The trend of SPAD values is increasing over the time from 23 DAP to 55 DAP (Table 2). This shows that the chlorophyll concentration is increasing and photosynthesis activity is increasing as the paddy plant grow into different growing stages (Liu et al., 2019). The NDVI values shows the lowest value at 23 DAP around 0.1 because at early rice growth stage, the canopy area of the paddy crops were small and the reflectance caused by the water and soil surface. The NDVI measurement is sensitive to the canopy structure and content of photosynthesis chemical (Gamon et al., 1995). The development of leaf canopy and the rate of photosynthesis were related (Gamon et al., 1995). However, NDVI increased as the plant growth at 36 DAP and 55 DAP. The increase value of NDVI is caused by the development of panicle on the top of the leaf canopy (Casanova et al., 1998). The increasing in value of NDVI is caused by the application of fertilizer and it is the management recommendation

to apply fertilizer (Guan et al., 2019) after seeding stage and at the tillering stage (Casanova et al., 1998).

Table 2

The average of SPAD and NDVI values for 23 DAP, 36 DAP and 55 DAP

	23 DAP	36 DAP	55 DAP
SPAD values	17.41	25.38	37.49
NDVI values	0.10	0.39	0.58

Rice Growth Stages Relationship with SPAD and NDVI Value

The values of SPAD increased as paddy growth into different stages (Figure 14). The highest SPAD reading was at 55 DAP, 37.49 due to high chlorophyll content in the leaf and its thickness (Miah et al., 1997). The values of NDVI increased as paddy growth into different stages (Figure 15). The NDVI value at tillering stages at 23 DAP was close to zero because of low vegetation cover at this stage as the paddy leaf area was still small that only water and soil could be seen (Wang et al., 2011). As the paddy grow, the chlorophyll content increased with its ability to produce high biomass and high photosynthesis rate (Hassan et al., 2009) and influenced by the NIR absorbance.

There are three main rice phenology stages vegetation stage, reproductive stage and ripening stage (Raza et al., 2019). At vegetation stage it takes about 20 days or it can be different depending on paddy variety. The process involved in these stages is germination to panicle initiation. Paddy height, leaf area will increase gradually and active in tillering. The transition from vegetation stage to reproductive stage after the number of tillers achieved its maximum. At reproductive stage, it takes 30 days from panicle initiation to flowering and plant during this time become sensitive to stress. After flowering process, it will enter the last stage of paddy development which is ripening stage. In this stage, filling process takes place until it becomes mature. It takes about 35 days and paddy leaf become dry and turn to yellow colour. The grains are fully developed, and the upper leaves become dry and the panicle will bend downward.

SPAD Value and NDVI Value Relationship

At early growth stages, the chlorophyll content increases and absorption of light energy is close to the correlation line due to strong absorption of red and strong reflection of NIR. The high reflectance energy transmitted on the uppermost layer of the leaf and reflected from the second layer of the leaf is not fully transmitted back to the first layer of the leaf (Roy, 1989). Based on statistical analysis, the regression values between NDVI with the SPAD value, fitting $R^2 = 0.4012$ (Figure 16). Meanwhile its correlation values are 0.6334 (Table 3)

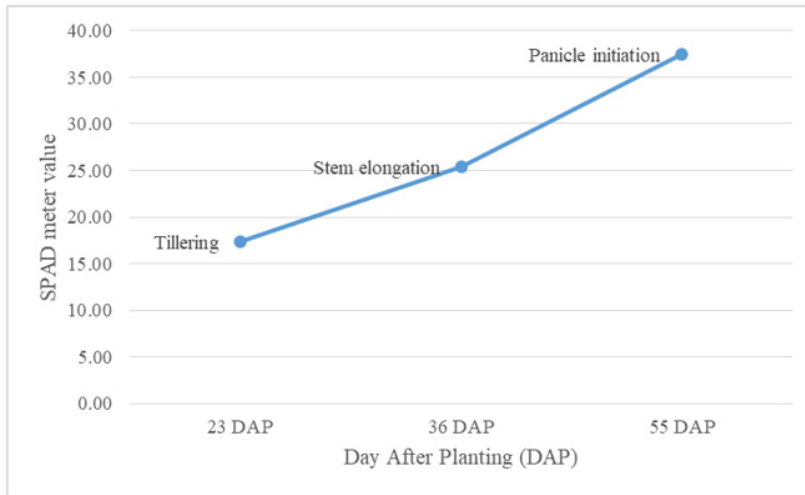


Figure 14. Rice growth stages and SPAD meter values.

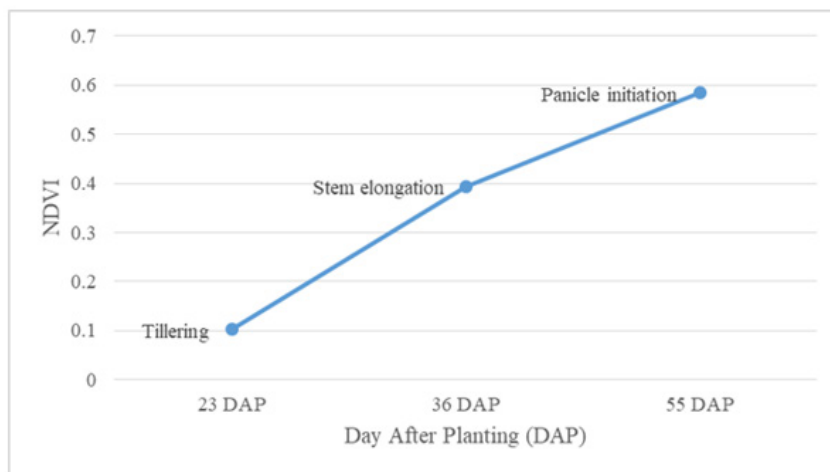


Figure 15. Rice growth stages and NDVI values.

with its root mean square error (RMSE), 0.0844. According to McHugh (2012), a correlation values ranged from 0.60 to 0.79 can be considerate as moderate with percentage of data that are reliable is 35% to 63%. We were unable to obtain a higher regression and correlation values due to small number of samples size and in future research it is recommended to increase the number of samples. From the analysis, NDVI show an acceptable correlation and regression value because NDVI sensitive to detect changes in serious water shortage, nutrient deficiencies and pest or disease attack. There is strong correlation between NDVI and SPAD at the early reproductive stage and reproductive stage (Guan et al., 2019).

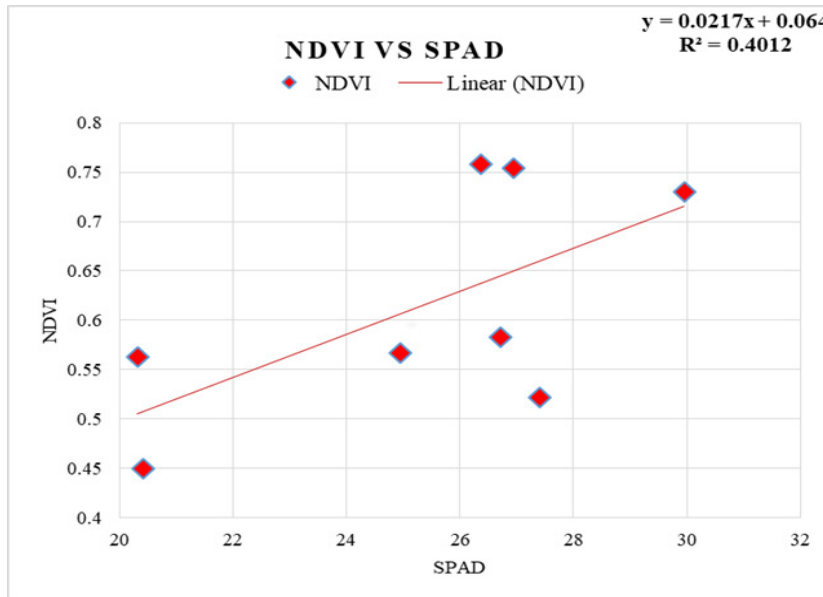


Figure 16. The regression values between SPAD with NDVI value.

Table 3

The correlation values between SPAD and NDVI

	SPAD	NDVI
SPAD	1	
NDVI	0.633418	1

Spatial Interpolation

The highest predicted value represents by green colour and for the lowest predicted value in red colour. Table 4 shows spatial interpolation of NDVI map and SPAD map of study area. The value range of NDVI value at 23 DAP, 36 DAP and 55 DAP was -0.64 - 0.73, -1 - 1 and -1 - 1. The value range of the SPAD value at 23 DAP, 36 DAP and 55 DAP was 0 - 24, 0 - 40 and 0-40.

Padi2U Mobile Application

The interface of the Padi2U smartphone app main menu successfully developed (Figure 17). Padi2U consisted of 14 menus. The different menu and their descriptions are outlined (Table 5).

Under the Drone Images menu, all the images were arranged according to the day of planting. To view the development of paddy growth for the whole season, the users just need to slide to the next image. The drone images were split into two types of images,

namely RGB images and the map of NDVI. The index reading referred to the health status of plants and RGB images represented the true colour images for the users to compare with the NDVI map so that they could correctly identify the location (Figure 18).

Table 4
Spatial interpolation of NDVI and SPAD map

DAP	NDVI map	SPAD map
23		
36		

Table 4 (Continued)

DAP	NDVI map	SPAD map
55		



Figure 17. Main menu of Padi2U mobile app.

Table 5
Padi2U menu description

Menu	Description
PadiU Putra	Details about PadiU Putra
Agriculture Agency	Details about KADA
Location	Details about project site
Planting Schedule	Details about activity
Drone Images	RGB and NDVI map
Field Problems	Field problem images
Paddy Diseases	List of diseases in paddy and method to control
Paddy Pests	List of pests in paddy and method to control
Paddy Weeds	List of weeds in paddy field and method to control
Weather Forecast	Details about weather
Supplier Details	Details about agriculture supplier
Yield Harvested	Total of yield harvested
Notification	Notify user of any latest news
Farmer Report	User can send report about crop or field



Figure 18. The drone images of multispectral image and RGB images.

Next, the pest and disease information menu provided photos of the disease, symptoms, and methods to control the disease (Figure 19). The pest images were difficult to find due to the poor weather condition and inappropriate time during the site visit. Thus, some of the images used were from reliable external sources. From the images, users were able to recognize the pest found in the field. Padi2U also provided recommendations on how to control the pest and disease in the field.

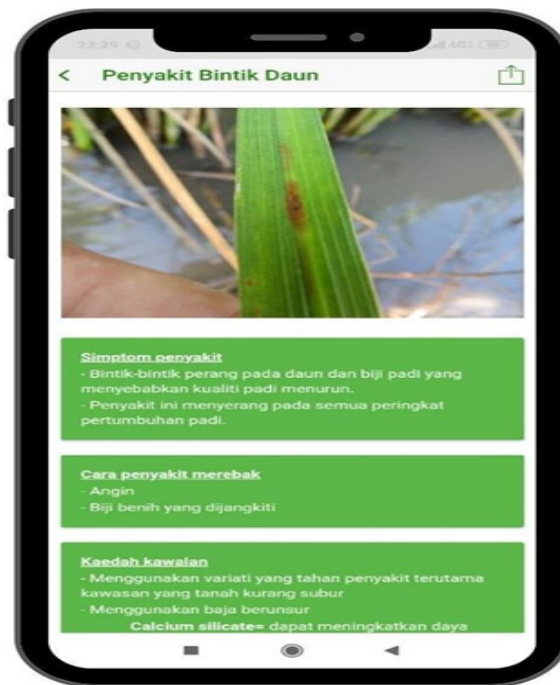


Figure 19. The symptoms caused by disease infection on the crop were shown in the app.

Lastly, the farmers could also send feedbacks or reports via the app (Figure 20). They just needed to fill in their personal information such as name before they submit the report details including photos. Therefore, Padi2U could be used by farmers and non-farmers who want to learn about agriculture and paddy crop management.

This study outlined the successful development of the Padi2U app for the Android operating system. In the future, the app can also be developed for the iOS system using the same method. The Padi2U app is different compared to existing mobile apps because it includes drone images of RGB images and multispectral images. Currently, there is no mobile app that contains drone images like Padi2U. Most of the mobile apps developed used satellite images for monitoring (Herrick et al., 2016). By comparison, images captured by drones are more realistic and clearer with higher resolution than satellite images (Khaliq

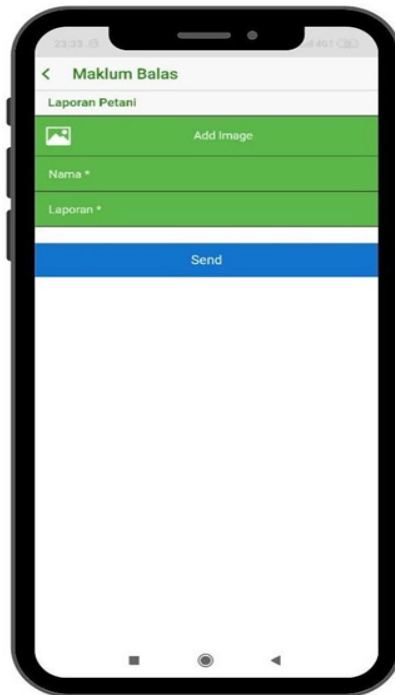


Figure 20. Farmer report form details in Padi2U app.

et al., 2019). In addition, most of the currently available mobile apps focus on one specific management aspect, such as pest and disease management (Johannes et al., 2017), irrigation system (Kodali & Sarjerao, 2017) and soil study (Stiglitz et al., 2017). In contrast, Padi2U covers all the management aspects in paddy cultivation, including planting schedule, pest management, disease management, weed management, weather forecast, farmer report, and yield harvest information.

Nowadays, with the widespread use of the smartphone, the development of a mobile app is becoming more relevant in many sectors, including agriculture. With this, farmers can rely on faster data transmission for decision making on the management of pests and diseases. The strength of Padi2U lies within the use of the local language, thus making it more user- friendly for local farmers. Its focus on paddy and the specific area in Kelantan would be beneficial for the users from that region. Nevertheless, there are some limitations to this app. Firstly, the mobile app software requires regular updates. This can be time-consuming and also interfere with the use until the update is complete. Secondly, it has a slow response and often lags, thus requiring further improvement on the app. Furthermore, the app is only applicable for paddy crop now.

User Acceptance Test

The UAT was carried out using a questionnaire. A total of 13 respondents were involved in this UAT. They were from various agricultural backgrounds such as agriculture extension officers, farmers, and interns. The questionnaire was divided into two sections (section A and section B). Section A captured personal information such as occupation, age, and smartphone usage. Section B consisted of features of the Padi2U app such as the colour, information, multispectral images, and satisfaction towards the app. The results were analysed using Microsoft Excel. It showed that all the users had no problems while using the Padi2U app. Thus, satisfaction rate was 100% (Figure 21).

Based on the result, the users showed a high percentage of satisfaction with the app. The GUI of the Padi2U app was straightforward, easy to use, and user-friendly. The farmers were also satisfied with the content regarding pests and diseases. However, they requested for the app developer to include more information about pesticides and its application in controlling pest and disease infection. Besides farmers, Padi2U can also be used by the agriculture officers. It can act as a tool to distribute and receive information in real-time. This app can also serve as a database on field information for paddy management.

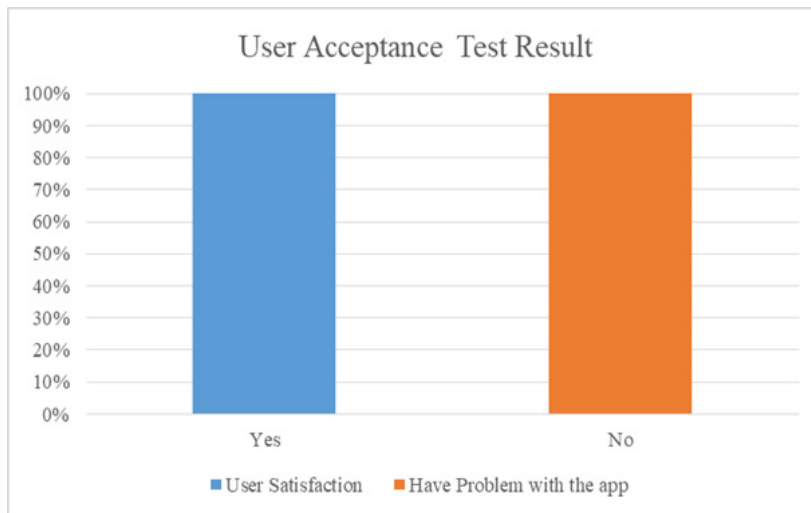


Figure 21. User Acceptance Test result showing 100% satisfaction rate towards the Padi2U app.

CONCLUSION

In conclusion, technologies used in agriculture such as UAV and smartphone app can helped farmer to manage the crop. The multispectral UAV can be used to generate NDVI map that can assist users to monitor and manage their field in effective way by using smart phone at anywhere and anytime. The Padi2U app was successfully developed for

paddy management based on the multispectral UAV and various data from the ground and agencies. It has a complete set of information about paddy cultivation from plantation up to harvesting. The NDVI map used in the app can assist crop growth monitoring and help farmers to increase production. Besides, farmers can overcome pest, disease, and weed problems at their farms by planning their actions based on the information in the app. This user-friendly app also makes it very convenient for the farmers to manage their paddy fields by providing information on the weather forecast, location to purchase agriculture equipment. With the 100 percentage of Padi2U acceptance by the users, this mobile app is proven to help the farmers manage their farms. By incorporating the feedbacks, Padi2U can be improved to be a better and more efficient mobile app. The use of Padi2U app will be useful to farmer and agriculture officer to manage the crop. However, this app needs real-time data of the paddy field and updated into the app. The NDVI value has correlation with the SPAD value, fitting $R^2 = 0.4012$ because NDVI sensitive to detect changes in serious water shortage, nutrient deficiencies and pest or disease attack. Thus, this information can be as an indicator to farmers to regularly monitor their crop status at the field in efficient way. For recommendation, the number of samples needs to be increased in order to increase the accuracies. This limitation due to less of human resources to collect the data. Future research can apply the same method to develop a mobile app for other types of crops such as corn and pineapple.

ACKNOWLEDGEMENT

The authors would like to acknowledge the assistance by Mr. Nasir Abd. Rani and Mr. Mohd Zalyah Shah B Noh from Kemubu Agricultural Development Authority (KADA). This study also received the financial support from the Ministry of Higher Education, Malaysia under the Long-term Research Grant Scheme (LRGS) for the program titled Development of Climate Ready Rice for Sustaining Food Security in Malaysia (Vote No: 5545000) and UPM GP-IPM (Vote No: 9611400).

REFERENCES

- Abdullah, S., Tahar, K. N., Rashid, M. F. A., & Osoman, M. A. (2019). Camera calibration performance on different non-metric cameras. *Pertanika Journal of Science & Technology*, 27(3), 1397-1406.
- Alam, M. J., Awal, M. A., & Mustafa, M. N. (2019). Crops diseases detection and solution system. *International Journal of Electrical and Computer Engineering*, 9(3), 2112-2120. <https://doi.org/10.11591/ijece.v9i3.pp2112-2120>
- Barkunan, S. R., Bhanumathi, V., & Sethuram, J. (2019). Smart sensor for automatic drip irrigation system for paddy cultivation. *Computers & Electrical Engineering*, 73, 180-193. <https://doi.org/10.1016/j.compeleceng.2018.11.013>

- Bueno-Delgado, M. V., Molina-Martínez, J. M., Correoso-Campillo, R., & Pavón-Mariño, P. (2016). Ecofert: An android application for the optimization of fertilizer cost in fertigation. *Computers and Electronics in Agriculture*, *121*, 32-42. <https://doi.org/10.1016/j.compag.2015.11.006>
- Casanova, D., Epema, G. F., & Goudriaan, J. (1998). Monitoring rice reflectance at field level for estimating biomass and LAI. *Field Crops Research*, *55*(1-2), 83-92. [https://doi.org/10.1016/S0378-4290\(97\)00064-6](https://doi.org/10.1016/S0378-4290(97)00064-6)
- Christensen, B. M. (2019). *Using mid-season NDVI data from drones to produce variable rate fertilizer maps in wheat* [Master Thesis]. The North Dakota State, United States. Retrieved July 17, 2020, from <https://search.proquest.com/docview/2330627605?accountid=27932>
- 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>
- Fountas, S., Carli, G., Sørensen, C. G., Tsiropoulos, Z., Cavalaris, C., Vatsanidou, A., Liakos, B., Canavari, M., Wiebensohn, J., & Tisserye, B. (2015). Farm management information systems: Current situation and future perspectives. *Computers and Electronics in Agriculture*, *115*, 40-50. <https://doi.org/10.1016/j.compag.2015.05.011>
- Gamon, J. A., Field, C. B., Goulden, M. L., Griffin, K. L., Hartley, A. E., Joel, G., Penuelas, J., & Valentini, R. (1995). Relationships between NDVI, canopy structure, and photosynthesis in three Californian vegetation types. *Ecological Applications*, *5*(1), 28-41. <https://doi.org/10.2307/1942049>
- Guan, S., Fukami, K., Matsunaka, H., Okami, M., Tanaka, R., Nakano, H., Sakai, T., Nakano, K., Ohdan, H., & Takahashi, K. (2019). Assessing correlation of high-resolution NDVI with fertilizer application level and yield of rice and wheat crops using small UAVs. *Remote Sensing*, *11*(2), Article 112. <https://doi.org/10.3390/rs11020112>
- Guo, J., Li, X., Li, Z., Hu, L., Yang, G., Zhao, C., Fairbairn, D., Watson, D., & Ge, M. (2018). Multi-GNSS precise point positioning for precision agriculture. *Precision Agriculture*, *19*(5), 895-911. <https://doi.org/10.1007/s11119-018-9563-8>
- Hassan, M. S., Khair, A., Haque, M. M., Azad, A. K., & Hamid, A. (2009). Genotypic variation in traditional rice varieties for chlorophyll content, SPAD value and nitrogen use efficiency. *Bangladesh Journal of Agricultural Research*, *34*(3), 505-515. <https://doi.org/10.3329/bjar.v34i3.3977>
- Hassan, S., Mohamed, Z. A. B., Abdullah, S. N. S., & Zaini, N. N. (2017). Personality traits and its relationship with work performance for majority group of paddy farmers in Malaysia. *Australian Academy of Business and Economics Review*, *2*(3), 234- 243.
- Hassan, S., Yussof, N., & Galadima, M. (2019). Farmers current agriculture practices on paddy cultivation and relationship with work performance in Iada Batang Lupar, Sarawak, Malaysia. *Asian Journal of Agricultural Extension, Economics & Sociology*, *31*(3), 1-14. <https://doi.org/10.9734/ajaees/2019/v31i330134>

- Herrick, J. E., Beh, A., Barrios, E., Bouvier, I., Coetzee, M., Dent, D., & Matuszak, J. (2016). The land-potential knowledge system (LandPKS): Mobile apps and collaboration for optimizing climate change investments. *Ecosystem Health and Sustainability*, 2(3), Article e01209. <https://doi.org/10.1002/ehs2.1209>
- Hudzari, R. M., Ishak, W. W. W., & Noorman, M. M. (2010). Parameter acceptance of software development for oil palm fruit maturity prediction. *Journal of Software Engineering*, 4(3), 244-256.
- Ibrahim, A. Z., & Alam, M. M. (2016). Climatic changes, government interventions, and paddy production: an empirical study of the Muda irrigation area in Malaysia. *International Journal of Agricultural Resources, Governance and Ecology*, 12(3), 292-304. <https://doi.org/10.1504/IJARGE.2016.078319>
- Ishak, W. I. W., & Hudzari, R. M. (2010). Image based modeling for oil palm fruit maturity prediction. *Journal of Food, Agriculture & Environment*, 8(2), 469-476.
- Jaiganesh, S., Gunaseelan, K., & Ellappan, V. (2017). IOT agriculture to improve food and farming technology. In *2017 Conference on Emerging Devices and Smart Systems (ICEDSS)*. IEEE Conference Publication.
- Johannes, A., Picon, A., Alvarez-Gila, A., Echazarra, J., Rodriguez- Vaamonde, S., Navajas, A. D., & Ortiz-Barredo, A. (2017). Automatic plant disease diagnosis using mobile capture devices, applied on a wheat use case. *Computers and Electronics in Agriculture*, 138, 200- 209. <https://doi.org/10.1016/j.compag.2017.04.013>
- Khaliq, A., Comba, L., Biglia, A., Ricauda Aimonino, D., Chiaberge, M., & Gay, P. (2019). Comparison of satellite and UAV-based multispectral imagery for vineyard variability assessment. *Remote Sensing*, 11(4), Article 436. <https://doi.org/10.3390/rs11040436>
- Khanna, A., & Kaur, S. (2019). Evolution of Internet of Things (IoT) and its significant impact in the field of precision agriculture. *Computers and Electronics in Agriculture*, 157, 218-231. <https://doi.org/10.1016/j.compag.2018.12.039>
- Kodali, R. K., & Sarjerao, B. S. (2017). A low cost smart irrigation system using MQTT protocol. In *2017 IEEE Region 10 Symposium (TENSYP)*. IEEE Conference Publication.
- Kularbphetong, K., Phoso, W., & Roonrakwit, P. (2019). The Automation of Mobile Application to Manage the Rice Fields. *TEM Journal*, 8(3), 866-871.
- Liu, C., Liu, Y., Lu, Y., Liao, Y., Nie, J., Yuan, X., & Chen, F. (2019). Use of a leaf chlorophyll content index to improve the prediction of above- ground biomass and productivity. *Peer J*, 6, Article e6240. <https://doi.org/10.7717/peerj.6240>
- Liu, S., Li, L., Gao, W., Zhang, Y., Liu, Y., Wang, S., & Lu, J. (2018). Diagnosis of nitrogen status in winter oilseed rape (*Brassica napus* L.) using in-situ hyperspectral data and unmanned aerial vehicle (UAV) multispectral images. *Computers and Electronics in Agriculture*, 151, 185-195. <https://doi.org/10.1016/j.compag.2018.05.026>
- Majid, K., Herdiyeni, Y., & Rauf, A. (2013). I-PEDIA: Mobile application for paddy disease identification using fuzzy entropy and probabilistic neural network. In *2013 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*. IEEE Conference Publication.
- Manso, G. L., Knidel, H., Krohling, R. A., & Ventura, J. A. (2019). A smartphone application to detection and classification of coffee leaf miner and coffee leaf rust. *Computer Vision and Pattern Recognition*, 1-36.

- McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, 22(3), 276-282.
- Miah, M. N. H., Yoshida, T., & Yamamoto, Y. (1997). Effect of nitrogen application during ripening period on photosynthesis and dry matter production and its impact on yield and yield components of semidwarf inica rice varieties under water culture conditions. *Soil Science and Plant Nutrition*, 43(1), 205-217. <https://doi.org/10.1080/00380768.1997.10414728>
- Muangprathub, J., Boonnam, N., Kajornkasirat, S., Lekbangpong, N., Wanichsombat, A., & Nillaor, P. (2019). IoT and agriculture data analysis for smart farm. *Computers and Electronics in Agriculture*, 156, 467-474. <https://doi.org/10.1016/j.compag.2018.12.011>
- Mwebaze, E., & Owomugisha, G. (2016). Machine learning for plant disease incidence and severity measurements from leaf images. In *2016 15th IEEE international conference on machine learning and applications (ICMLA)*. IEEE Conference Publication.
- Nam, W. H., Kim, T., Hong, E. M., Choi, J. Y., & Kim, J. T. (2017). A wireless sensor network (WSN) application for irrigation facilities management based on Information and Communication Technologies (ICTs). *Computers and Electronics in Agriculture*, 143, 185-192. <https://doi.org/10.1016/j.compag.2017.10.007>
- Nasir, H., Aris, A. N., Lajis, A., Kadir, K., & Safie, S. I. (2018). Development of android application for pest infestation early warning system. In *2018 IEEE 5th International Conference on Smart Instrumentation, Measurement and Application (ICSIMA)* (pp. 1-5). IEEE Conference Publication.
- Norasma, C. Y. N., Fadzilah, M. A., Roslin, N. A., Zanariah, Z. W. N., Tarmidi, Z., & Candra, F. S. (2019). Unmanned aerial vehicle applications in agriculture. In *IOP Conference Series: Materials Science and Engineering. Aceh, Indonesia*. IOP Publishing.
- Raza, S. M. H., Mahmood, S. A., Gillani, S. A., Hassan, S. S., Aamir, M., Saifullah, M., Basheer, M., Ahmad, A., Rehman, S. U., & Ali, T. (2019). Estimation of net rice production by remote sensing and multi source datasets. *Sarhad Journal of Agriculture*, 35(3), 955-965. <https://doi.org/10.17582/journal.sja/2019/35.3.955.965>
- Razali, M. H., Ismail, W. I. W., Ramli, A. R., Sulaiman, M. N., & Harun, M. H. (2009). Development of image based modeling for determination of oil content and days estimation for harvesting of fresh fruit bunches. *International Journal of Food Engineering*, 5(2), 1633-1637.
- Rosle, R., Che'Ya, N. N., Roslin, N. A., Halip, R. M., & Ismail, M. R. (2019). Monitoring early stage of rice crops growth using normalized difference vegetation index generated from UAV. In *IOP Conference Series: Earth and Environmental Science. Sulawesi, Indonesia*. IOP Publishing.
- Roy, P. S. (1989). Spectral reflectance characteristics of vegetation and their use in estimating productive potential. *Plant Sciences*, 99(1), 59-81. <https://doi.org/10.1007/BF03053419>
- Siahaan, A. P. U., & Wijaya, R. F. (2018). Smart farmer application in monitoring and learning of android-based rice cultivation. *International Journal of Scientific Research in Science and Technology* 4(11), 16-20. <https://doi.org/10.32628/IJSRST1840115>
- Simorangkir, G. D., Sarwoko, E. A., Sasongko, P. S., & Endah, S. N. (2018). Usability testing of corn diseases and pests detection on a mobile application. In *2018 2nd International Conference on Informatics and Computational Sciences (ICICoS)*. IEEE Conference Publication.

- Stiglitz, R., Mikhailova, E., Post, C., Schlautman, M., Sharp, J., Pargas, R., Glover, B., & Mooney, J. (2017). Soil color sensor data collection using a GPS- enabled smartphone application. *Geoderma*, 296, 108-114. <https://doi.org/10.1016/j.geoderma.2017.02.018>
- Sushanth, G., & Sujatha, S. (2018). IOT based smart agriculture system. *2018 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*. IEEE Conference Publication.
- Valdez-Morones, T., Pérez-Espinosa, H., Avila-George, H., Oblitas, J., & Castro, W. (2018). An android app for detecting damage on tobacco (*Nicotiana tabacum* L.) leaves caused by blue mold (*Penospora tabacina* Adam). In *2018 7th International Conference On Software Process Improvement (CIMPS)*. IEEE Conference Publication.
- Vesali, F., Omid, M., Kaleita, A., & Mobli, H. (2015). Development of an android app to estimate chlorophyll content of corn leaves based on contact imaging. *Computers and Electronics in Agriculture*, 116, 211-220. <https://doi.org/10.1016/j.compag.2015.06.012>
- Wang, F. M., Huang, J. F., & Lou, Z. H. (2011). A comparison of three methods for estimating leaf area index of paddy rice from optimal hyperspectral bands. *Precision Agriculture*, 12(3), 439-447. <https://doi.org/10.1007/s11119-010-9185-2>
- Watcharabutsarakham, S., Methasate, I., Watcharapinchai, N., Sinthupinyo, W., & Sriratanasak, W. (2016). An approach for density monitoring of brown planthopper population in simulated paddy fields. In *2016 13th International Joint Conference on Computer Science and Software Engineering (JCSSE)* (pp. 1-4). IEEE Conference Publication.
- Yusof, Z. M., Misiran, M., Baharin, N. F., Yaacob, M. F., Aziz, N. A. B. A., & Sanan, N. H. B. (2019). Projection of Paddy Production in Kedah Malaysia: A Case Study. *Asian Journal of Advances in Agricultural Research*, 1-6. <https://doi.org/10.9734/ajaar/2019/v10i330030>
- Yuzugullu, O., Marelli, S., Erten, E., Sudret, B., & Hajnsek, I. (2017). Determining rice growth stage with X-band SAR: A metamodel based inversion. *Remote Sensing*, 9(5), Article 460. <https://doi.org/10.3390/rs9050460>
- Zhang, F., & Cao, N. (2019). Application and research progress of geographic information system (GIS) in agriculture. In *2019 8th International Conference on Agro-Geoinformatics (Agro- Geoinformatics)* (pp. 1-5). IEEE Conference Publication.

Review Article

A Review Article on Software Effort Estimation in Agile Methodology

Pantjawati Sudarmaningtyas^{1,2} and Rozlina Mohamed^{1*}

¹*Faculty of Computing, Universiti Malaysia Pahang, 26300 UMP, Gambang, Kuantan, Pahang, Malaysia*
²*Department of Information System, Universitas Dinamika, 60298 Surabaya, Jawa Timur, Indonesia*

ABSTRACT

Currently, Agile software development method has been commonly used in software development projects, and the success rate is higher than waterfall projects. The effort estimation in Agile is still a challenge because most existing means are developed based on the conventional method. Therefore, this study aimed to ascertain the software effort estimation method that is applied in Agile, the implementation approach, and the attributes that affect effort estimation. The results showed the top three estimation that is applied in Agile, are machine learning (37%), Expert Judgement (26%), and Algorithmic (21%). The implementation of all machine learning methods used a hybrid approach, which is a combination of machine learning and expert judgement, or a mix of two or more machine learning. Meanwhile, the implementation of effort estimation through a hybrid approach was only used in 47% of relevant articles. In addition, effort estimation in Agile involved twenty-four attributes, where Complexity, Experience, Size, and Time are the most commonly used and implemented.

ARTICLE INFO

Article history:

Received: 18 July 2020

Accepted: 20 October 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.08>

E-mail addresses:

pantja@dinamika.ac.id (Pantjawati Sudarmaningtyas)

rozlina@ump.edu.my (Rozlina Mohamed)

*Corresponding author

Keywords: Agile, effort estimation attributes, expert judgement, hybrid approach, software effort estimation

INTRODUCTION

Software development projects widely use Agile software development method, especially Scrum methodology with iteration planning techniques. This is in accordance with the survey which showed that 94% of

respondents employ Agile, and 60% have more than three years of experience. Scrum is the most common methodology used by respondents' organizations (58%), while the top Agile techniques are iteration planning (90%) (VersionOne.com, 2017). Besides, statistics showed that the success rate of Agile is two times more likely to succeed, and one-third less likely to fail than waterfall projects (Mersino, 2018).

Based on a survey, it was found that 45% of IT projects transcend budget because it is not established according to the factual requirement (Bloch et al., 2012). The estimation of cost in software development is essential to avoid excessive costs. In general, the costs are based on effort estimation (Bloch et al., 2012). Therefore, effort estimation is a crucial part of the software development project.

The objective of this estimation process is to provide an approximation of the resources needed to complete a project, in order to deliver outputs in the form of products or services in accordance with the specified characteristics of functional and non-functional (Institute, 2017). Estimates are usually internally generated and periodically conducted. Meanwhile, early effort estimation, schedule, and cost are a repetitive work to be compromised and reviewed between stakeholders to reach an agreement regarding the requirement of resources and time to complete the projects (Bourque & Fairley, 2014).

The precision measurement of a single effort estimate is not straightforward; therefore, the value is better communicated in the interval. Although there is no connection between the implied effort intervals and confidence level, but there is estimator confidence of a possibility that the actual efforts will be within range (Jørgensen, 2016).

The estimate can be interpreted from the requirement of resources, such as people and tools (Bourque & Fairley, 2014). This effort is a composite of person and time, which indicates the number of thoroughly productive working hours necessary to get a work done. Also, the units of effort are typically stated in person-hours, person-days, person-months, and person-years (Trendowicz & Jeffery, 2014).

Agile substantially is quite different from the waterfall method in the manageability of software development. The collaborative and cooperative approach between all stakeholders is the main characteristic of the agile software development method that perform by involving users actively, empower the team to make decisions, capture requirements in lightweight and visual, focus on frequent delivery of products through developing small, incremental release, and iterate (Project-Management.com, 2019).

Following the characteristics of agile that focus on delivery products, effort estimation uses to determine the number of iterations to predict the date delivery project. In contrast, the other software development methodology uses effort estimation as a basis for calculating the cost. Thereby effort estimation remains the main challenge in agile software development projects because there were not yet commonly accepted standardized effort estimation techniques for agile software development. Therefore, this study aimed to review the

methods and approaches of software effort estimation, in order to ascertain the most appropriate method for Agile development.

This study was conducted to provide a basis for further development of effort estimation method to be used in Agile software development. Towards achieving the target as mentioned above, this paper was organized into five parts. Beginning with an introduction, followed by overview of effort estimation techniques, materials and methods, results, and discussion, as well as the conclusion.

OVERVIEW OF SOFTWARE EFFORT ESTIMATION METHOD

Currently, many software effort estimation methods are implemented in projects such as Constructive Cost Model (COCOMO), Function point Analysis, Source Line of Code (SLOC), SEER-SEM (Software Evaluation and Estimation of Resources-Software Estimation Model). Furthermore, they are implemented in Linear, Multiplicative, and Putnam Models, Brake Down Estimation, Artificial Neural Network, as well as Fuzzy. The existing methods are shown in Figure 1, and three researchers have classified them according to their perspective.

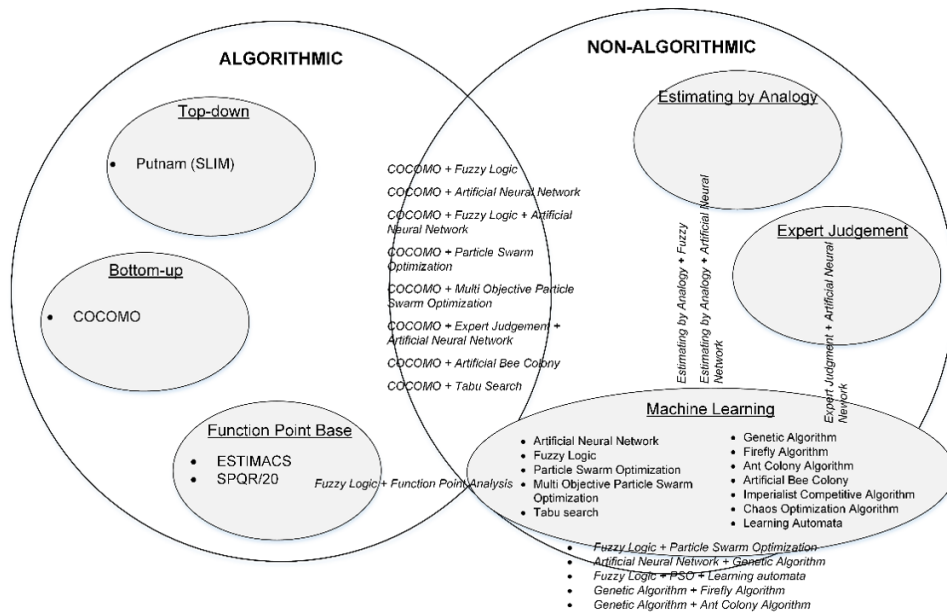


Figure 1. Existing Software Effort Estimation Method

The first researcher (Boehm, 1984) classified software estimation methods from the perspective of the technique used. Based on this, the classification was divided into seven categories, namely Algorithmic models, Expert Judgment, Analogy, Parkinson, Price-to-Win, Top-Down, and Bottom-Up.

The second researcher (Srivastava & Wadhwa, 2013) classified software estimation methods based on an algorithm. From that perspective, four categories of software effort estimation were obtained, which are Algorithmic, Non-Algorithmic, Parametric, and Machine learning. Some popular methods in Algorithmic models include Function point Analysis, Source Line of Code (SLOC), SEER-SEM (Software Evaluation and Estimation of Resources-Software Estimation Model), Linear Model, Multiplicative Model, Putnam Model, and Constructive Cost Model (COCOMO). Furthermore, Analogy and Expert Judgment are part of Non-algorithmic models, while Brake Down Estimation is one of the methods in Parametric. In addition, the estimation methods that include Machine learning models are Artificial Neural Network and Fuzzy.

The third researcher classified software estimation methods by the data input type and the principle of estimation that was employed (Trendowicz & Jeffery, 2014). Based on this, the methods were classified into three categories, namely Data-driven, Expert-based, and Hybrid. The Data-driven has two groups, which are Proprietary and Non-proprietary. Furthermore, the Non-proprietary is divided into three classes, namely Model-based, Memory-based, and Composite. The Model-based consists of Parametric, Non-parametric, and Semi-parametric.

Table 1 shows a summary of existing methods classification based on the above explanation.

Table 1
Existing Effort Estimation Methods Classification

Classification	(Boehm, 1984)	(Srivastava & Wadhwa, 2013)	(Trendowicz & Jeffery, 2014)
Algorithmic models	√	√	
Expert Judgment/Expert base	√		√
Analogy	√		
Parkinson	√		
Price-to-Win	√		
Top-Down	√		
Bottom-Up	√		
Non-algorithmic		√	
Parametric		√	√
Non-parametric			√
Semi-parametric			√
Machine Learning		√	
Data-driven			√
Hybrid			√

Based on the previous research, this study classified software effort estimation methods based on three aspects, namely (1) estimation principle that is employed (2) estimation strategy that is applied, and (3) data that are required. In general, each of these aspects are divided into two parts. The estimation principle issue consists of algorithmic and non-algorithmic. The facet of the strategy is divided into two types, namely Top-Down and Bottom-Up. While the data aspect indicates that some methods have a high dependency on historical data, but others do not. Figure 2 shows the classification that was used in this study and existing software effort estimation methods.

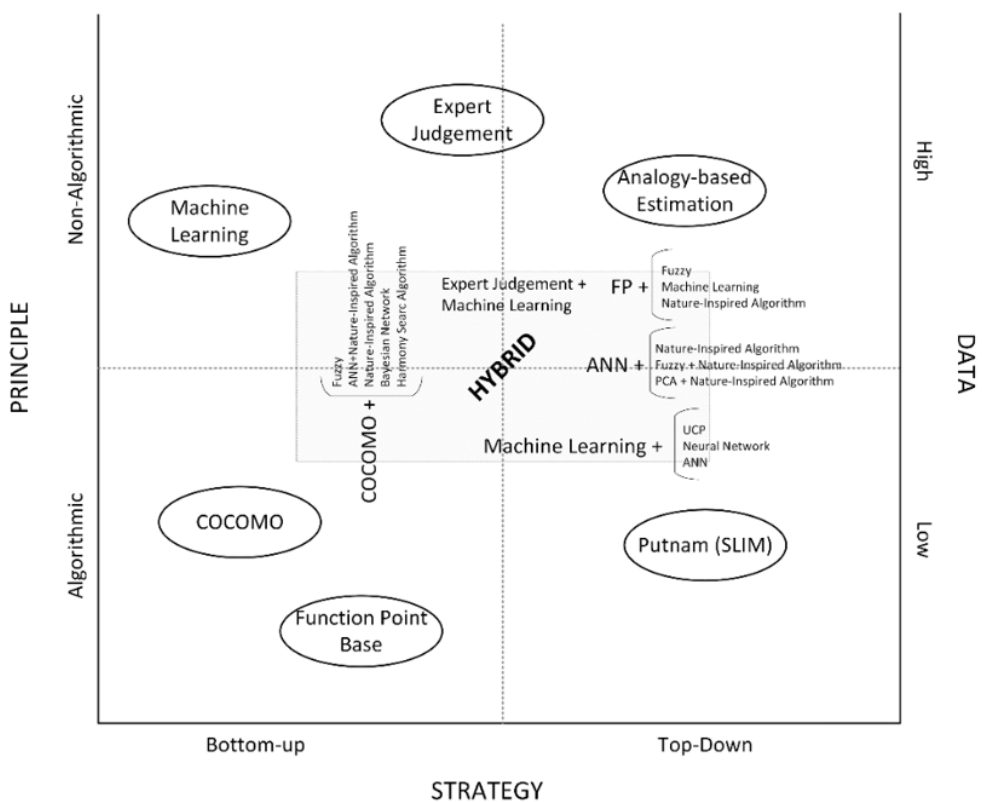


Figure 2. Software Effort Estimation Classification

In the development, the problem in estimation is better resolved when using a combination of several elements, this model is called hybrid. A hybrid model can combine components in a different aspect, and also combine attributes in the aspect of itself. There have been no perfect single estimation methods, therefore it is more suggested to use multiple methods. The confluence amongst the estimates generated by distinct methods shows that the estimation is most likely accurate. In fact, the discrepancy between the

estimates indicate that the possibility has been the neglect of certain factors. Therefore, it is essential to determine the factors that cause the difference, and then reexamine to converge result to preferable estimates (Bourque & Fairley, 2014).

MATERIALS AND METHODS

The method of conducting the literature review was based on Kitchenham & Charters (2007) that comprised three main phases, which include planning, conducting, and reporting. Therefore, the development of the systematic literature review protocol in this study consisted of planning and conducting, as shown in Figure 3.

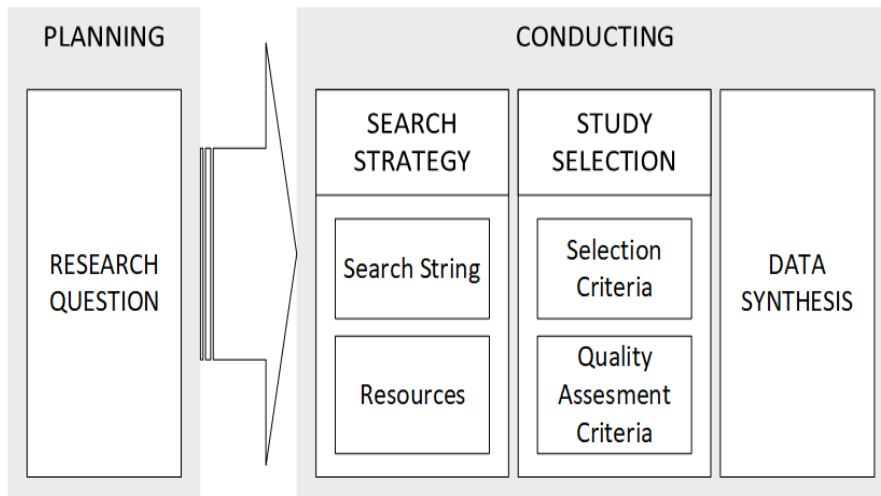


Figure 3. Systematic Literature Review Protocol

Planning

To achieve this research aim, the software effort estimation that is used in Agile project development needs to be known in advance. Next up is knowing how the estimation is implemented and what variables are involved.

The planning stage is the formulation of research questions related to Agile effort estimation as follows:

RQ#1: What kind of method is used for effort estimation in Agile?

RQ#2: How it works to implements estimation effort in Agile?

RQ#3: What are the attributes involved in the estimation of effort in Agile?

Conducting

The conducting phase comprises of three activities, namely search strategy, study selection,

and data synthesis. In contrast to data synthesis that is independent, the search strategy and study selection are each formed by two different activities.

On search strategy, a term that is used is known as search strings. This phase also determines resources, where the search will be executed. This process can be conducted after search strings and resources are identified. The strings used in the process is “Agile” (“effort” OR “cost”) and “Estimation” (“technique” OR “model”).

The process was conducted by embedding the search strings on seven resources as shown in Table 2.

Table 2
List of Resources

No.	Source Name	URL
1	IEEEExplore	www.ieeexplore
2	ACM Digital library	https://dl.acm.org/
3	Google Scholar	scholar.google.com
4	Inspec	https://digital-library.theiet.org/
5	ScienceDirect	www.sciencedirect.com
6	SpringerLink	www.springerlink
7	World Scientific	https://www.worldscientific.com

The study selection involved two activities. The first activity was filtering the result of the search process based on inclusion and exclusion criteria. The second activity was to implement quality assessment criteria on the first result.

The inclusion criteria consist of four aspects, which are publication year, the language used, the context of research assessed from title and abstract, and valid DOI. Meanwhile, the exclusion criteria is determined by two aspects, namely the type of work and kind of research. In this case, type of work may be in the form of a journal, conference proceeding, book, chapter, thesis, dissertation, or course material, while the kind of research consists of the study, literature review, comparative study, or survey. The inclusion and exclusion criteria are explained in Table 3.

After the first activity was performed, the next step was selection using quality assessment criteria. As seen in Table 4, the assessment criteria were derived from the research question. The relevant studies fulfilled the quality assessment criteria with an accumulated score value greater than 1.

Data synthesis is an activity to resume the selected studies evidence, in order to synchronize it with the research question. The primary purpose of this activity is to provide clear answers to the research questions.

Table 3
Inclusion and Exclusion Criteria

Inclusion Criteria	Exclusion Criteria
Publication Year: 2009 – 2019	Type of work: Book, Book Chapter, Thesis, Dissertation, course material
Language: English	Kind of Research: Literature Review, Comparative Study, Survey, Report
Title or Abstract: Contains search string	
DOI: valid DOI	

Table 4
Quality Assessment Criteria

ID	Question (Q)	Answer/ Score Points		
		Yes	Partially	No /
RQ#1	Does the study discuss the kind of technique that used to estimate effort in Agile?	1	0.5	0
RQ#2	Does the study explain the approach to implements the estimating effort methods in Agile?	1	0.5	0
RQ#3	Does the study declare attributes affecting the estimating of effort in Agile?	1	0.5	0

RESULTS AND DISCUSSION

Conducting Phase Result

The result of the conducting phase from seven resources obtained 171 articles. The number of articles for each resource is shown in Table 5. Furthermore, the top three articles are from Google Scholar (30.99%), SpringerLink (28.07%), and IEEEExplore (22.22%).

The inclusion criteria were implemented to obtain 118 articles and exclusion criteria to obtain 100. Table 6 shows the results of both.

Table 7 shows the number of articles of each resource generated from the quality assessment criteria.

Table 5
Result Conducting Phase

No.	Source Name	URL	Number of Articles
1	IEEEExplore	www.ieeexplore	38
2	ACM Digital library	https://dl.acm.org/	15
3	Google Scholar	scholar.google.com	53
4	Inspec	https://digital-library.theiet.org/	1
5	ScienceDirect	www.sciencedirect.com	14
6	SpringerLink	www.springerlink	48
7	World Scientific	https://www.worldscientific.com	2
TOTAL			171

Table 6
Result of Inclusion and Exclusion Criteria

No.	Source Name	URL	Number of Articles	
			Inclusion Criteria	Exclusion Criteria
1	IEEEExplore	www.ieeexplore	37	33
2	ACM Digital library	https://dl.acm.org/	15	15
3	Google Scholar	scholar.google.com	11	5
4	Inspec	https://digital-library.theiet.org/	1	0
5	ScienceDirect	www.sciencedirect.com	13	12
6	SpringerLink	www.springerlink	40	34
7	World Scientific	https://www.worldscientific.com	1	1
TOTAL			118	110

Table 7
Result of Quality Assessment Criteria

No.	Source Name	URL	Number of Articles
1	IEEEExplore	www.ieeeexplore	16
2	ACM Digital library	https://dl.acm.org/	6
3	Google Scholar	scholar.google.com	4
4	Inspec	https://digital-library.theiet.org/	0
5	ScienceDirect	www.sciencedirect.com	5
6	SpringerLink	www.springerlink	7
7	World Scientific	https://www.worldscientific.com	0
TOTAL			38

Data Synthesis Phase Results

Effort Estimation Method that Implemented in Agile. According to the relevant articles, effort estimation method that is implemented in Agile can be classified into four, which are Expert Judgement (EJ), Algorithmic (A), Machine Learning (ML), and Statistic (St).

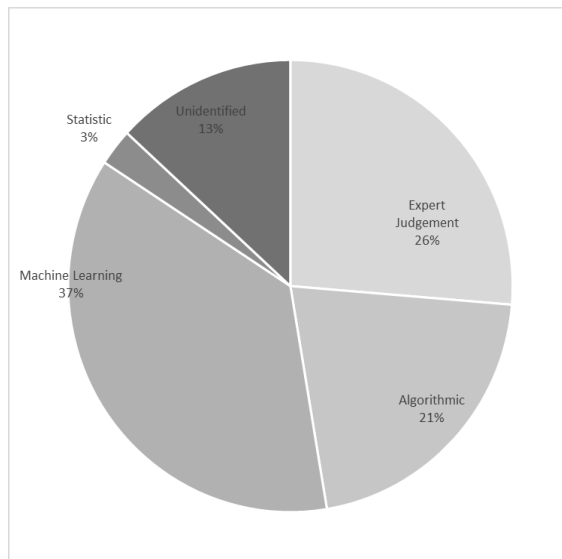


Figure 4. Distribution of Software Effort Estimation Method in Agile

Figure 4 shows the distribution of effort estimation method applied in Agile. From the relevant articles, most of them (87%) clearly stated the method that was used, and the rest (13%) did not describe the method. The top three estimation methods were Machine Learning (37%), Expert Judgement (26%), and Algorithmic (21%).

Expert Judgement Method. All the Expert Judgement method implemented in Agile is based on Planning Poker. Planning Poker is a group of estimation technique recommended for agile software development. The group discussion in this technique is assumed to provide higher accuracy and reduce excessive optimism which is the characteristic of expert judgment-based methods, although there is little empirical evidence about it. Mahnič and Hovelja (2012) conducted a research to fill the gaps by comparing the estimation effort based on the same user stories of the beginner groups and expert groups. The comparing result showed that the people involved in the group estimation process had the expertise, therefore the optimism bias arising from group discussion was diminished.

Lenarduzzi et al. (2015) stated that the estimation efforts conducted by the developers had higher accuracy than those achieved through measuring the functional size. Besides, estimation accuracy in Scrum cannot improve with the help of SiFP and IFPUG Function Points because of weak predictive power.

The two research results above showed that the expertise of professionals in software development companies is the primary aspect that influences the accuracy of estimation through Planning Poker. Furthermore, López-Martínez et al. (2017a) strengthened the statement by conducting a research that proved the existence of dissent between Scrum experts and beginners on the following factors, namely experience, time, effort, priority, and user stories value. From the five factors, experience and time have differences in opinions, while effort and priority have similar opinions. For the user stories value factor, many people doubted considering it as necessary, and some did not apply this factor in the planning phase. In line with this, the industry is likely not to consider this factor.

Although former research showed the existence of different opinions between Scrum experts and beginners, Chatzipetrou et al. (2018) showed that beginners were more engaged with the Planning Poker than estimates. During the Planning Poker activity, the beginners are more interested and involved. This showed that estimation with Planning Poker is a fun activity.

The user story is measured with relative value known as Story points (SP) that are commonly used as the base of calculation in the Planning Poker. Furthermore, Zahraoui and Idrissi (2015) improved the accuracy of effort and time estimates by proposing Adjusted Story Point (ASP) measure instead of SP to calculate the total effort of a scrum project. ASP used three adjustment factors, which are Priority Factor (PF), Story Size Factor (SF), and Complexity Factor (CF). Unfortunately, the use of ASP is not yet applicable to reel

scrum projects, therefore, further research is needed to improve the proper values of Story Point Adjustment Factors (SPF).

Although Planning Poker has a lot of benefits, the result was relying on observation experts. To overcome these issues, López-Martínez et al. (2017b) proposed a new model to improve Planning Poker by estimating the complexity and importance of user stories using Bayesian Network. Even though it showed good results primarily to facilitate newbie developers in deciding when to estimate user stories, this model still needed expert knowledge to build a Bayesian Network.

López-Martínez et al. (2018) validated the model that was built on previous studies using the Bayesian network that considered user stories based on its complexity and the level of importance. The validation result showed that effort estimation from professionals had a higher degree of correlation than beginner's estimates. This indicated that factors in the real-world application were also considered in the proposed model. Despite showing promising results, the model needs to be tested on a whole real project. Besides, it needs a mobile application development to ease its implementation.

Tanveer et al. (2017a) proposed an innovative hybrid method that incorporated expert knowledge and changes analysis to impact the process of estimation. This hybrid method was furnished by Tanveer et al. (2018) with a prototype tool built based on the previous framework. That hybrid method can improve the effectiveness of effort estimates.

Moharreri et al. (2016) complemented manual PP with an automated estimation through extracted Agile story cards and their actual effort time. The Auto-Estimate was developed by comparing alternate methods like Naive Bayes, Random Forest, J48, and Logistic Model Tree (LMT), whereas the better result was from the combination of J48+PP, J48, and LMT+PP.

Algorithmic Method. The algorithmic effort estimation method in Agile software development consists of the COSMIC, phase-wise algorithm, Function Point (FP), and Use Case Point (UCP). This part explains the relevant articles that implement the algorithmic estimation method.

COSMIC was proposed by Desharnais et al. (2011) to improve the guess estimate in Agile Project Management (APM). A new procedure was proposed that was built by consolidating COSMIC measurement method at the micro-level (User Stories) and the quality of the documentation for functional analysis deployment. Their study showed that this approach could help the planner to clearly understand why the global effort changed by the time.

The Phase-wise algorithm to compute the estimation effort is proposed by Choudhari and Suman (2012a; 2012b) which conducted an empirical study through a questionnaire to determine efforts of maintenance to compute phase-wise effort estimation. Their postulate provided more realistic results and worthwhile in estimating maintenance effort, especially

in extreme programming based on maintenance environments. Therefore, refinement is still needed on the proposed technique.

Function Point (FP) is a part of algorithmic effort estimation classification use by Kang et al. (2010) as an addition to the story point for Agile projects systematic estimation. FP is used to minimize fluctuations of value estimation that is caused by relative values from the user story. Based on a comparison with traditional methods, the addition story point with FP showed better performance.

Excellent performance of FP was applied by Silas et al. (2017) to improve cost estimation in Agile development by proposing the hybridization of Class Responsibility Collaborators models with FP. The process can be boosted through the hybridization adoption of Class Responsibility Collaborators models with FP.

The Use Case Point (UCP) is an effort estimation method that is included in the Algorithmic classification. Khatri et al. (2016) implemented UCP, as an estimation method for Agile development in early stage by emphasizing on main complexity factors like technical and environmental. In this context, the use of the UCP can adequately estimate effort and improve the accuracy based on environment and technical factors under agile development.

UCP can also collaborate with the Scrum Framework by connecting between User Stories on the Product Backlog and Use Case on UCP as performed by Yuliansyah et al. (2018). In their study, UCP adjusted User Stories on Product Backlog by adding the transactions attribute on a user story, and thereby transforming it to Functionality.

Popli and Chauhan (2013) proposed algorithmic estimation method by considering various factors to increase estimation accuracy of release date, cost, effort, and duration, especially for the project Scrum. The algorithm used Sprint-Point Based Framework for Agile and involved two factors, which are project and people-related factors that are proven to be effective and feasible. Although these factors show strong influence on the value of sprints point, in the future, other factors can be added that most affect the estimation.

Machine Learning Method. Most of the relevant articles have the objective of improving accuracy through machine learning as a crucial issue in Agile effort estimation. Some involve the application of machine learning to resolve the metric size of the user story, which is commonly used as the base of estimation in Agile.

To improve the accuracy of Agile effort estimation, a combination of Neural Network was proposed by Panda et al. (2015a; 2015b) and Rao et al. (2018). While Malgonde and Chari (2018) applied predictive algorithms with ensemble-based approaches that consisted of Support Vector Machine (SVM), Artificial Neural Networks (ANN), K-Nearest Neighbors (KNN), and Decision Trees (DT).

Beside Neural Network, the accuracy of Agile effort estimation is also improved using metaheuristic algorithms. Khuat and Le (2017) devised a hybrid model to enhance

accuracy in Agile estimation by applying Particle swarm optimization (PSO) and artificial bee colony (ABC). Meanwhile, a novel formula for Agile software effort estimation based on velocity and story points is built from two metaheuristic algorithms mentioned before. However, this hybrid model needs further research on its implementation.

The accuracy in Agile software estimation is also improved using the ontology model. Adnan and Afzal (2017) proposed the model to build a knowledge base by saving significant tacit knowledge during project development. Various agents of the estimation system access the existing knowledge base and autonomously perform a suitable estimate for the success of future projects.

Commonly, Planning Poker (PP) is one of the effort estimations in Agile, and it is based on a user story that is measured by relative values. Some studies overcome this problem by proposing models that extracted keywords from user stories automatically.

Abrahamsson et al. (2011) proposed an Agile effort prediction model that was based on user stories using keyword extraction tools. Meanwhile, Kowalska and Ochodek (2014) proposed a new approach through Semantic Web technologies. However, Choetkiertikul et al. (2018) introduced a novel combination of two robust deep learning architectures, which were long short-term memory and recurrent highway network to estimate story points.

Chongpakdee and Vatanawood (2017) applied document fingerprints to retrieve the similar issues from the public repository of project management assets. In addition to the extraction of the keyword on a user story, Dragicevic et al. (2017) proposed to implement the Bayesian Network model.

To handle frequent requirement changes in Agile software development, two studies had been conducted to resolve this issue. Bilgaiyan et al. (2018) applied Artificial neural networks (ANNs) in Agile effort estimation. Furthermore, the ANN-feedforward back-propagation neural network and Elman neural network are used to keep track, maintain, and estimate the whole product. Meanwhile, Soares (2018) proposed to embed autoencoders in automatic software development effort estimation as text classification.

Even though all those studies provided many advantages, the models still need more massive datasets and other features because user stories are not only written in English, and sometimes influenced by developers' demographics, story criticality, and other systems and framework aspects.

Improved Agile software estimation was conducted by proposing a model based on support vector regression (SVR) that is optimized by the grid search method (GS). In fact, empirical evaluation through the leave-one-out cross-validation method against 21 historical Agile projects demonstrated that this model increases the performance of the SVR technique (Zakrani et al., 2018).

Statistic Method. To make the effort estimation method from the conventional life cycle in accordance with Agile development, Garg and Gupta (2015) proposed a new model

by implementing Principal Component Analysis (PCA) to ascertain the key attributes of the development cost. Their study found that the proposed methodology was suitable for Agile projects.

After knowing software effort estimation methods and techniques applied in Agile, the next step is associating software effort estimation techniques with the agile characteristics.

Table 8 shows Planning Poker is a software effort estimation technique that fulfillment the Agile characteristics. Although the Planning Poker suitable for the Agile characteristics, there needs to be further developed to overcome some weaknesses such as a high level of reliance on experts and estimation based mostly on guesses or experience.

Table 8

Mapping software effort estimation method, technique, and characteristic of Agile

Method and Technique	Agile Characteristic				
	C1	C2	C3	C4	C5
EXPERT JUDGEMENT					
Planning Poker	✓	✓	✓	✓	✓
ALGORITHMIC					
Phase wise	✓		✓		
COSMIC	✓		✓		
FP	✓		✓		
UCP	✓		✓		
MACHINE LEARNING					
Document Fingerprints					
Bayesian Network					
Classification					
Ontology					
Neural Network Families					
Fuzzy Families					
Metaheuristic Algorithm (ABC & PSO)			✓		✓
STATISTIC					
PCA					
Impact Analysis					

Note:

C1: Involving users actively, C2: Empower the team to make decisions, C3: capture requirements in lightweight and visual, C4: focus on frequent delivery by developing small product, C5: incremental and iterate

Approach for Estimating Effort in Agile. The implementation of software effort estimation could be achieved through two approaches, which are Non-hybrid and Hybrid. The non-hybrid approach applies a single effort estimation method, whereas the hybrid implements a combination of several methods. Half of the relevant articles used the non-hybrid approach, 36.84% implemented the hybrid, while the rest did not mention the used approach.

The non-hybrid approach consists of effort estimation methods such as Planning Poker, Phase Wise, COSMIC, Function Points, Use Case Point, Document Fingerprints, Bayesian Network, Text Classification, Ontology Model, and Principal Component Analysis. Meanwhile, the Hybrid approach contains method combinations between Expert Judgement and Statistic, Expert Judgement and Machine Learning, and Machine Learning and others. Most of the hybrid approach is a mix of machine learning techniques. Table 9 showed the methods and approach applied in Agile.

Attributes that Affect Estimating Effort in Agile. Volatility and change of customer requirement in Agile software development (ASD) is a difficult and challenging task in estimation effort. Generally, estimation in ASD is mainly based on user story (US) that measures by story points (Zahraoui & Idrissi, 2015). The US is commonly estimated using group processes that improve accuracy compared to individual process (Moløkken-Østvold & Jørgensen, 2004).

Most relevant articles (84.21%) explained the attributes used on effort estimation, and 15.79% did not specifically mention the attributes. Based on the relevant articles, effort in Agile Software Development can be measured by user story, Use Case with sizing method story points, Use Case and function points. Nevertheless, the attribute of being used in the effort estimation can vary significantly.

According to Table 10, there are twenty-four attributes used in effort estimation. The Agile estimation classification that is most employed is Statistic, followed by Expert Judgment and Machine Learning. These attributes are grouped by three criteria, namely the frequency of use, implementation on the effort classification, the recent frequency of use (last three years). Figure 5 shows the mapping of attributes based on those criteria.

The top five highest frequency attributes are Complexity, Story Points, Experience, Size, User Story, Effort, and Time. The implementation of attributes on the Agile effort classification showed that Complexity is implemented on all estimation classification. Whereas, Experience, Function Point, Size, Task, Time, User Story, Weight are attributes applied to three estimation classification. In the last three years, the most attribute that was used in Agile is Complexity, followed by Experience, Size, Effort, and Time. Also, the attributes that fulfilled all criteria are Complexity, Experience, Size, and Time.

Table 9
Types of software effort estimation method and approach

Approach	Method	Technique	Author	
Non-hybrid	E J	Planning Poker	(Chatzipetrou et al., 2018), (López-Martínez et al., 2017a), (Lenarduzzi et al., 2015), (Mahnič & Hovelja, 2012), (Zahraoui & Idrissi, 2015)	
		Phase wise	(Choudhari & Suman, 2012a), (Choudhari & Suman, 2012b)	
	A	COSMIC	(Desharnais et al., 2011)	
		FP	(Kang et al., 2010)	
		UCP	(Yuliansyah et al., 2018)	
	M L	Document Fingerprints	(Chongpakdee & Vatanawood, 2017)	
		Bayesian Network Classification	(Dragicevic et al., 2017) (Soares, 2018)	
		Ontology	(Adnan & Afzal, 2017), (Kowalska & Ochodek, 2014)	
	E J E J	S t S t	PCA	(Garg & Gupta, 2015)
			EJ and Impact Analysis	(Tanveer et al., 2017b), (Tanveer et al., 2018)
M L		Planning Poker and Machine learning (J48, LMT, Bayesian Network)	(Moharreri et al., 2016), (López-Martínez et al., 2017b), (López-Martínez et al., 2018)	
		ABC and PSO	(Khuat & Le, 2017)	
		Adaptive Neuro-Fuzzy Modelling, Generalized Regression Neural Network and Radial Basis Function Networks (RBFNs)	(Rao et al., 2018)	
M L & M L	Combination of Neural Networks	(Bilgaiyan et al., 2018), (Panda et al., 2015a), (Panda et al., 2015b)		
	Combination two deep learning	(Choetkiertikul et al., 2018)		
	Combination of Machine Learning such as SVM, SVR, ANN, KNN, and DT.	(Abrahamsson et al., 2011), (Malgonde & Chari, 2018), (Zakrani et al., 2018)		

Table 10
Effort Estimation Attributes

No.	Attribute	SEE Method							Total						
		EJ	A	ML	St	2010	2011	2012		2013	2014	2015	2016	2017	2018
1	Application Type									1					1
2	Architecture									1					1
3	Complexity									3	1	3	5	2	14
4	Dependencies										1		1	1	3
5	Effort									1	2		2	2	7
6	Environmental													1	1
7	Experience										2	2	4	2	8
8	Function Point								1		2				3
9	Impact										1		1		2
10	Knowledge										1		1		2
11	Maturity										3				3
12	Platform										3				3
13	Previously Estimates										1		1		2
14	Priority												1	2	3
15	Size										1		1	5	8
16	Skill												1		1
17	Story Points												1	2	10
18	Sprint Points											5		1	4
19	Task											1		2	3

Table 10 (Continued)

No.	Attribute	SEE Method										Frequency of use				
		EJ	A	ML	St	2010	2011	2012	2013	2014	2015	2016	2017	2018	Total	
20	Time	█	█	█	█		1		1					3	1	6
21	Use Case Point	█	█	█									1		1	2
22	User Story	█	█	█	█		1		3		1		1	1		8
23	Velocity	█	█	█	█						1		2	1	1	5
24	Weight	█	█	█	█	█									2	3

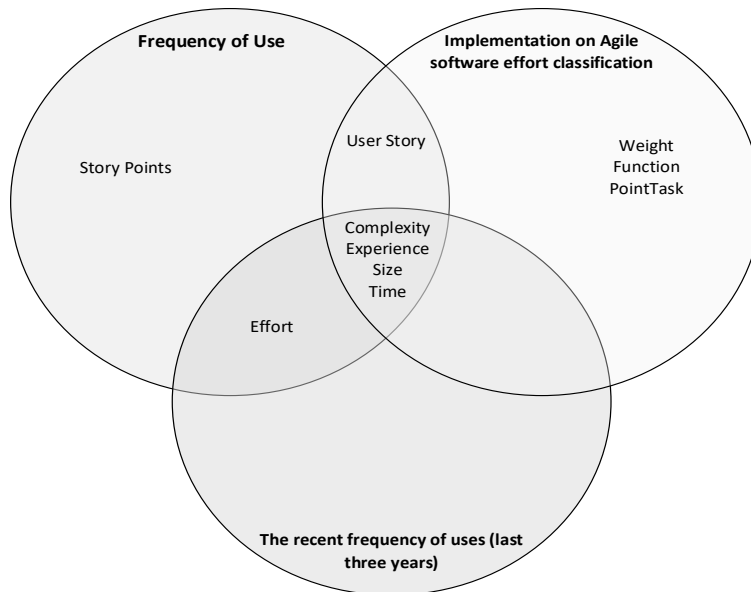


Figure 5. Mapping of Agile software effort estimation attributes

Complexity attribute is interpreted in different aspects, and most study understands it as the complexity of the project (Popli & Chauhan, 2014b; Garg & Gupta, 2015; Tanveer et al., 2016; Tanveer et al., 2017b; Bilgaiyan et al., 2018). Some research used it to represent technical complexity (Hamouda, 2014; Khatri et al., 2016; Yuliansyah et al., 2018). Also, environmental complexity is a part of the attribute that applies in two research (Hamouda, 2014; Khatri et al., 2016). In fact, the attribute is considered to represent form, function, report, and requirements complexities (Dragicevic et al., 2017).

Many studies used story point attribute that represent itself (Popli & Chauhan, 2014a; Hamouda, 2014; Panda et al., 2015a; Panda et al., 2015b; Khuat & Le, 2017; Zakrani et al., 2018; Rao et al., 2018). Meanwhile, others apply this to show baseline story-points, estimated story-points (ESP), and unadjusted value of story-points (Popli & Chauhan, 2014b).

Experience attribute in most studies is used to measure the implementation experience of the developer or programmer (Tanveer et al., 2016; Tanveer et al., 2017b; López-Martínez et al., 2017a; López-Martínez et al., 2017b; López-Martínez et al., 2018; Malgonde & Chari, 2018). Other studies assess this attribute based on developer's experience in making estimation (Tanveer et al., 2016; Tanveer et al., 2017b).

Some studies apply size attribute that represents the value of the user story (López-Martínez et al., 2017b; López-Martínez et al., 2018), the code metrics for each affected class (Tanveer et al., 2018), and the team size (Garg & Gupta, 2015). However, other studies do

not detail the size attribute (Kang et al., 2010; Bilgaiyan et al., 2018; Malgonde & Chari, 2018; Tanveer et al., 2018).

Effort estimation in Agile software development is mainly based on the user story attribute (Choudhari & Suman, 2012a; Choudhari & Suman, 2012b; Mahnič & Hovelja, 2012; Popli & Chauhan, 2014a; Zahraoui & Idrissi, 2015; Chongpakdee & Vatanawood, 2017; Choetkiertikul et al., 2018). In Abrahamsson et al. (2011), the user story was elaborated into keywords, namely length, and priority.

Effort can be assumed as effort per person (Popli & Chauhan, 2014a) or actual effort (Panda et al., 2015a; Panda et al., 2015b). However, the attribute also represents itself without supplement (López-Martínez et al., 2017a; López-Martínez et al., 2017b; López-Martínez et al., 2018; Malgonde & Chari, 2018).

Commonly, the attribute of time refers to the time it takes to complete a project (Popli & Chauhan, 2013; López-Martínez et al., 2017a; López-Martínez et al., 2017b; López-Martínez et al., 2018). In some studies, it is explicitly explained as Person-hours (Desharnais et al., 2011) or Working Hours (Dragicevic et al., 2017).

CONCLUSION

Machine learning is the most common effort estimation method used in Agile software development, followed by Expert judgment and Algorithmic. However, machine learning has limitation in its implementation because it needs very large dataset that source from expert's knowledge.

The implementation approach of the estimation effort in Agile software development does not differ between the non-hybrid and hybrid. The non-hybrid approach is applied on 50.00% of relevant articles and the hybrid is applied on 36.84%.

Twenty-four attributes are involved in Agile effort estimation, and those with the top five highest frequency of use are Complexity, Story Points, Experience, Size, User Story, Effort, and Time. Furthermore, the attribute that is implemented on all classification of the estimation is Complexity. Those with the highest frequency of use in the last three years are Complexity, Experience, Size, Effort, and Time. In addition, those that fulfilled all criteria are Complexity, Experience, Size, and Time.

ACKNOWLEDGEMENT

Pantjawati Sudarmaningtyas would like to thank her supervisor and her institution for providing support to this research, especially for Universitas Dinamika, which had provided financial support. Pantjawati Sudarmaningtyas is also grateful to colleagues and peer reviewers who saved her from many mistakes through comments, feedback, and constructive criticism. Pantjawati Sudarmaningtyas hope this research can give benefits and contribute to further research.

REFERENCES

- Abrahamsson, P., Fronza, I., Moser, R., Vlasenko, J., & Pedrycz, W. (2011, September 22-23). Predicting development effort from user stories. In *2011 International Symposium on Empirical Software Engineering and Measurement* (pp. 400-403). Banff, Canada. <https://doi.org/10.1109/ESEM.2011.58>
- Adnan, M., & Afzal, M. (2017). Ontology based multiagent effort estimation system for scrum agile method. *IEEE Access*, *5*, 25993-26005. <https://doi.org/10.1109/ACCESS.2017.2771257>
- Bilgaiyan, S., Mishra, S., & Das, M. (2018). Effort estimation in agile software development using experimental validation of neural network models. *International Journal of Information Technology*, *11*, 569-573. <https://doi.org/10.1007/s41870-018-0131-2>
- Bloch, M., Blumberg, S., & Laartz, J. (2012). *Delivering large scale IT.pdf*. Retrieved October 23, 2017, from <https://www.mckinsey.com/business-functions/digital-mckinsey/our-insights/delivering-large-scale-it-projects-on-time-on-budget-and-on-value>
- Boehm, B. W. (1984). Software engineering economics. *IEEE Transactions on Software Engineering*, *SE-10*(1), 4-21. <https://doi.org/10.1109/TSE.1984.5010193>
- Bourque, P., & Fairley, R. E. (2014). *Guide to the software engineering body of knowledge*. IEEE Computer Society Press. <https://doi.org/10.1234/12345678>
- Chatzipetrou, P., Ouriques, R., & Gonzalez-Huerta, J. (2018). Approaching the relative estimation concept with planning poker. In *Proceedings of the 7th Computer Science Education Research Conference (CSERC '18)* (pp. 21-25). Association for Computing Machinery. <https://doi.org/10.1145/3289406.3289409>
- Choetkiertikul, M., Dam, H. K., Tran, T., Pham, T., Ghose, A., & Menzies, T. (2018). A deep learning model for estimating story points. *IEEE Transactions on Software Engineering*, *45*(7), 637-656.
- Chongpakdee, P., & Vatanawood, W. (2017, November 24-26). Estimating user story points using document fingerprints. In *2017 8th IEEE International Conference on Software Engineering and Service Science (ICSESS)*, (pp. 149-152). Beijing, China. <https://doi.org/10.1109/ICSESS.2017.8342885>
- Choudhari, J., & Suman, U. (2012a). Phase wise effort estimation for software maintenance: An extended SMEEM model. In *Proceedings of the CUBE International Information Technology Conference* (pp. 397-402). Association for Computing Machinery. <https://doi.org/10.1145/2381716.2381790>
- Choudhari, J., & Suman, U. (2012b). Story Points based effort estimation model for software maintenance. *Procedia Technology*, *4*, 761-765. <https://doi.org/10.1016/j.protcy.2012.05.124>
- Desharnais, J. M., Buglione, L., & Kocatürk, B. (2011). Using the COSMIC method to estimate agile user stories. In *Proceedings of the 12th International Conference on Product Focused Software Development and Process Improvement* (pp. 68-73). Association for Computing Machinery. <https://doi.org/10.1145/2181101.2181117>
- Dragicevic, S., Celar, S., & Turic, M. (2017). Bayesian network model for task effort estimation in agile software development. *Journal of Systems and Software*, *127*, 109-119. <https://doi.org/10.1016/j.jss.2017.01.027>
- Garg, S., & Gupta, D. (2015, March 3-5). PCA based cost estimation model for agile software development projects. In *2015 International Conference on Industrial Engineering and Operations Management (IEOM)* (pp. 1-7). Dubai, United Arab Emirates. <https://doi.org/10.1109/IEOM.2015.7228109>

- Hamouda, A. E. D. (2014, July 28-August 1). Using agile story points as an estimation technique in CMMI organizations. In *2014 Agile Conference* (pp. 16-23). Kissimmee, FL, USA. <https://doi.org/10.1109/AGILE.2014.11>
- Institute, P. M. (2017). *A guide to the project management body of knowledge (PMBOK® Guide)-Sixth edition*. Project Management Institute, Inc.
- Jørgensen, M. (2016). The use of precision of software development effort estimates to communicate uncertainty. In *8th International Conference on Software Quality Days (SWQD)* (pp. 156-168). Springer. https://doi.org/10.1007/978-3-319-27033-3_11
- Kang, S., Choi, O., & Baik, J. (2010, August 18-20). Model-based dynamic cost estimation and tracking method for agile software development. In *2010 IEEE/ACIS 9th International Conference on Computer and Information Science* (pp. 743-748). Yamagata, Japan. <https://doi.org/10.1109/ICIS.2010.126>
- Khatri, S. K., Malhotra, S., & Johri, P. (2016, September 7-9). Use case point estimation technique in software development. In *2016 5th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)* (pp. 123-128). Noida, India. <https://doi.org/10.1109/ICRITO.2016.7784938>
- Khuat, T. T., & Le, M. H. (2017). A novel hybrid ABC-PSO algorithm for effort estimation of software projects using agile methodologies. *Journal of Intelligent Systems*, 27(3), 489-506. <https://doi.org/10.1515/jisys-2016-0294>
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering* (Version 2.3). Department of Computer Science, Keele University. <https://doi.org/10.1145/1134285.1134500>
- Kowalska, J., & Ochodek, M. (2014). Supporting analogy-based effort estimation with the use of ontologies. *E-Informatica Software Engineering Journal*, 8(1), 53-64. <https://doi.org/10.5277/e-Inf140104>
- López-Martínez, J., Ramírez-Noriega, A., Juárez-Ramírez, R., Licea, G., & Martínez-Ramírez, Y. (2017a). Analysis of planning poker factors between university and enterprise. In *2017 5th International Conference in Software Engineering Research and Innovation (CONISOFT)* (pp. 54-60). Conference Publishing Services. <https://doi.org/10.1109/CONISOFT.2017.00014>
- López-Martínez, J., Juárez-Ramírez, R., Ramírez-Noriega, A., Licea, G., & Navarro-Almanza, R. (2017b). Estimating user stories' complexity and importance in scrum with Bayesian networks. In *World Conference on Information Systems and Technologies* (pp. 205-214). Springer. https://doi.org/10.1007/978-3-319-56535-4_21
- López-Martínez, J., Ramírez-Noriega, A., Juárez-Ramírez, R., Licea, G., & Jiménez, S. (2018). User stories complexity estimation using Bayesian networks for inexperienced developers. *Cluster Computing*, 21(1), 715-728. <https://doi.org/10.1007/s10586-017-0996-z>
- Lenarduzzi, V., Lunesu, I., Matta, M., & Taibi, D. (2015). Functional size measures and effort estimation in agile development: A replicated study. In *International Conference on Agile Software Development* (pp. 105-116). Springer. https://doi.org/10.1007/978-3-319-18612-2_9
- Mahnič, V., & Hovelja, T. (2012). On using planning poker for estimating user stories. *Journal of Systems and Software*, 85(9), 2086-2095. <https://doi.org/10.1016/j.jss.2012.04.005>

- Malgonde, O., & Chari, K. (2018). An ensemble-based model for predicting agile software development effort. *Empirical Software Engineering*, 24, 1017-1055. <https://doi.org/10.1007/s10664-018-9647-0>
- Mersino, A. (2018). *Agile project success rates are 2x higher than traditional projects (2019)*. Retrieved January 27, 2020, from <https://vitalitychicago.com/blog/agile-projects-are-more-successful-traditional-projects/>
- Moharreri, K., Sapre, A. V., Ramanathan, J., & Ramnath, R. (2016). Cost-effective supervised learning models for software effort estimation in agile environments. In *2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC)* (pp. 135-140). Conference Publishing Services. <https://doi.org/10.1109/COMPSAC.2016.85>
- Moløkken-Østvold, K., & Jørgensen, M. (2004). Group processes in software effort estimation. *Empirical Software Engineering*, 9(4), 315-334. <https://doi.org/10.1023/B:EMSE.0000039882.39206.5a>
- Panda, A., Satapathy, S. M., & Rath, S. K. (2015a). Empirical validation of neural network models for agile software effort estimation based on story points. *Procedia Computer Science*, 57, 772-781. <https://doi.org/https://doi.org/10.1016/j.procs.2015.07.474>
- Panda, A., Satapathy, S. M., & Rath, S. K. (2015b). Neural network models for agile software effort estimation based on story points. *Proceedings of the International Conference on Advances in Computing, Control and Networking*, 57(57), 26-30. <https://doi.org/10.15224/978-1-63248-038-5-06>
- Popli, R., & Chauhan, N. (2013, March 9-10). A sprint-point based estimation technique in Scrum. In *2013 International Conference on Information Systems and Computer Networks* (pp. 98-103). Mathura, India. <https://doi.org/10.1109/ICISCON.2013.6524182>
- Popli, R., & Chauhan, N. (2014a, February 6-8). Cost and effort estimation in agile software development. In *2014 International Conference on Reliability Optimization and Information Technology (ICROIT)* (pp. 57-61). Faridabad, India. <https://doi.org/10.1109/ICROIT.2014.6798284>
- Popli, R., & Chauhan, N. (2014b, March 1-2). Estimation in agile environment using resistance factors. In *2014 International Conference on Information Systems and Computer Networks (ISCON)* (pp. 60-65). Mathura, India. <https://doi.org/10.1109/ICISCON.2014.6965219>
- Project-Management.com. (2019). *10 key principles of agile software development*. Retrieved January 11, 2020, from Project Management.com website: <https://project-management.com/10-key-principles-of-agile-software-development/>
- Rao, C. P., Kumar, P. S., Sree, S. R., & Devi, J. (2018). An agile effort estimation based on story points using machine learning techniques. In *Proceedings of the Second International Conference on Computational Intelligence and Informatics* (pp. 209-219). Springer. https://doi.org/10.1007/978-981-10-8228-3_20
- Silas, F. A., Yusuf, M., & Bijik, A. H. (2017). Hybridization of class responsibility collaborators model (HRCRM) with function point to enhance project estimation cost in agile software development. *Circulation in Computer Science*, 2(6), 20-24. <https://doi.org/10.22632/ccs-2017-252-32>
- Soares, R. G. F. (2018, July 8-13). Effort estimation via text classification and autoencoders. In *2018 International Joint Conference on Neural Networks (IJCNN)* (pp. 01-08). Rio de Janeiro, Brazil. <https://doi.org/10.1109/IJCNN.2018.8489030>

- Srivastava, B., & Wadhwa, M. (2013). Relative analysis of software cost and effort estimation techniques. *International Journal of Computer Science and Engineering (IJCSE)*, 2(3), 53-68.
- Tanveer, B., Vollmer, A. M., & Engel, U. M. (2017a). Utilizing change impact analysis for effort estimation in agile development. In *2017 43rd Euromicro Conference on Software Engineering and Advanced Applications (SEAA)* (pp. 430-434). Conference Publishing Services. <https://doi.org/10.1109/SEAA.2017.64>
- Tanveer, B., Guzman, L., & Engel, U. M. (2017b). Effort estimation in agile software development: Case study and improvement framework. *Journal of Software-Evolution and Process*, 29(11), 1-14. <https://doi.org/10.1002/smr.1862>
- Tanveer, B., Guzmán, L., & Engel, U. M. (2016). Understanding and improving effort estimation in agile software development. In *Proceedings of the International Workshop on Software and Systems Process - ICSSP '16* (pp. 41-50). Association for Computing Machinery. <https://doi.org/10.1145/2904354.2904373>
- Tanveer, B., Vollmer, A. M., & Braun, S. (2018). A hybrid methodology for effort estimation in agile development: An industrial evaluation. In *Proceedings of the 2018 International Conference on Software and System Process* (pp. 21-30). Association for Computing Machinery. <https://doi.org/10.1145/3202710.3203152>
- Trendowicz, A., & Jeffery, R. (2014). *Software project effort estimation: Foundation and best practice guidelines for success*. Springer. <https://doi.org/10.1007/978-3-319-03629-8>
- VersionOne.com. (2017). 11th Annual state of agile report. *VersionOne Agile Annual Report*, 1-16. <https://doi.org/10.1093/jicru/ndl025>
- Yuliansyah, H., Qudsiah, S. N., Zahrotun, L., & Arfiani, I. (2018). Implementation of use case point as software effort estimation in scrum framework. *IOP Conference Series: Materials Science and Engineering*, 403(1), 1-10. <https://doi.org/10.1088/1757-899X/403/1/012085>
- Zahraoui, H., & Idrissi, M. A. J. (2015, October 20-21). Adjusting story points calculation in scrum effort & time estimation. In *2015 10th International Conference on Intelligent Systems: Theories and Applications (SITA)* (pp. 1-8). Rabat, Morocco. <https://doi.org/10.1109/SITA.2015.7358400>
- Zakrani, A., Najm, A., & Marzak, A. (2018, October 21-27). Support vector regression based on grid-search method for agile software effort prediction. In *2018 IEEE 5th International Congress on Information Science and Technology (CiSt)* (pp. 1-6). Marrakech, Morocco. <https://doi.org/10.1109/CIST.2018.8596370>



Reducing Cognitive Impairment Among Dementia Users Through Mobile Application

Nur Atheera Mohd Hassan¹, Aslina Baharum^{1*}, Zaidatul Haslinda Abdullah Sani¹, Kent Chau² and Noorsidi Aizuddin Mat Noor³

¹Faculty of Computing and Informatics, Universiti Malaysia Sabah, 88400 UMS, Kota Kinabalu, Sabah, Malaysia

²Head of Comfort Aged Care Center, 377, Jalan Likas, 88450 Kota Kinabalu, Sabah, Malaysia

³Faculty of Built Environment and Surveying, Universiti Teknologi Malaysia, 81310 Skudai, Johor, Malaysia

ABSTRACT

Cognitive impairment includes the lacking ability to remember things, disorientation in remembering the current location, and the struggle to find the correct word. People with dementia (PwD) are often involved in this impairment. With that being said, this project proposes the use of a mobile application to help in improving their cognitive issues. To tackle this problem, features and functionality of a mobile application specifically for dementia users are identified which contributes to the development of a diary application. Identifying and gathering features from previous studies was the initial method. Development of the diary application followed the software development life cycle (SDLC) waterfall method and evaluation of the application was experimented with identified dementia users. The

findings of this project are the application set of guidelines gathered from literature into the diary application. Four verified dementia people were involved in the evaluation of the effectiveness of the application. The evaluation of the application includes some good points. Some parts of the application are pointed out for their unsuitable design and suggestions are given to improve the application in the later future.

ARTICLE INFO

Article history:

Received: 4 April 2020

Accepted: 27 July 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.09>

E-mail addresses:

mi1822022t@student.ums.edu.my (Nur Atheera Mohd Hassan)

aslina@ums.edu.my (Aslina Baharum)

linda.sani@ums.edu.my (Zaidatul Haslinda Abdullah Sani)

kjchau88@gmail.com (Kent Chau)

noorsidi@utm.my (Noorsidi Aizuddin Mat Noor)

*Corresponding author

Keywords: Cognitive, diary, early-stage, features, memory loss, technology

INTRODUCTION

The human-computer interaction (HCI) is described as the interaction between human and computers. HCI is considered as cognitive in nature as it heavily involves in problem solving, navigation as well as language processing. Now, it is undeniable that people interact with technologies and the internet every day. They utilize these for communication, information searching and entertainment.

In a world where computing is pretty much everywhere now, it is important that all users are catered for when designing an application with good features and functionality. Research suggests that people with dementia are keen to utilize latest technology even though they find it difficult to follow or understand the way to use any mobile application (Holsapple, 2005).

Dementia is a degenerative brain condition in which it is not a natural part of the aging process (Ancient & Good, 2013). People with dementia typically have problems with language impairments, memory loss and mood changes. According to the Alzheimer's Disease Foundation Malaysia (2020), there are about 127,000 people in Malaysia with dementia. This number is projected to be 261,000 by 2030 and will continue to increase to 590,000 people in 2050.

Memory decline is the most noticeable initial symptom of dementia and people with dementia struggle to remember recent events. If this memory degradation continues, their long-term memory would be affected (Lian et al., 2017). This further affects their cognitive ability. Cognitive function impairments are usually accompanying and every so often it can even precede the disease (Awada et al., 2018). Symptoms of dementia may vary, when mental functions such as the decline in memory and thinking, orientation in space and time, and emotional control are affected, they are suffering from dementia.

The world's population is continuously aging and people living with dementia mostly come from age groups that are above 60 years old. Considering how the estimated number of dementia patients might reach 82 million by 2030, it is unfortunate that there is a lack of understanding and awareness in terms of supporting them to living independently and coping with daily activities through technology (WHO, 2012).

Not only does current technology support people living with dementia to manage their daily activities, but it can also help in improving their life quality by enhancing their independence and self-confidence (Awada et al., 2018). There are numerous kinds of applications nowadays but only a few of these are specifically aimed at people with dementia. There are hardly any applications that have been designed with the viewing requirements of people with dementia in mind (Freeman et al., 2005). Designing interfaces for people with dementia is certainly not an easy task. The challenges lie in the designing of interfaces which can satisfactorily meet the requirements of people with dementia. Research has shown that people with dementia has nothing against wanting to adopt to new

technologies and learn new skills provided that the interfaces are designed well and are easily accessible for people living with dementia (Claire et al., 2000; Awada et al., 2018). Computer interfaces can and should be designed in a way that maximizes their accessibility and enables them to benefit from this.

Ancient and Good (2013) stated that designing interfaces should consider three possible areas where dementia would have an impact on. They are cognitive impairments, motor impairments and visual impairments. They have found that people with dementia struggle with understanding the pinch-zoom functionality of touch screen applications. This involves motor impairments where people with dementia may have slower movements. So, it is wise to take in the consideration of increasing the response time of people with dementia.

Gowans et al. (2007) proposed several criteria should be addressed to obtain a suitable interface for people with dementia. These criteria include support easy start-up and failure-free activity. Niklasson and Sandström (2016) listed some points needed to be considered when designing a user interface for people with dementia and they included avoiding displaying too much information in one page or screen, utilizing strong contrast between colors and using simple icons.

This research involves the development of a diary application for people with dementia which can encourage them to record diaries of their experiences, achievements, and challenges or just basically their daily life activities. The aim is to offer support to people with dementia in managing and coping with their daily activities.

The remaining parts of this paper are organized as follows. In the Methods section, the three phases of methods which are the research, development and evaluation phases are discussed in detail. The results and discussion section give the implementation outputs of the application and the outcome of application evaluation. The design guideline for the application is also summarized. Conclusion section concludes the research with summary of findings and future works.

RELATED WORKS

Dementia describes various symptoms of cognitive decline that hinder with normal life activities (Kilmova & Semradova, 2016). Cognitive function includes the ability in learning, reasoning, language, spatial ability, and orientation, and handling complex tasks (Group Health Cooperative, 2013). The decline in cognitive function can result in difficulty when using technology (Williams et al., 2013).

Niklasson and Sandström (2016) mentioned that cognitive impairments were one of the symptoms that were linked to our mind and intelligence. They further stated some cognitive symptom examples such as memory loss, orientation, speech, and concentration (Niklasson & Sandström, 2016).

Memory loss makes it difficult for PwD to recall previous events that happened. For instance, they would forget whether they have eaten breakfast earlier or not or they would not remember where they put certain things. It is common for PwD to lose their orientation, especially the ones with moderate and severe dementia. Orientation describes the orientation of both time and place. Finding places and keeping track of the time is quite an impossible task for those suffering from dementia. This is because, as the disease progresses further, time conception is affected and becomes distorted making it tough to even tell which part of the day it is day or night (Niklasson & Sandström, 2016). Speech ability still works alright in the early stages of dementia, but it deteriorates in later stages, therefore causing PwD to lose the ability to speak normally and keeping track of the line of conversation gets more difficult. Cognitive symptom also affects concentration, and this describes how PwD often lose their focus. As mentioned earlier, conversing on one topic with them is hard as they lose concentration.

There is a lack of study regarding the impact of cognitive impairment on their everyday life experience. However, there is one that focuses on the approach that people with dementia use to handle their difficulties. They use their sight, hearing, and touch senses to cope with cognitive decline, and they also try to count on old patterns and routines of everyday life (Johansson et al., 2015). It is suggested that they utilize this approach to obtain a sense of control over their lives more than to be as effective as before. The Table 1 shows the different stages of dementia and how each stage affects PwD in their daily lives (Soufneyestani et al., 2021).

Table 1
Effects of cognitive impairment to PwD daily activities

	Phase 1: Very Mild Dementia	Phase 2: Mild Dementia	Phase 3: Moderate Dementia	Phase 4: Severe Dementia
Cognitive impairment	-Forgets names -Slight memory problems -Difficulty in finding the right words	-Learning ability -Problem solving -Make decision -Finding the right words -Recalling memory	-Short-term memory -Orientation in time and location	-Language impairment -Often lose concentration
Activities of daily living (ADL) functions	-Misplaces familiar objects	-Driving -Grocery shopping	-Ability to handle money -Ability to cook -Ability to be independent	-Ability to get dressed -Ability to walk -Incontinence (lack voluntary control over urination and defecation)

As the quality of life for people in most developed countries increases due to health care and medicine being very much advanced nowadays, it does not guarantee that elderly people will not be affected by any illnesses that are common with their age (Coppola et al., 2013). Rather, they are more prone to get diseases that can decrease their cognitive ability. This makes them more dependent on being assisted by other people, mostly in this case, their family members.

Attempts at communication using technology can help lessen these symptoms their severity. As symptoms grow more severe, PwDs become less able to function on their own and would be more and more dependent on the care of those around them. After some time and as their disease further, worsen, they would have to be placed in nursing homes (Coppola et al., 2013). This is because in nursing homes, there would be staff and specialists around all the time and they are more experienced, thus family members would think that that is the best option for better care.

It is often that products make a notable request on cognitive abilities. However, technologies that are designed for elderly people, especially the ones that are living with dementia, need to take into consideration the number of demands being put. Measures should be taken to minimize the effect on a cognitive impaired (Czaja et al., 2012). Elder people often experience a decrease in their ability to note and remember things and for those living in a situation with dementia will have it worse because of the nature of the condition.

For people with dementia to use and interact with a new technology, the impairment needs to mitigate. This can be done by using cognitive cues or by lessening the memory load. One of the main symptoms of dementia is short term memory loss (Peterson et al., 2009) and that makes interacting with technologies difficult for people living with this condition. Difficulty in remembering will make navigation very challenging as they would find it hard to recall previous information (Etcheverry et al., 2012). Not only does current technology support people living with dementia to manage their daily activities, but it can also help in improving their life quality by enhancing their independence and self-confidence (Awada et al., 2018).

When it comes to designing interfaces, these two approaches should be taken into consideration; (1) a design that is universal which is accessible and manageable to those with disabilities, (2) an easy-to-use technology that is developed applying user-centered technique (Peterson et al., 2009). Ancient and Good (2013) thinks that the factors that should be regarded are accessibility, usability, and user experience. On the contrary, they further proposed that personalization and user acceptance are the main approach in designing interfaces for dementia patients. The concept of personalization refers to the modification of user interface to meet the requirements of the user. However, user acceptance refers to user experience and technology implementation. User acceptance is where the user's previous experience with technology is taken into serious consideration (Awada et al.,

2018). The decline in memory for people with dementia means that they are less capable in navigating through complicated design structures and instructions. Therefore, this should be given attention as to reduce memory load in designing by using, for instance simpler instructions and shorter messages.

Designing interfaces for people with dementia is certainly not an easy task. The challenges lie in the designing of interfaces which can satisfactorily meet the requirements of people with dementia. Research has shown that people with dementia has nothing against wanting to adopt to new technologies and learn new skills provided that the interfaces are designed well and are easily accessible for people living with dementia (Claire et al., 2000; Awada et al., 2018). Computer interfaces can and should be designed in a way that maximizes their accessibility and enables them to benefit from this. Peterson et al. (2009) proposed several criteria should be addressed to obtain a suitable interface for people with dementia. The set of criteria is briefed in the Table 2.

Table 2

Interface criteria suitable for people with dementia

Challenges	
1	Support easy start-up
2	Support intuitive navigation for a novice-user demographic
3	Elicit memories to prompt and support reminiscence, communication and social contact
4	Promote ‘non-immersive’ engagement
5	Support an enjoyable shared experience
6	Support the cognitively impaired person in having a more proactive and equitable role in shared conversations
7	Relieve caregivers of the pressure of constantly needing to ‘prop up’ conversations
8	Promote ‘failure-free’ activity
9	Support customization of content
10	Promote and support good practice in reminiscence intervention

Niklasson and Sandström (2016) had also stated in their study some points that needed to be considered when designing a user interface for PwD. Those points are listed in the Table 3.

Table 3

User interface suitable for people with dementia

	Challenges
1	Utilize a strong contrast between colors.
2	Support intuitive navigation for a novice-user demographic
3	Use simple icons.
4	Avoid displaying too much information in one page or screen.
5	Avoid adding abstract patterns.
6	Avoid using similar colors on background and buttons.

Based on the research conducted by Ancient and Good (2013), designing interfaces should consider three possible areas where dementia will have an impact on. They are cognitive impairments, motor impairments and visual impairments. They had found that people with dementia struggled with understanding the pinch-zoom functionality of touch screen applications. This involves motor impairments where people with dementia may have slower movements. So, it is wise to take in the consideration of increasing the response time of people with dementia. Say for example, an interface which expects a response within a specific length of time such as ‘time-outs’ when completing an online form. Hence, it is recommended that interfaces must minimize the number of information which the people with dementia are required to take in. This is because the need to remember certain information or position within the system may cause the users to feel disoriented if they are not able to recall their location (Awada et al., 2018).

It is also a fact that people with dementia may have problems with perceiving colors, shapes, and movements. It is important to ensure that the readability is improved for all users with dementia, with suggestions of reducing the density of the text (Ziefel, 2010). Poor interfaces such as small font, button and icon sizes can discourage them from further use (Rodrigues et al., 2014).

As suggested by Freeman et al. (2005), interfaces should be designed bearing in mind that the user may forget the task at hand. Recognition is better than recall, thus they suggest that retrieval cues should be used whenever appropriate. So, to aid with this recognition memory, visual cues such as pictures or icons with the combination of verbal cues should be fully utilized. It is also likely that learning to navigate the site will

be helped by keeping different pages as structurally similar as possible, so that even if the surfer does not consciously remember how to navigate each page, implicit memory derived from viewing previous pages will help to promote easier navigation of the current page (Freeman et al., 2005).

An older study by Morris and McKiernan (1994) suggested that the language on the interface should not be complex and should avoid more abstract or metaphorical language. For instance, the customary use of “home” to signify a link to the home page might be confusing and difficult for them to understand. People with dementia may have difficulties in concentrating and giving attention to multiple tasks at one time. Hence, when designing interfaces, the number of items on the page that requires attention are to be cut down.

METHODS

Research Phase

Definition of dementia, the symptoms of dementia, the needs, and requirements of people with dementia in mobile application features and functionality were studied through literature and preliminary studies. The study of literature was utilized to obtain information and more knowledge regarding dementia and the design considerations for people with dementia. The sources for these literatures were mainly from ResearchGate, which was an academic social networking site designed to facilitate access to academic research. The preliminary study was done by interviewing an expert in the field of dementia in Kota Kinabalu which was also the Head of the Comfort Aged Care Center (Hassan et al., 2020). A design guideline was obtained from the information gathered from this phase.

Development Phase

The development of the application for this research followed the Software Development Life Cycle (SDLC) Waterfall method. Figure 1 shows the SDLC waterfall phases. How the stages of this method are adapted into the development of the application is briefed as follows.

Planning. This phase for this project started with studying existing literature and conducting preliminary research, which is described in research phase. Here was where user requirements and design features were identified and searched for.

Analysis. This was where all important information regarding design features and functions for mobile application for people with dementia was picked out and gathered to produce a set of design guidelines. These guidelines were utilized in the design and implementation phase.

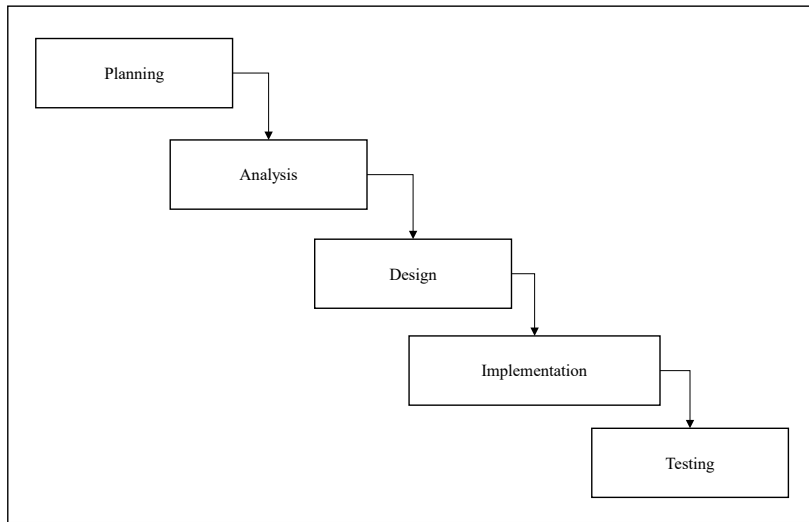


Figure 1. Phases in SDLC waterfall method

Design. Following the guidelines obtained from the previous phase, the design framework for this application was planned and created in this phase.

Implementation. The development of the Diary application was done in this phase.

Testing. This was the stage where the diary application was being tested by dementia patients to evaluate its effectiveness towards their cognitive problems and at the same time identified if there were any crashes or defects.

Evaluation Phase. Quantitative method was used to evaluate the effectiveness of the diary application and whether its design and features were suitable for people with dementia. Figure 2 shows the flow procedure of the evaluation process.

Participants for the application testing and evaluation were recruited from the Comfort Aged Care Center and the main criteria for selecting them included people who were diagnosed with an early stage of dementia. To identify them, a questionnaire developed by Maki et al. (2013) was given to them to answer. Among 20 of the people in the center, only four were identified to be at early stage of dementia and they were willing to participate.

The evaluation process was conducted in two different sessions: one with traditional diary and the other with the diary application. The four selected participants were divided into these two groups and each group consisted of two people. They were divided randomly among these two groups. They were asked to each write diaries for two days with the assistance and guidance of one caregiver in each phase. This means that two caregivers

were assigned to each phase and they were responsible in observing and following up the participants' progress on writing their diaries.

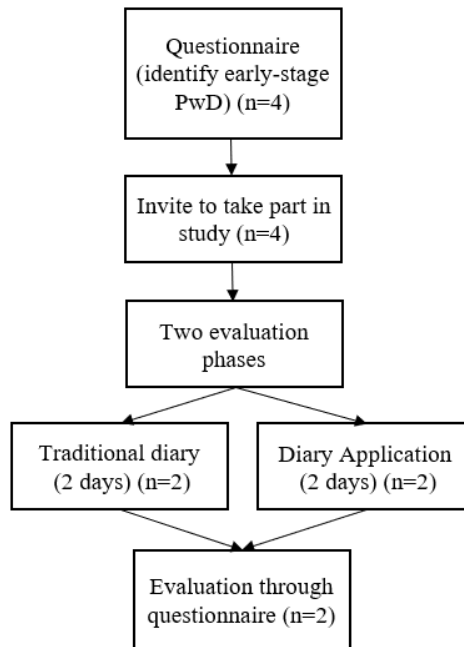


Figure 2. Flow procedure of evaluation process

At the end of each phase, participants completed a set of close-ended questionnaires to measure the effectiveness between both phases. Participants described their experience and their evaluation on using the traditional diary and the diary application using a 5-point Likert scale (strongly agree = 5 to strongly disagree = 1). To get feedback on the design of features and functions of the application, the caregivers that assisted the participants were given a set of open-ended questions. This was done to receive comments and suggestions for the improvement of the application in the future.

RESULTS AND DISCUSSION

All information collected and gathered from literature study and preliminary research are summarized as seen in Table 4 which is then used as design guideline in developing the application. There are not many researchers studying on the use of mobile applications for dementia patients. Most studies only focus on the design consideration and creating interfaces based on those guidelines (Awada et al, 2018; Peterson et al., 2009; Ancient & Good, 2013; Claire et al., 2000; Gowans et al., 2007; Niklasson & Sandström, 2016;

Ziefel, 2010; Rodrigues et al., 2014) and existing applications (Dementia/Digital Diary/Clock, My House of Memories: Dementia & Alzheimer's App, CogniCare - Support for Dementia Care and Alzheimer's Daily Companion) only focus on the information for awareness, audio for diary input, emergency help, and automatic link on calendar sync. They are limited research touching on developing applications that caters patients' cognitive disability by applying those design guidelines.

General Design of Application

This Diary application has five functionalities which are time, date, music, diary, and mood tracker. However, mood tracker is not fully functioning thus, this feature is omitted for testing and evaluation.

Start Page. This page is where the time and date functions are put. It is placed at the very top of the screen, as seen in Figure 3. Following the design guideline, the date and time functions are included on this page. This is because PwD loses track of time often. This helps tackle their problem with time orientation. The background has an image of flowers in it (Freeman et al., 2005), and the buttons at the bottom of the page are not placed too close to each other (Yamagata et al., 2013; Coppola et al., 2013). Different images are set as the background following the time of the day. For example, when it is morning, an image that indicates that it is now the morning is set as background, and an afternoon-looking image is set when the time of the day is currently in the afternoon. PwD, find it entertaining if there are images of flowers or animals inserted (Freeman et al., 2005).

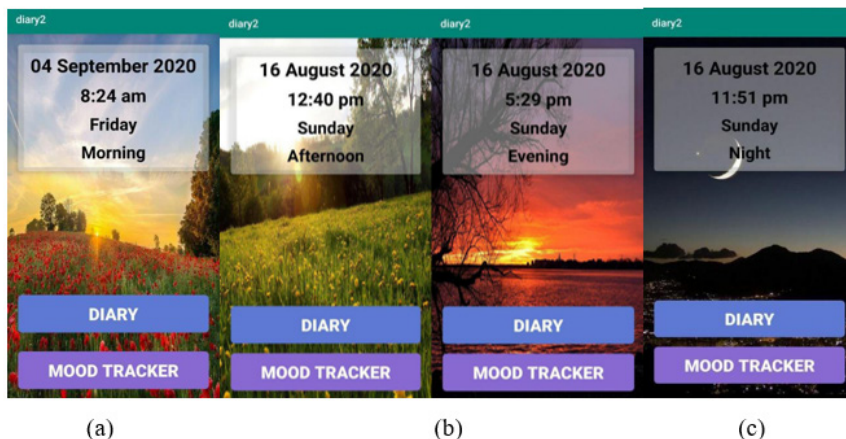


Figure 3(a), (b) & (c). Start page with different background images following the time of the day: morning, afternoon, evening, and night.

Table 4

Design guideline for application development

No	Design guideline
1	Reduce number of items to remember (Ancient & Good, 2013)
2	Use simpler instructions (Awada et al., 2018)
3	Minimize number of information. Only add necessary ones (Awada et al., 2018)
4	Reduce density of text (Ziefel, 2010)
5	Keep different pages as structurally similar as possible (Freeman et al., 2005)
6	Avoid using abstract or metaphorical language (Morris & McKiernan, 1994)
7	Avoid placing buttons close to each other (Yamagata et al., 2013)
8	Avoid clustered images or abstract patterns and elements (Hassan et al., 2020)
9	Use familiar and simple icons (Niklasson & Sandström, 2016)
10	Avoid any hidden message or pop-ups which can confuse PwD (Gowans et al., 2007; Hassan et al., 2020)
11	Use strong contrast between colors. PwD can easily see with red and green colors best (Niklasson & Sandström, 2016; Hassan et al., 2020)
12	Avoid very deep and difficult navigation (Hassan et al., 2020)
13	Place important information at the top of the screen as that location is seen first when opening an application (Niklasson & Sandström, 2016)
14	Use images like animals or flower as they can be entertaining for PwD (Freeman et al., 2005)
15	Buttons and fonts should be large in size (Hassan et al., 2020)
16	Use music (Hassan et al., 2020)
17	If music is added to the application, make sure sound is clear (Hassan et al., 2020)
18	Increase response time for applications that include time-outs (Awada et al., 2018)
19	Include functions like a calendar with today's date and time (Hassan et al., 2020)

Below the time and date, there are the Diary and Mood Tracker buttons. Diary button brings users to the Diary feature and the Mood Tracker button brings users to the Mood Tracker feature. Also, music is automatically played the moment users open the application. Though, there is no function to stop the music.

Diary Feature. Figure 4(a) shows the Diary main page. Use of contrast between colours are utilized (Niklasson & Sandström, 2016) and the buttons and fonts are large (Hassan et al., 2020).

Figure 4(b) gives the Write Diary page. The buttons and fonts here are also large (Hassan et al., 2020). On the title and content input area, there are texts giving description to indicate where to write. These are called hints. This is to give the users an easy understanding. Once diary entries are written and saved, they are added to the Diary main page like in Figure 5. Other pages of the application include the Detail page and the Edit Diary, and they are shown in Figure 6(a) and 6(b) respectively. Note that every page except the start page, is similarly structured (Freeman et al, 2005).

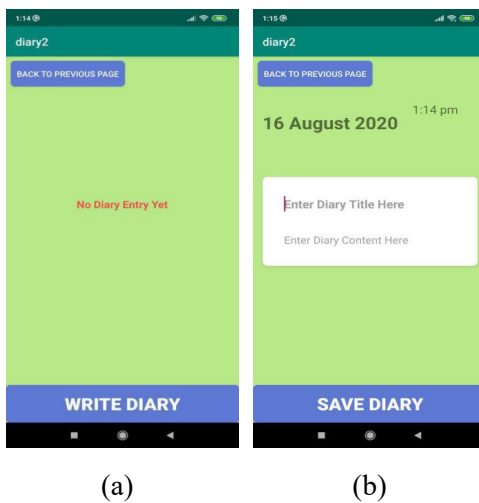


Figure 4(a) & (b). Diary main page and Write Diary page

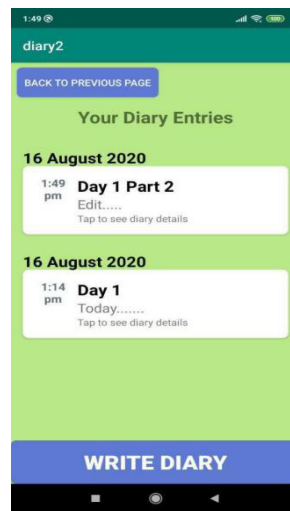


Figure 5. Diary main page with new added diary entries

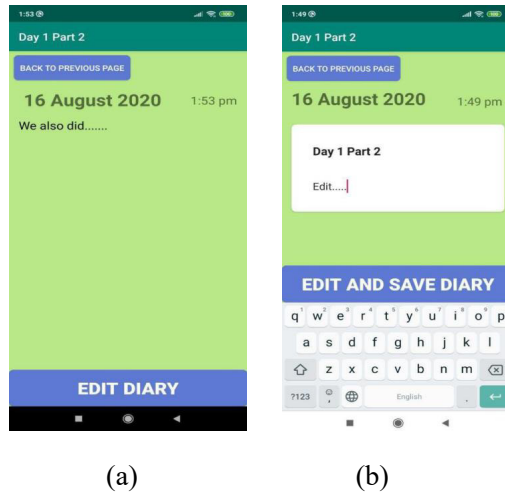


Figure 6(a) & (b). Details page and Edit Diary page

Evaluation of Application

Table 5 shows the brief detail of the four participants that participated in the evaluation process. The four participants were aged between 60 to 70. As mentioned, the main criteria for selecting participants were early stage of PwD. From Table 5, it is shown that three of them were male and one was female. All four of them were married and their final education was being undergraduates. Participants 1, 2 and 3 had very mild severity dementia and participant 4 had mild severity, all of which were considered as early-stages of dementia (Surr et al., 2017).

Table 5
Details of participants

	Participant 1	Participant 2	Participant 3	Participant 4
Age	64	65	65	67
Gender	Male	Male	Female	Male
Marriage	Married	Married	Married	Married
Education	Undergraduate	Undergraduate	Undergraduate	Undergraduate
Severity	Very mild	Very mild	Very mild	Mild
Years living with dementia	2	3	3	3
Evaluation phase	Traditional Diary	Traditional Diary	Diary Application	Diary Application

Evaluation and Measurement of Effectiveness. The set of questionnaires to evaluate and to measure the effectiveness of both traditional diary and diary application are adapted from Zhou et al. (2019) and Al-Aidarooos and Abdul Mutalib (2015). They are closed-ended questions where the respondents' answers are limited to a fixed set of evaluation questions. Table 6 gives the findings of the close-ended questionnaires by the mean average for participants using a traditional diary and Table 7 shows the result for participants using diary application.

A comparison between the means of both data is done to evaluate their effectiveness. This comparative analysis is adopted from a study by Turumogan et al. (2019) and the result is shown in Table 8.

Table 6
Participants' evaluation towards traditional diary(n=2)

No.	Evaluation question	Strongly Agree 5	Agree 4	Neutral 3	Disagree 2	Strongly disagree 1	Mean
1	Was it convenient to use?				2		2.0
2	Was it time-consuming?		1	1			3.5
3	Was it easy to use?			1	1		2.5
4	Was it enjoyable?		1	1			3.5
5	Did you ever lose focus while writing?		2				4.0
6	Did you experience any mood changes while writing?				2		2.0
7	Did you find it difficult to write/type?	1	1				4.5
8	Can this increase your efficiency in your everyday life?	1	1				4.5
9	Does this encourage you to write diary in the future?		2				4.0

Table 7

Participants' evaluation towards diary application (n=2)

No.	Evaluation question	Strongly Agree 5	Agree 4	Neutral 3	Disagree 2	Strongly disagree 1	Mean
1	Was it convenient to use?		2				4.0
2	Was it time-consuming?		1		1		3.0
3	Was it easy to use?	2					5.0
4	Was it enjoyable?	1	1				4.5
5	Did you ever lose focus while writing?		1		1		3.0
6	Did you experience any mood changes while writing?				2		2.0
7	Did you find it difficult to write/type?		1	1			3.5
8	Can this increase your efficiency in your everyday life?	1	1				4.5
9	Does this encourage you to write diary in the future?	2					5.0

Table 8

Means of traditional diary and diary application according to the evaluation questions

Evaluation question	Mean (Traditional Diary)	Mean (Diary Application)
Was it convenient to use?	2.0	4.0
Was it time-consuming?	3.5	3.0
Was it easy to use?	2.5	5.0
Was it enjoyable?	3.5	4.5
Did you ever lose focus while writing?	4.0	3.0
Did you experience any mood changes while writing?	2.0	2.0
Did you find it difficult to write/type?	4.5	3.5
Can this increase your efficiency in your everyday life?	4.5	4.5
Does this encourage you to write diary in the future?	4.0	5.0

Based on Table 7, the chart of the comparison of means between the two diaries is made. It is found that the mean of diary application for the positive questions (convenient, time-consuming, easy, enjoyable and encouragement) are seen to be at a more beneficial level than the mean of traditional diary. Diary application is deemed to be more convenient, less time-consuming, easy to be used, enjoyable and gives encouragement to write diaries in later future. From the Table 7, it can also be concluded that PwD tend to lose focus while writing. This could be affected by how PwD are starting to find it less easy to write, as seen on the mean result of traditional dairy for “difficult to write” (Mean=4.5). For both traditional and application, it is discovered that the participants did not experience any change of moods. The mean result for “efficient” is equal for both as well (Mean=4.5). They thought that writing diaries might help them became more efficient in everyday life which was a very positive response, and this could give them more confidence and support in improving their cognitive impairments especially memory loss.

Evaluation of Design and Functions of Diary Application

Open-ended questionnaire was given to the caregiver that were assigned to assist and help the participants during application testing and evaluation process. The findings of it are listed as follows.

1. Font sizes are large enough. However, the fonts for the time and date can be made larger.
2. Image background can be distracting. Makes the date and time to be seen less clear. Should change to brighter color to make a contrast or make it simpler with plain background with singlecolor.
3. Music is nice to listen to.
4. Buttons do not seem like buttons. Need to tell the users in advance that they are buttons to be clicked. Better to add shadow or simple animation to make it seem like they are clickable. Wordings on button can be seen clearly, simple and easy to understand.
5. Background color for Diary pages are contrast with the button colors. Can see texts on page clearly.
6. Scrolling function in Diary main page is difficult for them. It is better to put up and down button to act as scroller.
7. Does not contain too much information which is good.
8. Back buttons are at the same position in every Diary page which is also good.
9. Better to include record voice option in case PwD would prefer to speak and record themselves rather than typing.
10. Also, better to include a reminder function that connects with calendar.

CONCLUSION

The evaluation of the application was obtained through close-ended and open-ended questionnaires for the participants and caregivers to answer, respectively. It is discovered that the based on the result to measure effectiveness, the diary application gave a higher result on the positive questions, where it is inferred to be more convenient, easy, and enjoyable. From the result of the open-ended questionnaires, there were some good points and suggestions were given to further improve the application. For further development, the mood tracker can be completely developed and a voice recording function can be included for the possibility that people with dementia would prefer speaking directly rather than typing on the device. Future works can include in developing a similar application that targets on the caregivers instead.

ACKNOWLEDGEMENT

Researchers thank Universiti Malaysia Sabah (UMS) for the support of resources and facilities needed to prepare research. Also, appreciation is extended to the participants and the caregivers of Comfort Aged Care Center for helping and participating in the process of completing this project.

REFERENCES

- Al-Aidaroos, A. S. A., & Mutalib, A. A. (2015). Design of the usability measurement tool for multimodal mobile applications. *Jurnal Teknologi*, 77(29). <http://doi.org/10.11113/jt.v77.6810>
- Alzheimer's Disease Foundation Malaysia. (2020). *Alzheimer's - We never think how great a gift is to think*. Retrieved July 4, 2020, from <http://adfm-imu.com/>
- Ancient, C., & Good, A. (2013). Issues with designing dementia-friendly interfaces. In *International Conference on Human-Computer Interaction* (pp. 192-196). Springer. https://doi.org/10.1007/978-3-642-39473-7_39
- Awada, I. A., Mocanu, I., Nastac, D. I., Benta, D., & Radu, S. (2018). Adaptive user interface for healthcare application for people with dementia. In *2018 17th RoEduNet Conference: Networking in Education and Research (RoEduNet)* (pp. 1-5). IEEE Conference Publishing. <https://doi.org/10.1109/ROEDUNET.2018.8514150>.
- Clare, L., Wilson, B., Carter, G., & Breen, K. (2000). Intervening with everyday memory problems in dementia of Alzheimer type: An errorless learning approach. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 132-146. [https://doi.org/10.1076/1380-3395\(200002\)22:1;1-8;FT132](https://doi.org/10.1076/1380-3395(200002)22:1;1-8;FT132)
- Coppola, J. F., Kowtko, M. A., Yamagata, C., & Joyce, S. (2013). Applying mobile application development to help Dementia and Alzheimer patients. *Wilson Center for Social Entrepreneurship*. Retrieved July 4, 2020, from <https://digitalcommons.pace.edu/wilson/16>
- Czaja, S. J., Sharit, J., Lee, C. C., Nair, S. N., Fu, S. H., & Herna, M. A. (2012). Factors influencing use of an e-health website in a community sample of older adults. *Journal of the American Medical Informatics Association*, 20(2), 277-284. <https://doi.org/10.1136/amiajnl-2012-000876>
- Etcheverry, I., Terrier, P., & Marquié, J. (2012). Are older adults less efficient in making attributions about the origin of memories for web interaction? *European Review of Applied Psychology*, 62(2), 93-102. <https://doi.org/10.1016/j.erap.2010.11.002>
- Freeman, E. D., Clare, L., Savitch, N., Royan, L., Litherland, R., & Lindsay, M. (2005). Improving website accessibility for people with early-stage Dementia: A preliminary investigation. *Aging & Mental Health*, 9(5), 442-448. <https://doi.org/10.1080/13607860500142838>.
- Gowans, G., Dye, R., Alm, N., Vaughan, P., Astell, A., & Ellis, M. (2007). Designing the interface between Dementia patients, caregivers and computer-based intervention. *The Design Journal*, 10(1), 12-23. <https://doi.org/10.2752/146069207789318018>
- Group Health Cooperative. (2013). *Dementia and cognitive impairment diagnosis and treatment guideline*. Retrieved July 4, 2020, from <https://www.universityhealthsystem.com/~media/files/clinical-pathways/dementia-cognitive-impairment-guideline-2012.pdf?la=en>
- Hassan, N. M., Baharum, A., Ismail, R., Zain, N. M., Daruis, D. D. I., & Noor, N. A. M. (2020). Improving cognitive impairment among dementia users using mobile application: An initial study. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(1 Spec), 44-48.
- Holsapple, C. W. (2005). A website interface design framework for the cognitively impaired: A study in the context of Alzheimer's disease. *Journal of Electronic Commerce Research*, 6(4), 291-303.

- Johansson, M. M., Marcusson, J., & Wressle, E. (2015). Cognitive impairment and its consequences in everyday life: Experiences of people with mild cognitive impairment or mild dementia and their relatives. *International Psychogeriatrics*, 27(6), 949-58. <https://doi.org/10.1017/S1041610215000058>.
- Klimova, B., & Semradova, I. (2016). Cognitive decline in dementia with special focus on language impairments. *European Proceedings of Social and Behavioural Sciences Cyprus: Nicosia*, 86-90. <https://doi.org/10.15405/epsbs.2016.05.02.9>.
- Lian, Y., Xiao, L. D., Zeng, F., Wu, X., Wang, Z., & Ren, H. (2017). The experiences of people with Dementia and their caregivers in Dementia diagnosis. *Journal of Alzheimer's Disease*, 59(4), 1203-1211. <https://doi.org/10.3233/JAD-170370>.
- Maki, Y., Yamaguchi, T., & Yamaguchi, H. (2013). Symptoms of early Dementia-11 questionnaire (SED-11Q): A brief informant-operated screening for dementia. *Dementia and Geriatric Cognitive Disorders Extra*, 3(1), 131-142. <http://doi.org/10.1159/000350460>
- Morris, R. G., & McKiernan, F. (1994). Neuropsychological investigations of Dementia. In Burns, A. & Levy, R. (Eds.) *Dementia* (pp. 327-354). Springer. https://doi.org/10.1007/978-1-4615-6805-6_19
- Niklasson, A., & Sandström, E. (2016). Iterative design of a user interface adapted for people with Dementia. *Ergonomics and Aerosol Technology*. Retrieved July 4, 2020, from <https://lup.lub.lu.se/student-papers/search/publication/8883307>
- Peterson, C. B., Mitseva, A., Mihovska, A., Prasad, N. R., & Prasad, R. (2009). The phenomenological experience of dementia and user interface development. In *2009 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies* (pp. 1-5). IEEE Conference Publications. <https://doi.org/10.1109/ISABEL.2009.5373697>
- Rodrigues, É., Carreira, M., & Gonçalves, D. (2014). Developing a multimodal interface for the elderly. *Procedia computer science*, 27, 359-368. <https://doi.org/10.1016/j.procs.2014.02.040>
- Soufneyestani, M., Khan, A., & Sufneyestani, M. (2021). Impacts of music intervention on Dementia: A review using meta-narrative method and agenda for future research. *Neurology international*, 13(1), 1-17. <https://doi.org/10.3390/neurolint13010001>
- Surr, C. A., Gates, C., Irving, D., Oyeboode, J., Smith, S. J., Parveen, S., Drury, M., & Dennison, A. (2017). Effective Dementia education and training for the health and social care workforce: A systematic review of the literature. *Review of Educational Research*, 87(5), 966-1002. <https://doi.org/10.3102/0034654317723305>
- Turumogan, P., Baharum, A., Ismail, I., Noh, N. A. M., Ab Fatah, N. S., & Noor, N. A. M. (2019). Evaluating users' emotions for Kansei-based Malaysia higher learning institution website using Kansei checklist. *Bulletin of Electrical Engineering and Informatics*, 8(1), 328-335. <http://doi.org/10.11591/eei.v8i1.1448>
- Williams, D., Alam, M. A. U., Ahamed, S. I., & Chu, W. (2013). Considerations in designing human-computer interfaces for elderly people. In *proceedings of 13th International Conference on Quality Software (QSIC)* (pp. 372-377). IEEE Conference Publications. <https://doi.org/10.1109/QSIC.2013.36>
- WHO. (2012). *Dementia: A public health priority*. World Health Organization. Retrieved July 4, 2020, from <https://apps.who.int/iris/handle/10665/75263>

- Yamagata, C., Coppola, J. F., Kowtko, M., & Joyce, S. (2013, May). Mobile app development and usability research to help dementia and Alzheimer patients. In *2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT)* (pp. 1-6). IEEE Conference Publishing. <https://doi.org/10.1109/LISAT.2013.6578252>
- Ziefle, M. (2010). Information presentation in small screen devices: The trade-off between visual density and menu foresight. *Applied Ergonomics*, *41*(6), 719-730. <https://doi.org/10.1016/j.apergo.2010.03.001>.
- Zhou, L., Bao, J., Setiawan, A., Saptono, A., & Parmanto, B. (2019). The mHealth app usability questionnaire (MAUQ): Development and validation study. *JMIR mHealth and uHealth*, *7*(4), Article e11500. <https://doi.org/10.2196/11500>



Photonics Rib Waveguide Dimension Dependent Charge Distribution and Loss Characterization

Angie Teo Chen Chen, Mohammad Rakib Uddin* and Foo Kui Law

Electrical and Electronic Engineering Programme Area, Faculty of Engineering, Universiti Teknologi Brunei, BE1410 Brunei Darussalam

ABSTRACT

The simulation of behaviour of the charge distribution and the loss characteristic for rib-waveguide is demonstrated by using silicon-on-insulator (SOI). In this simulation, the rib waveguide is designed at a core width of 450nm, core height of 250nm, rib height of 50nm and buried oxide height of 100nm. These dimensions are set as reference. The aspiration of designing rib waveguide instead of other type of waveguide such as ridge waveguide is from the higher light confinement that can be accomplished by rib waveguide as the refractive index difference is huge and the designing of an active device can be realized. In this analysis, free carrier-injection effect was implemented in the first part of the simulation to study the distribution charges of rib-based waveguide structure based on basic dimensions. In this analysis, electrical voltage was varied from 0V to 1.2V in steps of 0.2V for the analysis of distribution of electron. In the second part of the simulation, four design parameters had been amended which included the core width and height, rib height and buried oxide height. Physical dimensions of the waveguide were altered to achieve smaller device footprint with optimized performance affecting large Free Spectral Range (FSR) and high Q-factor. With proper waveguide physical dimensions design, a good performance Micro-Ring Resonator (MRR) exhibits the principles of wide FSR and Q-factor can be achieved.

ARTICLE INFO

Article history:

Received: 28 September 2020

Accepted: 06 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.10>

E-mail addresses:

angie_teo236@hotmail.com (Angie Teo Chen Chen)

rakib.uddin@utb.edu.bn (Mohammad Rakib Uddin)

p20171002@student.utb.edu.bn (Foo Kui Law)

* Corresponding author

Keywords: Carrier injection; distribution of free charges; physical dimensions; rib-based waveguide structure

INTRODUCTION

In the past ten years, silicon photonics have become well known as an encouraging technology for the integration of optical

functionalities within microelectronic chips in which various distinct components have already been presented (Lim et al., 2013; VenNatesh et al., 2016). These include optical filters, modulators, receivers, and converter of wavelengths (Marchetti et al., 2017; Thomson et al., 2016). Nowadays, an optical MRR has become one of the constitutional building blocks in the development of Si photonic circuit integration such as optical filters based on tunability. Optical filter is one of the outstanding devices that is used in Wavelength-Division-Multiplexed (WDM) systems. The function of an optical filter is to select the desired wavelengths which pass only the wanted channel and leave other channels uninterrupted. In WDM systems, a high information capacity is achieved by a large bandwidth and with the use of optical filter based on MRRs, this can be further improved (Aziz et al., 2016). In addition to that, low power consumption and small device size can be achieved by MRR based optical filter. The main two important principal elements that persuade the advancement of silicon photonics are (i) the feasibility in realization of structure with enormous refractive index difference; and (ii) the exploitation of technology on silicon processing that have been readily developed and validated for microelectronic device industry which have the capability of delivering low cost silicon photonic components, at very large volumes. Silicon has a refractive index of as high as 3.5 (Haroon et al., 2012). A rib-based waveguide structure is built on silicon layer which is deposited on top of a silica layer (Bogaerts et al., 2014; Lin et al., 2007). In addition, silicon oxide can be structured with high thermal conductivity silicon. It is known both silicon refractive index and the number of free carrier charges are directly proportional to each other. With high refractive index, free carriers are also increased. With high refractive index, light modulation can take place at a faster speed with low consumption of power in silicon waveguide (Liu et al., 2007). Modulation scheme of waveguide can be categorized into three types; carrier injection, depletion and accumulation based electro-optic modulator on silicon-on-insulator (Debnath et al., 2018; Lim et al., 2015; Mulyanti et al., 2014; Qin et al., 2016). In here, carrier injection is chosen to carry out the simulation of charge distribution and the loss characteristics of rib waveguide. Previously in other reports, rib waveguide dimension studies had performed analysis on each dimension separately (such as varying only core width or core height). However, this study includes the analysis of all the dimensions (core width, core height, rib height and buried oxide height) under a single study to determine the loss relationship.

In this paper, the simulation and analysis of charge distribution and loss characteristics for SOI based rib waveguide based on free carrier-injection effect. The analysis covers the study of the distribution charges of rib-based waveguide structure based on basic dimensions. The rib waveguide analysis is chosen due to higher light confinement as the refractive index difference is high and an active device can be realized. With this basic design, further design dimension variations investigation on waveguide performance are

accomplished such as varying the width and height of core, height of rib and height of buried oxide.

Rib Waveguide

Rib waveguide is designed with a similar structure as strip or ridge waveguide, yet the strip and the planer layer beneath the strip will share the same high index and is part of the core waveguide. This type of waveguide is known as rectangular waveguides in which the x direction is determined by thickness d and the y direction is determined by width w , however their shapes are not exactly rectangular.

The design of the rib waveguide is shown in Figure 1 which consists layers of silicon and silicon oxide. As seen from Figure 1, there are W , H , H_r and H_o , respectively.

- Where, W is the core waveguide width,
- H is the core waveguide height,
- H_r is the rib waveguide height and
- H_o is the buried oxide height

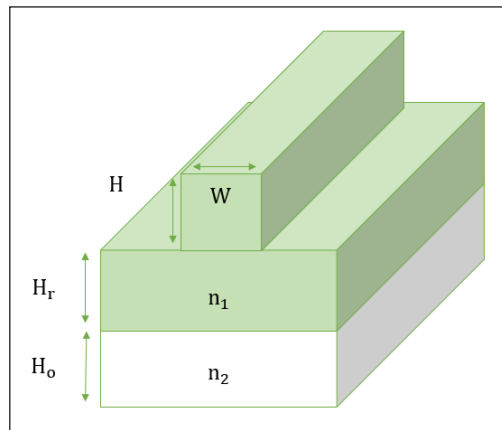


Figure 1. Rib Waveguide

METHODS

Design and Simulation of Rib Waveguide Structure

The rib-based waveguide structure on SOI platform shown in Figure 2 are designed as follows and is set as a basic reference in the entire simulation.

1. Width at 450nm – core waveguide
2. Height at 250nm – core waveguide
3. Height at 50nm – rib waveguide
4. Height at 1000nm – buried oxide

In here, rib waveguide carrier charge distribution is investigated by using a software named Lumerical DEVICE whereas the loss with respect to various dimensions is accomplished by using MODE software.

Figure 3 demonstrates the cross section of the rib waveguide where the connection of voltage source is shown. The addition of an electrical voltage source is to enhance electro-optic to occur in the waveguide as there are electric fields. In this design, for free carriers to diffuse within the core waveguide, there must be different in concentration of doping. Thus, the charge concentration is higher at highly doped p and n region as compared to

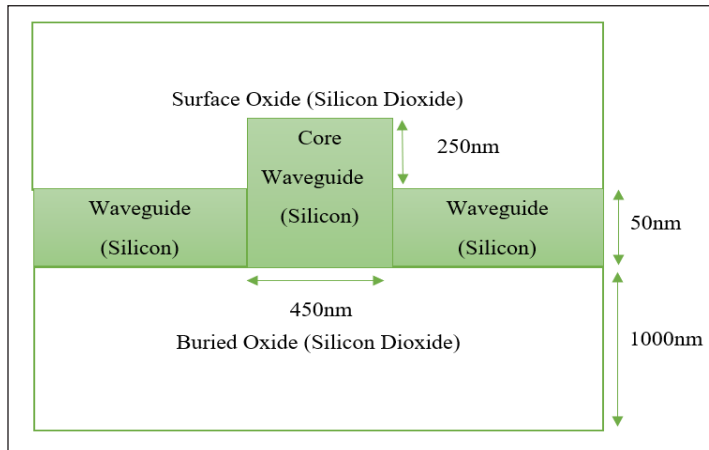


Figure 2. Waveguide block structure

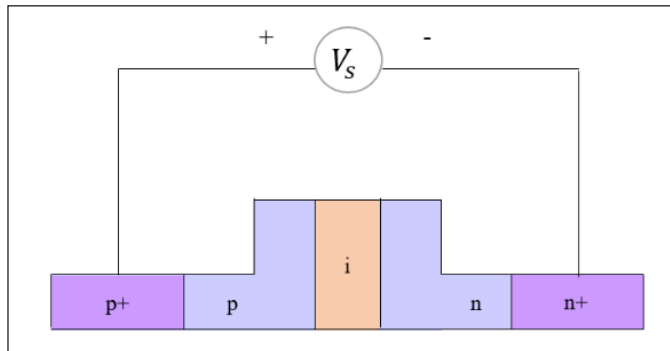


Figure 3. Cross sectional of Rib-based waveguide

the lightly doped p and n regions. Diffusion of free carriers leads to the core waveguide effective index changes.

RESULTS AND DISCUSSION

Analysis of Rib Waveguide Structure

In the first part of the simulation, the analysis of distribution of electron was carried out by varying voltage. The voltage was altered from 0 V to 1.2 V in the steps of 0.2. From the result based on Figure 4, it was observed that when there was no voltage applied, there would be no free carrier diffusion. It was seen that n free charges (red colour) slowly diffused into the p doped region (blue colour) as voltage increases. The reason of diffusion of n free charges is mainly due to the positive voltage applied to p doped region. As a result, n free charges diffused into p doped region.

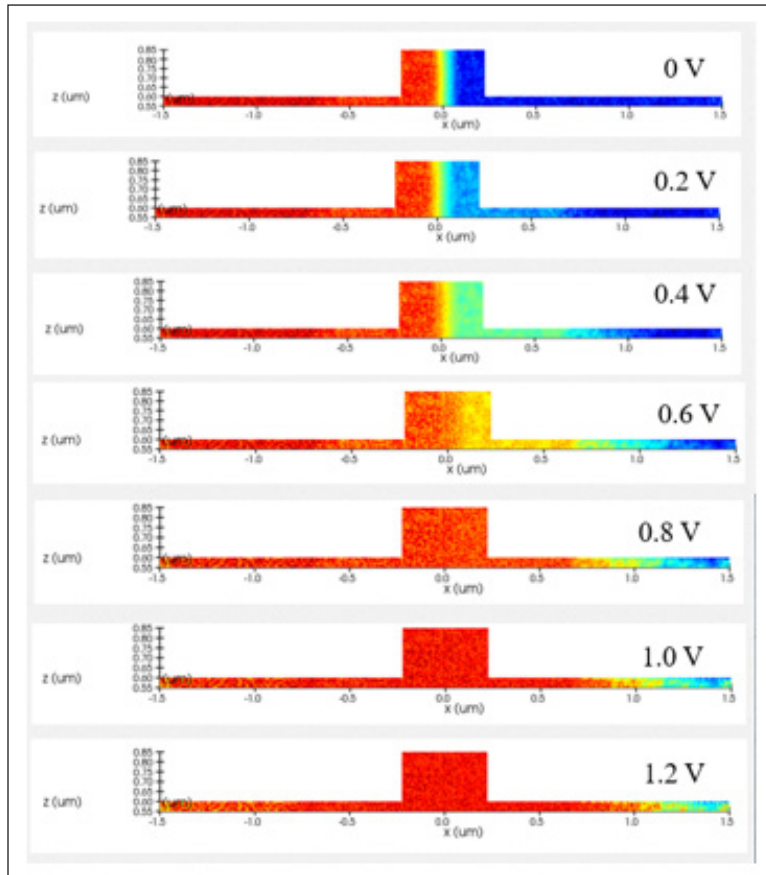


Figure 4. Distribution of charge inside the waveguide at various voltage

Simulation and Analysis Rib Waveguide Dimension Change

In the second section of the simulation, the rib-based waveguide dimensions would be varied one by one to study the performance of dimension change which included core waveguide width and height, rib waveguide height and rib waveguide height. From this, the loss characteristic graph with respect to various dimensions were obtained and analyzed.

Core Waveguide Width Change

In this part, the width of core was altered from 500nm to 700nm at the steps of 100nm in which the height of the core, rib and buried oxide were constant at 250nm, 50nm and 1000nm, respectively. Based on the results shown in Figure 5, as the core width increased, loss would also be increased. Thus, to achieve smaller loss, it is recommended to choose the core width of less than 700nm.

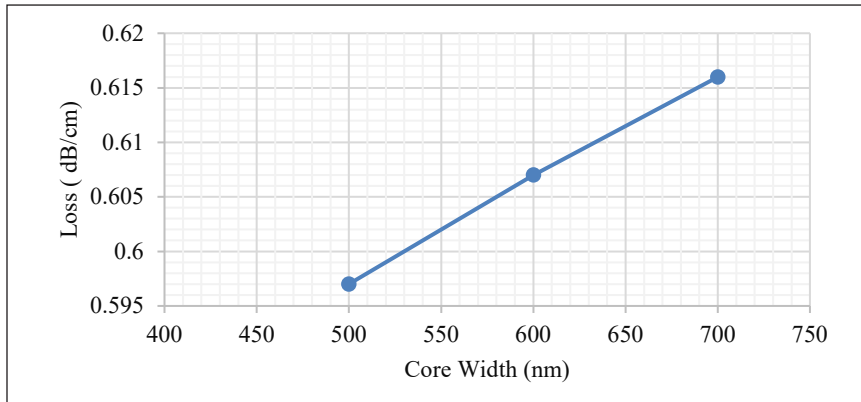


Figure 5. Loss characteristics with respect to different core width

Core Waveguide Height Change

The height of core was altered from 190nm to 230nm at the steps of 20nm in which the other dimensions remain at the basic reference value. It was noticeable that losses were directly proportional to the height of the core waveguide as shown in Figure 6. Thus, it can be determined that as we increased the height of the core waveguide, losses also increased. For this reason, it is advisable to choose a smaller core height for lower loss.

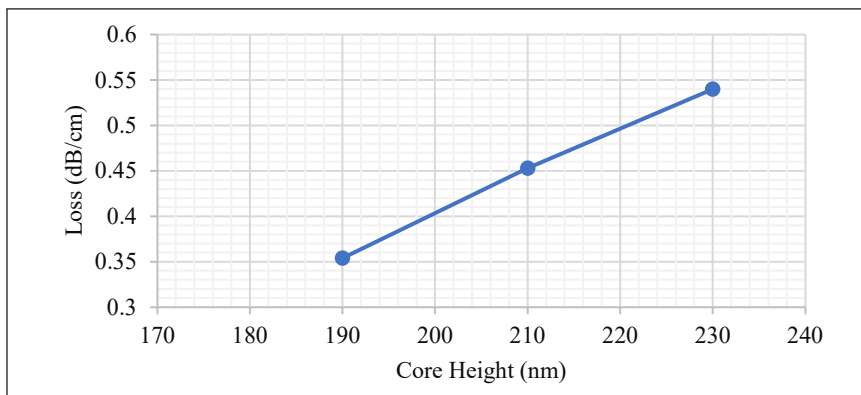


Figure 6. Loss characteristics with respect to different core height

Rib Waveguide Height Change

In this part, the height of the rib waveguide would be varied from 50nm to 90nm at the interval of 20nm in which the height of the core, rib and buried oxide were constant at 250nm, 50nm and 1000nm, respectively. The simulated results of losses with respect to rib height is shown in Figure 7. From the simulated results, it was shown that increasing the rib height would results in lower losses.

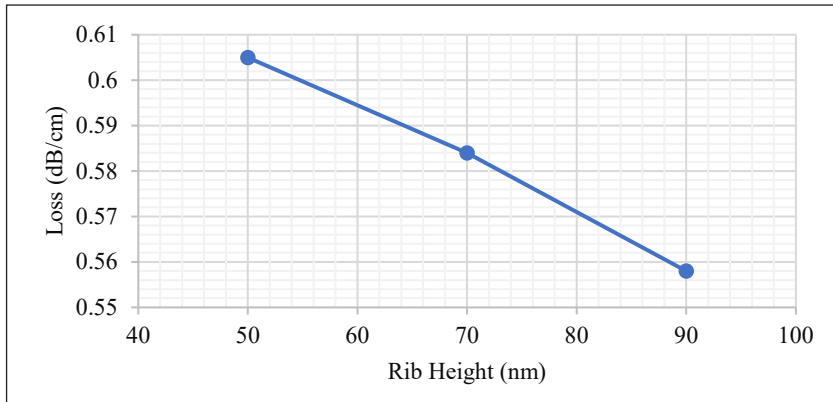


Figure 7. Loss characteristics with respect to different rib height

Buried Oxide Height Change

In this part, the height of buried oxide would be varied from 500nm to 1500nm at the interval of 500nm with all other parameters remaining constant at reference. Based on the simulated results shown in Figure 8, it was observed that there were no changes in loss as the height of buried oxide was increased from 500nm to 1500nm at an interval of 500nm.

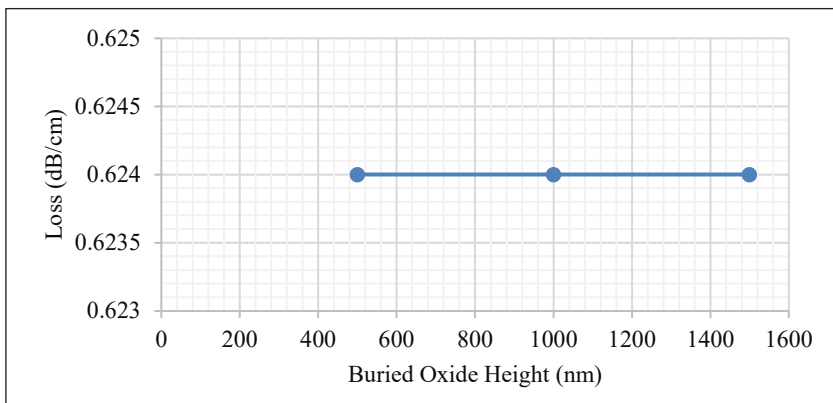


Figure 8. Loss characteristics with respect to different buried oxide height

CONCLUSION

The simulation of behaviour of the distribution of charge and the loss characteristics of rib waveguide is demonstrated by using silicon-on-insulator platform. Free carrier-injection effect is used in the simulation of distribution of charge by the means of varying electrical voltage. In here, the range of voltage changes in the steps of 0.2 V from 0V to 1.2V for examining distribution of electron. Once charge distribution is achieved by Lumerical DEVICE solution, parameters such as loss is further obtained by Lumerical MODE solution.

From there on, further design dimension variations investigation on waveguide performance are accomplished such as varying the width and height of core, height of rib and height of buried oxide. Simulation of rib waveguide dimension change have been accomplished, and graphs haven been plotted to present the results. From the graphs plotted, we can clearly see how each dimension can affect the loss characteristic. It is seen that as core width and core height increases, more losses will be achieved. On the other hand, smaller rib height will lead to lower losses. Simulations have also shown that height of buried oxide does not have significant effect on loss of the waveguide, this means that there is no change in loss as we vary the height of buried oxide. As we know the performance of the waveguide not only depends on loss but there are also other factors. Therefore, in terms of loss, it is deduced that smaller rib height, larger width core and larger height core will lead to more losses. As a conclusion, with this analysis we can determine the variables to maximize the efficiency of rib wave guide. This study can be used as a reference for future research designers to make fundamental decisions relating to the results obtained from these simulations so they can be more effective in designing reliable photonic devices.

ACKNOWLEDGEMENT

We would like to thank Universiti Teknologi Brunei (UTB) for financial support to carry on this research.

REFERENCES

- Aziz, N. N. A., & Haroon, H., & Razak, H. A. (2016). Compact optical filter based on microring resonator. *ARPN Journal of Engineering and Applied Sciences*, 11(5), 3320-3323.
- Bogaerts, W., Fiers, M., & Dumon, P. (2014). Design challenges in silicon photonics. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(4), Article 8202008. <https://doi.org/10.1109/JSTQE.2013.2295882>.
- Debnath, K., Thomson, D. J., Zhang, W., Khokhar, A. Z., Littlejohns, C., Byers, J., Mastronardi, L., Husain, M. K., Ibukuro, K., Gardes, F. Y., Reed, G. T., & Saito, S. (2018). All-silicon carrier accumulation modulator based on a lateral metal-oxide-semiconductor capacitor. *Photonics Research*, 6(5), 373-379. <https://doi.org/10.1364/PRJ.6.000373>
- Haroon, H., Shaari, S., Menon, P., Mardiana, B., Hanim, A., Arsad, N., Majlis, B., Mukhtar, W., & Abdullah, H. (2012). Design and characterization of multiple coupled microring based wavelength demultiplexer in silicon-on-insulator (SOI). *Journal of Nonlinear Optical Physics & Materials*, 21(1), Article 1250004. <https://doi.org/10.1142/S021886351250004X>
- Lim, A. E. J., Song, J., Fang, Q., Li, C., Tu, X., Duan, N., Chen, K. K., Tern, R. P. C., & Liow, T. Y. (2013). Review of silicon photonics foundry efforts. *IEEE Journal of Selected Topics in Quantum Electronics*, 20(4), 405-416. <https://doi.org/10.1109/JSTQE.2013.2293274>
- Lim, S. T., Sun, M. J., & Png, C. E. (2015). Silicon optical modulator simulation. *Frontiers in Physics*, 3, Article 27. <https://doi.org/10.3389/fphy.2015.00027>

- Lin, Q., Painter, O. J., & Agrawal, G. P. (2007). Nonlinear optical phenomena in silicon waveguides: Modeling and applications. *Optics express*, *15*(25), 16604-16644. <https://doi.org/10.1364/OE.15.016604>
- Liu, A., Liao, L., Rubin, D., Nguyen, H., Ciftcioglu, B., Chetrit, Y., Izhaky, N., & Paniccia, M. (2007). High-speed optical modulation based on carrier depletion in a silicon waveguide. *Optics express*, *15*(2), 660-668. <https://doi.org/10.1364/OE.15.000660>
- Marchetti, R., Vitali, V., Lacava, C., Cristiani, I., Giuliani, G., Muffato, V., Fournier, M., Abrate, S., Gaudino, R., Temporiti, E., Carroll, L., & Minzioni, P. (2017). Low-loss micro-resonator filters fabricated in silicon by CMOS-compatible lithographic techniques: design and characterization. *Applied Sciences*, *7*(2), Article 174. <https://doi.org/10.3390/app7020174>
- Mulyanti, B., Menon, P. S., Shaari, S., Hariyadi, T., Hasanah, L., & Haroon, H. (2014). Design and optimization of coupled Microring Resonators (MRRs) in silicon-on-insulator. *Sains Malaysiana*, *43*(2), 247-252.
- Qin, Y., Yu, Y., Wu, W., & Zhang, X. (2016). Integrated tunable optical add/drop filter for polarization and wavelength multiplexed signals. *Optics Express*, *24*(7), 7069-7078. <https://doi.org/10.1364/OE.24.007069>
- Thomson, D., Zilkie, A., Bowers, J. E., Komljenovic, T., Reed, G. T., Vivien, L., & Nedeljkovic, M. (2016). Roadmap on silicon photonics. *Journal of Optics*, *18*(7), Article 073003. <https://doi.org/10.1088/2040-8978/18/7/073003>
- VenNatesh, B. S., Britto, E. R. A., & Balaji, S. S. (2016). *Overview of Silicon Photonic Signal Processing*. Citeseer Sastra University. Retrieved October 4, 2020, from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.857.5778&rep=rep1&type=pdf>



Optimizing Placement of Field Experience Program: An Integration of MOORA and Rule-Based Decision Making

Okfalisa Okfalisa^{1*}, Rizka Hafsari¹, Gusman Nawanir², Saktioto Toto³ and
Novi Yanti¹

¹Informatics Engineering Department, Universitas Islam Negeri Sultan Syarif Kasim Riau, 28293 Pekanbaru, Riau, Indonesia

²Faculty of Industrial Management, Universiti Malaysia Pahang, 26300 UMP, Kuantan, Pahang, Malaysia

³Physic Department in Universitas Riau, 28293 Pekanbaru, Riau, Indonesia

ABSTRACT

The lack of optimality in the Field Experience Program (FEP) placement has affected universities' educational services to the stakeholders. Bringing together the stakeholders' needs, university capacities, and participants' willingness to quality and quantity is not easy. This study tries to optimize the placement of FEP by considering the interests of multiple perspectives through the application of Multi-Objective Optimization on the Basic of Ratio Analysis (MOORA) and Rule-Based methods in the form of a decision-making model. MOORA ranked the students based on the FEP committee's perspective and other criteria, such as micro-teaching grades, final GPAs, study programs, number of credits, and student addresses. Meanwhile, the school perspective was ordered based on its accreditations, levels, types, facilities, and performances. To achieve the optimal recommendation of FEP placement, the integration of MOORA and Rule-based intertwined the requirement of such perspectives. A prototype of the system recommendation is then acquired to simplify the decision-making model. As adjudications, a survey from twenty stakeholders evidenced around 86.92% of system user acceptances. The confusion matrix testing defines the

accuracy of this method reaches 78.33%. This paper reveals that the recommendation model has been successfully increasing the effectiveness of decision making in FEP placement under the needs and expectations of the entire stakeholders.

Keywords: Decision support system, multiple perspectives, optimization, recommendation system, rule-based

ARTICLE INFO

Article history:

Received: 28 October 2020

Accepted: 03 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.11>

E-mail addresses:

okfalisa@gmail.com (Okfalisa Okfalisa)

rizkahafsari@gmail.com (Rizka Hafsari)

gusman@ump.edu.my (Gusman Nawanir)

saktioto@yahoo.com (Saktioto Toto)

novi_yanti@uin-suska.ac.id (Novi Yanti)

* Corresponding author

INTRODUCTION

The FEP is a program to improve students' quality, a mandatory requirement before taking the final assignment course at the Faculty of Education and Teacher Training. In this program, the students carry out teaching practices and other academic activities at junior and senior high school levels set by the FEP committee for two months. Based on interviews with FEP committees, several obstacles were found, including ineffective administrative procedures, which took a long time to access and overload student placements in one school. Besides, some complaints are considered related to the quality and imbalance of student competencies delivered by school needs. The school mileage factor also becomes an obstacle for the students in applying discipline and finance in implementing programs. This equip became a significantly affecting issue, especially for initial teachers in establishing their professional identities (Gang et al., 2020). Feelings of professional unease and discomfort during the first year of teaching regarding the teacher educators' substantial and situational selves kindly influence new pedagogies and the confident practitioner towards achieving new professional identity (Julie & Katie, 2017). The initial teachers require adequate induction support to analyze the knowledge, interpret and understand the education environment setting, mission statement, curriculum, and precise nature of work; and understand the teaching. These processes help the new teacher educator evolve their epistemic, overview knowledge of schooling, the establishment and extended of pedagogy, and provide a potential platform for evaluating what to teach, when, and how (Jones et al., 2020). The world of education continues to grow and is greatly influenced by the development. Thus, it directly affects the educational system. Teachers are responsible for operating the educational system and ensuring the sustainable achievement of schools' objectives and curriculum. To date, a strong and efficient professional competency of teachers should be redefined on the development of human life and education as field competencies, pedagogical competencies, research competencies, curriculum competencies, lifelong learning competencies, social-cultural competencies, emotional competencies, communication competencies, information and communication technologies competencies and environmental competencies (Selvi, 2010). The above competencies imply developing students' competencies towards acquiring self-confidence, learning motivation, and social skills (Chan & Yeung, 2019). This issue aims to allow academics in the field to study and assess their lessons and skills critically. Yeo et al. (2008) found that teachers' efficacy in instruction, classroom management, and student engagement influenced the teacher attributes and the teacher-student relationship. Zendarski et al. (2020) examined the teacher characteristics, thus relating to the social culture competencies that significantly contributed to the child-teacher relationship quality.

In a nutshell, this research tries to respond to the importance of the teacher's roles on the student education in schools, mostly related to the formation of professionalism of the new

teachers; a high demand of policy and effective mechanism during the placement of students as prospective teachers in this FEP; and the consideration of stakeholders' perspectives at the side of students' needs, FEP committees, and schools. These three perspectives give a new contribution to a more objective assessment and decision-making by developing a Decision Support System (DSS) towards FEP placement effectiveness. Thus, it provides opportunities for the professional development of new practical teachers. DSS is an application that aids in providing management recommendations in making decisions more objective by considering various alternatives and criteria defined. Understanding the knowledge-based components in decision-making (performance, attitude, and behavior) will enhance managers' influential role and key actors in making decisions (Bonjar et al., 2019). The application of DSS in assisting management decision making, especially in education, has been widely carried out (Delen et al., 2020). The EVALOE- DSS framework for teacher professional development has been recently designed by Gràcia et al. (2020) with the intention of increasing diverse student linguistic skills. Pardiyono and Indrayani (2019) applied the DSS concept in choosing private higher education based on the marketing mix model criteria. Ardana et al. (2016) developed a DSS model to select blended learning platforms for Mathematic and Information Communication Technology (ICT) learning.

Meanwhile, Ibrahim et al. (2014) integrated the DSS framework for strategic planning in Higher Education Institutions. The above research flourished the significant and successful DSS roles in solving the management, technical, professional, social and culture problems related to decision-making in education fields. Herein, the performance of the FEP program through the practice of DSS approach is strongly optimistic.

Unfortunately, most of the raised studies deploy a single perspective in solving the complexity of DSS problems. In reality, a synthesis of broad worldviews is essentially developed rather than a single perspective to recognize complex social problems' connectedness. The development of multiple perspectives generates open, honest, effective dialogues and trust among the relevant stakeholders affected by the decision. Hsu et al. (2020) explored the multiple perspectives of experts' evaluation in generating a set of influence criteria in promoting the healthcare industry innovative technologies. Meanwhile, Petkov et al. (2007) studied how the multiple perspective representations of complex managerial problems could support the integration of Multi-Attribute Decision Making (MADM) and soft systems methodologies. El-Gayar and Fritz (2010) presented a web-based multi-perspective DSS for information security planning. Yazdani et al. (2017) had successfully developed a group decision making approach in assisting the multiple perspectives of decision makers and customers values in selecting third-party logistic providers. The result revealed that the multiple perspectives substantially rose in the efficient decision support system towards the quality and reliability of decisions. The multiple perspective approach above has successfully bridged qualitative value judgments

with the quantitative data relative criteria to address inquiring organizations' needs. The analysis provided captures the relevant stakeholders' subjective preferences for dealing with conflict priority and presents the trade-offs in a decision fairly.

Moreover, group decision making let the participants boost their ability to learn and stimulate their cognitive level (Carneiro et al., 2020). Therefore, the adoption of three multiple perspectives in this research (i.e., of students, FEP committees, and schools) enhances decision-making effectiveness. More importantly, organizational learning continues towards sustainable gains in productivity and organizational excellence. Many methods have been applied in figuring out the problems related to MADM, including the simple additive weighting method (SAW) (Engel et al., 2017), the analytic hierarchy process (AHP) (Fox et al., 2015; Okfalisa et al., 2018; Leny & Okfalisa, 2019), and the technique for order preference by similarity to ideal solution (TOPSIS) (Conejero et al., 2020), ELECTRE (Mishra et al., 2020) and the latest is MOORA (Okfalisa et al., 2020). Besides, Martusorn et al. (2019) tried to compare the effectiveness of SAW, AHP, and TOPSIS in selecting a suitable warehouse location. The existing techniques have exclusively addressed their benefits and highlighted the attributes of various analyses such that their value can be clarified by using the technique-based of case completion efficiently. As comparatively, the AHP approach and its derivatives, including ANP, Fuzzy AHP, and Fuzzy ANP have intensity emphasizes resolving the bias of weight assessment and analysis sensitivity (Okfalisa et al., 2021). The reviews of MADM methods are disclosed at Table 1.

MOORA was first introduced by Brauers and Zavadskas (2006) to work out the various complex and conflicting decision-making issues for optimal decision purposes. MOORA responds to the representative alternatives concerning that particular objective by calculating the square root of each alternative's sum of squares per objective chosen (Brauers, 2008). MOORA accommodates multiple criteria in simple computational procedures. Thus, a particular single equation is required for decision matrix normalization irrespective of the nature criteria (Madić et al., 2015). This technique has been successfully showing the perfect correlation for order preference to the ideal solution. It is not affected by introducing any additional parameters (Stanujkic et al., 2012) and undefined of the criteria weights (Chakraborty, 2011). MOORA can simultaneously consider the numbers of quantitative and qualitative selection attributes (Gadakh et al., 2013).

The MOORA application in various scientific fields has been carried out, including MOORA and Goal Programming for solving credit lending decision-making problems for real-time commercial banking environment (Yusuf, 2019), AHP-MOORA for solving the composite material selection for structural component development. Herein, the ration system, multiplicative, and the reference point of MOORA used to compare and rank the proposed materials (Patnaik et al., 2020), integrated MOORA and DEMATEL-ANP to provide optimum sensitivity analysis and consistency decision-makers' priorities

Table 1
The reviews of MADM approaches

MADM Approach	Advantages	Method Extension	Reference
AHP	<ol style="list-style-type: none"> 1. Focus on weighting of evaluation criterion. 2. Determine the significant weight and prioritize of criterion 3. Dealing with human bias during the knowledge transformation of decision makers 	AHP-Entropy; AHP-OMAX; Fuzzy AHP; Fuzzy C Mean (FCM) – Fuzzy AHP; Quality Function Deployment (QFD) – Fuzzy ANP; Fuzzy AHP; ANP	Du et al. (2020); Okfalisa et al. (2018); Mangla et al. (2017); Kazemi et al. (2020); Mistarihi et al. (2020); Li et al. (2020)
SAW	<ol style="list-style-type: none"> 1. Commonly applied for converted negative criteria into positive value. 2. Determine the weighted summation values for each alternative and assessment criteria. 3. Finding the complete ranking of the alternatives 	SAW-HFL Additive Ratio Assessment (ARAS); AHP – SAW; Fuzzy SAW and Fuzzy TOPSIS; Fuzzy SAW	Büyüközkan & Güler (2020); Kumar et al. (2019); Roszkowska & Kacprzak (2016); Mukodimah et al. (2018); Engel et al. (2017)
TOPSIS	<ol style="list-style-type: none"> 1. Can be used to determine the weights of decision makers 2. Commonly applied for ranking the alternatives and to select the best one 3. Considering both positive-ideal and negative-ideal solutions in decision-making. 	Fuzzy TOPSIS; TOPSIS; TOPSIS with Hesitant Pythagorean Fuzzy Sets.	Gündoğdu & Kahraman (2019); Kacprzak (2019); Liang & Xu (2017); Memari et al. (2019); De Farias Aires & Ferreira (2019); Martusorn et al. (2019); Conejero et al. (2020).
ELECTRE	<ol style="list-style-type: none"> 1. Can be used to calculate and rank the criteria and alternatives. 2. Famous for its outranking relations to rank a set of alternatives. 3. Applied to determine the concepts of concordance and discordance relations among alternatives. 4. Effective application in group decision-making environment 	Intuitionistic Fuzzy (IF)-DEMATEL and IF-ELECTRE; ELECTRE III with weighted Borda rule; ELECTRE; ELECTRE and Pythagorean Fuzzy Sets	Kilic et al. (2020); Liao et al. (2020); Fei et al. (2019); Mishra et al. (2020); Akram et al. (2019).
MOORA	<ol style="list-style-type: none"> 1. Overing some of the available decision-making methods include fewer mathematical computations, less computational time, more simplicity, and more stability compared with other MADM 2. Proposing optimum materials and stable ranking result. 3. One of the latest MADM methods that covers the weakness of other older methods. 4. Has been applied in many selections case study and fields of research background. 	FCM – MOORA; Fuzzy MOORA; Fuzzy MOORA – FMEA; MOORA; MULTIMOORA; AHP-MOORA; MOORA and Goal Programming; MOORA and DEMATEL-ANP; MOORA-based Taguchi	Dabbagh & Yousefi (2019); Emovon et al. (in Press); Arabsheybani et al. (2018); Shihab et al. (2018); Omrani et al. (2019); Yusuf (2019); Patnaik et al. (2020); Dinçer et al. (2019); Liang et al. (2020).

for the recommendation of the financial service in E7 economy evolution (Dinçer et al., 2019), application of MOORA-based Taguchi method for predicting the optimal welding parameters by considering the multiple quality of perspectives (Liang et al., 2020), MOORA and FCM as a hybrid decision-making system for prioritizing Occupational Health and Safety (OHS) risks based on the proposed weight of FCM approach (Dabbagh & Yousefi, 2019). As a result, the previous works derived that MOORA can consider all the attributes essential and provided a better accurate evaluation of the alternatives. Arabsheybani et al. (2018) also found that MOORA accommodated the optimum decision-making methods with uncomplicated mathematical computations, low execution time, more simplicity, and revealed stable ranking result compared with others MADM techniques.

Moreover, the sensitivity analysis approach of MOORA continues to evolve and be enhanced through the integration process with various MADM approaches.

Therefore, this study applies the MOORA method in prompting the proposed alternatives by each perspective. Every perspective has measurable attributes to quantitatively well-defined the alternative solutions, viz., micro-teaching grades, final GPA, study programs, number of credits and student address, the school accreditation, the school level, the school type, the school facilities, and the school performance. The objective consists of the optimization model as well as maximization or minimization of an attribute. The satisfaction of all perspectives becomes the primary consideration that must be revealed in this research. Herein, MOORA defines the robustness in connection with multiple objectives and conditions set in this case.

Furthermore, to inextricably link the MOORA optimization ranking from both perspectives, a rule-based concept with forwarding chaining inference is applied. Rule-based is capable of selecting the minimal and representative criteria objectively and reliably for forming the MADM model and overlooking the inter-relationships among the involved criteria towards the continuous improvement and the measurement of underestimated effect of non-additive aggregators (You et al., 2019). By deploying the rule bases, the MADM method can translate decision-makers' knowledge and explanation facilities into quantitative and qualitative analytical functions. The integration of body knowledge in MADM and the expert system paradigm is represented by a MOORA and Rule-based embodiment, thus explicitly articulating the stakeholders' knowledge about a particular decision-making issue. Thus, the recommendations given are expected to meet each perspective's needs and interests by considering the defined criteria' value.

The suitable pair proposed is, without a doubt, the most incredible combination according to the parameters' significance. As a limitation, this research was conducted on the execution of the FEP at Education and Teacher Training Faculty of Universitas Islam Negeri Sultan Syarif Kasim Riau, which involved a total number of 1,036 students in the year 2018/2019 from seven study programs, including the Islamic Religious Education,

Arabic Language Education, English Language Education, Economic Education, Chemical Education, Mathematics Education, and Counseling Guidance.

MATERIALS AND METHODS

Several activities were carried out this research. First, the problem identification process was put into practice qualitatively through interviews and focus group discussion with one FEP manager, two heads of departments, two deputy dean, two schools' management, and thirty-five students as participants. The FEP managers, head departments, deputy dean, and schools' management were asked due to their responsibility in handling this program. Meanwhile, the representative students from the post participants were discussed their problematics and expectation of this program. As a result, a pattern of the FEP procedure was defined. The emerging of various obstacles encountered during the execution from both perspectives. Several possible variables were also formulated and proposed as criteria, such as the students perceives criteria viz., the value of micro-teaching (C1), final GPA (C2), program study (C3), credit numbers (C4), and student address (C5). Concurrently, the school apprehends criteria expressly school accreditation (C1), school level (C2), school type (C3), school facilities-Wifi (C4), school facilities-labour (C5), school facilities-library (C6), school facilities-air conditioner (C7), and school performance-Adiwiyata (C8). The determination of criteria weight for the FEP placement was set during the discussion (details are described in the Result and Discussion section). Quantitative literature reviews have been structured to reinforce the formulation of standards. As the main activity, the optimization analysis of current program participants delegated by forty students and ten schools was preliminarily examined through MOORA.

A series of MOORA flow processes is followed, namely the formation of matrices (Equation 1), the determination of normalized matrices (Equation 2), the determination of weighted normalization matrices (Equation 3), and the determining preference values (Equation 4).

1. Matrix Formation

$$X_{ij} = \begin{bmatrix} X_{11} & X_{12} & X_{1n} \\ X_{21} & X_{22} & X_{2n} \\ X_{m1} & X_{m2} & X_{mn} \end{bmatrix} \quad (1)$$

x_{ij} = the result of matrix formation

x = the value of each criterion

i = the criterion value

j = the alternative value

m = the criterion value to m

n = the alternative value to n

2. Normalize matrix determination

$$x_{ij}^{\bar{}} = \frac{x_{ij}}{\sqrt{\sum_{j=1}^m x_{ij}^2}} \quad (2)$$

($j = 1, 2, \dots, n$)

$x_{ij}^{\bar{}}$ = the average of the i -th criteria to the value of the j -th criteria

x_{ij} = the matrix formation

i = the value derived from the number of criteria

j = the value derived from the number of alternatives

n = the number of alternative values up to n

3. Weighted normalize matrix determination

$$y_i = \sum_{j=1}^g X_{ij} * \sum_{j=g+1}^n W_j \quad (3)$$

y_i = the result of weighted matrix multiplication

w_j = the weight value of the j -th criterion

x_{ij} = values of each matrix formation

4. Preference value determination

$$y_i = \sum_{j=1}^g W_j X_{ij} - \sum_{j=g+1}^n W_j X_{ij} \quad (4)$$

y_i = the result of weighted matrix multiplication

w_j = the weight value of the j -th criterion

x_{ij} = values of each matrix formation

Finally, the highest rank of student and school alternatives was merged using the rule-based to track the appropriate recommendation pair. A decision tree diagram is then developed based on the association rules defined by the stakeholders. To automate the integration of MOORA and rule-based calculation, a prototype DSS was constructed. Object-Oriented and Unified Modeling Language (UML) tools were applied in system analysis and design. These tools have demonstrated a promising future as a pragmatic methodology in modelling DSS, including user interface, architectural design, analysis and design, programming, data management, and model management (Liu & Stewart, 2004). The visualization derives from this powerful approach is capable in interactively express the immersive learning of information and knowledge of the respondent, as examples Kumar et al. (2019) designed web-based object-oriented decision support system for coastal

water quality prediction and Sztubeka et al. (2020) deployed Geographic Information Systems as innovative DSS in identifying the potential location for energy efficiency improvement. Thus, the adoption of the object-oriented technique has been successfully generated the dynamic analytical and development of multi-attribute decision making components of DSS. Furthermore, Blackbox Testing, User Acceptance Testing (UAT), and Confusion Matrix were systematically organized in testing the work out functionality, the user responses, and the accuracy comparison for 120 test data simulation. Blackbox testing and UAT are commonly and effectively applied for software development testing (Pressman & Maxim, 2020).

Below is the description of the Confusion Matrix. It is acceptable and straightforward for classification algorithms and area estimation models (Lewis & Brown, 2001; Han et al., 2011) (Equation 5 & 6).

$$\text{Accuracy} = \frac{TP+TN}{P+N} \times 100\% \quad (5)$$

$$\text{Error-rate} = \frac{FP+FN}{P+N} \times 100\% \quad (6)$$

TP (True Positive) = The amount of correctly classified data (*Actual class (yes), Predicted class (yes)*).

TN (True Negative) = The amount of correctly classified data (*Actual class (no), Predicted class (no)*).

FN (False Negatif) = The amount of incorrectly classified data (*Actual class (yes), Predicted class (no)*).

FP (False Positif) = The amount of incorrectly classified data (*Actual class (no), Predicted class (yes)*).

P = Total of *TP* and *FN*

N = Total of *FP* and *TN*

Blackbox Testing is a structural software evaluation technique used to analyze system functionality, requirement analysis validation, system integration relates to the codification review, acknowledging the customer's requirement analysis phase, and the system regression test (Nidhra, 2012; Copeland, 2004). Meanwhile, UAT is conducted by disseminating the questionnaire with five Linkert scales to measure the respondents' agreement related to the system utilization, procedures, and the recommendations offered by this prototype DSS system (Davis & Venkatesh, 2004). UAT is essential in grabbing information and profound knowledge relevant to the validation and acceptance of the solution.

RESULTS AND DISCUSSIONS

Problem Identification

Based on the observations and interviews, the flow of FEP student placement activities is described in Figure 1. The current system was manually performed and randomly positioned without considering any particular criteria. Students enrolled for the FEP program at the front desk of committee services. Next, the committee checked the validity of documents' requirements by referencing the faculty data, including students' status, micro-teaching score, final GPA, study programs, and credits. The committee applied word and excel programming to make the recapitulation and data classification. Finally, the committee randomly placed the students for the FEP program. The FEP placement report will be announced and conducted based on the proposed date and schools. As a result, manually operated of registration, data validation, and data recapitulation triggered the un-effective time management on the side of committee services. Lack of students' discipline during the registration process caused the target's achievement not to set as expected. The committee services were disappointed on the schools' side, especially regarding time management, students' quality and competencies, students' discipline, students' requirements, and quota allocation. By referring to Indonesia's government regulation No. 13 the Year 2015 about national education standard, Indonesia Ministry of Environment No. 05 the Year 2013, and focus group discussion summary during the interview session, student performance can be assessed through the score value of micro-teaching, final GPA, program study, numbers of credits, and students' location. Microteaching provides a space for the pre-service teacher to evaluate their teaching (Saban & Coklar, 2013) by considering the timing, planning, asking questions, management of the class, materials usage, and physical appearance in front of the class. The value of micro-teaching practices contributes to the pre-service teachers' qualifications and experiences. Hemdi et al. (2016) found that the cumulative GPA recognized as the holistic student academic assessment mechanism clearly described their actual ability, knowledge, skills, and attitude. Understanding GPA is beneficial for the students, faculty, university, and external stakeholders in identifying the students' strengths

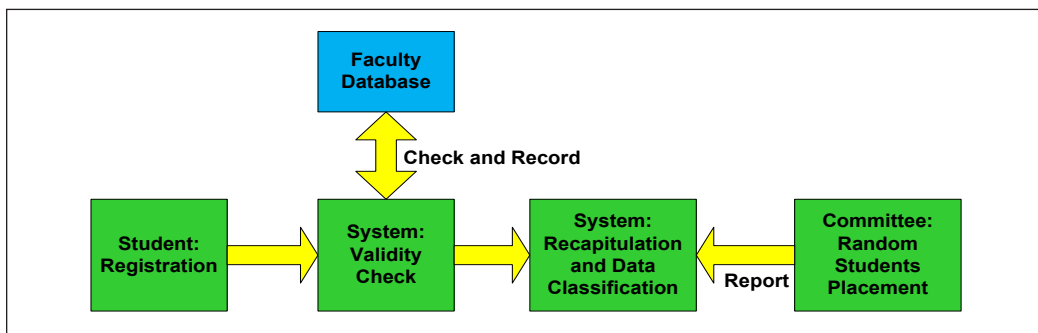


Figure 1. Flow diagram for FEP student placement

and weaknesses for vacancy allocation. Concurrently, the school appraisal encompasses the school accreditation, the school level, the school type, the school facilities, and the school performance. School accreditation evaluates the school's quantitative educational quality judgment of facilities, resources, and teaching. The accreditation mirrors the broad aspects of teacher competencies and values during the educational program (Davis & Ringsted, 2006). Cherchye et al. (2010) exhibited that environmental characteristics in terms of the school type private and public sector positively impact the educational output in efficiency and equity. The Indonesian government introduced the Adiwijaya program to encourage knowledge creation and school awareness of nature (Krisnawati et al., 2015). The program embodies the environmental policy, environmental-based curriculum, participative activities, and environmental facilities management. In a nutshell, the school performance is valuably measured through Adiwijaya program availability.

Criteria Formulation

The formulation of criteria is defined in Tables 2 and 3. The weights of its priority are put in place following the summing up of focus group discussion. The discussion revealed the weighted of students' criteria as in Table 2, thus C1-The Value of Micro-teaching set as the most priority criteria at 30% weighted value following by C2-Final GPA, C3-Program Study, C4- Credit Numbers, and C5- Student Address, respectively. The values of Max and Min for each criterion is referred by MOORA regulation for optimum and minimum values. The value of micro-teaching subsumes in grade A, A-, B+, B, B-, C+, and C with weight precedence into score 4.00, 3.70, 3.30, 3.00, 2.70, 2.30, and 2.00, respectively. The final GPA is specified as more than equal to 2.75. Program study is covered into seven programs, viz., Chemistry Education (P1), Economic Education (P2), English Language Education (P3), Mathematics Education (P4), Counseling Education (P5), Islamic Religious Education (P6), and Arabic Language Education (P7) with the weight emphasis from seven to one accordingly. Credit numbers are designated not less than 120 credits. Criteria for student address encompasses the distance from school 0 to 500 meter, 500 to 2000 meter, and more than 2000 meter. The weights are generated into the value of three, two, and one, respectively.

Table 3 explained the weight of school criteria whereby C1-School Accreditation is denoted as the weightiest priority criteria at 30%. It is then pursued by C2-School Level, C3-School Type, C4-School Facilities (Wifi), C5-School Facilities (labour), C6-School Facilities (Library), C7-School Facilities (Air Conditioner), and C8-School Performance (Adiwiyata), respectively. The school's accreditation (C1) is measured in grades A and B thus, the weights in each case are specified in subject values of four and three. The school-level (C2) criteria contain the sub level in senior high school (TS1), Islamic senior high school (TS2), vocational high school (TS3), junior high school (TS4), an Islamic junior

Table 2
Weight of student criteria

No.	Criteria	Weight	Value
C1	The Value of Micro-teaching	30%	Max
C2	Final GPA	25%	Max
C3	Program Study	20%	Max
C4	Credit Numbers	15%	Max
C5	Student Address	10%	Min

Table 3
Weight of school criteria

No.	Criteria	Weight	Value
C1	School Accreditation	30%	Max
C2	School Level	25%	Max
C3	School Type	15%	Max
C4	School Facilities (Wifi)	5%	Max
C5	School Facilities (labor)	5%	Max
C6	School Facilities (Library)	5%	Max
C7	School Facilities (Air Conditioner)	5%	Min
C8	School Performance (Adiwiyata)	10%	Min

high school (TS5) with the circumscribed weight from five to one, respectively. The school type (C3) is explained in public and private schools by considering the weighted priority in two public and private schools. Concurrently, the school facilities (C4-C7) and performance (C8) are determined by their availability, one weighted for provided, and zero for the rest.

MOORA Analysis

Following Equation 1, Tables 4 for forty (A1 to A40) students' matrix formation against the values of each criterion (C1 to C5) are elucidated. Subsequently, Equation 2 calculation for students' normalization matrix is spelt out as in Table 5. Table 6 is then declared according to the MOORA estimation in Equation 3 to denote the students' weighted matrix's normalization across criteria. Besides, Table 7 indicates the value of student's preference conforming to the Equation 4. Thus, it ranks the students from the highest score of Y_i at student-A6 (0.1468), student-A25 (0.1466), student-A27(0.1462), student-A28 (0.1461), student-A20 (0.1455), and student-A22 (0.1452). Furthermore, a similar calculation from Equation 1 to 4 is applied to the side of the school perspective. As a final result, Table 8 shows the schools' place appertaining to the values of Y_i where school-A4 as the highest score at 0.2713, following by school-A3 (0.2610), school-A2 (0.2474), school-A1 (0.2372), school-A8 (0.2067), school-A6 (0.1998), school-A10 (0.1986), school-A5 (0.1833), school-A7 (0.1759), and school-A9 (0.1728) accordingly.

Table 4
Student matrix formation

A	C ₁	C ₂	C ₃	C ₄	C ₅
A ₁	3.3	3.50	5	136	3
A ₂	3.3	3.17	5	132	1
A ₃	2.7	3.36	5	136	2
A ₄	3.3	3.37	5	136	3
A ₅	4	3.30	5	136	2
....
A ₃₆	3	3.18	3	132	1
A ₃₇	3.3	3.55	3	132	3
A ₃₈	2.7	3.73	3	134	1
A ₃₉	4	3.49	3	132	2
A ₃₉	4	3.49	3	132	2
A ₄₀	2	3.26	3	134	3
$\sum C_n$	20.2776	21.4116	28.5482	833.9017	13.8203

Table 5
Student normalization matrix

A	C ₁	C ₂	C ₃	C ₄	C ₅
A ₁	0.1627	0.1635	0.1751	0.1631	0.2171
A ₂	0.1627	0.1481	0.1751	0.1583	0.0724
A ₃	0.1332	0.1569	0.1751	0.1631	0.1447
A ₄	0.1627	0.1574	0.1751	0.1631	0.2171
A ₅	0.1973	0.1541	0.1751	0.1631	0.1447
....
A ₃₆	0.1479	0.1485	0.1051	0.1583	0.0724
A ₃₇	0.1627	0.1658	0.1051	0.1583	0.2171
A ₃₈	0.1332	0.1742	0.1051	0.1607	0.0724
A ₃₉	0.1973	0.1630	0.1051	0.1583	0.1447
A ₄₀	0.0986	0.1523	0.1051	0.1607	0.2171

Table 6
Students normalization weighted matrix

A	C ₁	C ₂	C ₃	C ₄	C ₅
A ₁	0.0488	0.0409	0.0350	0.0245	0.0217
A ₂	0.0488	0.0370	0.0350	0.0237	0.0072
A ₃	0.0399	0.0392	0.0350	0.0245	0.0145
A ₄	0.0488	0.0393	0.0350	0.0245	0.0217
A ₅	0.0592	0.0385	0.0350	0.0245	0.0145
....
A ₃₆	0.0592	0.0333	0.0280	0.0212	0.0145
A ₃₇	0.0444	0.0371	0.0210	0.0237	0.0072
A ₃₈	0.0488	0.0414	0.0210	0.0237	0.0217
A ₃₉	0.0399	0.0436	0.0210	0.0241	0.0072
A ₄₀	0.0592	0.0407	0.0210	0.0237	0.0145

Table 7
Student preference value

A	Max(C ₁ +C ₂ +C ₃ +C ₄)	Min(C ₅)	Y _i = Max - Min
A ₁	0.1492	0.0217	0.1275
A ₂	0.1446	0.0072	0.1374
A ₃	0.1387	0.0145	0.1242
A ₄	0.1477	0.0217	0.1260
A ₅	0.1572	0.0145	0.1427
A ₆	0.1540	0.0072	0.1468
....
A ₃₆	0.1263	0.0072	0.1190
A ₃₇	0.1350	0.0217	0.1133
A ₃₈	0.1286	0.0072	0.1214
A ₃₉	0.1447	0.0145	0.1302
A ₄₀	0.1128	0.0217	0.0911

Table 8
School preference value

A	Y _i	C ₁	C ₂	C ₃	C ₄	C ₈
A ₄	0.2713	4	5	2	1	1
A ₃	0.2610	4	5	1	0	0
A ₂	0.2474	4	4	2	1	1
A ₁	0.2372	4	4	1	0	0
A ₈	0.2067	3	3	1	0	0
A ₆	0.1998	4	2	2	1	1
A ₁₀	0.1986	4	3	2	0	1
A ₅	0.1833	4	1	1	0	0
A ₇	0.1759	4	1	2	1	1
A ₉	0.1728	4	2	1	0	0

Rule-based Application

Rule-based construction was carried out by considering the variables of the study program (C3) from the students' perspective and school level (C2) from the school's perspective. Decision Tree Diagram can be depicted in Figure 2.

Rule-based development is presented as follows:

Rule 1: IF [P1] OR [P2] THEN [TS1, TS2]

Rule 2: IF [P6] OR [P7] THEN [TS2, TS5]

Rule 3: IF [P3] OR [P4] OR [P5] then [TS1, TS2, TS3, TS4, TS5].

In a nutshell, the selected students at the side of committee perspectives are merged with the preferred schools from the school perspective. Thus, it is pursuing the rule-based

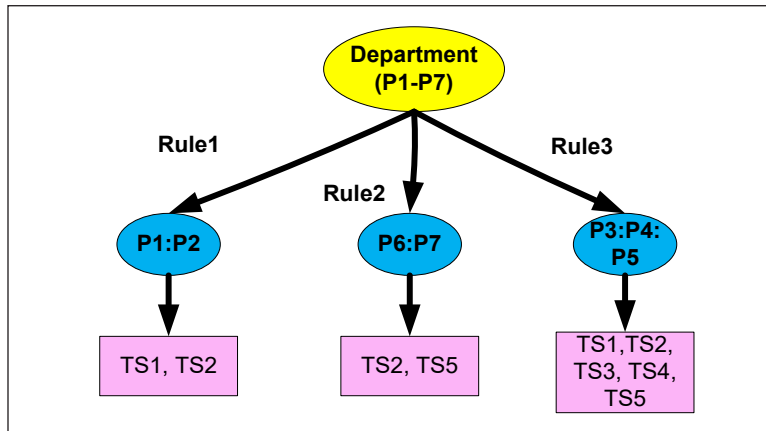


Figure 2. Decision tree diagram for FEP rule-based

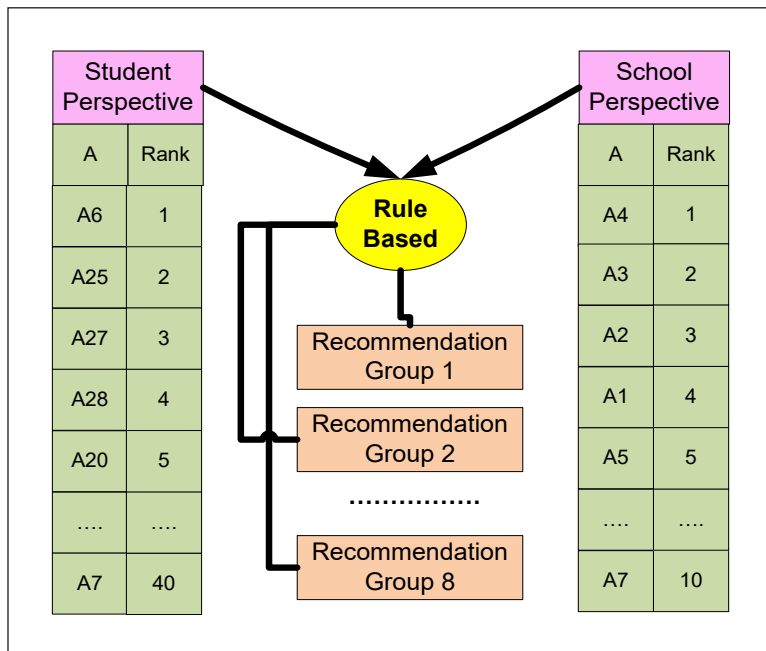


Figure 3. Integration of MOORA and Rule-based

formula as above (Rule 1 to 3). As a result, eight groups recommendation as in Table 9 is then suggested in Figure 3. Figure 3 elucidated that group one's recommendation is revealed from the first five rank students (A6, A25, A27, A28, A20) to be placed in the first three order schools (A4, A3, A2). The recommendation group two puts the sixth position students (A22, A24, A26, A5, A2) into school-A3, school-A2, school-A1, and reciprocally. For the detailed integration of MOORA and rule-based in eight groups, the detailed recommendation is interpreted in Table 9.

Table 9
FEP student placement recommendation

Recommendation Group 1					
Alternatives (Student) by Rank		Department	Alternatives (School) by Rank		School Level
A ₆	1	P3	A ₄	1	TS1
A ₂₅	2	P1	A ₃	2	TS1
A ₂₇	3	P1	A ₂	3	TS2
A ₂₈	4	P1			
A ₂₀	5	P2			
Recommendation Group 2					
A ₂₂	6	P2	A ₃	2	TS1
A ₂₄	7	P2	A ₂	3	TS2
A ₂₆	8	P1	A ₁	4	TS2
A ₅	9	P7			
A ₂	10	P7			
Recommendation Group 3					
A ₃₀	11	P1	A ₂	3	TS2
A ₂₁	12	P2	A ₁	4	TS2
A ₃₂	13	P4	A ₅	5	TS3
A ₁₄	14	P6			
A ₃₉	15	P5			
...
Recommendation Group 8					
A ₁₁	36	P7	A ₉	8	TS5
A ₁₀	37	P7	A ₁₀	9	TS5
A ₄₀	38	P5	A ₇	10	TS4
A ₁₃	39	P6			
A ₇	40	P7			

This study has generally established how the simple mathematical calculation in MOORA reveals the best decision-making by taking potential attributes into account (Arabsheybani et al., 2018) and prioritizing the alternative from several perspectives, namely students and schools' sides. This result is reinforced by Patnaik et al. (2020) findings that used the MOORA to juxtapose the material alternatives and correctly order them. Further, this study's evidence stated how a rule's functionality based on the IF-THEN feature's logical use provided the optimum recommendations from entangled rank alternatives of two sides perceives. In that sense, the rule base exhibited the ability in dealing with the complexity of possible cases and uncertainty in students placement intentions. Indeed, the algorithm of the decision tree, therefore, ensures the transparency of the student placement.

System Development

A prototype of MOORA-Rulebased-DSS system as a management information system is designed and developed to automate the procedure of FEP, starting from the registration, management data of students and schools, management data of criteria and alternatives provided, management stakeholders internal and external as users, and integrated MOORA and Rule-based calculation procedures. The MOORA-Rulebased-DSS system architecture can be seen in Figure 4. Two key figures, namely the FEP committee and administrator, have engaged in this application to date. The administrator has access to the whole process of knowledge based. In the meantime, the committee received the recommendations suggested by the application.

Testing

To evaluate the MOORA-Rulebased-DSS, a UAT assessment’s survey of twenty respondents came from five of the committees, eight study program leaders, two management schools, and five management faculty. The respondents are asked their perceives on the acceptance of MOORA-Rulebased-DSS application by considering the interactive design of system

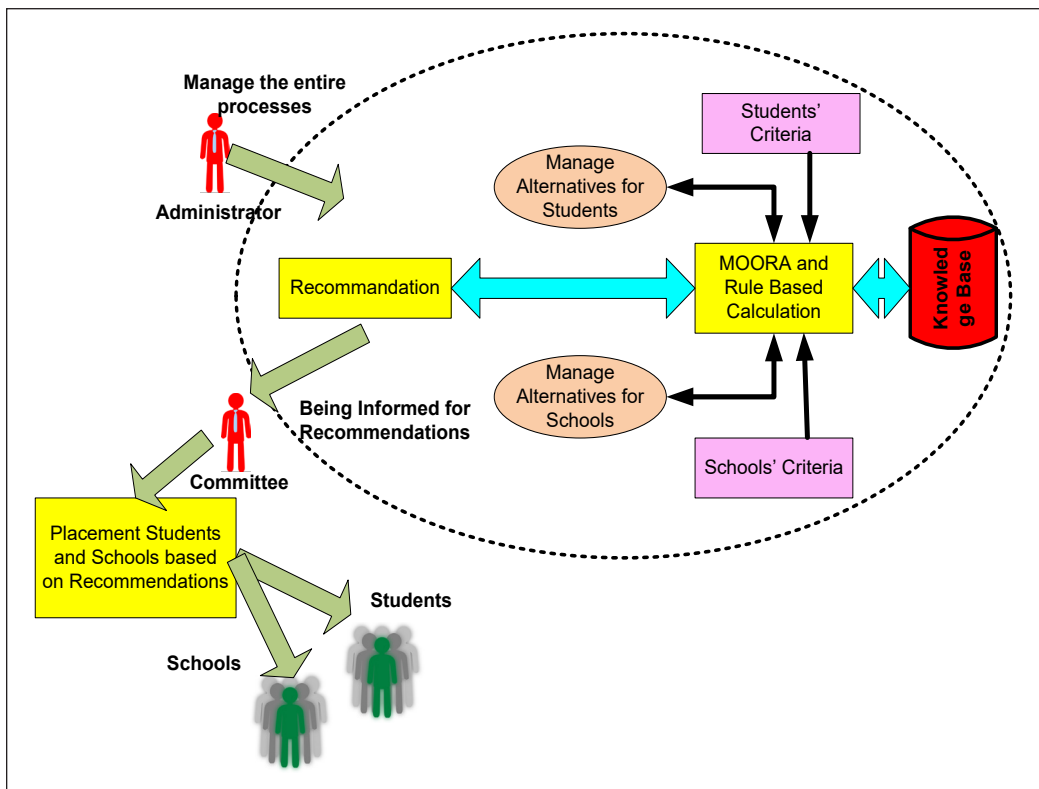


Figure 4. System architecture of MOORA-Rulebased-DSS

interface (5 questions), the easy use of DSS system (5 questions), the system utilization in aiding management decision making (5 questions), and the agreement on the proposed DSS system recommendation (5 questions). As a result, it obtained a very agreeable response of 86.92% covenant. Subsequently, Blackbox testing is conveyed into several DSS system functions, including login, user updated data, students updated data alternatives, schools updated data alternative, MOORA execution, Rule-based execution, and MOORA-Rulebased execution process. The testing revealed that 100% functionality and codifying test following the requirements analysis phase and user expectations. A Confusion Matrix is then generated by referring to Equation 5 and 6. As a result, the Confusion Matrix achieved an accuracy value of 78.33% from 94 data tests with status "True" and 21.67% error rate from 26 data sets with conditional status "False". The calculation shows that this recommendation system succeeded in providing the most optimal advice in students' placement in the FEP.

CONCLUSION

This research has prevailed in propounding the optimal students' placement in the FEP. The multiple perspectives based on committee and schools have been accommodated by considering criteria and the adherence of MOORA calculation and rule based. The inclusion of rule-based has been successfully intensified the role of MOORA in optimizing the decisions. This recommendation system puts forward eight group suggestions as alternatives in placing the students for the FEP. The evaluation reveals the stakeholders' satisfaction and acceptance of the procedures and alternatives proposed. Hence, the emerging obstacles during the FEP can be minimized, and the stakeholders' amusement will be increased. Besides, the advancement of traineeship procedure and the suggested solution's objectivity impact the committee's performance towards the optimal, effective, and efficient services, especially in decision-making. Due to the constraints of user capabilities and knowledge from the committee, students, and schools' side on DSS system operation, the administration's role is playing significant values in managing the knowledge-based and data modelling. Therefore, future studies are encouraged to design the dynamic and smart DSS system. Therefore, all stakeholders will be directly involved and manage the knowledge based on the criteria and alternatives provided. Consequently, the efficiency of rule-based in complex environment tracking will linearly increase and powerful. Besides, the smart DSS system can accommodate more valuable perspectives.

ACKNOWLEDGEMENT

The authors fully acknowledged Faculty Science and Technology and Faculty Education and Teacher Training Universitas Islam Negeri Sultan Syarif Kasim Riau for data collection. It makes a high contribution to this research to be more viable and effective.

REFERENCES

- Akram, M., Ilyas, F., & Garg, H. (2019). Multi-criteria group decision making based on ELECTRE-I method on Pythagorean fuzzy information. *Soft Computing*, 24(5), 3425-3453. <https://doi.org/10.1007/s00500-019-04105-0>
- Arabsheybani, A., Paydar, M. M., & Safaei, A. S. (2018). An integrated fuzzy MOORA method and FMEA technique for sustainable supplier selection considering quantity discounts and supplier's risk. *Journal of Cleaner Production*, 190, 577-591. <https://doi.org/10.1016/j.jclepro.2018.04.167>
- Ardana, I. M., Ariawan, I. P. W., & Divayana, D. G. H. (2016). Development of decision support system to selection of the blended learning platforms for mathematics and ICT learning at SMK TI Udayana. *International Journal of Advanced Research in Artificial Intelligence*, 5(12), 15-18.
- Bonjar, H. S., Salajegheh, S., Pourkiani, M., & Sayadi, S. (2019). Identifying the dimensions and sustainable supply chain operations for knowledge-based decision making in Iran's tax affairs organization. *International Journal of Supply Chain Management*, 8(2), 735-743.
- Brauers, W. K. M. (2008). Multi-objective contractor's ranking by applying the MOORA method. *Journal of Business Economics and Management*, 9(4), 245-255.
- Brauers, W. K. M., & Zavadskas, E. K. (2006). The Moora method and its application to privatization in a transition economy. *Control and Cybernetics, Systems Research Institute of the Polish Academy of Sciences*, 35(2), 445-469.
- Büyüközkan, G., & Güler, M. (2020). Smart watch evaluation with integrated hesitant fuzzy linguistic SAW-ARAS technique. *Measurement*, 153(March 2020), Article 107353. <https://doi.org/10.1016/j.measurement.2019.107353>
- Carneiro, J., Alves, P., Marreiros, G., & Novais, P. (2020). Group decision support systems for current times: Overcoming the challenges of dispersed group decision-making. *Neurocomputing*, 423, 735-746. <https://doi.org/10.1016/j.neucom.2020.04.100>
- Chakraborty, S. (2011). Applications of the MOORA method for decision making in manufacturing environment. *International Journal of Advanced Manufacturing Technology*, 54(9-12), 1155-1166. <https://doi.org/10.1007/s00170-010-2972-0>
- Chan, C. K., & Yeung, N. C. J. (2019). Students' approach to develop in holistic competency: An adaption of the 3P model. *An International Journal of Experimental Educational Psychology*, 40(5), 1-21. <https://doi.org/10.1080/01443410.2019.1648767>
- Cherchye, L., De Witte, K., Ooghe, E., & Nicaise, I. (2010). Efficiency and equity in private and public education: A nonparametric comparison. *European Journal of Operational Research*, 202, 563-573. <https://doi.org/10.1016/j.ejor.2009.06.015>
- Conejero, J. M., Preciado, J. C., Prieto, A. E., Bas, M. C., & Bolós, V. J. (2020). Applying data driven decision making to rank vocational and educational training programs with TOPSIS. *Decision Support Systems*, 142, 113470. <https://doi.org/10.1016/j.dss.2020.113470>
- Copeland, L. (2004). *A practitioner's guide to software test design*. Artech House.

- Dabbagh, R., & Yousefi, S. (2019). A hybrid decision-making approach based on FCM and MOORA for occupational health and safety risk analysis. *Journal of Safety Research*, 71, 111-123. <https://doi.org/10.1016/j.jsr.2019.09.021>
- Davis, D. J., & Ringsted, C. (2006). Accreditation of undergraduate and graduate medical education: How do the standards contribute to quality? *Advances in Health Sciences Education*, 11(3), 305-313. <https://doi.org/10.1007/s10459-005-8555-4>
- Davis, F. D., & Venkatesh, V. (2004). Toward pre prototype user acceptance testing of new information systems: Implications for software project management. *IEEE Transactions on Engineering Management*, 51(1), 31-46. <https://doi.org/10.1109/TEM.2003.822468>
- De Farias Aires, R. F., & Ferreira, L. (2019). A new approach to avoid rank reversal cases in the TOPSIS method. *Computers & Industrial Engineering*, 132, 84-97. <https://doi.org/10.1016/j.cie.2019.04.023>
- Delen, D., Topuz, K., & Eryarsoy, E. (2020). Development of a Bayesian belief network-based DSS for predicting and understanding freshmen student attrition. *European Journal of Operational Research*, 281(3), 575-587. <https://doi.org/10.1016/j.ejor.2019.03.037>
- Dinçer, H., Yüksel, S., & Martínez, L. (2019). Interval type 2-based hybrid fuzzy evaluation of financial services in E7 economies with DEMATEL-ANP and MOORA methods. *Applied Soft Computing*, 79, 186-202. <https://doi.org/10.1016/j.asoc.2019.03.018>
- Du, Y., Zheng, Y., Wu, G., & Tang, Y. (2020). Decision-making method of heavy-duty machine tool remanufacturing based on AHP-entropy weight and extension theory. *Journal of Cleaner Production*, 252, Article 119607.
- El-Gayar, O. F., & Fritz, B. D. (2010). A web-based multi-perspective decision support system for information security planning. *Decision Support Systems*, 50(1), 43-54.
- Emovon, I., Okpako, O. S., & Edjokpa, E. (In Press). Application of fuzzy MOORA method in the design and fabrication of an automated hammering machine. *World Journal of Engineering*. <https://doi.org/10.1108/WJE-07-2020-0250>
- Engel, M. M., Utomo, W. H., & Purnomo, H. D. (2017). Fuzzy multi attribute decision making – simple additive weighting (MADM - SAW) for information retrieval (IR) in e - commerce recommendation. *International Journal of Computer Science and Software Engineering*, 6(6), 136-145.
- Fei, L., Xia, J., Feng, Y., & Liu, L. (2019). An Electre-based multiple criteria decision-making method for supplier selection using Dempster-Shafer theory. *IEEE Access*, 7, 84701-84716. <https://doi.org/10.1109/ACCESS.2019.2924945>
- Fox, W. P., Ormond, B., & Williams, A. (2015). Ranking terrorist targets using a hybrid AHP-TOPSIS methodology. *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 13(1), 77-93. <https://doi.org/10.1177/1548512914563619>
- Gadakh, V. S., Shinde, V. B., & Khemnar, N. S. (2013). Optimization of welding process parameters using MOORA method. *International Journal of Advanced Manufacturing Technology*, 69(9-12), 2031-2039. <https://doi.org/10.1007/s00170-013-5188-2>

- Gang, Z., Juanjo, M., & Carmen, J. (2020). Chinese student teachers' teaching practicum experiences: Insights from transformative learning, third space, and dialogical-self theory. *International Journal of Educational Research*, 103, Article 101638.
- Gràcia, M., Casanovas, J., Riba, C., Sancho, M. R., Jarque, M. J., Casanovas, J., & Vega, F. (2020). Developing a digital application (EVALOE-DSS) for the professional development of teachers aiming to improve their students' linguistic competence. *Computer Assisted Language Learning*, 1-26. <https://doi.org/10.1080/09588221.2019.1707690>
- Gündoğdu, F. K., & Kahraman, C. (2019). A novel fuzzy TOPSIS method using emerging interval-valued spherical fuzzy sets. *Engineering Applications of Artificial Intelligence*, 85, 307-23. <https://doi.org/10.1016/j.engappai.2019.06.003>
- Han, J., Pei, J., & Kamber, M. (2011). *Data mining techniques* (3rd Ed.). British Library.
- Hemdi, M. A., Salehuddin, M. Z. M., & Hafiz, M. H. (2016, December 1-2). Integrated cumulative grade point average (iCGPA): Concepts, challenges, and reality. In *Proceedings of the Conference Heads of Academic Department (CHAD)*. ILD Nilai, UiTM, Malaysia.
- Hsu, W. C. J., Liou, J. J. H., & Lo, H. W. (2020). A group decision-making approach for exploring trends in the development of the healthcare industry in Taiwan. *Decision Support Systems*, 141, Article 113447. <https://doi.org/10.1016/j.dss.2020.113447>
- Ibrahim, O., Sundgren, D., & Larsson, A. (2014). An integrated decision support system framework for strategic planning in higher education institutions. In *Proceedings of Joint International Conference on Group Decision and Negotiation: Group Decision and Negotiation. A Process-Oriented View* (pp. 237-238). Springer. https://doi.org/10.1007/978-3-319-07179-4_27
- Jones, A., Korson, S. J., Murray-Everett, N. C., Kessler, M. A., Harrison, D., Cronenberg, S., Parrish, M. L., & Parsons, M. J. (2020). Teacher candidate tensions with the edTPA: A neoliberal policy analysis. In *The Educational Forum* (pp. 1-19). Routledge. <https://doi.org/10.1080/00131725.2020.1790705>
- Julie, B., & Katie, Z. (2017). A good news story: Early-career music teachers' accounts of their flourishing professional identities. *Teaching and Teacher Education*, 68, 241-251. <https://doi.org/10.1016/j.tate.2017.08.009>
- Kacprzak, D. (2019). A doubly extended TOPSIS method for group decision making based on ordered fuzzy numbers. *Expert System with Applications*, 116, 243-254. <https://doi.org/10.1016/j.eswa.2018.09.023>
- Kazemi, F., Bahrami, A., & Sharif, J. A. (2020). Mineral processing plant site selection using integrated fuzzy cognitive map and fuzzy analytical hierarchy process approach: A case study of gilsonite mines in Iran. *Minerals Engineering*, 147, Article 106143. <https://doi.org/10.1016/j.mineng.2019.106143>
- Kilic, H. S., Demirci, A. E., & Delen, D. (2020). An integrated decision analysis methodology based on IF-DEMATEL and IF-ELECTRE for personnel selection. *Decision Support Syst*, 137, Article 113360. <https://doi.org/10.1016/j.dss.2020.113360>
- Krisnawati, Y., Irawati, M., Rochman, F., & Budiasih, E. (2015). The implementation of students' campaign program to form Adiwiyata school in Malang, Indonesia. *International Journal of Research Studies in Education*, 4(4), 53-65.

- Kumar, N., Singh, T., Grewal, J. S., Patnaik, A., & Fekete, G. (2019). A novel hybrid AHP-SAW approach for optimal selection of natural fiber reinforced non-asbestos organic brake friction composites. *Materials Research Express*, 6(6), Article 065701.
- Leny, N., & Okfalisa. (2019). Measuring the sustainability performance of Islamic banking in Indonesia. *Pertanika Journal Social Science & Humanities*, 27(2), 1073-1090.
- Lewis, H. G., & Brown, M. (2001). A generalized confusion matrix for assessing area estimates from remotely sensed data. *International Journal of Remote Sensing*, 22(16), 3223-3235. <https://doi.org/10.1080/01431160152558332>
- Li, Z., Chow, D. H. C., Ding, D., Ying, J., Hu, Y., Chen, H., & Zhao, W. (2020). The development and realization of a multi-faceted system for green building planning: A case in Ningbo using the fuzzy analytical hierarchy process. *Energy & Buildings*, 226, Article 110371. <https://doi.org/10.1016/j.enbuild.2020.110371>
- Liang, D., & Xu Z. (2017). The new extension of TOPSIS method for multiple criteria decisions making with hesitant Pythagorean fuzzy sets. *Applied Soft Computing*, 60, 167-179. <https://doi.org/10.1016/j.asoc.2017.06.034>
- Liang, Z. L., Yun, T. J., Oh, W. B., Lee, B. R., & Kim, I. S. (2020). A study on MOORA-based Taguchi method for optimization in automated GMA welding process. *Materials Today: Proceedings*, 22, 1778-1785. <https://doi.org/10.1016/j.matpr.2020.03.011>
- Liao, H. C., Wu, X., Mi, X., & Herrera, F. (2020). An integrated method for cognitive complex multiple experts' multiple criteria decisions making based on ELECTRE III with weighted Borda rule. *Omega*, 93(C), Article 102052. <https://doi.org/10.1016/j.omega.2019.03.010>
- Liu, D., & Stewart, T. J. (2004). Object-oriented decision support system modelling for multicriteria decision making in natural resource management. *Computers & Operations Research*, 31(7), 985-999. [https://doi.org/10.1016/S0305-0548\(03\)00047-9](https://doi.org/10.1016/S0305-0548(03)00047-9)
- Madić, M., Radovanović, M., & Petković, D. (2015). Non-conventional machining processes selection using multi-objective optimization on the basis of ratio analysis method. *Journal of Engineering Science and Technology*, 10(11), 1441-1452.
- Mangla, S. K., Govindan, K., & Luthra, S. (2017). Prioritizing the barriers to achieve sustainable consumption and production trends in supply chains using fuzzy Analytical Hierarchy Process. *Journal of Cleaner Production*, 151, 509-525. <https://doi.org/10.1016/j.jclepro.2017.02.099>
- Martusorn, K. H., Chattrarat, K. B., & Shaharudin, M. R. (2019). Comparative analysis of multiple criteria decision making (MCDM) approach in warehouse location selection of agricultural products in Thailand. *International Journal of Supply Chain Management*, 8(5), 168-175.
- Memari, A., Dargi, A., Jokar, M. R. A., Ahmad, R., & Rahim, A. R. A. (2019). Sustainable supplier selection: A multi-criteria intuitionistic fuzzy TOPSIS method. *Journal of Manufacturing Systems*, 50, 9-24. <https://doi.org/10.1016/j.jmsy.2018.11.002>
- Mishra, A. R., Singh, R. K., & Motwani, D. (2020). Intuitionistic fuzzy divergence measure-based ELECTRE method for performance of cellular mobile telephone service providers. *Neural Computing and Applications*, 32, 3901-3921. <https://doi.org/10.1007/s00521-018-3716-6>

- Mistarihi, M. Z., Okour, R. A., & Mumani, A. A. (2020). An integration of a QFD model with Fuzzy-ANP approach for determining the importance weights for engineering characteristics of the proposed wheelchair design. *Applied Soft Computing Journal*, 90, Article 106136. <https://doi.org/10.1016/j.asoc.2020.106136>
- Mukodimah, S., Muslihudin, M., Andoyo, A., Hartati, S., & Maselena, A. (2018). Fuzzy simple additive weighting and its application to toddler healthy food. *International Journal of Pure and Applied Mathematics*, 118(7), 1-7.
- Nidhra, S. (2012). Black box and white box testing techniques-A literature review. *International Journal of Embedded Systems and Applications*, 2(2), 29-50. <https://doi.org/10.5121/ijesa.2012.2204>
- Okfalisa, O., Anugrah, S., Anggraini, W., Absor, M., Fauzi, S. S. M., & Saktioto, S. (2018). Integrated analytical hierarchy process and objective matrix in balanced scorecard dashboard model for performance measurement. *Telecommunication Computing Electronics and Control*, 16(6), 2703-2711.
- Okfalisa, O., Hidayati, R., Dwi, U. I., Pranggono, B., Elin, H., & Saktioto, S. (2021). Decision support system for smartphone recommendation: The comparison of fuzzy AHP and fuzzy ANP in multi-attribute decision making. *Journal SINERGI*, 25(1), 101-110. <https://doi.org/10.22441/sinergi.2021.1.013>
- Okfalisa, O., Zulfahri, Iswavigra, D. U., Megawati, & Saktioto. (2020). Online shop recommendations: decision support system based on multi objective optimization on the basis of ratio analysis. In *Proceeding in 8th International Conference on Cyber and IT Service Management (CITSM)* (pp. 1-6). IEEE Conference Publication. <https://doi.org/10.1109/CITSM50537.2020.9268848>
- Omrani, H., Alizadeh, A., & Amini, M. (2019). A new approach based on BWM and MULTIMOORA methods for calculating semi-human development index: An application for provinces of Iran. *Socio-Economic Planning Sciences*, 70(C), Article 100689. <https://doi.org/10.1016/j.seps.2019.02.004>
- Pardiyono, R., & Indrayani, R. (2019). Decision support system to choose private higher education based on marketing mix model criteria in Indonesia. In *IOP Conference Series: Materials Science and Engineering* (Vol. 508, No. 1, p. 012112). IOP Publishing. <https://doi.org/10.1088/1757-899X/508/1/012112>
- Patnaik, P. K., Swain, P. T. R., Mishra, S. K., Purohit, A., & Biswas, S. (2020). Composite material selection for structural applications based on AHP-MOORA approach. *Materials Today: Proceedings*, 33(8), 5659-5663. <https://doi.org/10.1016/j.matpr.2020.04.063>
- Petkov, D., Petkova, O., Andrew, T., & Nepal, T. (2007). Mixing multiple criteria decisions making with soft systems thinking techniques for decision support in complex situations. *Decision Support Systems*, 43(4), 1615-1629. <https://doi.org/10.1016/j.dss.2006.03.006>
- Pressman, R., & Maxim, B. (2020). *Software engineering: Practitioner's approach* (9th Ed.). McGraw-Hill.
- Roszkowska, E., & Kacprzak, D. (2016). The fuzzy SAW and fuzzy TOPSIS procedures based on ordered fuzzy numbers. *Information Science*, 369, 564-584. <https://doi.org/10.1016/j.ins.2016.07.044>
- Saban, A., & Coklar, A. N. (2013). Pre-service teachers' opinions about the micro-teaching. *The Turkish Online Journal of Educational Technology*, 12(2), 234-240.
- Selvi, K. (2010). Teachers' Competencies. *International Journal of Philosophy of Culture and Axiology*, 7(1), 167-176. <https://doi.org/10.5840/cultura20107133>

- Shihab, S. K., Khan, N. Z., Myla, P., Upadhyay, S., Khan, Z. A., & Siddiquee, A. N. (2018). Application of MOORA method for multi optimization of GMAW process parameters in stain-less steel cladding. *Management Science Letters*, 8(4), 241-246. <https://doi.org/10.5267/j.msl.2018.2.002>
- Stanujkic, D., Magdalinovic, N., Stojanovic, S., & Jovanovic, R. (2012). Extension of ratio system part of MOORA method for solving decision-making problems with interval data. *Informatica*, 23(1), 141-154.
- Sztubecka, M., Skiba, M., Mrówczyńska, M., & Bazan-Krzywoszańska, A. (2020). An innovative decision support system to improve the energy efficiency of buildings in urban areas. *Remote Sensing*, 12(2), Article 259. <https://doi.org/10.3390/rs12020259>
- Yazdani, M., Zarate, P., Coulibaly, A., & Zavadskas, E. K. (2017). A group decision making support system in logistics and supply chain management. *Expert Systems with Applications*, 88, 376-392. <https://doi.org/10.1016/j.eswa.2017.07.014>
- Yeo, L. S., Ang, R. P., Chong, W. H., Huan, V. S., & Quek, C. L. (2008). Teacher efficacy in the context of teaching low achieving students. *Current Psychology*, 27(3), 192-204. <https://doi.org/10.1007/s12144-008-9034-x>
- You, Y., Sun, J., Jiang, J., & Tan, Y. (2019). Belief-rule-base model with attribute reliability for multiple attribute decision making. In *Proceedings of Chinese Control and Decision Conference (CCDC)* (pp. 4770-4775). IEEE Conference Publication. <https://doi.org/10.1109/CCDC.2019.8832405>
- Yusuf, T. I. (2019). A multi-objective credit evaluation model using MOORA method and goal programming. *Arabian Journal of Science and Engineering*, 45, 2035-2048. <https://doi.org/10.1007/s13369-019-03930-7>
- Zendarski, N., Haebich, K., Bhide, S., Quek, J., Nicholson, J. M., Jacobs, K. E., Efrom, D., & Sciberras, E. (2020). Student-teacher relationship quality in children with and without ADHD: A cross-sectional community-based study. *Early Childhood Research Quarterly*, 51, 275-284. <https://doi.org/10.1016/j.ecresq.2019.12.006>

Assessing Malaysian University English Test (MUET) Essay on Language and Semantic Features Using Intelligent Essay Grader (IEG)

Wee Sian Wong* and Chih How Bong

Faculty of Computer Science & Information Technology, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia

ABSTRACT

Automated Essay Scoring (AES) refers to the Artificial Intelligence (AI) application with the “intelligence” in assessing and scoring essays. There are several well-known commercial AES adopted by western countries, as well as many research works conducted in investigating automated essay scoring. However, most of the products and research works are not related to the Malaysian English test context. The AES products tend to score essays based on the scoring rubrics of a particular English text context (e.g., TOEFL, GMAT) by employing their proprietary scoring algorithm that is not accessible by the users. In Malaysia, the research and development of AES are scarce. This paper intends to formulate a Malaysia-based AES, namely Intelligent Essay Grader (IEG), for the Malaysian English test environment by using our collection of two Malaysian University English Test (MUET) essay dataset. We proposed the essay scoring rubric based on its language and semantic features. We analyzed the correlation of the proposed language and semantic features with the essay grade using the Pearson Correlation Coefficient. Furthermore, we constructed an essay scoring model to predict the essay grades. In our result, we found that the language featured such as vocabulary count and advanced part of speech were highly correlated with the essay grades, and the language features showed a greater influence on essay grades than the semantic features. From our prediction model, we observed that the model yielded

better accuracy results based on the selected high-correlated essay features, followed by the language features.

Keywords: Artificial intelligence, automated essay scoring, intelligent system in education, machine learning, MUET, natural language processing

ARTICLE INFO

Article history:

Received: 30 November 2020

Accepted: 10 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.12>

E-mail addresses:

weesian.wong@gmail.com (Wee Sian Wong)

chbong@unimas.my (Chih How Bong)

* Corresponding author

INTRODUCTION

Automated Essay Scoring (AES) is a measurement technology in which computers evaluate written work (Shermis & Burstein, 2003). The purpose of the AES is to automate the essay scoring process, and thus reducing labor-intensive marking activities, overcoming time, cost, and yet assuring high assessment reliability. From the perspective of the Computer Science discipline, AES is a Natural Language Processing (NLP) application – a subfield of Artificial Intelligence (AI) focusing on the interactions between computers and human (natural) languages. The main task of AES is to “*intelligently*” classify the given essays into the discrete classes corresponding to the grades defined in the assessment; in another context, it is a document classification problem. The classification of essays is achievable based on the rationale that essays with similar grades shall possess similar characteristics in term of their style and content, which can be measured and quantified by essay characteristics such as words use, syntax structure, and content similarity which in turn can be generalized by employing computer algorithms and statistical methods.

Since the inception of the first AES, named Project Essay Grader (PEG) (Page, 1966), there were numerous commercial AES applications developed as the means of essay assessment in the western world. The typical examples of AES were Project Essay Grader (PEG) (Measurement Incorporated, n.d.), Intelligent Essay Assessor (IEA) (Pearson Education, 2010), IntelliMetric (Vantage Learning, n.d.), and e-rater (Educational Testing Service, n.d.). They employed several methods to assess different aspects of an essay. PEG mainly utilizes the observable essay features denoted as “*proxes*” to evaluate the writing style; and IEA focuses on assessing the essay content by Latent Semantic Analysis (Landauer et al., 1998). E-rater relies on NLP techniques to evaluate essays; while IntelliMetric claimed itself as the first AES that fully leverages various AI techniques in essay scoring. Both the e-rater and IntelliMetric focus on assessing the content and writing style of the essay.

Besides the commercial AES application, there are many research works carried out in investigating automated essay scoring. A large quantity of AES research works involves feature engineering to score a particular dimension of an essay. For example, Crossley and McNamara (2016) used the Tool for the Automatic Analysis of Cohesion (TAACO) to investigate the relationship of text cohesion and essay quality. Zupanc and Bosnić (2017) implemented AES by evaluating the semantic aspect of essays. Cozma et al. (2018) used word embedding based on a pre-trained large corpus to represent n-gram features for scoring essays’ lexical attributes. Other aspects of essay scoring can be found in works such as grammatical and mechanic errors detection (Crossley et al., 2019a), prompt adherence modeling (Persing & Ng, 2014), essay stance classification (Persing & Ng, 2016), argumentative structure modeling (Nguyen & Litman, 2018).

In the Malaysian context, there are very few active research works of AES, especially from the discipline of Computer Science. For all we know, there is one AES-related research works focusing on grammar checking of tenses in essays by using a heuristic approach. This work can be found in the literature by Maasum et al. (2012), and Omar et al. (2009). Besides the tenses error detection, other aspects of the essays such as content, language, and organization, are ignored in their work. Apart from this work, Wong and Bong (2019) carried out a feasibility study of adopting AES in the Malaysian English test context based on its reliability and validity requirement. They assessed a well-known AES called LightSide (LightSide, 2019) for determining its feasibility in scoring the Malaysian University English Test (MUET) essays. On the other hand, all other AES works in Malaysia are mainly related to adopting AES as the pedagogical or instructional tool, for studying the AES impact on students' writing skills. Such works include the impact study of Criterion - the instructional application of e-rater (Darus et al., 2003), MY Assess - the instructional interface of Intellimetric (Govindasamy et al., 2013), PaperRate.com (Manap et al., 2019), and Automated Essay Scorer with Feedback – AESF (Ng et al., 2019).

To the best of our knowledge, all the well-established AES, such as PEG, IEA, e-rater, and Intellimetric were emerged and grown in the western world, especially in the United State. They are all western software products, tied with the native English speaking environment and specific test setting. Such AES by their nature may not be suitable when applied to the Malaysian context, and probably of producing an invalid score. The inapplicability of adopting those AES in the Malaysian educational setting can be described by the characteristic below:

- (i) The Association of the AES Scoring Rubrics to a particular Test Setting (Wong & Bong, 2019)

The assessment criteria of the AES tend to be associated with the specific scoring rubrics of a particular English assessment context. For example, IntelliMetric's scoring rubric is formulated mostly based upon GMAT; while the e-rater's rubric is designed based on TOEFL.

- (ii) The Proprietary Nature of the AES (Wong & Bong, 2019)

Most, if not all of the AES is commercial software, incorporating the proprietary scoring algorithm, which is not accessible and understandable by the users.

- (iii) The lack of dataset standardization

All the well-known AES are developed by educational assessment institutions, using their essay dataset which is not available for public use. The Automated Student Assessment Prize (ASAP) corpus (Kaggle, 2012) is the only free available essay dataset that is widely used by most AES research works. However, this ASAP corpus incorporates different assessment criteria and does not contain any

paragraph information, which makes them inappropriate to be used for formulating AES based on the Malaysian educational context.

To address the issues, i.e. the absence of specific AES tailored for the Malaysian English educational setting, this research work is carried out to formulate a Malaysia-based AES, namely Intelligent Essay Grader (IEG). We characterized this IEG as the AES that is custom-made for Malaysian educational assessment and distinguish itself from other commercial proprietary AES software in the following aspects:

- (i) The IEG assesses essays based on its unique set of scoring rubrics associated with the specific Malaysian English test context.
- (ii) The IEG's scoring rubrics shall be transparent, understandable, and accessible by the end-users.
- (iii) The IEG employs a scoring function that is formulated by using the local Malaysian English test dataset. This local context dataset is essential for demonstrating the validity of our developed AES, i.e. it assesses what it claims to assess.

MATERIALS AND METHODS

Malaysian University English Test (MUET)

For the provision of the Malaysian English test context, we have chosen the Malaysian University English Test (MUET) as our use case study. MUET is the English test designed to measure the English language proficiency of pre-university students for entry into tertiary education (Malaysian Examination Council, 2014). It tests all the four language skills of Listening, Speaking, Reading, and Writing; and assess candidates' level of proficiency based upon an aggregated score range of zero to 300, which correlates with a banding system ranging from Band 1 to Band 6 (Malaysian Examination Council, 2014). The MUET writing paper comprises two writing tasks: transferring information from a non-linear source to a linear text and a piece of extended writing which may cover the rhetorical style of analytical, descriptive, persuasive, and argumentative. This study is carried out particularly to tackle the essay assessment of the second task in the MUET writing paper, i.e. the extended writing based on a given topic.

Scoring Rubrics of MUET Essay

In the MUET writing paper, students are assessed on their ability to write various types of text covering a range of rhetorical styles. As specified in the document of "*MUET Regulations, Test Specifications, Test Format and Sample Questions*" (Malaysian Examination Council, 2014), the scoring rubric of the MUET essay covers the aspect of accuracy, appropriacy, coherence and cohesion, use of language functions, and task fulfilment. Table 1 shows the specification of each scoring rubrics based on the document.

Table 1
Scoring rubrics of MUET essay

Scoring Rubrics	Test Specifications
Accuracy	<input type="checkbox"/> Using correct spelling and mechanics
	<input type="checkbox"/> Using correct grammar
	<input type="checkbox"/> Using correct sentence structures
Appropriacy	<input type="checkbox"/> Using varied vocabulary & expression
	<input type="checkbox"/> Using clear varied sentences
	<input type="checkbox"/> Using language appropriate for intended purpose and audience
	<input type="checkbox"/> Observing conventions appropriate to a specific situation or text type
Coherence and Cohesion	<input type="checkbox"/> Develop and organising ideas
	<input type="checkbox"/> Using appropriate markers and linking devices
	<input type="checkbox"/> Using anaphora appropriately together with cohesive devices
Use of Language Functions	<input type="checkbox"/> Defining, describing, explaining
	<input type="checkbox"/> Comparing and contrasting
	<input type="checkbox"/> Classifying
	<input type="checkbox"/> Giving reasons
	<input type="checkbox"/> Giving opinions
	<input type="checkbox"/> Expressing relationships
	<input type="checkbox"/> Make suggestion and recommendations
	<input type="checkbox"/> Expressing agreement and disagreement
	<input type="checkbox"/> Persuading
	<input type="checkbox"/> Interpreting information from non-linear texts
	<input type="checkbox"/> Drawing conclusion
	<input type="checkbox"/> Stating & justifying points of view
	<input type="checkbox"/> Presenting an argument
Task Fulfilment	<input type="checkbox"/> Presenting relevant ideas
	<input type="checkbox"/> Providing adequate content
	<input type="checkbox"/> Show a mature treatment of topic

Source: Malaysia Examination Council (2014)

Mapping of Essay Dimension with MUET Scoring Rubrics

In addressing such wide-scope and sometimes ambiguous MUET essays scoring rubrics as specified in Table 1, this work tries to formulate a scoring scheme for MUET essays mainly from two aspects, namely the language and semantic dimensions. As illustrated by Figure 1, we hypothesized that our proposed language and semantic dimensions should be correlated with the scoring rubrics of the MUET essays in Table 1. We justified our hypothesis as below:

- (i) It is common knowledge that the language dimension, such as spelling error, grammatical error, lexical richness, and parts of speech shall reflect the rubric of accuracy (e.g. using correct spelling and grammar), and appropriacy (e.g. using varied vocabulary and sentences) as specified in the MUET Test Specification.
- (ii) The semantic dimension is highly related to the coherence and cohesion rubrics in the MUET Test Specification. The close relationship of semantic dimensions

with coherence and cohesion is described in works by Foltz (2007), Janda, et al. (2019), Zupanc and Bosnić (2014).

- (iii) The semantic dimension which is defined as relating to the meaning of words, sentences, and texts can be associated with the use of language functions rubric (e.g. defining, describing and explaining, presenting argument), and the task fulfilment rubric which is particularized as content and topic treatment in the MUET Test Specification.

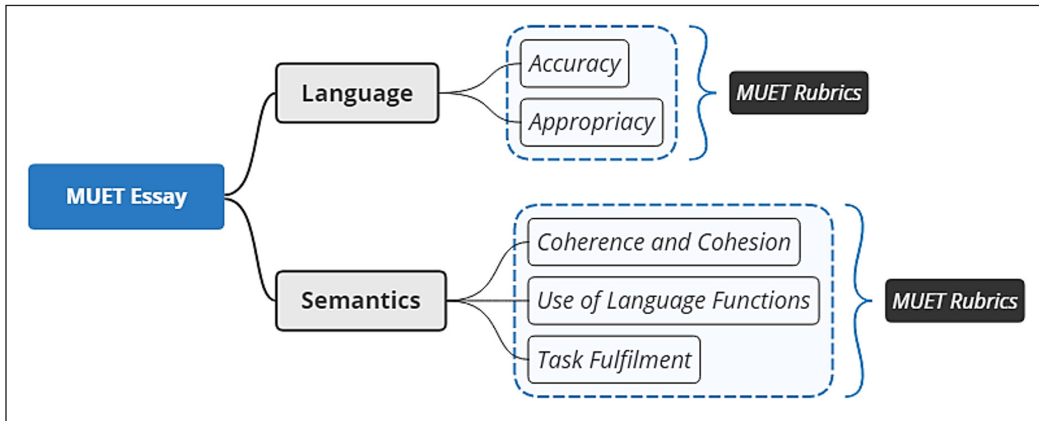


Figure 1. Mapping of IEG dimension with MUET scoring rubrics

Essay Features

As depicted by Figure 1, we formulated our IEG through the perspective of the language and the semantic dimension of the essay. We then represented the essay dimensions using three groups of features: language, local semantics, and global semantics. Each feature quantifies a particular attribute of the dimension, which overall devises a fine-grained and quantifiable approach in grading the essays.

Language. In our IEG model, the language dimension represents the language features and surface features of an essay. The language features cover the linguistic attributes of writing such as lexical richness, parts of speech, spelling error, and grammatical error; while the surface features are related to the attributes of length and “count” of the essay such as word count, sentence count, paragraph count, and average words per sentence. Table 2 specifies the 14 language features used in IEG.

Local Semantics. We categorized the essay semantic dimension into local and global semantics. Derived from the works of Crossley et al. (2019b), Crossley and McNamara (2016), we defined the local semantic dimension as the semantic similarity at the span of

Table 2
Language features of IEG

Features	Acronym	Description
Word Count	word_count	Number of Words
Sentence Count	sent_count	Number of Sentences
Paragraph Count	para_ccount	Number of Paragraphs
Average Words Per Sentence	word_sent	Word Count divided by Sentence Count
Average Words Per Paragraph	word_para	Word Count divided by Paragraph Count
Average Sentences Per Paragraph	sent_para	Sentence Count divided by Paragraph Count
Vocabulary Count	vocab_count	Count of Distinct Word
Lexical Richness	lexical_rich	Vocabulary Count divided by Word Count
Spelling Error	spell_err	Number of Spelling Error
Spelling Error Rate	spell_err_rate	Spelling Error divided by Word Count
Grammatical Error	grammar_err	Number of Grammatical Error
Grammatical Error Rate	grammar_err_rate	Grammatical Error divided by Word Count
Advanced Parts of Speech	adv_pos	Accumulated count of Adjective, Adverb, Past Participle, Present Participle Parts of Speech
Advanced Parts of Speech Rate	adv_pos_rate	Advanced Parts of Speech / Word Count

sentence-level (e.g. noun synonym overlap between adjacent sentences). We used eight features in representing the local semantics of an essay, which included the co-occurrence of WordNet Synonym (Miller, 1995) between sentences, the sentences similarity computed by LSA (Landauer et al., 1998), LDA (Blei et al., 2003) and Word2Vec (Mikolov et al., 2013), with the span of one and two-sentence chunk. Table 3 shows the eight local semantic features used in IEG.

Global Semantics. Based on the works of Crossley et al. (2019b), Crossley and McNamara (2016), we defined the global semantic dimension as semantic similarity at the paragraph level (such as the noun overlap between subsequent paragraphs). Similar to local semantics, we adopted eight features in representing the global semantics of an essay; including the co-occurrence of WordNet Synonym (Miller, 1995) between paragraph, the paragraph similarity computed by LSA (Landauer et al., 1998), LDA (Blei et al., 2003) and Word2Vec (Mikolov et al., 2013), with the span of one and two-paragraph chunk. Table 4 shows the eight global semantic features used in IEG.

Dataset

As one of the emphases in our study is to use the Malaysian English test dataset, it is essential that we trained and tested our IEG based on the local context dataset, instead of any public available essay dataset. In the experiment, we used the MUET dataset collected from several schools in Kuching, Sarawak. All the essays collected were written by Form Sixth

Table 3
Local semantic features of IEG

Features	Acronym	Description
Synonym Overlap (Sentence, Noun)	syn_overlap_sent_noun	Average sentence to sentence overlap of noun synonyms based on WordNet
Synonym Overlap (Sentence, Verb)	syn_overlap_sent_verb	Average sentence to sentence overlap of verb synonyms based on WordNet
LSA Cosine Similarity (Adjacent Sentences)	lsa_1_all_sent	Average Latent Semantic Analysis cosine similarity between all adjacent sentences
LSA Cosine Similarity (Two Adjacent Sentences)	lsa_2_all_sent	Average Latent Semantic Analysis cosine similarity between all adjacent sentences (with a two-sentence span)
LDA Divergence (Adjacent Sentences)	lda_1_all_sent	Average Latent Dirichlet Allocation divergence score between all adjacent sentences
LDA Divergence (Two Adjacent Sentences)	lda_2_all_sent	Average Latent Dirichlet Allocation divergence score between all adjacent sentences (with a two-sentence span)
Word2Vec Similarity (Adjacent Sentences)	word2vec_1_all_sent	Average word2vec similarity score between all adjacent sentences
Word2Vec Similarity (Two Adjacent Sentences)	word2vec_2_all_sent	Average word2vec similarity score between all adjacent sentences (with a two-sentence span)

Table 4
Global semantic features of IEG

Features	Acronym	Description
Synonym Overlap (Paragraph, Noun)	syn_overlap_para_noun	Average paragraph to paragraph overlap of noun synonyms based on WordNet
Synonym Overlap (Paragraph, Verb)	syn_overlap_para_verb	Average paragraph to paragraph overlap of verb synonyms based on WordNet
LSA Cosine Similarity (Adjacent Paragraphs)	lsa_1_all_para	Average Latent Semantic Analysis cosine similarity between all adjacent paragraphs
LSA Cosine Similarity (Two Adjacent Paragraphs)	lsa_2_all_para	Average Latent Semantic Analysis cosine similarity between all adjacent paragraphs (with a two-paragraph span)
LDA Divergence (Adjacent Paragraphs)	lda_1_all_para	Average Latent Dirichlet Allocation divergence score between all adjacent paragraphs
LDA Divergence (Two Adjacent Paragraphs)	lda_2_all_para	Average Latent Dirichlet Allocation divergence score between all adjacent paragraphs (with a two-paragraph span)
Word2Vec Similarity (Adjacent Paragraphs)	word2vec_1_all_para	Average word2vec similarity score between all adjacent paragraphs
Word2Vec Similarity (Two Adjacent Paragraphs)	word2vec_2_all_para	Average word2vec similarity score between all adjacent paragraphs (with a two-paragraph span)

students, which were then graded by qualified graders. In addition, to correctly represent the real-word MUET context, we used the real essay prompts between July and November 2014. The datasets consist of two essay prompts labeled as Dataset-1 and Dataset-2, with 259 essay samples in Dataset-1 and 200 essay samples in Dataset-2. The essay samples

Table 5
Grade distribution of MUET essay dataset

	Dataset-1	Dataset-2
Topic	The imbalance between the number of boys and girls pursuing university education creates social problems. To what extent is this statement true? Discuss.	Playing computer games is beneficial for everyone. Discuss.
Grade 1	20	0
Grade 2	73	17
Grade 3	130	103
Grade 4	15	67
Grade 5	11	12
Grade 6	10	1
Total Sample	259	200

are normally distributed with the highest frequency in the Grade-3 essay. Table 5 shows the corresponding grade distribution of the essay samples in Dataset-1 and Dataset-2.

Experiment Setting

There are two ultimate goals in this study:

- (i) To identify, analyze, and rank each essay feature that infers the essay grade.
- (ii) To attempt a machine learning framework to predict the essay grade based on the different feature groups.

To achieve the goals, we attempted to answer two specific questions of our automated essay scoring construct for MUET:

- (i) What are the essay features that are influencing the essay grades?
- (ii) What are the optimal essay features to be formulated in IEG for scoring MUET essays?

The answers to the questions are detailed in the following section, namely:

- (i) The Correlation of IEG Essay Features with MUET Essay Grades
- (ii) The Prediction of MUET Essay Grades using IEG

Figure 2 illustrates the overall process flow of the experiment carried out in this IEG work.

RESULTS AND DISCUSSIONS

The Correlation of Intelligent Essay Grader (IEG) Essay Features with MUET Essay Grades

To answer the question regarding the essay features which influence the MUET essay grades, we examined the correlation of each feature value with the essay grade. We hypothesize that the significant features will have a higher magnitude of correlation. By

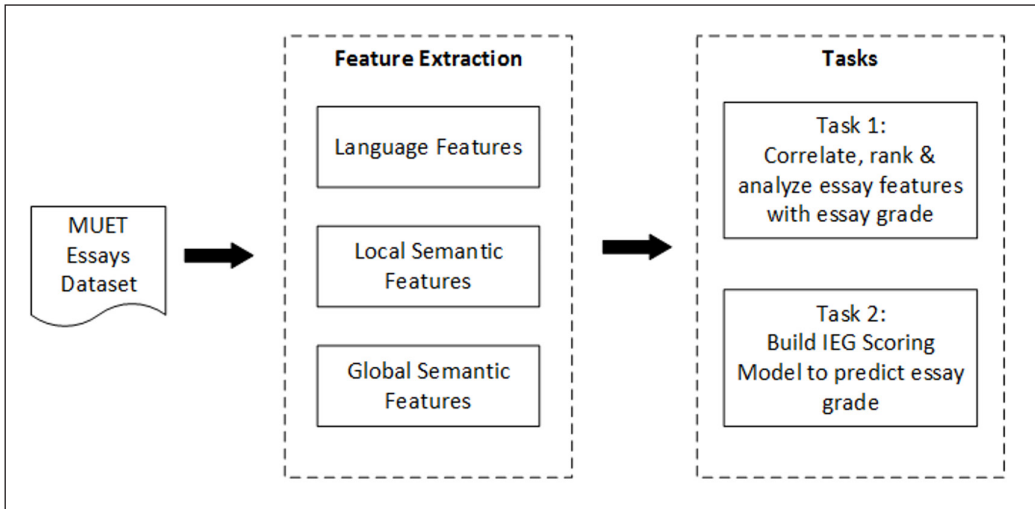


Figure 2. IEG process flow

using the Pearson Correlation Coefficient (Benesty et al., 2009), we investigated and discussed the relationship between the three essay dimensions of interest and the essay grades in the following aspects:

- (i) Essay features ranking by Pearson Correlation Coefficient
- (ii) Comparison of language dimension and semantic dimension effect on essay grades
- (iii) Comparison of local semantic dimension and global semantic dimension effect on essay grades

Evaluation Metric – Pearson Correlation Coefficient, r . We used the Pearson Correlation Coefficient, r (Benesty et al., 2009) as the metric to evaluate the correlation of essay features and essay grades. Pearson Correlation Coefficient measures the effect size r , i.e. the strength of the linear relationship between two variables. The coefficient value ranges from -1 to 1, with a value greater than 0 indicates a positive correlation, while a value less than 0 indicates a negative correlation. The larger the absolute value of the coefficient, the stronger the relationship between the variables. As represented by Equation 1 in the following, we quantified the strength of association between the essay features and the essay grades with this Pearson Correlation Coefficient. We used the Pearson Correlation Coefficient in this work as all the essay features (independent variables x) are continuous variable, whereby the essay grades (dependent variable y) is a continuous interval variable.

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad [1]$$

where

x_i = value of a particular feature for the i th essay

y_i = grade of the i th essay

\bar{x} = mean value of a particular essay feature

\bar{y} = mean value of essay grade

Essay Features Ranking by Pearson Correlation Coefficient. The sorted Pearson Correlation Coefficient, *r-value*, between each essay feature and essay grade is attached in Appendix A (Table A1). To further analyze and interpret the relationship of the essay features and the essay grades, we plotted the data into a line chart to illustrate the effect size of each essay feature against the essay grade. As shown in Figure 3, we can notice that the language features such as vocabulary count, advanced parts of speech, word count, and paragraph count are highly correlated to the essay grades. On the other hand, language features such as grammatical error and spelling error rate demonstrated a strong negative correlation to the essay grades.

From the plot, we found out that:

- (i) Essay features such as vocabulary count, advanced part of speech, word count, average words per paragraph, sentence count, and average sentences per paragraph are strongly positive-correlated with the essay grades. The higher of these features value indicates the higher essay grades. We also computed the average score for

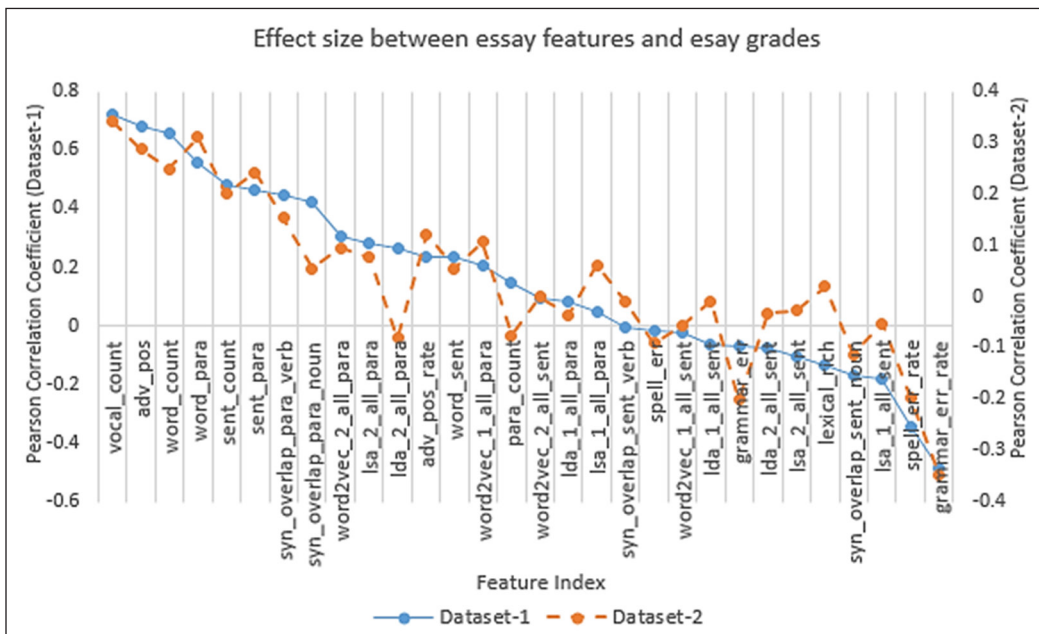


Figure 3. Effect size between essay features and essay grades

Table 6
Average score of highly correlation features in (a) Dataset-1, and (b) Dataset-2

(a)						
Essay Grade (y)	Essay Features (x)					
	vocal_count	adv_pos	word_count	word_para	sent_count	sent_para
1	79.7	31.2	190.5	41.7	12.3	12.3
2	132.8	58.0	351.1	66.5	21.0	21.0
3	157.8	70.3	430.4	73.2	25.4	25.4
4	190.9	86.3	509.1	88.8	28.3	28.3
5	270.6	128.7	646.5	112.2	33.6	33.6
6	293.3	142.6	752.8	142.1	35.3	35.3

(b)						
Essay Grade (y)	Essay Features (x)					
	vocal_count	adv_pos	word_count	word_para	sent_count	sent_para
1	N/A	N/A	N/A	N/A	N/A	N/A
2	183.2	85.6	488.8	83.1	25.1	4.3
3	192.2	86.7	480.8	86.5	26.3	4.7
4	209.7	95.4	530.6	95.4	28.6	5.2
5	249.4	124.7	626.6	110.3	31.8	5.6
6	235.0	88.0	508.0	127.0	27.0	6.8

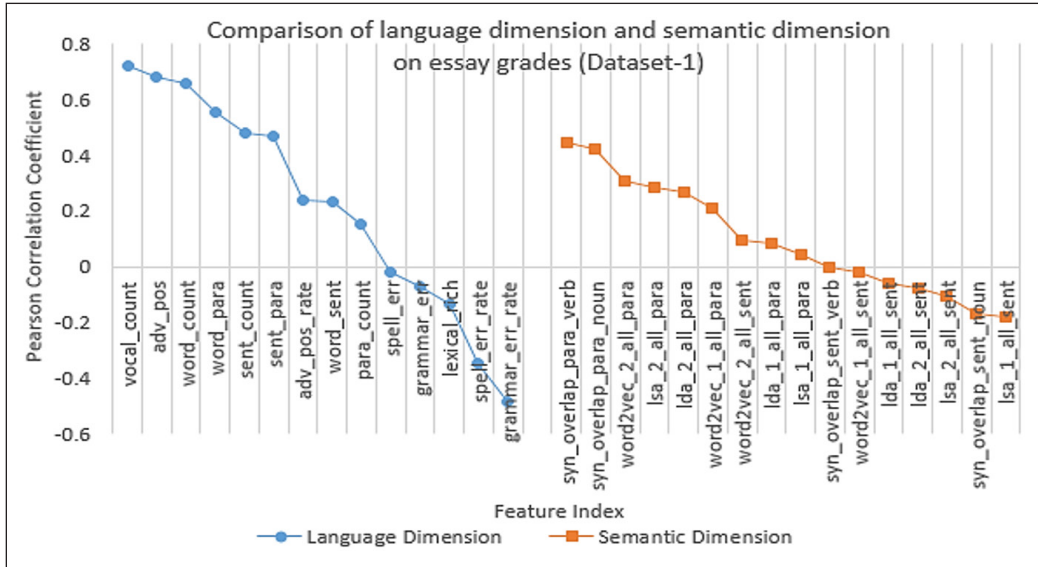
each feature as the following according to the essay grades. As shown in Table 6, we can notice that each average score shows an uptrend from the lower to higher essay grades.

- (ii) Essay features such as grammatical error rate and spelling error rate are strongly negative-correlated with the essay grades. The stronger these features value; the lower the essay grades. The error rate (grammatical error rate and spelling error rate), which shows a higher *r-value*, compared with the error count (grammatical error count and spelling error count), indicates the use of error rate instead of its absolute value is more effective in predicting essay grade.
- (iii) Holistically, the *r-value* of both Dataset-1 and Dataset-2 shows a consistent trend where the ranking of the sorted essay features are similar and close to each other for both the dataset. This observation justifies that our result of ranking the influential essay features are consistent and reliable.

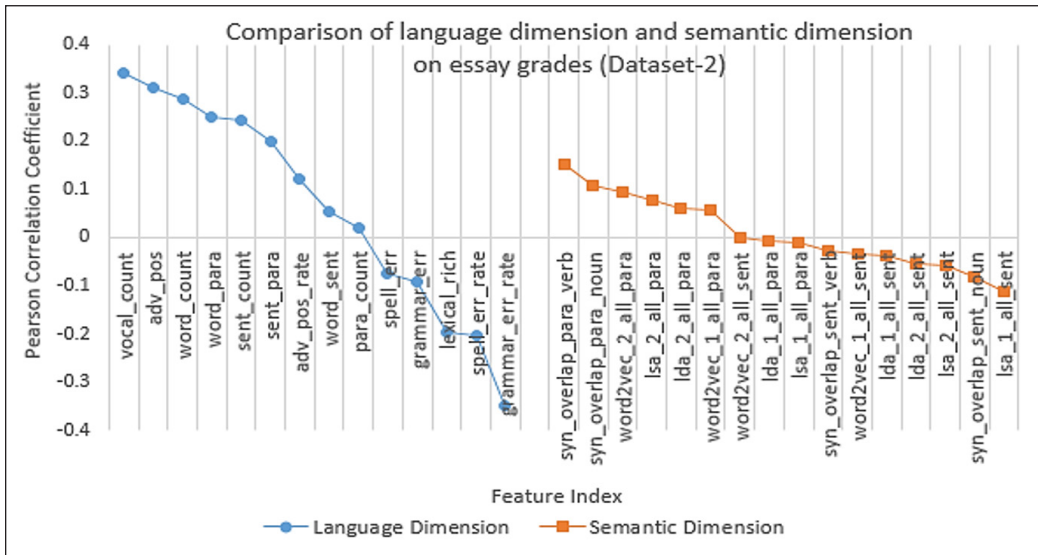
Comparison of Language Dimension and Semantic Dimension Effect on Essay Grades.

Figure 4 shows the comparison of language dimension and semantic dimension effect on essay grades in Dataset-1 and Dataset-2. From the plot, we observed that the language

dimension plot exhibits a greater gradient value with a steeper slope, compared with the semantic dimension plot. This greater gradient value denotes the language dimension covers a wider scope of effect size in both the positive and negative correlation, and thus indicate a greater influence on essay grades than the semantic dimension. This result of the higher impact of language features than semantic features in determining essay grades has been reported in works done by Crossley and McNamara (2011) and McNamara et al. (2010).



(a)

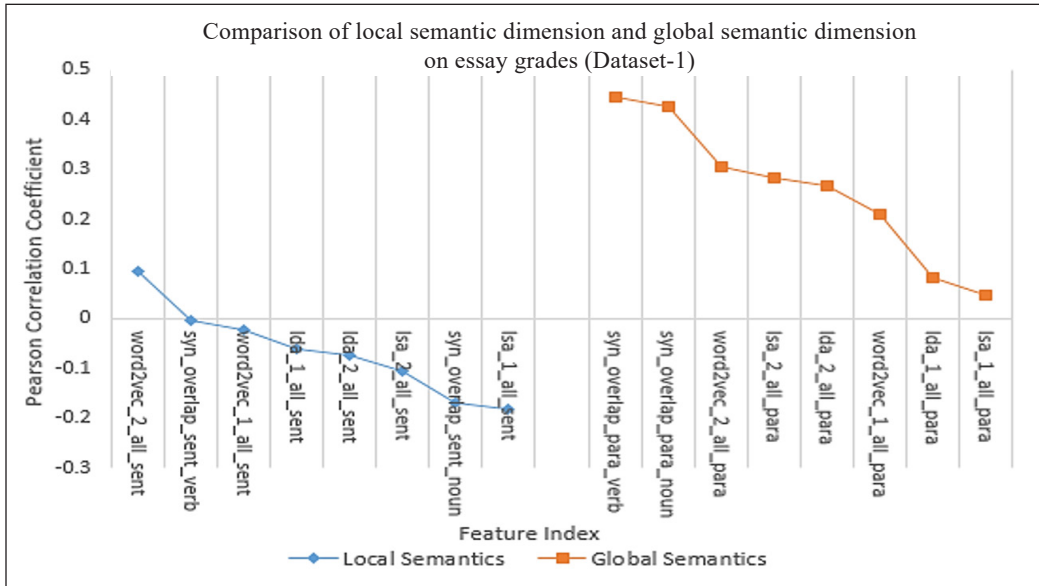


(b)

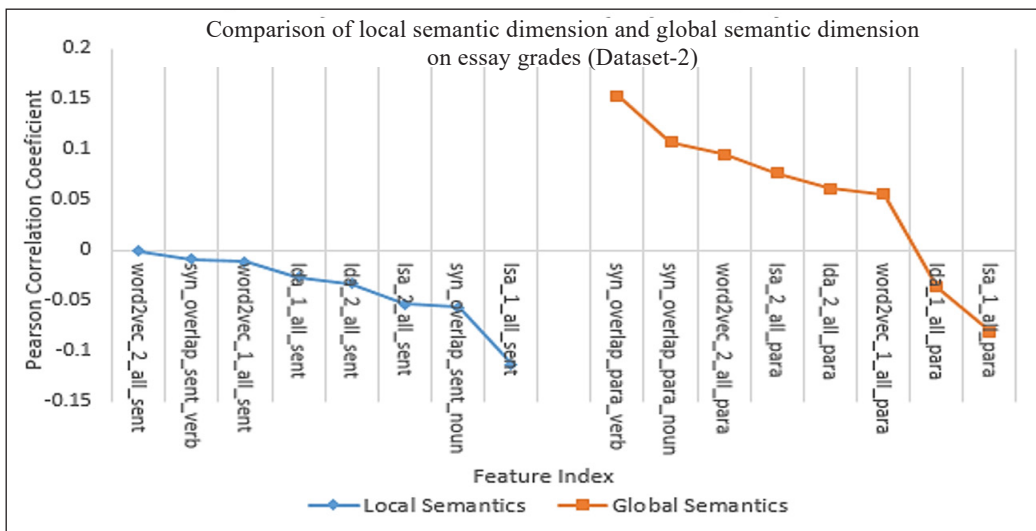
Figure 4. Comparison of language dimension and semantic dimension on essay grades in (a) Dataset-1, and (b) Dataset-2

Comparison of Local Semantic Dimension and Global Semantic Dimension Effect on Essay Grades. Figure 5 shows the comparison of local semantic dimension and global semantic dimension effect on essay grades in Dataset-1 and Dataset-2. From the plot, we observed that:

- (i) The global semantic dimension plot exhibits a greater gradient value with a steeper slope, compared with the local semantic dimension plot. This greater gradient



(a)



(b)

Figure 5. Comparison of local semantic dimension and global semantic dimension on essay grades in (a) Dataset-1, and (b) Dataset-2

value denotes the global semantic dimension covers a wider scope of effect and thus indicates a greater impact on essay grades than the local semantic dimension. This little to no effect of local semantic features on essay grade is supported by the studies carried out by Crossley and McNamara (2011) and McNamara et al. (2010).

- (ii) Most of the global semantic features show the *r-value* greater than 0. This finding denotes the global semantic dimension (corresponding to the paragraph scope) is positively correlated with the essay grades. The higher semantic similarity between paragraphs tends to produce higher essay grades. The work done by Crossley and McNamara (2016) exhibits the same observation of positive-correlation between global semantic features and essay quality.
- (iii) Most of the local semantic features show the *r-value* less than 0. This finding denotes the local semantic dimension (associated with the sentence scope) is negatively correlated with the essay grades. The higher semantic overlap between sentences tends to produce lower essay grades. The same finding of the negative-correlation between local semantic features and essay quality has been reported in the work done by Crossley and McNamara (2016).

The Prediction of MUET Essay Grades using Intelligent Essay Grader (IEG)

To investigate the optimal essay features in scoring MUET essays, we attempted a machine learning framework to predict the essay grade based on the different feature groups defined in our method. We treated the problem of essay grading as a multiclass classification task where each grade was represented as a class.

Intelligent Essay Grader (IEG) Scoring Model. Figure 6 illustrates the framework of our IEG Scoring Model. The gold standard of our experiment was the two MUET datasets which we collected. Each essay in the collection is preprocessed by removing all punctuations and stop-words, changing all words to lower case, and lemmatizing the words. We then computed and extracted the essay features from each essay. The extracted essay features were then fed into various machine learning classifiers for training the essay scoring model.

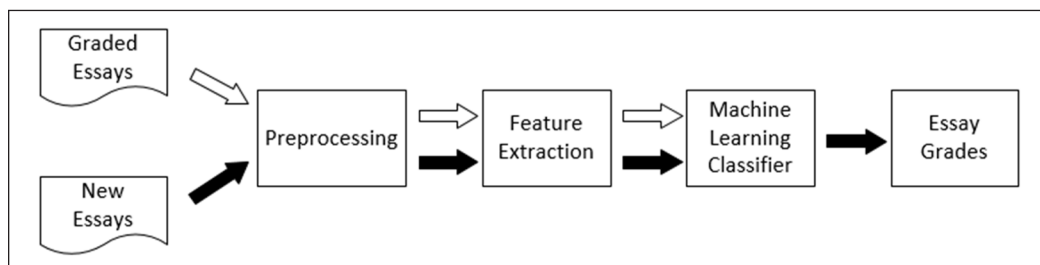


Figure 6. IEG scoring model

In the task, we employed four well-known machine learning classifiers, namely Logistic Regression (Cramer, 2002), Neural Network (Rumelhart et al., 1985), Random Forest (Breiman, 2001), and Support Vector Machine (Cortes & Vapnik, 1995), for building the IEG scoring model. Finally, we used the Leave-One-Out Cross Validation (Leave-One-Out Cross-Validation, 2011) to evaluate each of the machine learning classifiers in predicting the unseen essay grade.

To further investigate the classification accuracy based on different essay features, we built the IEG scoring model by using different features group as specified below:

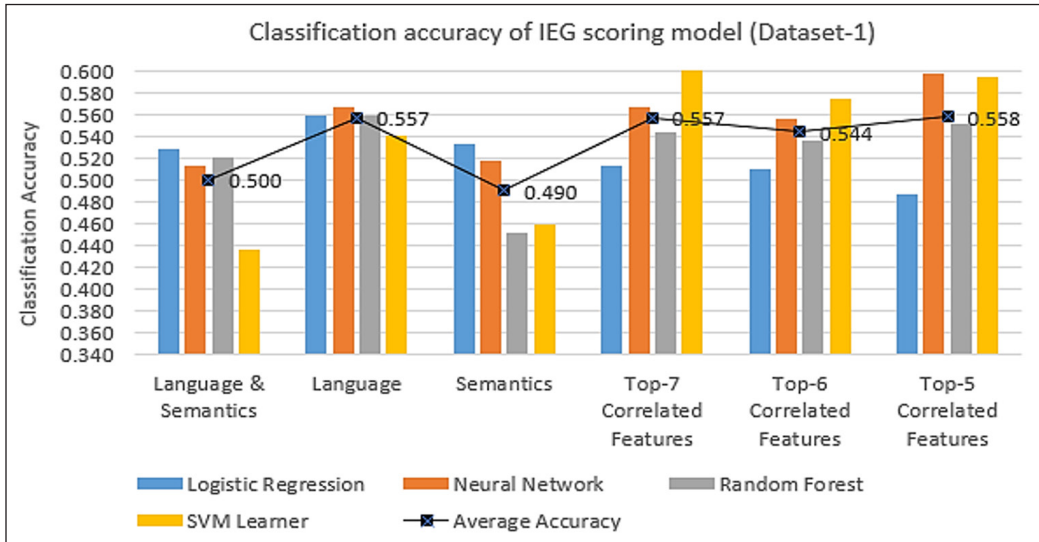
- (i) Language and Semantic Dimensions
All essay features as specified in Tables 2, 3, and 4.
- (ii) Language Dimension Only
The language features group as specified in Table 2.
- (iii) Semantic Dimension Only
The semantic features group as specified in Tables 3, and 4.
- (iv) Top-7 Correlated Features
The Top-7 essay features with the highest absolute correlation value as listed in Appendix A (Table A1) (*vocal_count*, *adv_pos*, *word_count*, *word_para*, *grammar_err_rate*, *sent_count*, *sent_para*).
- (v) Top-6 Correlated Features
The Top-6 essay features with the highest absolute correlation value as listed in Appendix A (Table A1) (*vocal_count*, *adv_pos*, *word_count*, *word_para*, *grammar_err_rate*, *sent_count*).
- (vi) Top-5 Correlated Features
The Top-5 essay features with the highest absolute correlation value as listed in Appendix A (Table A1) (*vocal_count*, *adv_pos*, *word_count*, *word_para*, *grammar_err_rate*).

Evaluation Metric – Classification Accuracy, CA. Classification accuracy, *CA* (Accuracy, 2017) measures the ratio of numbers of the correct predictions to the total numbers of input samples. We evaluated our formulated IEG scoring model based on this classification accuracy as expressed in Equation 2.

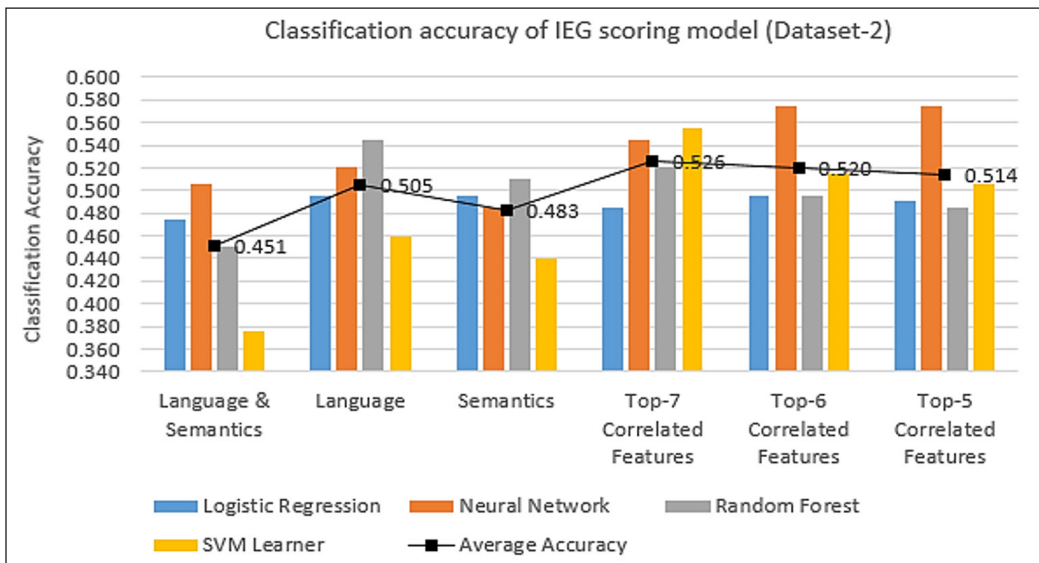
$$CA = \frac{\text{Numbers of correct prediction}}{\text{Total numbers of essays to be predicted}} \quad [2]$$

Results. The Classification Accuracy, *CA* of each essay feature group by employing various machine learning classifiers on Dataset-1 and Dataset-2 is attached in Appendix

B (Table B1). To facilitate effective visualization of the various CA values obtained by different essay feature groups, we plotted two combo charts (Figure 7); to show the CA values obtained in Dataset-1 and Dataset-2. In Figure 7, the CA values of each feature group with different machine learning classifiers are shown by the column chart; while the mean CA-values of a particular feature group averaged by the corresponding machine classifiers are shown in the line chart.



(a)



(b)

Figure 7. Classification accuracy of IEG scoring model in (a) Dataset-1, and (b) Dataset-2

Based on the average accuracy value in the plot, we observed the following results:

- (i) The IEG scoring model based on the selected high correlated essay features (either Top-7, Top-6, or Top-5 highly correlated features) shows better CA results, compared with the prediction based on the language and semantic dimension alone or the combination of both. This result is consistent with the essay feature ranking by effect size shown in Figure 3.
- (ii) The IEG scoring model based on the language dimension performed better than the scoring model based on the semantic dimension or the combination of both language and semantic dimensions. This result is consistent with the comparison of language dimension and semantic dimension effect size observed in Figure 4.
- (iii) Among the different feature groups, the semantic dimension and the combination of both semantic and language dimensions produce lower CA results. The semantic dimension yields the lowest CA value in Dataset-1; while the combination of both semantic and language dimensions produces the lowest CA value in Dataset-2.

CONCLUSION

This paper discusses our work in formulating an Automated Essay Scoring, namely Intelligent Essay Grader (IEG), based on the English assessment context in Malaysia. In our work, we employed a total of 30 essay features based on language and semantic dimensions in assessing the MUET essays. We studied the relationship of each essay feature with essay grade and built an essay scoring function in predicting the essay grade. Based on our experiment, we summarized our findings and contribution as below:

- (i) Instead of relying upon the publicly available corpus which does not reflect the essay writing in Malaysia, we collected and used the dataset of two essay prompts from MUET past year papers. This real-world dataset is essential in ensuring the validity of the constructed IEG.
- (ii) Compared with other features, the essay language features such as vocabulary count, advanced part of speech, word count, average words per paragraph are significantly positive-correlated with essay grades; while grammatical error and spelling error are significantly negative-correlated with essay grades.
- (iii) The essay language features show a higher effect on essay grades, compared with essay semantic features.
- (vi) The global semantic features of the essay indicate a greater impact on essay grades, compared with the local semantic features.
- (v) The global semantic features of the essay are mostly positively-correlated with the essay grades; while the local semantic features are mostly negatively-correlated with the essay grades.

- (vi) Our formulated IEG scoring function produces the classification accuracy results, which is consistent with our findings observed in the correlation of the essay features and the essay grades. The scoring function yields better accuracy results based on the selected high-correlated essay features, following by the language features.

Based on the findings, we perceived the essential need to improve the classification accuracy of the IEG scoring model. In our future work, we are planning to:

- (i) Refine the language and semantic features by removing the insignificant and redundant feature indexes to reduce the multicollinearity problem.
- (ii) Identify other feasible high-impact language and semantic features on essay grades.
- (iii) Incorporate other aspects of essay features such as essay content, organization, and argument strength for providing a comprehensive scope of IEG essay scoring.

ACKNOWLEDGEMENT

This study is supported and made possible by the RACE/B(6)/1098/2014(06) and ERGS/ICT07(01)/1018/2013(15) by the Malaysia Ministry of Higher Education. The authors would like to extend their greatest gratitude to the Malaysia Ministry of Higher Education for granting the funding to conduct this research.

REFERENCES

- Accuracy. (2017). Accuracy. In C. Sammut & G. I. Webb (Eds.) *Encyclopedia of machine learning and data mining* (pp. 1-48). Springer. https://doi.org/10.1007/978-1-4899-7687-1_3
- Benesty, J., Chen, J., Huang, Y., & Cohen, I. (2009). Pearson correlation coefficient. In *Noise reduction in speech processing* (pp. 1-4). Springer. https://doi.org/10.1007/978-3-642-00296-0_5
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273-297. <https://doi.org/10.1007/BF00994018>
- Cozma, M., Butnaru, A. M., & Ionescu, R. T. (2018). Automated essay scoring with string kernels and word embeddings. *Computation and Language*, 2018, 1-7.
- Cramer, J. S. (2002). The origins of logistic regression. *Tinbergen Institute Working Paper No. 2002-119/4*. <https://doi.org/10.2139/ssrn.360300>
- Crossley, S. A., & McNamara, D. S. (2011). Understanding expert ratings of essay quality: Coh-Metrix analyses of first and second language writing. *International Journal of Continuing Engineering Education and Life Long Learning*, 21(2-3), 170-191.

- Crossley, S. A., & McNamara, D. S. (2016). Say more and be more coherent: How text elaboration and cohesion can increase writing quality. *Journal of Writing Research, 7*(3), 351-370.
- Crossley, S. A., Bradfield, F., & Bustamante, A. (2019a). Using human judgments to examine the validity of automated grammar, syntax, and mechanical errors in writing. *Journal of Writing Research, 11*(2), 251-270.
- Crossley, S. A., Kyle, K., & Dascalu, M. (2019b). The tool for the automatic analysis of cohesion 2.0: Integrating semantic similarity and text overlap. *Behavioral Research Methods, 51*(1), 14-27. <https://doi.org/10.3758/s13428-018-1142-4>
- Darus, S., Stapa, S. H., & Hussin, S. (2003). Experimenting a computer-based essay marking system at Universiti Kebangsaan Malaysia. *Jurnal Teknologi, 39*(E), 1-18.
- Educational Testing Service. (n.d.). *About the e-rater® scoring engine*. Retrieved October 30, 2020, from <https://www.ets.org/erater/about>
- Foltz, P. W. (2007). Discourse coherence and LSA. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 167-184). Lawrence Erlbaum Associates.
- Govindasamy, P. N., Tan, B. H., & Yong, M. F. (2013). Lower six students' preferred mode of feedback for essay revision. *Malaysian Journal of ELT Research, 9*(2), 82-104.
- Janda, H. K., Pawar, A., Du, S., & Mago, V. (2019). Syntactic, semantic and sentiment analysis: The joint effect on automated essay evaluation. *IEEE Access, 7*, 108486-108503. <https://doi.org/10.1109/ACCESS.2019.2933354>
- Kaggle (2012). *The Hewlett foundation: Automated essay scoring*. Retrieved October 30, 2020, from <https://www.kaggle.com/c/ASAP-AES>
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes, 25*, 259-284. <https://doi.org/10.1080/01638539809545028>
- Leave-One-Out Cross-Validation. (2011). Leave-One-Out Cross-Validation. In C. Sammut, & G. I. Webb (Eds.) *Encyclopedia of machine learning*. Springer. https://doi.org/10.1007/978-0-387-30164-8_469
- LightSide. (2019). *LightSide researcher's workbench*. Retrieved January 11, 2021, from <http://ankara.lti.cs.cmu.edu/side>
- Malaysian Examination Council. (2014). *Malaysian University English Test (MUET) - regulations, test specifications, test format and sample questions*. Retrieved October 30, 2020, from https://www.mpm.edu.my/images/dokumen/calon-peperiksaan/muet/regulation/Regulations_Test_Specifications_Test_Format_and_Sample_Questions.pdf
- Manap, M. R., Ramli, N. F., & Kassim, A. A. M. (2019). Web 2.0 automated essay scoring application and human ESL essay assessment: A comparison study. *European Journal of English Language Teaching, 5*(1), 146-161. <https://doi.org/10.5281/zenodo.3461784>
- McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*(1), 57-86.
- Measurement Incorporated. (n.d.). *Automated Essay Scoring - Project Essay Grade (PEG®)*. Retrieved October 31, 2020, from <https://www.measurementinc.com/products-services/automated-essay-scoring>

- Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. *Advances in Neural Information Processing Systems*, 26, 3111-3119.
- Miller, G. A. (1995). WordNet: A lexical database for English. *Communications of the ACM*, 38(11), 39-41.
- Ng, S. Y., Bong, C. H., Hong, K. S., & Lee, N. K. (2019). Developing an automated essay scorer with feedback (AESF) for Malaysian University English Test (MUET): A design-based research approach. *Pertanika Journal of Social Science & Humanities*, 27(3), 1451-1468.
- Nguyen, H., & Litman, D. (2018). Argument mining for improving the automated scoring of persuasive essays. *Proceedings of the AAAI Conference on Artificial Intelligence*, 32(1), 5892-5899.
- Omar, N., Razali, N. A. M., & Darus, S. (2009) Automated grammar checking of tenses for ESL writing. In P. Wen, Y. Li, L. Polkowski, Y. Yao, S. Tsumoto, & G. Wang (Eds.), *Lecture notes in computer science, Vol 5589: Rough Sets and Knowledge Technology* (pp. 475-482). Springer. https://doi.org/10.1007/978-3-642-02962-2_60
- Page, E. B. (1966). The imminence of grading essays by computer. *The Phi Delta Kappan*, 47(5), 238-243.
- Pearson Education. (2010). *Intelligent Essay Assessor (IEA)™ Fact Sheet* [Fact sheet]. Retrieved October 31, 2020, from <https://images.pearsonassessments.com/images/assets/kt/download/IEA-FactSheet-20100401.pdf>
- Persing, I., & Ng, V. (2014). Modeling prompt adherence in student essays. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, 1, 1534-1543.
- Persing, I., & Ng, V. (2016). Modeling stance in student essays. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, 1, 2174-2184.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1985). *Learning internal representations by error propagation* (No. ICS-8506). California Univ San Diego La Jolla Inst for Cognitive Science.
- Shermis, M. D., & Burstein, J. (2003). Introduction. In M. D. Shermis & J. Burstein (Eds.), *Automated essay scoring: A cross-disciplinary perspective* (pp. xiii-xvi). Lawrence Erlbaum Associates.
- Maasum, T. N. R. T. M., Stapa, S. H., Omar, N., Aziz, M. J. A., & Darus, S. (2012). Development of an automated tool for detecting errors in tenses. *GEMA Online Journal of Language Studies*, 12(2), 427- 442.
- Vantage Learning, (n.d.). *Intellimetric®*. Retrieved October 31, 2020, from <http://www.intellimetric.com/direct>
- Wong, W. S., & Bong, C. H. (2019). A study for the development of automated essay scoring (AES) in Malaysian English test environment. *International Journal of Innovative Computing*, 9(1), 69-78. <https://doi.org/10.11113/ijic.v9n1.220>
- Zupanc, K., & Bosnic, Z. (2014). Automated essay evaluation augmented with semantic coherence measures. In R. Kumar, H. Toivonen, J. Pei, J. Z. Huang, & X. Wu (Eds.), *2014 IEEE International Conference on Data Mining* (pp. 1133-1138). IEEE Conference Publication. <https://doi.org/10.1109/ICDM.2014.21>
- Zupanc, K., & Bosnić, Z. (2017). Automated essay evaluation with semantic analysis. *Knowledge-Based Systems*, 120, 118-132. <https://doi.org/10.1016/j.knosys.2017.01.006>

APPENDIX A

Ranking of IEG essay features and MUET essay grades by Pearson Correlation Coefficient

Table A1

Ranking of IEG essay features and MUET essay grades by Pearson Correlation Coefficient

Feature Index	Essay Dimension	Pearson Correlation Coefficient, r		Feature Ranking	
		Dataset-1	Dataset-2	Dataset-1	Dataset-2
vocal_count	Language	0.724	0.342	1	1
adv_pos	Language	0.679	0.288	2	3
word_count	Language	0.658	0.249	3	4
word_para	Language	0.556	0.312	4	2
sent_count	Language	0.480	0.201	5	6
sent_para	Language	0.467	0.243	6	5
syn_overlap_para_verb	Global Semantic	0.445	0.153	7	7
syn_overlap_para_noun	Global Semantic	0.425	0.056	8	13
word2vec_2_all_para	Global Semantic	0.306	0.095	9	10
lsa_2_all_para	Global Semantic	0.283	0.077	10	11
lda_2_all_para	Global Semantic	0.268	-0.080	11	25
adv_pos_rate	Language	0.238	0.121	12	8
word_sent	Language	0.233	0.054	13	14
word2vec_1_all_para	Global Semantic	0.209	0.107	14	9
para_count	Language	0.150	-0.076	15	24
word2vec_2_all_sent	Local Semantic	0.096	0.000	16	16
lda_1_all_para	Global Semantic	0.082	-0.036	17	21
lsa_1_all_para	Global Semantic	0.047	0.062	18	12
syn_overlap_sent_verb	Local Semantic	-0.002	-0.011	19	18
spell_err	Language	-0.018	-0.091	20	26
word2vec_1_all_sent	Local Semantic	-0.021	-0.056	21	23
lda_1_all_sent	Local Semantic	-0.061	-0.008	22	17
grammar_err	Language	-0.069	-0.202	23	29
lda_2_all_sent	Local Semantic	-0.074	-0.033	24	20
lsa_2_all_sent	Local Semantic	-0.106	-0.027	25	19
lexical_rich	Language	-0.135	0.020	26	15
syn_overlap_sent_noun	Local Semantic	-0.167	-0.113	27	27
lsa_1_all_sent	Local Semantic	-0.183	-0.053	28	22
spell_err_rate	Language	-0.345	-0.197	29	28
grammar_err_rate	Language	-0.484	-0.349	30	30

APPENDIX B**Classification accuracy of IEG scoring model**

Table B1

Classification accuracy of IEG scoring model in (a) Dataset-1, and (b) Dataset-2

(a)						
	Language and Semantics	Language	Semantics	Top-7 Correlated Features	Top-6 Correlated Features	Top-5 Correlated Features
Logistic Regression	0.529	0.560	0.533	0.514	0.510	0.486
Neural Network	0.514	0.568	0.517	0.568	0.556	0.598
Random Forest	0.521	0.560	0.452	0.544	0.537	0.552
SVM Learner	0.436	0.541	0.459	0.602	0.575	0.595
Average Accuracy	0.500	0.557	0.490	0.557	0.544	0.558
(b)						
	Language & Semantics	Language	Semantics	Top-7 Correlated Features	Top-6 Correlated Features	Top-5 Correlated Features
Logistic Regression	0.475	0.495	0.495	0.485	0.495	0.490
Neural Network	0.505	0.520	0.485	0.545	0.575	0.575
Random Forest	0.450	0.545	0.510	0.520	0.495	0.485
SVM Learner	0.375	0.460	0.440	0.555	0.515	0.505
Average Accuracy	0.451	0.505	0.483	0.526	0.520	0.514



Artificial Neural Network Intelligent System on the Early Warning System of Landslide

Aghus Sofwan^{1*}, Sumardi¹, Najib² and Indrah Wendah Atma Bhirawa¹

¹Department of Electrical Engineering, Diponegoro University, Jalan Prof. Soedarto SH, Semarang, Indonesia

²Department of Geological Engineering, Diponegoro University, Jalan Prof. Soedarto SH, Semarang, Indonesia

ABSTRACT

Landslide is a natural sloping ground movement disaster that can occur due to several factors such as high rainfall, soil moisture in the depth of the soil of an area, vibrations experienced in the region, and the slope of the ground structure. A system that can deliver these factor values into the levels of vulnerability of landslide disasters is needed. The system uses Arduino Mega 2560 to process the level of vulnerability. It can predict the moment and the probability of the disaster occurring as an early warning system. The artificial neural network (ANN) intelligent system can expect an event of a disaster. The designed ANN used five parameters causing landslide as input data: rainfall, slope, soil moisture on the surface, soil moisture in the ground's depth, and soil vibration. The ANN system output delivered three-level conditions: the safe, the standby, and the hazardous. The feed-forward backpropagation (FFBP) and the cascade forward backpropagation (CFBP) methods were analyzed. The performance of both methods was compared in terms of minimum square error (MSE). The MSE results of FFBP and CFBP in the safe, the standby, and the hazardous

conditions were 0.017076 and 0.034952; 0.049597 and 0.046764; 0.062105 and 0.060355; respectively. The results point to the supremacy of CFBP to FFBP in standby and hazardous conditions. Therefore, the CFBP is implemented into the hardware of the early warning system.

ARTICLE INFO

Article history:

Received: 13 November 2020

Accepted: 04 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.13>

E-mail addresses:

asofwan@elektro.undip.ac.id (Aghus Sofwan)

sumardi@elektro.undip.ac.id (Sumardi)

najib@ft.undip.ac.id (Najib)

indrahwendah591996@gmail.com (Indrah Wendah Atma Bhirawa)

* Corresponding author

Keywords: Artificial neural network, cascade-forward backpropagation, feed-forward backpropagation, landslides

INTRODUCTION

Landslides are natural disasters that occur as a result of sloping ground movements on a hillside or mountain. Indonesia also has a tropical monsoon marine climate with high rainfall. The tropical monsoon climate causes the reason for the landslides. The tropical monsoon climate has extreme changes in air pressure in almost resulting in high rainfall. Furthermore, it increases the soil water content, coupled with the soil vibration, which will trigger landslides (Huang & Lin, 2002). According to released data from the National Disaster Management Agency Indonesia, the country has 274 cities with landslide potential (BNPB, 2019). The total number of natural disasters in this country from 2017 to 2019 is 1326 events, resulting in 642 fatalities and injuries (BNPB, 2019). Recognizing the changing pattern of each parameter will make it easier to predict when a landslide disaster occurs.

In the previous research, the utilization of artificial neural network (ANN) methods to recognize data patterns and forecasting have been carried out in many applications, such as biological (Gu et al., 2012), food (Stangierski et al., 2019), chemical (Radfard et al., 2018), environment (Li & Jiang, 2010; Ul-Saufie et al., 2011), and disaster (Borujeni & Nateghi, 2019; Elsafi, 2014; Pradhan & Lee, 2010; Tsakiri et al., 2018). Notably, Pradhan and Lee (2010), used the backpropagation neural network to analyze landslide susceptibility. The result suggests the effect of the topographic slope. In Borujeni and Nateghi (2019), the ANN model was selected as a computational model for landslide forecasting, which only measured the soil layer displacement. Moreover, much practical research on the landslide warning system has been studied, such as in Chaturvedi et al. (2017), Chen et al. (2017), Hemalatha et al. (2019), and Sofwan et al. (2017). Sofwan et al. (2017) developed a wireless sensor network that applied Internet of Things architecture for landslide warnings. Chen et al. (2017) used a probability regression model to detect landslide and applied it in practical cases. Chaturvedi et al. (2017) used sensor data to develop a landslide early warning system (EWS). Hemalatha et al. (2019) applied machine learning algorithms, which supported vector regression for wireless sensor networks for real-time monitoring and early warning of a landslide. Nguyen et al. (2019) proposed the hybrid machine learning models, such as Particle Swarm Optimization for Adaptive Neuro Fuzzy Inference System and Particle Swarm Optimization for Artificial Neural Networks for landslide spatial prediction. The proposed models provide MSEs with value of 0.225 and 0.312, respectively. Hereafter, in our research, we considered implementing the ANN model. The model used either the feed-forward backpropagation (FFBP) or cascade-forward backpropagation (CFBP) method. The chosen method would then be embedded into an early warning hardware system to identify a landslide disaster level. Our research contribution is more focused on developing the ANN model with FFBP and CFBP methods. Whereas, Sofwan et al. (2018a)

published our hardware system development. The hardware is the node, which consists of sensors, microcontroller Arduino Mega 2560, communication module, and solar cell power supply. The sensors measure rainfall, soil moisture, temperature, and air humidity. The microcontroller processes the measurement sensor results with ANN. Furthermore, the processing result is submitted to the server as information for the stakeholder. The node is placed at remote area so that it is equipped with the solar cell power supply.

METHODS

An ANN is a system inspired by human neuron information processing to learn a particular procedure. The ANN construction is achieved by giving excitement into the neuronal fashion, computing the output, and adjusting the weights until the expected result is obtained. In this research, the ANN was used to identify conditions and assessed a landslide vulnerability. The main steps of the conducted research were as follows. Firstly, the 10,000 static training data was collected as the input of a model with computational programming. The data was referred to the parameters of landslide causes, based on the Ministry of Public Works Regulation of Indonesia (PMPU) No.22/PRT/M/2007 (Ministry of Public Work Regulation, 2007), obtained from sensor measurement except for vegetation. The regulation is used as a basis for regulating landslide-prone areas in Indonesia. Therefore, we use this regulation to apply the weight of landslide parameters. The weight values, i.e., 1 to 3, and weighting percentage values, are applied to the parameters shown in Table 1. Percentage of weighting for rainfall, slope, soil moisture, vibration, and vegetation are 30%, 15%, 22%, 23%, and 10%, respectively. The main factor of landslide cause is rainfall so that the regulation applies the weight value of 30%. The total weighting percentage is 100 per cent, hereafter, and is utilized as the manual calculation for the landslide vulnerabilities resulting in some output levels. The output is classified into three ranges: 1.00-1.69, 1.70-2.39, and 2.40-3.00 refer to safe, standby, and hazard situations.

Table 1
Parameters of landslide causes

Parameter (unit)	Reading (Weight)	Percentage of Weighting	Reading (Weight)	Percentage of Weighting	Reading (Weight)	Percentage of Weighting
Rainfall (mm ³ /hour)	0-20 (1)	30%	20-40 (2)	30%	>40 (3)	30%
Slope (degree)	0-30 (1)	15%	31-70 (2)	15%	>70 (3)	15%
Soil Moisture (%)	0-30 (1)	22%	31-35 (2)	22%	>35 (3)	22%
Vibration (Richter's scale)	0-2 (1)	23%	3-5 (2)	23%	>5 (3)	23%
Vegetation	(2)	10%	(2)	10%	(2)	10%

Furthermore, the vulnerability of the landslide based on (Ministry of Public Work Regulation, 2007) can be calculated using the following Equation 1.

$$\begin{aligned}
 \text{Level of vulnerability} &= \text{rainfall weight} * 30\% + \text{slope weight} * 15\% \\
 &+ \text{soilmoisture1 weight} * 11\% + \text{soilmoisture2 weight} * 11\% \\
 &+ \text{vibration weight} * 23\% + 2 * 10\% \quad [1]
 \end{aligned}$$

Secondly, the data is applied to the algorithm of the ANN intelligence system model. The system model consists of three layers: the input layer, the hidden layer, and the output layer. The feed-forward backpropagation (FFBP) and the cascade-forward backpropagation (CFBP) methods are analyzed to predict the possibility of a landslide occurrence. Furthermore, the ANN intelligent system gives three statuses, namely safe, standby, and hazard situations. After that, the chosen result, either the FFBP or the CFBP method, will be embedded in the hardware system.

The FFBP and CFBP methods structure are shown in Figure 1(a) and 1(b), respectively. It consists of three layers: the input layer, the hidden layer, and the output layer. In the first layer, there are five input variables used, which are rainfall, slope, soil moisture with two different depths, and vibration. The second layer uses the hyperbolic tangent activation function to get a more accurate system output in predicting landslides symptoms. The activation function is applied to each neuron. Afterwards, there is an output in the scaling

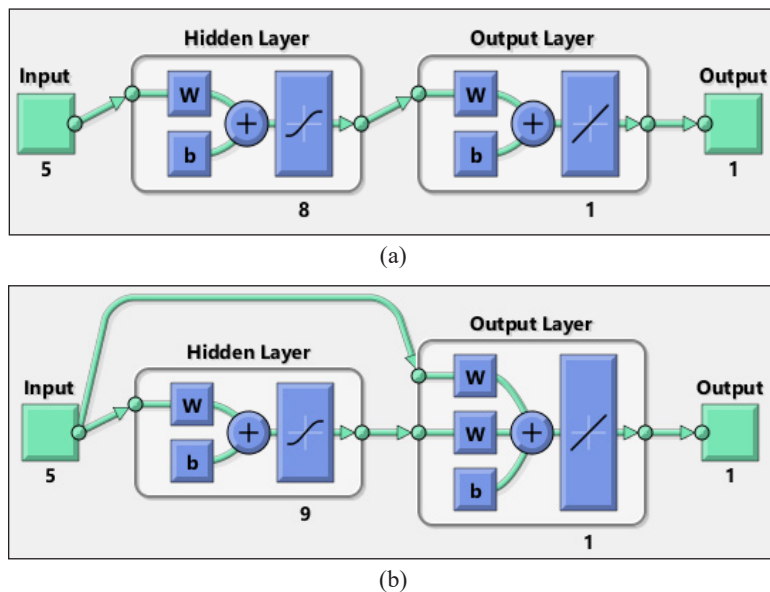


Figure 1. Structure of ANN with FFBP (a) and CFBP (b)

form of landslide possibility level in the third layer, from scale 1 for safe situation up to scale 3 for hazard situation. Later, this ANN intelligent system will be embedded into a developed hardware system shown in Figure 2. It consists of sensors, such as a reed switch, YI-69, MPU 6050, 801S, and DHT22 (Sofwan et al., 2018a; Sofwan et al., 2018b). The sensors measure physical parameters, such as rainfall, slope, soil moisture, and vibration. The tipping bucket calculates the number of rainfall. Sensor MPU 6050 measures slope change, especially when the landslide happens. Soil moisture sensor quantifies the level of groundwater. The 801S vibration sensor gauges the amount of ground vibration. The DHT22 sensor detects temperature and air humidity. The hardware system uses the Arduino Mega 2560 as the data processing centre, which holds an appropriate ANN intelligence. The Real Time Clock (RTC) module functions as an electronic clock, counting the system clock and keeping data in real-time. The utilization of relay and fan has a purpose of keeping cooling the hardware node, especially the microcontroller. Furthermore, the processing result is sent to the database and web server by using the SIM900A module.

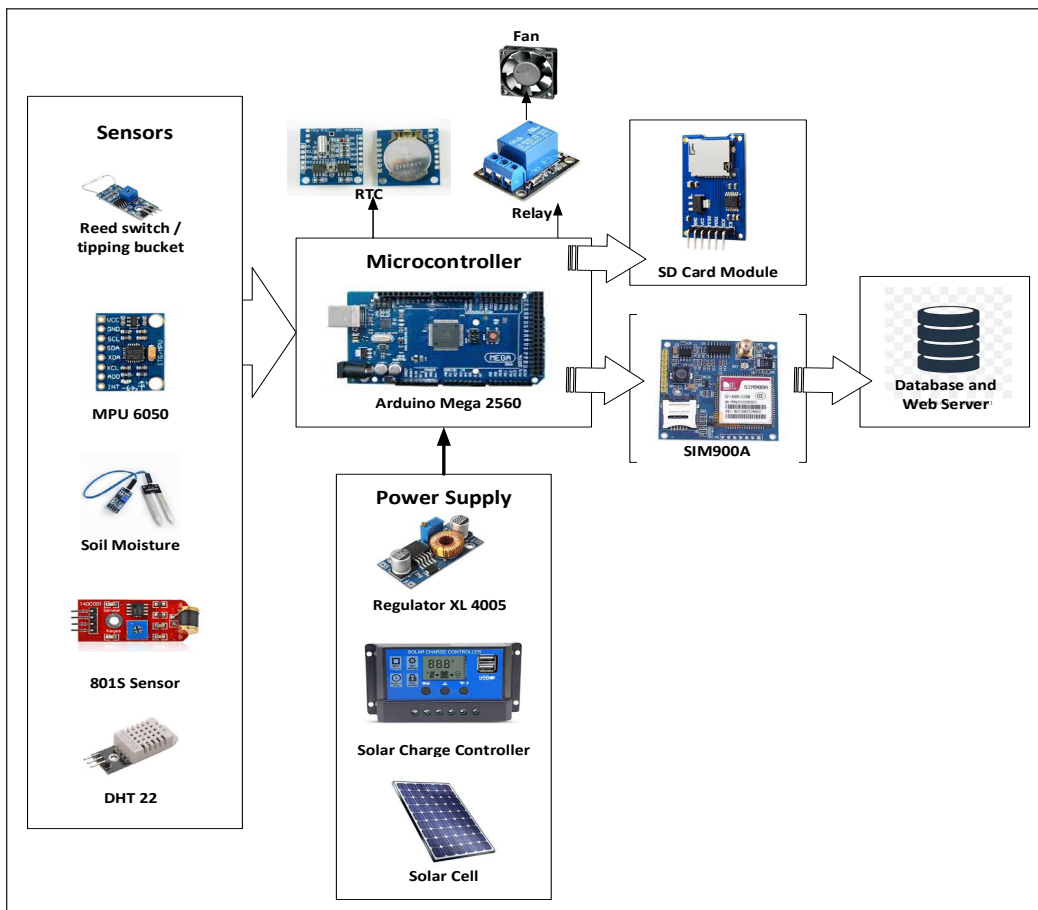


Figure 2. Overall system block diagram

Artificial Neural Network for Landslide Early Warning System

In this section, a design of the ANN intelligent system for the early detection system of landslides using the FFBP and CFBP is described. The designed system utilized a 3-layer or multi-layer perceptron structure, namely the input layer, hidden layer, and output layer, shown in Figure 1. The following subsections describe the design of each layer for both methods in detail.

Design of Input Layer. The input layer uses five data parameter inputs: rainfall, slope, two depths of soil moisture, and vibration parameters. The rainfall parameter is obtained from the reed switch sensor in the tipping bucket. The slope parameter value is delivered by the MPU6050 sensor. The soil moisture parameters at two different depths are received from the YL-69 sensors. The aim of using two sensors in different depths is to obtain soil moisture characteristics in two levels, on the surface and in the ground. The two-level humidity strengthens the prediction of the landslide. And the last parameter, which is the vibration, is taken from the 801S sensor. Furthermore, these parameters are named rainfall, slope, soilmoisture1, soilmoisture2, and vibration variables.

The ANN can process the variables if they have been normalized. Normalization is performed to simplify the calculation of weights in ANN and reduce the distance between the input data values. It also speeds up the learning process and leads to faster convergence. Therefore, the variables must be normalized before processing with ANN. In this layer, the min-max function is used to normalize the input data expressed in the following Equation 2.

$$y = \frac{(x-x_{min})(y_{max}-y_{min})}{x_{max}-x_{min}} + y_{min} \quad [2]$$

The variables y , y_{max} , and y_{min} denote the obtained, the maximum, and the minimum of the normalized data. While the variables x , x_{max} , and x_{min} refer to the original, the maximum, and the minimum value of measurement data from input data per minute.

From the formula above, the minimum and maximum values can be referred to as parameter value as follows.

Rainfall	= 0 – 150 (mm3 per hour)
Slope	= 0 – 100 (%) which 0 is 0 degree and 100 is 45 degree
Soilmoisture1	= 0 – 100 (%)
Soilmoisture2	= 0 – 100 (%)
Vibration	= 0 – 7 (Richter's Scale)

Design of Hidden Layer. The hidden layer design is performed by applying variations of neurons, which provides a fast computation, and the calculation results with a little value or close to zero of error. The variety of neurons, which are applied in the hidden layer with

a number of 5, 8, 9, 10, and 15 neurons. The reason for experimenting with variations in neurons is to determine the running system response with several variations of the neuron. Later on, one of the numbers of neurons with the minimum training error, i.e., close to zero, will be applied in this layer.

The hidden layer performs computation by weighting to each input value. In the FFBP method, the hidden layer's input is obtained by performing a mathematical calculation using the Equation 3.

$$W_f = \sum_{k=1}^8 ((i_1 + i_2 + i_3 + i_4 + i_5)W_{1k} + (W_{1k} \cdot Bias_1)) \quad [3]$$

where W_f denotes the weight in the hidden layer of FFBP, i_k refers to input k-th, and $Bias_1$ indicates the bias in the hidden layer. After some epochs, the best value of weight and bias, which results in minimum training error, is chosen in the hidden layer with eight neurons. The weight and bias are shown in Tables 2 and 3, respectively.

Furthermore, the computation in the hidden layer is conducted by multiplying the weights and biases, which are obtained from the training set, with the hyperbolic tangent activation function. The activation function is essential for the neural network model to learn and understand complex non-linear functions. They allow the introduction of non-linear features to the network. The hyperbolic tangent activation function is shown in the following Equation 4.

$$f(x) = \tanh\left(\frac{x}{2}\right) = \frac{1 - e^{-x}}{1 + e^{-x}} \quad [4]$$

Table 2
Weight in the hidden layer of FFBP with eight neurons

i(input)	W_i (weight)							
	1	2	3	4	5	6	7	8
1	0.12430	23.87400	0.014135	0.06201	-35.34480	-0.03811	-36.2062	43.35870
2	0.04945	7.74800	0.007768	0.02425	-0.14958	0.036918	-0.15933	10.01040
3	0.02935	34.99580	-0.037923	0.01069	3.24120	-0.09094	3.29970	14.36040
4	-0.25145	40.67150	-0.051283	-0.13615	2.76630	0.09815	2.82470	13.82550
5	0.19447	0.073166	0.017005	0.08366	-1.07030	-0.02005	-1.08850	3.13340

Table 3
Bias in the hidden layer of FFBP with eight neurons

Bias ₁	W_i (weight)							
	1	2	3	4	5	6	7	8
B1	-0.02017	91.4655	0.00797	-0.007346	-31.1158	0.064688	-31.8737	69.6398

where x denotes the input value of weight and bias. Furthermore, the obtained weight and bias values will be forwarded to the output layer.

In the CFBP method, the input of the hidden layer is calculated using a formula, which operates nine neurons instead of eight, as Equation 5.

$$W_c = \sum_{k=1}^9 ((i_1 + i_2 + i_3 + i_4 + i_5)W_{1k} + (W_{1k} \cdot Bias_1)) \quad [5]$$

where W_c denotes the weight in the hidden layer of CFBP. The weight and bias values are shown in Tables 4 and 5, respectively. Furthermore, similar to the FFBP method, computation in this layer also is performed by multiplying with the tanh activation function.

Design of the Output Layer. The output layer design uses one neuron, with input from the hidden layer multiplied by the second weighting. The output layer is then added by weighting the input from the input layer to obtain new weight and bias values. In the FFBP and CFBP methods, the new weight and bias values are shown in Tables 6 and 7, respectively.

Table 6 shows that the weight and bias values in the output layer of FFBP are obtained by performing mathematical calculations using the following Equation 6.

$$x = \sum_{k=1}^8 ((W_f \cdot W_{2k}) + (W_{2k} \cdot Bias_2)) \quad [6]$$

where x refers to the output layer's result, W_2 denotes weight in the output layer, and $Bias_2$ indicates the bias in the output layer.

Table 4
Weight in the hidden layer of the CFBP with nine neurons

i (input)	W ₁ (weight)								
	1	2	3	4	5	6	7	8	9
1	1.1647	-1.1811	0.065471	0.50684	1.0310	1.3121	0.16789	-1.0794	-1.0587
2	-1.2763	1.4417	-0.91833	0.89618	-0.7197	1.3052	-0.58401	-0.7623	-0.76754
3	0.50112	-0.19541	-0.56255	0.81184	0.22198	0.0666	1.616	-0.5104	0.74318
4	1.1795	-0.65703	0.30079	-1.166	-1.182	0.0955	1.066	1.0743	-0.9955
5	0.28405	-0.17847	-1.7925	-0.38268	-0.8963	0.9124	-0.7575	-0.1675	-0.94622

Table 5
Bias in the hidden layer of CFBP with nine neurons

Bias ₁	W ₁ (weight)								
	1	2	3	4	5	6	7	8	9
B1	-2.1337	1.7003	-1.0764	-0.38706	0.047341	0.49982	1.0936	-1.6464	-2.1233

Table 7 shows that the weight and bias values in the output layer of CFBP are obtained by performing mathematical calculations using the Equation 7.

$$x = \sum_{k=1}^9 \left((W_c \cdot W_{3k}) + ((i_1 + i_2 + i_3 + i_4 + i_5) \cdot W_{2k}) + (W_{3k} \cdot Bias_2) \right) \quad [7]$$

where x refers to the output layer result, W_2 and W_3 denote weight in the output layer, and $Bias_2$ indicates the bias in the output layer.

The output layer result must be multiplied by the activation function. The activation function used is a linear function, as shown in the following Equation 8.

$$y = f(x) = x \quad [8]$$

After getting one output variable from the output layer, the next process is to normalize the data again, to return the data to its original form using the min-max reverse function (Equation 9).

Table 6
Weight and bias in the output layer of FFBP

No.	W_2 (Weight)	$Bias_2$
1	-1.3237	
2	0.013994	
3	-15.0164	
4	11.3335	
5	-0.76339	-0.098634
6	4.0832	
7	0.76187	
8	-0.01352	

Table 7
Weight and bias in the output layer of CFBP

No.	W_2 (weight)		$Bias_2$
	2	3	
1	0.11734	-5.5917e-17	
2	0.25938	-5.2874e-17	
3	0.30466	-4.2555e-17	
4	-0.02882	-4.1597e-17	
5	0.35765	-5.9389e-17	-0.0097987
6		-1.7423e-17	
7		-6.4548e-18	
8		2.5679e-17	
9		2.0381e-17	

$$x = \frac{(y - y_{min})(x_{max} - x_{min})}{y_{max} - y_{min}} + x_{min} \quad [9]$$

where x is the postprocessing output, y is the input of the output layer, x_{min} is the minimum output value, x_{max} is the maximum output value, y_{min} is the minimum weighting value, and y_{max} is the maximum weighting value.

Program Design at Arduino Mega 2560. The hardware system utilizes Arduino Mega 2560 microcontroller as the mainboard of the hardware system. The microcontroller runs the program with a designed ANN method. The algorithm performs some steps similar to the ANN model system, which obtain five input data from sensors, performs calculation of weight and bias, calculates the hidden layer's value, and conducts postprocessing of the output ANN. The flowchart of the ANN program in the microcontroller is shown in Figure 3. Furthermore, the implementation program in the microcontroller is using Arduino IDE software.

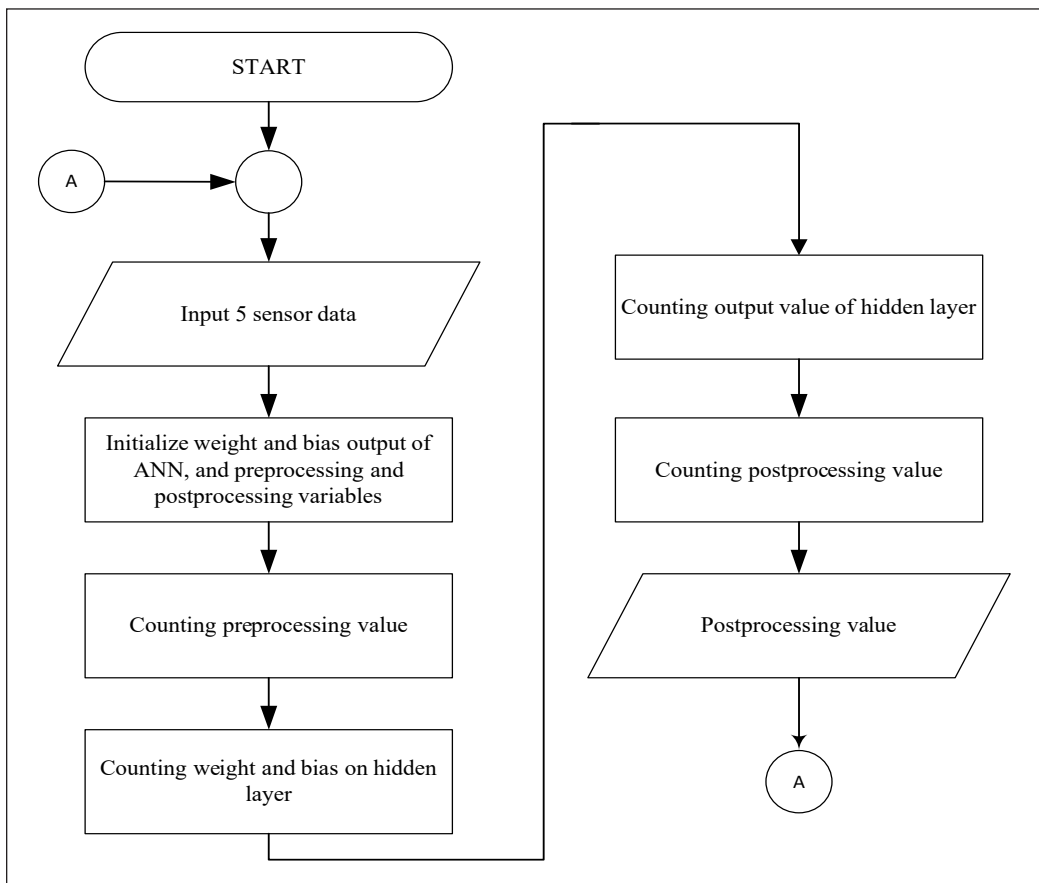


Figure 3. Flow chart of the ANN system

RESULTS AND DISCUSSION

This section evaluates and discusses the ANN intelligent system performance in three levels of vulnerability of a landslide: safe, standby, and hazard conditions. The evaluation testing was applied to the FFBP and CFBP methods. The results of testing are complemented with detailed discussions.

ANN System Testing on the Safe Condition

The system testing was conducted to see the FFBP and CFBP methods' performance in terms of minimum square error (MSE). The five parameters, which are rainfall, slope, soil moisture 1, soil moisture 2, and vibration, are delivered to the ANN methods. Table 8 exposes the input data samples. The testing was carried out with ten samples as input data to examine both methods.

Furthermore, the output values of both methods' output values are compared to that of the manual calculation. The manual calculation is obtained by using a mathematical equation, which is expressed in Equation 1. The score range of safe condition, which is according to Equation 1, is in range with a value of 1 to 1.69. According to data in Table 8, the manual calculation results in an amount of 1.32, which is in the safe condition level. Table 9 exposes the manual calculation and results of the ANN system with the FFBP and CFBP methods in safe conditions.

The error is obtained from the difference between the value of manual calculation and the value of a method. Hereafter, in the safe condition from Table 10, it can be seen the MSE of the FFBP and CFBP are 0.017076 and 0.034952, respectively. The former method performance is better than that of the latter, with a different value of 0.017876. This difference in error value is caused by differences in structure and weight values used in each method.

Tabel 8
The input of ANN system testing on safe condition data

No.	Rainfall (mm ³)	Slope (%)	Soil Moisture 1 (%)	Soil Moisture 2 (%)	Vibration
1	0	2.12	42	8.3	0
2	0	3.34	45	8.6	0
3	0	0.88	43	8.6	0
4	0	2	42	8.5	0
5	0	3.1	41	8.5	0
6	0	1.8	43	8.1	0
7	0	1.9	42	7.98	0
8	0	2.06	43	7.90	0
9	0	3.76	45	9.1	0
10	0	4.15	47	8.98	0

Table 9
Result of ANN system output on safe condition

No.	Manual Calculation	Status	FFBP	Status	Error	CFBP	Status	Error
1	1.32	Safe	1.350	Safe	0.030	1.2896	Safe	0.0304
2	1.32	Safe	1.331	Safe	0.011	1.2897	Safe	0.0303
3	1.32	Safe	1.331	Safe	0.011	1.2884	Safe	0.0316
4	1.32	Safe	1.342	Safe	0.022	1.2862	Safe	0.0338
5	1.32	Safe	1.3176	Safe	0.0026	1.2771	Safe	0.0429
6	1.32	Safe	1.330	Safe	0.010	1.2848	Safe	0.0352
7	1.32	Safe	1.342	Safe	0.022	1.2870	Safe	0.0330
8	1.32	Safe	1.344	Safe	0.024	1.2854	Safe	0.0346
9	1.32	Safe	1.330	Safe	0.010	1.2811	Safe	0.0389
10	1.32	Safe	1.3249	Safe	0.0049	1.2832	Safe	0.0368

ANN System Testing on Standby Condition

The system testing in the alert condition was also conducted to see the performance of both FFBP and CFBP methods. The input data that is used as the samples can be seen in Table 10. The testing was held by using ten samples to examine the output testing of both methods.

Similar to the testing of the safe condition, the result values of the output layer of FFBP and CFBP methods are also compared to that of the manual calculation. The score range of standby condition, which is according to Equation 1, is in range with a value of 1.7 to 2.39. According to data in Table 10, the manual calculation results in variation values in the standby condition level. Table 11 exhibits the manual calculation results and results of the ANN system testing with FFBP and CFBP methods.

Table 11 shows the MSE of the FFBP and CFBP are 0.049597 and 0.046764, respectively. The former method performance is less than that of the latter, with a different value of 0.002833.

ANN System Testing on the Hazardous Condition

The ANN system testing on the hazardous condition was also held with goals to perceive the achievement of the FFBP and the CFBP methods. The input data samples that are utilized in this testing can be seen in Table 12. The testing was held by using those samples to examine the output testing of both methods. Meanwhile, the range of scores that state the hazard condition is in a value of 2.4 to 3, which is referred to Equation 1.

Referring to data in Table 12, the manual calculation results in variation values in a range between 2.40-3.00, which are in the standby condition level. Table 13 exhibits the results of the manual calculation and results of the FFBP and CFBP methods. Table 13 shows the error of the FFBP and the CFBP in each sample. Furthermore, the MSE of the

Tabel 10
The input of the ANN system testing with standby condition

No.	Rainfall (mm ³)	Slope (%)	Soil Moisture 1 (%)	Soil Moisture 2 (%)	Vibration
1	56	16.12	42.2	32.3	4
2	61	39.34	45.5	33.6	4
3	34	40.88	43.7	18.6	7
4	32	52.2	42.6	38.5	2
5	48	53.1	41.6	48.2	2
6	67	23.8	43.2	33.1	5
7	69	40.9	42.3	32.98	7
8	66	32.06	43.2	49.9	3
9	60	33.76	45.8	59.1	2
10	57	14.15	47.1	48.98	4

Tabel 11
Result of the ANN system testing on standby condition

No.	Manual Calculation	Status	FFBP	Status	Error	CFBP	Status	Error
1	2.15	Standby	2.1974	Standby	0.0474	2.1642	Standby	0.0142
2	2.03	Standby	2.0366	Standby	0.0066	2.0281	Standby	0.0019
3	2.27	Standby	2.2174	Standby	0.0526	2.2547	Standby	0.1153
4	2.18	Standby	2.151	Standby	0.029	2.2048	Standby	0.0248
5	2.29	Standby	2.2952	Standby	0.0052	2.2869	Standby	0.0031
6	2.11	Standby	2.196	Standby	0.086	2.0163	Standby	0.0937
7	2.11	Standby	2.1532	Standby	0.0432	2.1747	Standby	0.0647
8	2.22	Standby	2.1885	Standby	0.0315	2.2706	Standby	0.0506
9	2.07	Standby	2.1006	Standby	0.0306	2.1394	Standby	0.0694
10	1.88	Standby	1.7935	Standby	0.0865	1.8585	Standby	0.0615

Tabel 12
The input of the ANN system testing with hazard condition

No.	Rainfall (mm ³)	Slope (%)	Soil Moisture 1 (%)	Soil Moisture 2 (%)	Vibration
1	128	76.12	61.23	58.30	5
2	61	59.34	59.21	58.60	7
3	94	60.88	33.00	31.60	7
4	72	42.20	47.00	38.50	6
5	88	73.10	61.00	39.20	6
6	87	83.80	73.00	78.10	6
7	99	74.90	62.00	47.98	6
8	106	82.06	33.30	32.90	5
9	90	83.76	48.00	49.10	7
10	97	74.15	77.3	68.98	6

Table 13
Result of the ANN system on hazard condition

No.	Manual Calculation	Status	FFBP	Status	Error	CFBP	Status	Error
1	2.67	Hazard	2.7026	Hazard	0.0326	2.7297	Hazard	0.0597
2	2.52	Hazard	2.4822	Hazard	0.0378	2.4835	Hazard	0.0365
3	2.68	Hazard	2.5614	Hazard	0.1186	2.6186	Hazard	0.0614
4	2.67	Hazard	2.7044	Hazard	0.0344	2.7527	Hazard	0.0827
5	2.67	Hazard	2.6153	Hazard	0.0547	2.6297	Hazard	0.0403
6	2.90	Hazard	2.8291	Hazard	0.0709	2.8799	Hazard	0.0201
7	2.67	Hazard	2.7196	Hazard	0.0496	2.7374	Hazard	0.0674
8	2.45	Hazard	2.4917	Hazard	0.0417	2.4894	Hazard	0.0394
9	2.67	Hazard	2.633	Hazard	0.037	2.6674	Hazard	0.0026
10	2.90	Hazard	2.8149	Hazard	0.0851	2.7869	Hazard	0.1131

FFBP and CFBP are 0.062105 and 0.060355, respectively. These MSE values indicate that the former method performance is less than that of the latter, with a different amount of 0.00175.

Based on the testing in three levels of vulnerability of the landslide, both methods performance of each level is obtained. The FFBP method performance is better than that of the CFBP method in terms of safe condition level. In contrast, the CFBP method is superior to the FFBP method in standby and hazard conditions. CFBP model provides a connection from the input layer to the hidden layer and a direct connection to the output layer. Data from the input layer then will be weighted in the hidden layer and output layer. With this direct connection, the weighting process is more quickly performed. So that the CFBP method more appropriate for dynamic input parameters. Therefore, we deploy the ANN CFBP method into the developed hardware system. The hardware system was installed in the middle of the hill, a landslide-prone area with coordinates -7.3406467, 110.3829545.

CONCLUSION

This paper has performed and analyzed the ANN intelligent system with the FFBP and the CFBP methods for a landslide early warning system. The intelligent system uses sensors to obtain landslide causative parameters: rainfall, the slope of the ground, soil moisture, and vibration. The ANN system output delivers three landslide vulnerability levels: safe, standby, and hazardous conditions. The MSE results of FFBP and CFBP in the safe, the standby, and the hazardous conditions are 0.017076 and 0.034952; 0.049597 and 0.046764; 0.062105 and 0.060355; respectively. Based on the performed testing evaluation, the MSE of the CFBP method is superior to that of the FFBP method in terms of standby and hazardous conditions.

ACKNOWLEDGMENTS

This work was supported by the Superior Applied Research of Higher Education (Penelitian Terapan Unggulan Perguruan Tinggi - PTUPT) Diponegoro University grant in 2019. The authors are grateful for this support.

REFERENCES

- BNPB. (2019). Indonesia's disaster data information. *National Disaster Management Agency*. Retrieved November 11, 2020, from <http://dibi.bnpb.go.id/>
- Borujeni, M. S., & Nateghi, S. (2019, April 11-14). Modeling of sensor network for autonomous landslide monitoring based on neural network. In *Proceeding of the IEEE SoutheastCon 2019* (pp.1-5). Huntsville, Alabama, USA. <https://doi.org/10.1109/SoutheastCon42311.2019.9020350>
- Chaturvedi, P., Srivastava, S., & Kaur, P. B. (2017, April 13). Landslide early warning system development using statistical analysis of sensors' data at tangni landslide, Uttarakhand, India. In *Proceedings of Sixth International Conference on Soft Computing for Problem Solving* (pp. 259-270). Singapore. https://doi.org/10.1007/978-981-10-3325-4_26
- Chen, F., Yu, B., Xu, C., & Li, B. (2017). Landslide detection using probability regression, a case study of Wenchuan, northwest of Chengdu. *Applied Geography*, *89*, 32-40. <https://doi.org/10.1016/j.apgeog.2017.10.001>
- Elsafi, S. H. (2014). Artificial neural networks (ANNs) for flood forecasting at Dongola Station in the River Nile, Sudan. *Alexandria Engineering Journal*, *53*(3), 655-662. <https://doi.org/10.1016/j.aej.2014.06.010>
- Gu, S., Sun, X., Wu, Y., & Cui, Z. (2012, May 29-31). An approach to forecast red tide using generalized regression neural network. In *Proceeding of the 8th International Conference on Natural Computation 2012* (pp. 194-198). Sichuan, China. <https://doi.org/10.1109/ICNC.2012.6234545>
- Hemalatha, T., Ramesh, M. V., & Rangan, V. P. (2019). Effective and accelerated forewarning of landslides using wireless sensor networks and machine learning. *IEEE Sensors Journal*, *19*(21), 9964-9975. <https://doi.org/10.1109/JSEN.2019.2928358>
- Huang, L. J., & Lin, X. S. (2002). Study on landslide related to rainfall. *Journal of Xiangtan Normal University*, *24*(4), 55-62.
- Li, Y. X., & Jiang, L. C. (2010). Application of ANN algorithm in tree height modeling. *Applied Mechanics and Materials*, *20-23*, 756-761. <https://doi.org/10.4028/www.scientific.net/AMM.20-23.756>
- Ministry of Public Work Regulation. (2007). *Guidelines for spatial planning for disaster-prone areas, regulation of the minister of public works No. 22/PRT/M/2007*. Retrieved November 11, 2020, from http://landspatial.bappenas.go.id/komponen/peraturan/the_file/permen22_2007.pdf
- Nguyen, V., Pham, B., Vu, B., Prakash, I., Jha, S., Shahabi, H., Shirzadi, A., Ba, D. N., Kumar, R., Chatterjee, J. M., & Bui, D. T. (2019). Hybrid machine learning approaches for landslide susceptibility modeling. *Forests*, *10*(2), 157. <https://doi.org/10.3390/f10020157>

- Pradhan, B., & Lee, S. (2010). Regional landslide susceptibility analysis using backpropagation neural network model at Cameron Highland, Malaysia. *Landslides*, 7(1), 13-30. <https://doi.org/10.1007/s10346-009-0183-2>
- Radfard, M., Soleimani, H., Nabavi, S., Hashemzadeh, B., Akbari, H., Akbari, H., & Adibzadeh, A. (2018). Data on estimation for sodium absorption ratio: Using artificial neural network and multiple linear regressions. *Data in Brief*, 20, 1462-1467. <https://doi.org/10.1016/j.dib.2018.08.205>
- Sofwan, A., Sumardi, R. M. I., Goni, A., & Najib. (2017, October 18-19). Wireless sensor network design for landslide warning system in IoT architecture. In *Proceeding of the 4th International Conference on Information Technology, Computer, and Electrical Engineering (ICITACEE) 2017* (pp. 280-283). Semarang, Indonesia. <https://doi.org/10.1109/ICITACEE.2017.8257718>
- Sofwan, A., Sumardi, R. M. I., & Najib. (2018a, October 16-17). *Measurement design of sensor node for landslide disaster early warning system* [Paper presented]. 2nd International Conference on Electrical Engineering and Informatics (ICon EEI) 2018 (pp. 75-80). Batam, Indonesia. <https://doi.org/10.1109/ICon-EEI.2018.8784341>
- Sofwan, A., Sumardi, R. M. I., & Ulwiyati, N. (2018b, September 26-27). Filtering for data acquisition on wireless sensor network. In *Proceeding of the 5th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE) 2018* (pp.180-184). Semarang, Indonesia. <https://doi.org/10.1109/ICITACEE.2018.8576940>
- Stangierski, J., Weiss, D., & Kaczmarek, A. (2019). Multiple regression models and artificial neural network (ANN) as prediction tools of changes in overall quality during the storage of spreadable processed Gouda cheese. *European Food Research and Technology*, 245(11), 2539-2547. <https://doi.org/10.1007/s00217-019-03369-y>
- Tsakiri, K., Marsellos, A., & Kapetanakis, S. (2018). Artificial neural network and multiple linear regression for flood prediction in Mohawk River, New York. *Water - Open Access Journal*, 10(9), 20. <https://doi.org/10.3390/w10091158>
- Ul-Saufie, A. Z., Yahaya, A. S., Nor, Y., Hazrul, A., & Hamid, H. A. (2011). Comparison between multiple linear regression and feed forward back propagation neural network models for predicting PM 10 concentration level based on gaseous and meteorological parameters. *International Journal of Applied Science and Technology*, 1(4), 42-49.

Performance Evaluation of Different Membership Function in Fuzzy Logic Based Short-Term Load Forecasting

Oladimeji Ibrahim^{1,2*}, Waheed Olaide Owonikoko¹, Abubakar Abdulkarim³, Abdulrahman Okino Otuoze¹, Mubarak Akorede Afolayan¹, Ibrahim Sani Madugu⁴, Mutiu Shola Bakare¹ and Kayode Elijah Adedayo^{1,5}

¹Department of Electrical and Electronics Engineering, University of Ilorin, 240103 Ilorin, Nigeria

²Department of Electrical and Electronics Engineering, Universiti Teknologi PETRONAS, Bandar Seri Iskandar, 32610 UTP, Perak, Malaysia

³Department of Electrical Engineering, Ahmadu Bello University, Zaria, Nigeria

⁴Department of Electrical Engineering, Kano University of Science & Technology Wudil, Nigeria

⁵Department of Electrical and Electronics Education, Kwara State College of Education (Technical) Lafiagi, Nigeria

ABSTRACT

A mismatch between utility-scale electricity generation and demand often results in resources and energy wastage that needed to be minimized. Therefore, the utility company needs to be able to accurately forecast load demand as a guide for the planned generation. Short-term load forecast assists the utility company in projecting the future energy demand. The predicted load demand is used to plan ahead for the power to be generated, transmitted, and distributed and which is crucial to power system reliability and economics. Recently, various methods from statistical, artificial intelligence, and hybrid methods have been widely used for load forecasts with each having their merits and drawbacks. This paper

investigates the application of the fuzzy logic technique for short-term load forecast of a day ahead load. The developed fuzzy logic model used time, temperature, and historical load data to forecast 24 hours load demand. The fuzzy models were based on both the trapezoidal and triangular membership function (MF) to investigate their accuracy and effectiveness for the load forecast. The obtained low Mean Absolute Percentage Error (MAPE), Mean Forecast Error (MFE), and Mean Absolute Deviation

ARTICLE INFO

Article history:

Received: 28 March 2020

Accepted: 27 July 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.14>

E-mail addresses:

ibrahim.o@unilorin.edu.ng (Oladimeji Ibrahim)
waheedowonikoko@gmail.com (Waheed Olaide Owonikoko)
abkzarewa@yahoo.com (Abubakar Abdulkarim)
otuoze.ao@gmail.com (Abdulrahman Okino Otuoze)
mubakorede@yahoo.com (Mubarak Akorede Afolayan)
maduguemir@gmail.com (Ibrahim Sani Madugu)
bkjashola@gmail.com (Mutiu Shola Bakare)
kayodeelijah@yahoo.com (Kayode Elijah Adedayo)

* Corresponding author

(MAD) values from the forecasted load results showed that both models are suitable for short-term load forecasting, however the trapezoidal MF showed better performance than the triangular MF.

Keywords: Artificial intelligence, fuzzy logic, load forecasting, mean absolute percentage error, MF, short-term

INTRODUCTION

The electrical load forecast involves the process of predicting future short, medium, and long term energy demand. This aids the optimal deployment of available resources to accurately generate power to match the load demand. In either a regulated or deregulated market, it is necessary for the supply authority to carry out load forecasting to balance the load demand with supply. The load forecasting is considered as an important aspect of electrical power systems operation as it allows the utility company to make proper planning for the power generation, transmission, and distribution of electrical energy. With accurate load forecasting, an operation such as unit commitment, economic dispatch, load balancing, power quality, and scheduled maintenance can be carried out effectively (Lei et al., 2019).

Load forecasting is broadly classified into three categories: the short-term load forecasting which forecasts from one hour to seven days, the medium-term load forecast starting from a week to several months, and long-term load forecasting starting from a year to several years. The short-term load forecasting is used by the supply authority for operational purposes such as unit commitment, economic dispatch, load flow, frequency control, security, and reliability of the system (Srivastava et al., 2016). The medium-term load forecasting predicts the load demand that provides information for system planning and operation while long term load forecasting is mainly used for power system planning (Ganguly et al., 2019; Peng et al., 2019).

This paper presents a short-term electrical load forecast using the fuzzy logic approach. Two fuzzy MF approaches were investigated: the trapezoidal MF and the triangular MF for their efficiency and accuracy in the short-term load forecast. The developed model performance was evaluated using typical load data and different error analysis techniques were adopted to ascertain the model degree of accuracy. The error analysis results showed that the fuzzy was suitable for short-term load forecast and the trapezoidal MF provided improved performance compared to the triangular MF.

RELATED WORK

There are three broad classification of approaches for carrying out electrical load forecasting in powers system planning and operational. The conventional approach is referred to as the classical statistical methods which include Autoregressive (AR), Moving Average (MA), Autoregressive Integrated Moving Average (ARIMA), Seasonal Auto Regression Integrated Moving Average (SARIMA) and Multi-Linear Regression. The second approach

involves the application of artificial intelligent techniques such as fuzzy logic, Artificial Neural Network (ANN), and Support Vector Machine (SVM) to predict future load demand. In recent times, there is an evolving approach termed hybrid methods which involves a combination of two of the aforementioned methods. This could be a combination of classical statistic method and artificial intelligent method or combination of two artificial intelligent methods.

Several methods have been investigated on short-term load forecasting including the conventional statistical methods such as autoregression moving average, multiple linear regression method, and single exponential smoothing (Dudek, 2016; Fan et al., 2016; Mi et al., 2018). Literatures have shown that classical statistical methods such as the AR and MA work seamlessly on non-seasonal load forecasts while the ARIMA and SARIMA accepts seasonality of the load pattern. The SARIMA requires deep knowledge of forecaster to determine the appropriate seasonal parameters (Chikobvu & Sigauke, 2012; Singhal et al., 2020).

In 2013, Ding et al. proposed a new method to implement short-term load forecasting in which statistical time series prediction methods of Autoregressive Integrated Moving Average Model with Exogenous Inputs (ARIMAX) and machine learning-based regression were used to forecast load demand. In the error result analysis, the ARIMAX method produce a more accurate forecast when compared with machine learning-based regression as it has the least MAPE in each of the ten selected smart meters (Ding et al., 2013). Razak et al. (2012) forecasted load by developing five different technique of SARIMA for different days of the week. The data from Peninsular Malaysia was used as the input to the developed techniques. The forecast produced high accuracy given MAPE that ranged between 1% to 3% for each of the model.

Also, in 2017, Bozkurt et al. compared the result of the ANN and SARIMA model for the electric power load of Turkish electricity market. The input to the model was the essential factors that affected the load forecasting such as the previous load data, electricity price, weather, and currency exchange rate. The ANN had a MAPE of 1.80% while SARIMA had MAPE of 2.60% (Bozkurt et al., 2017). In 2015, Cui and Peng proposed an improved ARIMAX by combining the series with regression analysis to forecast electric load for short-term period. The model filled the gaps of external effects on electric load and data from some selected city in China were used as the input. The improved ARIMAX model had least MAPE of 0.37% when compared with common time series methods like AR, MA and ARIMA (Cui & Peng, 2015). Generally, statistical methods are acceptable techniques for short-term load forecast even though the evolving artificial intelligent methods are seen performing better. Surveyed literature showed that ARIMA and SARIMA are among the best statistical approach for load forecast because they work efficiently with time series model and can consider the seasonality of load. The main challenge is that they involve complex computations that makes their implementation a bit difficult.

In recent time, artificial intelligent methods such as ANN, genetic algorithm, and fuzzy logic are being widely used and most recently the application of the hybrid system has been investigated (Jetcheva et al., 2014; Kumar et al., 2016; Siri, 2018; Yu & Xu, 2014). The ANN and fuzzy logic forecast methods are commonly used for load forecasting as they do not require rigorous mathematical modelling. Their implementation is much simplified and often results in high load forecasting accuracy. They often result in minimum error when forecast is compared with actual electrical load demand. Besides, these methods allow the forecaster to consider numerous exogenous factors such as temperature and humidity (Kuster et al., 2017). As reported in the literature, several researchers have investigated the performance of fuzzy logic on the electrical load forecasting efficiency and accuracy (Al-Kandari et al., 2004; Danladi et al., 2016; Faysal et al., 2019; Silva et al., 2017) and today fuzzy logic is widely used for load forecast.

Rizwan et al. (2012) investigated daily hourly load demand using historical load data as input to the fuzzy logic model. This model did not consider exogenous factor such as temperature and error analysis have low MAPE value of 1.39% (Rizwan et al., 2012). In 2014, Gohil and Gupta presented short-term load forecasting using fuzzy logic with temperature, humidity, wind speed, and historical load data as inputs. The model was used to forecast hourly load of weekend and weekdays. The MAPE of the day's ranged from 10.55% to 11.74%. The high MAPE value was as a result of the inability of the model to accurately respond to daily abrupt load changes for a few hours of the days (Gohil & Gupta, 2014). Ganguly et al. (2017) presented two fuzzy logic models for short-term load forecast. The first model took day type and time as the input while the second model took day type, previous day load, and the peak forecasted load as its input. The models had MAPE of between 2.37 % to 2.53 % between the actual and forecasted load for three consecutive days.

It is worthy to note that the most available work on fuzzy logic load forecast focused on the application of specific MF which is an important phase in the fuzzy logic process. The MF defines how the input points are mapped to a fuzzy membership value ranging between 0 and 1. This stage does have a significant effect on the model performance as it plays a significant role in the decision making that directly influences the output. Among the common MF reported in the literature is the triangular and trapezoidal but researcher focuses more on one of these MFs for a specific application without necessarily investigate the performance of other available MFs. This paper tends to explore this hole by investigating the efficiency and performance accuracy of both triangular and trapezoidal MFs on a fuzzy model for short-term electrical load forecast.

In recent times, load forecast research is gearing towards the development of hybrid methods to take the advantages of some of the basic methods to bridge the gap of one approach with the others. The objective of hybrid methods is to take the merits of statistical and artificial intelligent methods to obtain an optimal load forecasting. With limited

information, the hybrid method has the capacity to optimize the available information, integrate the individual model information, and make use of the merits of the multiple forecasting models to improve prediction accuracy. Although the techniques are usually complex to implement, they often give a more accurate result. Some of the investigated hybrid approach includes a combination of ANN and modified neural network with the fuzzy logic (Černe et al., 2018; Emarati et al., 2019; Sadaei et al., 2019; Wen et al., 2020). Electrical load forecasting from these approaches has contributed immensely to resources management and economics of power industries and indirectly on global warming mitigation.

METHODS

Fuzzy Logic Approach

Fuzzy logic is an approach of computation based on the degree of truth rather than the usual true or false. The fuzzy logic process comprised the inputs, the controller or fuzzy inference engine, the rules base, and the defuzzification stage as presented in Figure 1. The input stage is the point where the forecaster selects the number of inputs such as load data, time, or exogenous factors like the weather parameters for electrical load forecasting. At the input stage, the data is fuzzified where they are formulated to fuzzy sets without any crisp values or elements. The combination of the fuzzification output and the fuzzy rule base serves as input to the fuzzy inference engine or controller which is the heart of the system. The fuzzy controller processes the input data to produce an output. The controller implements the rules prepared by the electrical load forecaster based on the input data classification. The rule base is prepared by the forecaster for the fuzzy inference system and the accuracy of

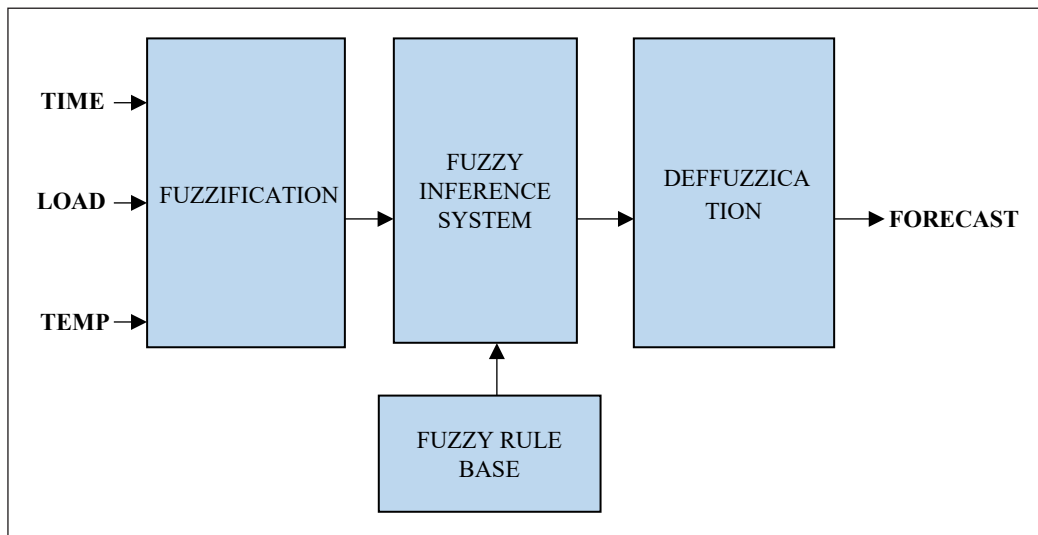


Figure 1. Fuzzy logic process

the system depends on these rules. The more the rules the better the accuracy of the system. In this application, the defuzzification block converts the processed fuzzy set into a crisp output that can be displayed on the graph as the forecasted load or load curve. The fuzzy short-term electrical load forecast uses the actual load, time, and temperature as input data to the fuzzy system. The block diagram in Figure 1 shows the basic fuzzy logic model for the load forecast starting with the fuzzification of input data to the fuzzy inference system with the rules and the defuzzification stage before the output display.

Fuzzification of Input Data

The historical load, time, and temperature are considered as the input, while the forecast load is considered as the output. The historical load data were first examined to study the parameter trends before the fuzzification of the input data. The maximum and minimum values of the parameters including the temperature and the load data were obtained and used for the fuzzification process. From the collected data, Table 1 shows the time fuzzy set, Table 2 presents the temperature fuzzy set, and Table 3 shows the fuzzy set for the load.

Table 1
Time fuzzy set

Time of the day	Notation	Time range
Midnight	MN	1-6
Morning	MG	4-10
For-Noon	FN	9-13
Afternoon	AN	12-17
Evening	EV	15-20
For-Night	FT	19-24

Table 2
Temperature fuzzy set

Temperature	Notation	Temperature range
Temperature very low	LT	20°C-29° C
Temperature low	HT	28°C-36°C
Temperature High	VT	35° C - 42°C

Table 3
Fuzzy set for the load

Load	Notation	Load range
Low load	LL	1MW-30MW
High load	HL	10MW-50MW
Very High load	VH	30MW -60MW

Fuzzy Logic MF

The MF is the curve that defines how each point in the input space is mapped to a membership value between 0 and 1. The common fuzzy logic MFs are the triangular, trapezoidal, and the Gaussian MF. In this study, the triangular and trapezoidal MFs were investigated to ascertain their computational efficiency and effectiveness in the electrical load forecast. The inputs to the fuzzy model for the load forecast are the time, load, and temperature which are fuzzified to data sets between 0 and 1 for the ranges specified in “Fuzzification of Input Data”. The triangular MF allows the forecaster to fix the element of the time, load, and temperature into a fuzzy set with triangular shape as shown in Figures 2, 3, and 4, respectively. In this regard, the forecaster will be able to assign the elements of the fuzzy set between the boundary element and the middle element with the

highest membership value. Also, for the trapezoidal MF, the boundary of the MF is selected and allowed to increase within the range of the elements with the highest membership value in the fuzzy set. The trapezoidal MF mapping for the three input: time, load, and temperature into membership value is presented in Figures 5, 6, and 7, respectively.

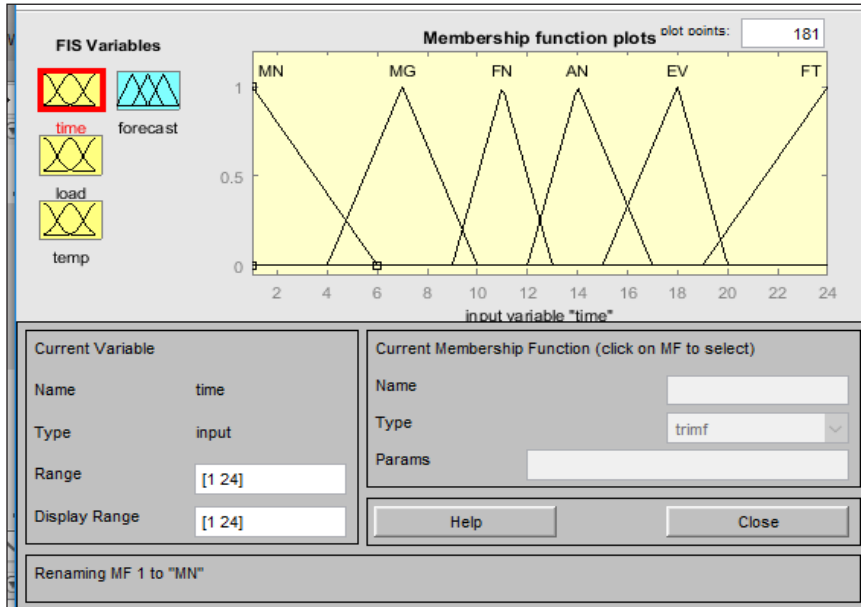


Figure 2. Fuzzy logic triangular MF of time input

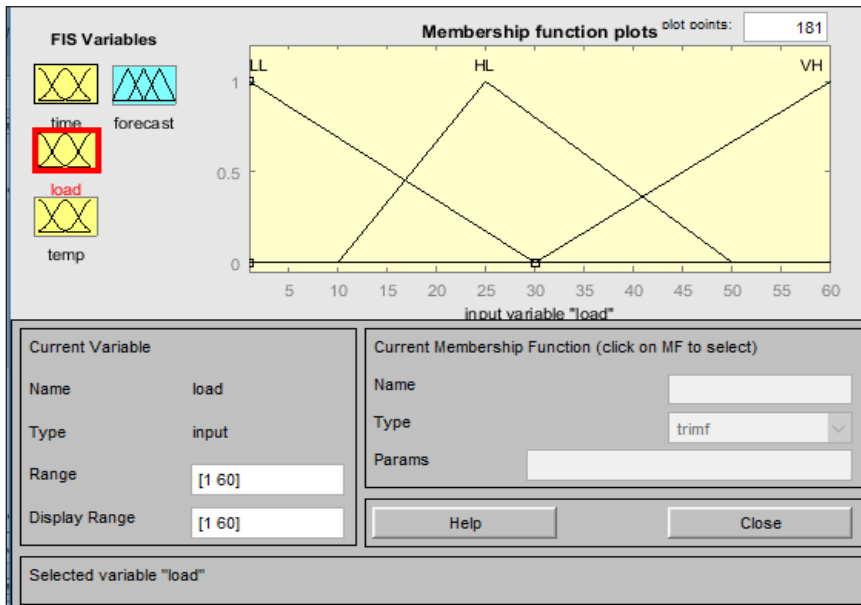


Figure 3. Fuzzy logic triangular MF of load input

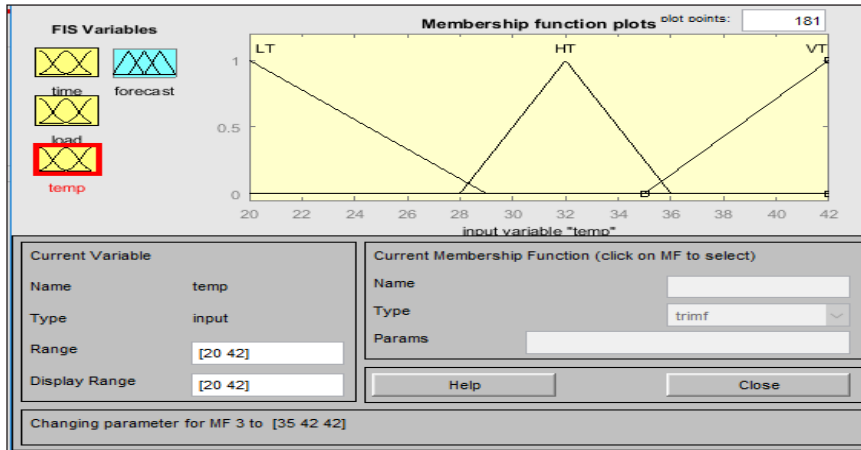


Figure 4. Fuzzy logic triangular MF of temperature input

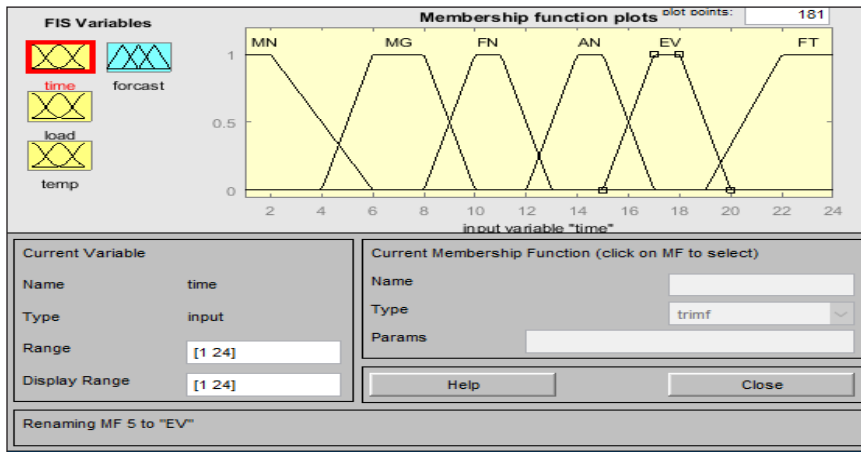


Figure 5. Fuzzy logic trapezoidal MF of time input

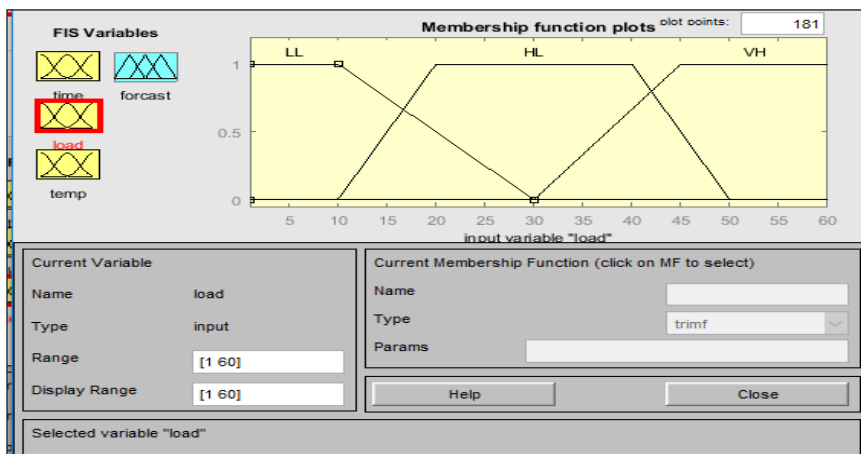


Figure 6. Fuzzy logic trapezoidal MF of load input

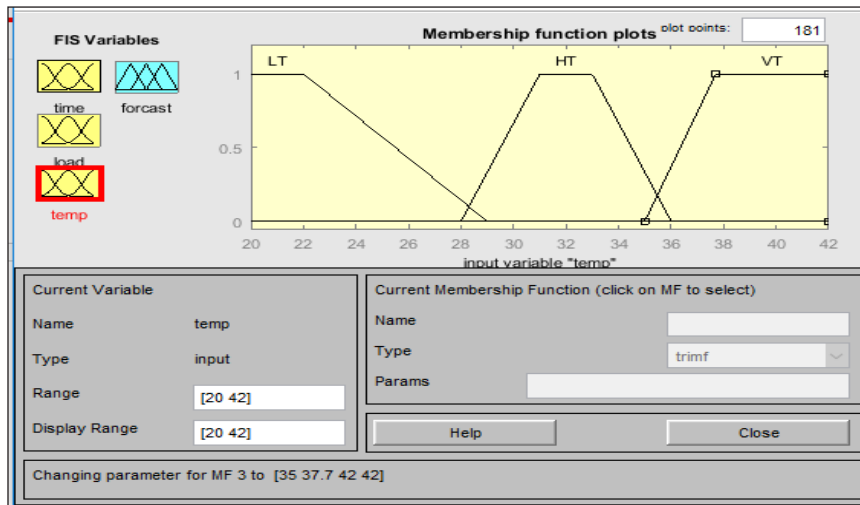


Figure 7. Fuzzy logic trapezoidal MF of temperature input

Fuzzy Rule Base

The fuzzy rule base is a crucial aspect of the fuzzy process because the fuzzy output largely depends on the set rules. In forecasting, the antecedents (input variables) are fed to the fuzzy controller and the rules are applied, then the controller acts on the antecedent and produces the consequences (output). Where there are multiple inputs as applicable in most cases and in this work, the fuzzy operators such as AND, OR and NOT, are used to combine the variable to form a fuzzy judgment. In this study, based on the fuzzy set for the input parameters of time, historic load data, and temperature of the day; sample of the formulated rules are as follows:

- *If*(time is MN) *and* (load is LL) *and* (temperature is LT) *then* (forecast is low)
- *If*(time is MN) *and* (load is HL) *and* (temperature is HT) *then* (forecast is high)
- *If*(time is AN) *and* (load is LL) *and* (temperature is LT) *then* (forecast is low)
- *If*(time is AN) *and* (load is HL) *and* (temperature is VT) *then* (forecast is high)
- *If*(time is FT) *and* (load is LL) *and* (temperature is LT) *then* (forecast is low)
- *If*(time is FT) *and* (load is VH) *and* (temperature is VT) *then* (forecast is very high)

Having three inputs of time, load and temperature that have been fuzzified into data sets as described in “Fuzzy Logic MF”, the fuzzy inference engine or controller combine these inputs into fuzzy results or fuzzified output which is in turn defuzzified into single crisp output. As stated in the first sample formulated rules, when the time is morning with low load demand, and the temperature is low, the three input combination with *if-and-then* statement by the inference engine fuzzified the output to a shape with higher potential for predicting the forecast output load to be low. In the second statement, even though it is morning time, the load demand is high, and the temperature also high, thereby the fuzzy

inference engine consequently fuzzified the membership output to have significant potential for high output load forecast. The fuzzified output MF for the sample rule statement for both the triangular and trapezoidal are presented in Figure 8 (a) and (b), respectively. This serves as the basis for the defuzzification stage to obtain a single crisp output load.

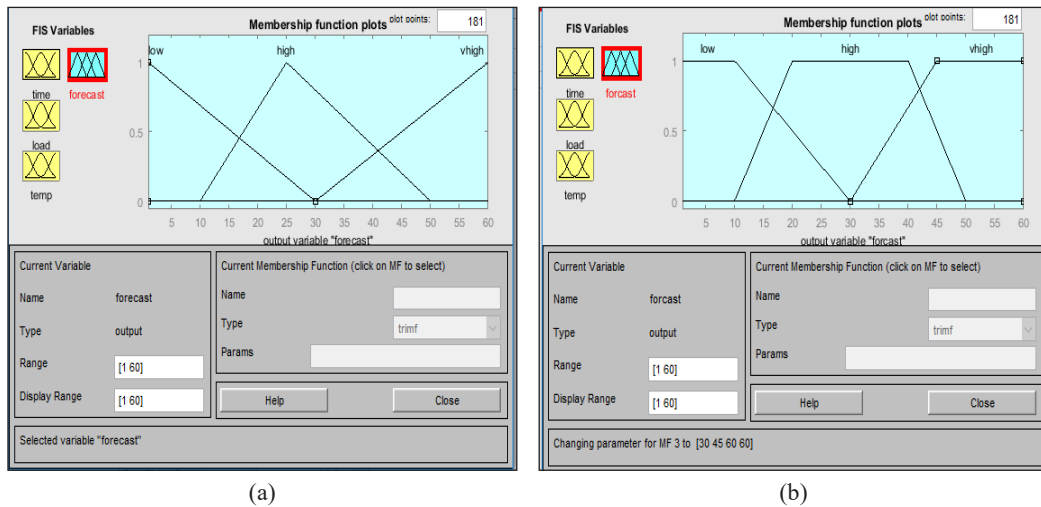


Figure 8. Fuzzy logic (a) triangular MF load output (b) trapezoidal MF load output

Defuzzification

The final stage of the fuzzy process is the defuzzification that converts the fuzzy output to a quantifiable crisp value that can be displayed on the graph. There are different techniques for fuzzy defuzzification including the weighted average method, bisector method, mean of maximum methods, and center of area method (Sivanandam et al., 2007). The centroid of the area is the most commonly used method due to better outcomes, but its main disadvantage is computationally difficulty for complex MFs. In this work, the centroid method was chosen for the defuzzification stage and centroid position is calculated from the Equation 1:

$$W^* = \frac{\sum_{i=1}^n j_i A_i}{\sum_{i=1}^n A_i} \quad (1)$$

Where W^* is the centroid position, j_1 are the center of the area of the fuzzy set in consideration, while A_i is the area of the fuzzy set, and $i = 1, 2, \dots, n$

Fuzzy Logic Interface

The fuzzy logic programming interface is represented in Figure 9 shows the load forecast implementation environment. The time is one of the inputs where the time fuzzy set is classified, the load input is where the load is classified into three fuzzy sets and temperature is an input where the temperature is also classified into three sets. The Mamdani fuzzy typed was employed in the implementation forming the heart of the model where all the rules were actualized. The output of the process is the forecasted load.

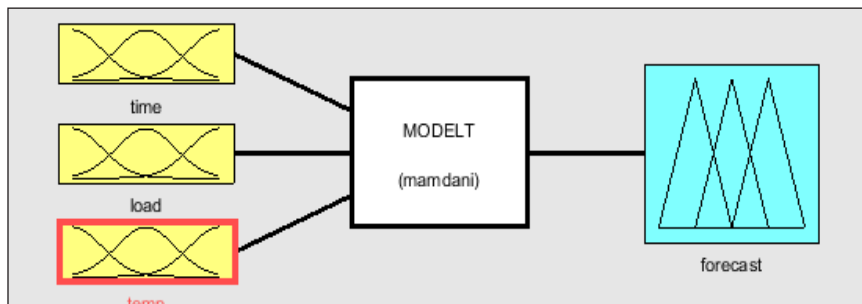


Figure 9. Fuzzy logic interface

Transmission Station Load Data

The electrical load data of the transmission station and the temperature data were collected for one month in March 2019 which served as input to the fuzzy logic controller for the electrical load forecast. The weather temperature was obtained from the world weather for the load center, while the historical load data were recorded at the transmission station on an hourly basis (WeatherOnline, 2019). The Ganmo Transmission Company of Nigeria (TCN) was used as case study and it operates on 330/132/33kV. It received High Voltage Alternating Current (HVAC) of 330kV from Jebba generating station and Oshogbo National Grid station. The Ganmo work center is an Area Control Centre (ACC) with a capacity of 240MW directly serving two sub-transmission stations under its command: Sawmill and Omu-Aran, both in kwara state Nigeria. The two sub-stations operate on 132/33kV from where 33kV is fed to the district's distribution companies and the eligible or special customers.

Presented in Figure 10 is a weekly load demand record between 1st to 7th March 2019 with 55MW maximum and 6MW minimum load demand. This hourly load data are the demands of customers from Sawmill substation comprising of commercial, industrial, and residential customers. As observed from load demand profile, there were quiet irregularity in the hourly load demand pattern as the huge difference between the same hour demand for different days does not represent true load demand. These were due to issues like low generation, non-picking of load by distribution companies and frequent fault on the

transmission lines. To have adequate load data input to the fuzzy controller, averaged weekly hourly load data were computed for the one-month load data representing day 1 to day 4 load profile. Therefore, the average weekly hourly load data was then used as the input load of the station. Similarly, the hourly temperature daily data obtained from 1st to 7th March 2019 is presented in Figure 11.

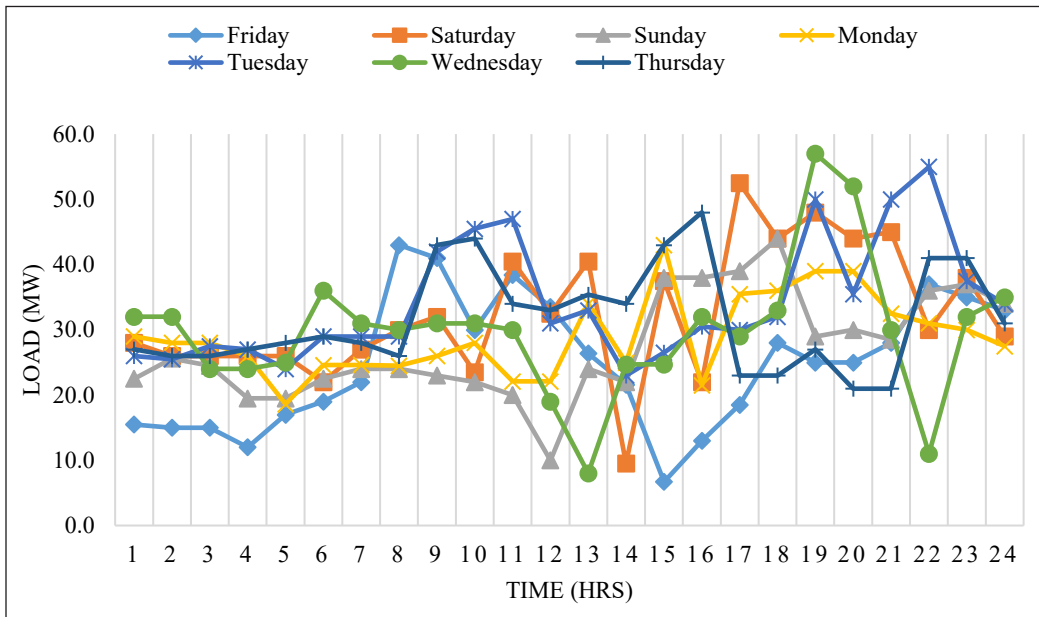


Figure 10. Typical one week daily demand profile

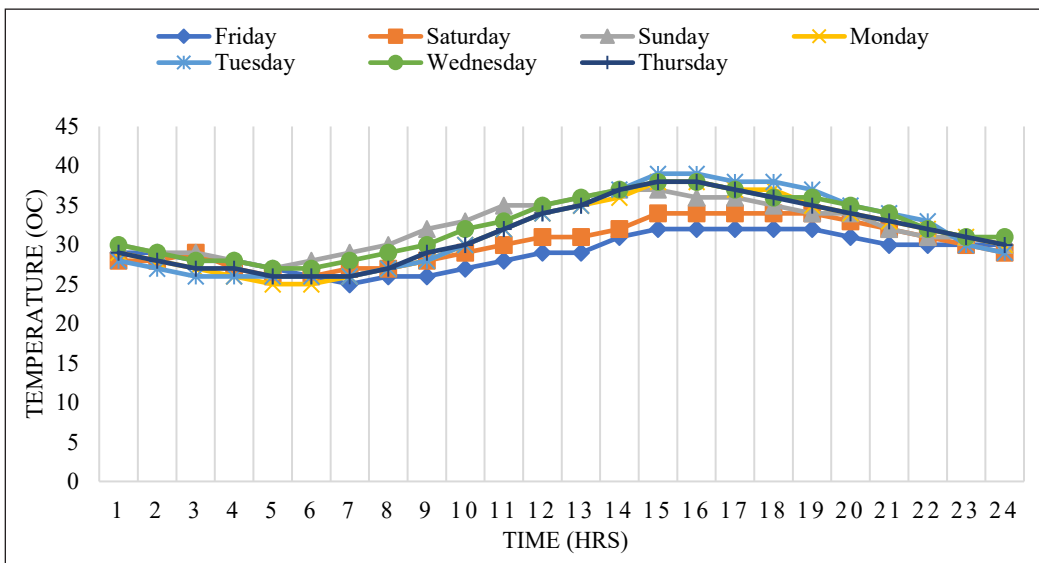


Figure 11. Typical one week daily temperature

Model Error Analysis

Forecast model validation is an important task when forecasting or predicting the future behavior of a system. Comparing the forecasted values with the actual discloses information on the efficiency and accuracy of a forecast model. The testing or validation process involves evaluating the deviation error between the forecasted and the actual data. The statistical approach such as absolute percentage error (APE), MAPE, MFE, and MAD are the most popular method for testing model performance. The aforementioned statistical error analysis methods were used in this study to compute the percentage deviation between the forecasted electrical load and the actual load. The error analysis further shows the effectiveness of the two investigated MFs; triangular and trapezoidal. The APE and MAPE are computed using Equation 2 and 3, respectively. Also, the MFE and MAD were computed using the Equation 4 and 5, respectively.

$$\text{APE (\%)} = \left| \frac{\text{actual}(i) - \text{forecast}(i)}{\text{actual}(i)} \right| \times 100 \quad (2)$$

$$\text{MAPE (\%)} = \frac{1}{n} \sum_{i=1}^n \left| \frac{\text{actual}(i) - \text{forecast}(i)}{\text{actual}(i)} \right| \times 100 \quad (3)$$

$$\text{MFE (\%)} = \frac{1}{n} \sum_{i=1}^n (\text{actual}(i) - \text{forecast}(i)) \times 100 \quad (4)$$

$$\text{MAD (\%)} = \frac{1}{n} \sum_{i=1}^n |\text{actual}(i) - \text{forecast}(i)| \times 100 \quad (5)$$

RESULTS AND DISCUSSION

The fuzzy logic-based electrical load forecast for a transmission substation considering the time of the day, historic load profile, and temperature as input data has been implemented in the MATLAB Simulink environment. The fuzzy logic load forecast model was used to predict the next day load demand and compared with the actual load to ascertain the model performance. Two fuzzy MFs: triangular and trapezoidal were investigated and the percentage error difference between the actual load and forecasted load showed the effectiveness of the forecast model for a day ahead load demand.

Presented in Table 4 is the results of the Day 1 load forecast for 24 hours ahead period with the computed APE and MAPE between the forecast and actual load for both the triangular and trapezoidal MF. It was observed that the APE between the actual and

forecasted load for the triangular MF varied between 1.55 % minimum to a maximum of 17.35 %. The trapezoidal MF was between 0 % minimum to a maximum of 14.48 %. The trapezoidal MF produced improved load forecast with less deviation as this was obvious at 14:00 hours where the forecasted load of 30MW was exact of the 30MW actual load for the day. The MAPE obtained for the triangular membership was 8.93% while that of the trapezoidal membership was 6.34% for the 24 hours.

The MFE measures the average deviation of the forecast from actual loads while MAD measures the average absolute deviation of the forecast from the actual value. The model accuracy depends on the closeness of the MFE and MAD to zero, where the closer to zero the more accurate the model. As presented in Table 4, the error analysis shows that the

Table 4
Time, temperature, actual and forecast load for Day 1

Actual load (MW)	Time (hour)	Temperature (°C)	Forecast Triangular MF (MW)	FE	APE (%)	Forecast Trapezoidal MF (MW)	FE	APE (%)	
25.7	1	28	25.3	0.4	1.55642	27.1	-1.4	5.447471	
25.4	2	28	24.9	0.5	1.968504	26.4	-1	3.937008	
24.4	3	28	23.6	0.8	3.278689	25	-0.6	2.459016	
23.1	4	28	22.9	0.2	0.865801	24.1	-1	4.329004	
22.6	5	27	23.3	-0.7	3.097345	24.1	-1.5	6.637168	
26	6	27	23.8	2.2	8.461538	24.6	1.4	5.384615	
26.5	7	28	25.8	0.7	2.641509	27.8	-1.3	4.90566	
29.5	8	29	25.8	3.7	12.54237	28.1	1.4	4.745763	
34	9	30	28.1	5.9	17.35294	30.3	3.7	10.88235	
32	10	32	26.5	5.5	17.1875	29.1	2.9	9.0625	
33.1	11	34	27.9	5.2	15.70997	30.3	2.8	8.459215	
25.9	12	36	26.4	-0.5	1.930502	28.8	-2.9	11.19691	
29	13	38	30	-1	3.448276	33.2	-4.2	14.48276	
30	14	39	27.4	2.6	8.666667	30	0	0	
31.3	15	41	26.5	4.8	15.33546	29.1	2.2	7.028754	
29.3	16	40	27.6	1.7	5.802048	30	-0.7	2.389078	
32.5	17	39	27.5	5	15.38462	30	2.5	7.692308	
30	18	39	24.1	5.9	19.66667	26.3	3.7	12.33333	
28	19	38	24.3	3.7	13.21429	25.2	2.8	10	
29	20	37	24.1	4.9	16.89655	26	3	10.34483	
29	21	36	25.9	3.1	10.68966	28.4	0.6	2.068966	
28	22	32	25.4	2.6	9.285714	27.9	0.1	0.357143	
27	23	30	26.3	0.7	2.592593	28.8	-1.8	6.666667	
28	24	29	26.1	1.9	6.785714	28.4	-0.4	1.428571	
				2.49167	8.93172			0.42917	6.34330
				(MFE)	(MAPE)			(MFE)	(MAPE)

MFE for trapezoidal MF is 0.4292 which is closer to zero compared to that of the triangular MF with a value of 2.4917. Similarly, the MAD computation from the absolute values of the FE for trapezoidal MF gives a lower error value compared to the Triangular MF. This indicates that the trapezoidal MF load forecast was much closer to actual load demand and produced better performance on short-term load forecasting than the triangular MF.

The Day 2 load forecast is presented in Table 5. It was observed that the load forecast with the trapezoidal MF showed greater linearity with the actual load as compared to the triangular MF. This was further established by the computed MAPE where trapezoidal MF was 6.24% compared to that of a triangular MF of 11.70%. Consistent high daily load demand was observed between the hours of 13:00 hours to 16:00 hours which was due to

Table 5
Time, temperature, actual and forecast load for Day 2

Actual load (MW)	Time (hour)	Temperature (°C)	Forecast Triangular MF (MW)	FE	APE (%)	Forecast Trapezoidal MF (MW)	FE	APE (%)	
31.9	1	29	26.8	5.1	15.9875	29.3	2.6	8.1505	
32.5	2	28	26.7	5.8	17.8462	29	3.5	10.7692	
34.1	3	28	26.2	7.9	23.1672	28.4	5.7	16.7155	
28.8	4	27	25.2	3.6	12.5	27.1	1.7	5.90278	
31.1	5	27	24.5	6.6	21.2219	26.7	4.4	14.1479	
30.7	6	27	27.1	3.6	11.7264	29.4	1.3	4.23453	
30	7	28	28.4	1.6	5.33333	29.7	0.3	1	
30	8	28	30.8	-0.8	2.66667	31	-1	3.33333	
35.5	9	29	30.8	4.7	13.2394	34.7	0.8	2.25352	
34.7	10	31	31.3	3.4	9.7985	32	2.7	7.7810	
30.7	11	33	25.5	5.2	16.9381	33.3	-2.6	8.46906	
36.1	12	34	28.4	7.7	21.3296	29.2	6.9	19.1136	
31.7	13	36	28.5	3.2	10.0946	30.5	1.2	3.78549	
34.2	14	37	29.2	5	14.6199	31.7	2.5	7.30994	
32.4	15	39	28.3	4.1	12.6543	31.7	0.7	2.16049	
34.2	16	39	31.2	3	8.7719	31.1	3.1	9.06433	
31.5	17	38	31.1	0.4	1.26984	32.3	-0.8	2.53968	
31.2	18	38	28.7	2.5	8.01282	31	0.2	0.64103	
32.5	19	36	27.9	4.6	14.1539	30	2.5	7.69231	
29	20	35	28.5	0.5	1.72414	31.2	-2.2	7.58621	
30.8	21	34	28.5	2.3	7.46753	30.7	0.1	0.32466	
30.8	22	32	27.8	3	9.74026	30	0.8	2.59740	
30.3	23	31	27.3	3	9.9010	30	0.3	0.99001	
31	24	31	27.7	3.3	10.6452	30	1	3.22581	
				3.72080	11.70041			1.48750	6.24118
				(MFE)	(MAPE)			(MFE)	(MAPE)

an increase in temperature and more energy was required for temperature control through the use of air-conditioning systems. Considering the lower APE and MAPE errors between the forecasted and real load obtained for both trapezoidal and triangular MF implies that the Fuzzy model performed moderately for a day ahead load forecast.

The error analysis results for Day 2 presented in Table 5 also show that the MFE of triangular MFE was 3.7208 which much greater than zero when compared to that MFE of Trapezoidal MF 1.4875. Likewise, the computed MAD for the triangular MF load forecast had a large value of 3.7875 when compared to that of trapezoidal MF with 2.0375 value.

Table 6
Time, temperature, actual and forecast load for Day 3

Actual load (MW)	Time (hour)	Temperature (°C)	Forecast Triangular MF (MW)	FE	APE (%)	Forecast Trapezoidal MF (MW)	FE	APE (%)	
31	1	29	27.4	3.6	12	30	1	3.225806	
32	2	28	26.5	5.5	17.57188	28.7	3.3	10.3125	
33	3	28	25.8	7.2	25.62278	27.8	5.2	15.75758	
25.8	4	27	24.7	1.1	3.914591	26.5	-0.7	2.713178	
30	5	27	26.8	3.2	9.69697	29.5	0.5	1.666667	
31.3	6	27	26.5	4.8	14.95327	28.6	2.7	8.626198	
28.1	7	28	26.6	1.5	5.050505	29	-0.9	3.202847	
28.1	8	28	26.7	1.4	3.966006	29.1	-1	3.558719	
33	9	29	28.6	4.4	13.7931	29.5	3.5	10.60606	
32.1	10	30	28.4	3.7	10.27778	30.6	1.5	4.672897	
29.7	11	32	28.1	1.6	4.705882	30.4	-0.7	2.356902	
35.3	12	35	30.2	5.1	15.45455	32.5	2.8	7.932011	
31.9	13	36	30.6	1.3	4.193548	33.4	-1.5	4.702194	
36	14	37	30.8	5.2	16.25	32.6	3.4	9.444444	
34	15	38	29.8	4.2	13.54839	34.7	-0.7	2.058824	
33	16	38	33.1	-0.1	0.344828	36.4	-3.4	10.30303	
31	17	37	28.1	2.9	9.666667	30.2	0.8	2.580645	
32	18	37	30.4	1.6	5.16129	32.2	-0.2	0.625	
31	19	35	30.2	0.8	2.666667	32.5	-1.5	4.83871	
29	20	34	27.4	1.6	5.333333	30	-1	3.448276	
30	21	33	27.7	2.3	7.666667	30	0	0	
31	22	32	29	2	6.451613	31.3	-0.3	0.967742	
30	23	31	28.7	1.3	4.333333	31.3	-1.3	4.333333	
30	24	30	28.9	1.1	3.666667	31.4	-1.4	4.666667	
				2.80417	9.01210			0.42083	5.10834
				(MFE)	(MAPE)			(MFE)	(MAPE)

The results from the error analysis on the prediction accuracy of the two investigated triangular and trapezoidal fuzzy MF showed that the triangular MF had better performance.

To further evaluate the performance of the fuzzy load forecast model, a third-day load forecast was performed using the Day 2 actual load, and the results are presented in Table 6. The trapezoidal MF MAPE is 5.11% and triangular MF has a 9.01% error between the predicted and the actual load consumption for the day. Observing the error trend between the actual and forecasted load from Day 1 to Day 3, it can be inferred that the trapezoidal MF performed better than the triangular MF for the fuzzy input mapping. In Table 6, the MFE of triangular MF is 2.8041 while that of trapezoidal MF is 0.4208. This clearly shows that the trapezoidal MF model load forecast prediction is much more accurate compared to the triangular MF. Also, MAD values follows the same pattern where the trapezoidal MF model having 2.8125 and lower than the triangular 1.6375.

The summary of the MAPE, MFE, and MAD error variations between the triangular MF and trapezoidal MF is presented in Figure 12. The error analysis for both MFs is moderate with that of trapezoidal MF producing the lowest error between the actual and forecasted load as compared with that of triangular MF. It is appropriate to note that forecasting errors reported for various short-term load forecasting models in the literature range from 1% to 20%. Having the MAPE, MFE, and MAD values lower than 12% confirm the suitability of the developed fuzzy logic model for short-term load forecasting. With a high degree of closeness between the forecasted and actual load means that utility companies can rely on such outcomes in planning and operating their power generating plants to meet the load demand.

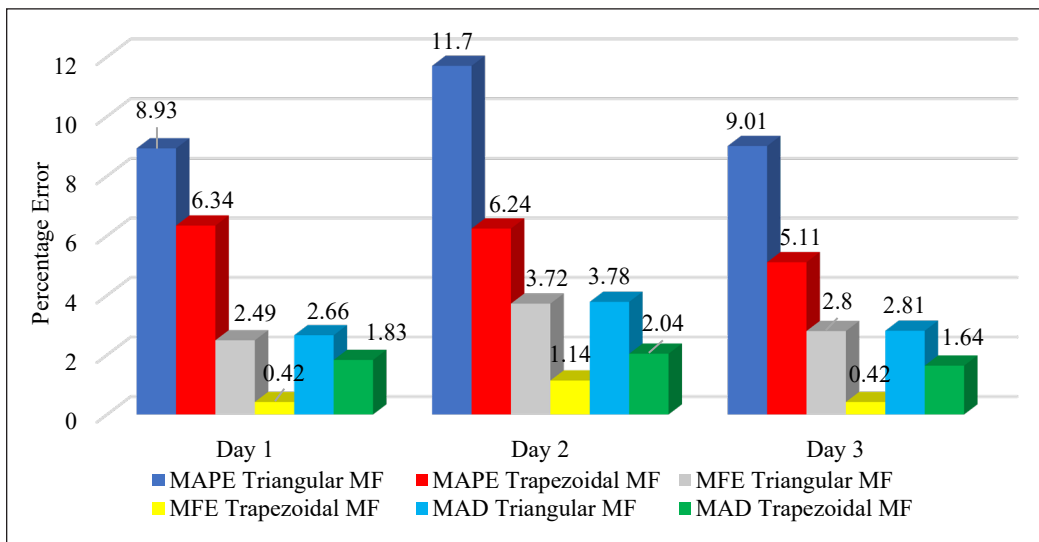


Figure 12. MAPE, MFE and MAD for the Fuzzy logic load forecast model

CONCLUSION

Electrical load forecast models play vital role in the planning and operation of electric power system and consequently lead to improved network reliability. In this work, a fuzzy logic forecast model was developed for short-term load forecast of a day ahead and two MFs; triangular and trapezoidal MFs were investigated. The error analysis between the actual and forecasted load shows that the fuzzy logic is suitable for short-term load forecast of power transmission stations. The fuzzy logic load forecast model performed efficiently with triangular MF having MAPE of 8.93%, 11.7% and 9.01% and trapezoidal 6.34%, 6.24% and 5.11% for the three days tested. The MAPE error for both MF is considerably low, and less than 20% benchmark been reported in the literature. The results show the effect of input data MF mapping on the efficiency and accuracy of the fuzzy logic process as the two investigated functions produce various degree of prediction accuracy between the actual and forecasted load. It was observed from the error analysis results that the trapezoidal MF produced better performance with improved accuracy of load forecast than the triangular MF. To this end, other MFs such as Gaussian, piece-wise, and quadratic polynomial MFs can also be investigated. Aside the MFs, effect of other parts of the fuzzy logic process should be further investigated on short-term load forecast such as rules for the hierarchical fuzzy system. The results in this study have shown that transmission stations can rely on the use of fuzzy model for accurate load forecast.

ACKNOWLEDGEMENT

The authors would like to acknowledge both the University of Ilorin, Nigeria for supporting this work and the Transmission Company of Nigeria (TCN) for making the data available for this research.

REFERENCES

- Al-Kandari, A. M., Soliman, S. A., & El-Hawary, M. E. (2004). Fuzzy short-term electric load forecasting. *International Journal of Electrical Power & Energy Systems*, 26(2), 111-122. [https://doi.org/10.1016/S0142-0615\(03\)00069-3](https://doi.org/10.1016/S0142-0615(03)00069-3)
- Bozkurt, Ö. Ö., Biricik, G., & Tayşi, Z. C. (2017). Artificial neural network and SARIMA based models for power load forecasting in Turkish electricity market. *PLOS ONE*, 12(4), Article e0175915. <https://doi.org/10.1371/journal.pone.0175915>
- Černe, G., Dovžan, D., & Škrjanc, I. (2018). Short-term load forecasting by separating daily profiles and using a single fuzzy model across the entire domain. *IEEE Transactions on Industrial Electronics*, 65(9), 7406-7415. <https://doi.org/10.1109/TIE.2018.2795555>.
- Chikobvu, D., & Sigauke, C. (2012). Regression-SARIMA modelling of daily peak electricity demand in South Africa. *Journal of Energy in Southern Africa*, 23(3), 23-30.

- Cui, H., & Peng, X. (2015). Short-term city electric load forecasting with considering temperature effects: An improved ARIMAX model. *Mathematical Problems in Engineering*, 2015, Article 589374. <https://doi.org/10.1155/2015/589374>
- Danladi, A. D., Puwu, M. I., Michael, Y., & Garkida, B. M. (2016). Use of Fuzzy Logic To Investigate Weather Parameter Impact on Electrical Load Based on Short Term Forecasting. *Nigerian Journal of Technology*, 35(3), 562-567. <http://dx.doi.org/10.4314/njt.v35i3.14>
- Ding, Y., Neumann, M. A., Da Silva, P. G., & Beigl, M. (2013). A framework for short-term activity-aware load forecasting. In *Joint Proceedings of the Workshop on AI Problems and Approaches for Intelligent Environments and Workshop on Semantic Cities* (pp. 23-28). Association for Computing Machinery. <https://doi.org/10.1145/2516911.2516919>
- Dudek, G. (2016). Pattern-based local linear regression models for short-term load forecasting. *Electric Power Systems Research*, 130, 139-147. <https://doi.org/10.1016/j.epsr.2015.09.001>
- Emarati, M., Keynia, F., & Askarzadeh, A. (2019). Application of hybrid neural networks combined with comprehensive learning particle swarm optimization to short-term load forecasting. *Computational Intelligence in Electrical Engineering*, 10(1), 33-40. <http://dx.doi.org/10.22108/isee.2017.21744>
- Fan, G. F., Peng, L. L., Hong, W. C., & Sun, F. (2016). Electric load forecasting by the SVR model with differential empirical mode decomposition and auto regression. *Neurocomputing*, 173, 958-970. <https://doi.org/10.1016/j.neucom.2015.08.051>
- Faysal, M., Islam, M. J., Murad, M. M., Islam, M. I., & Amin, M. R. (2019). Electrical load forecasting using fuzzy system. *Journal of Computer and Communications*, 7(9), 27-37. <http://doi.org/10.4236/jcc.2019.79003>
- Ganguly, A., Goswami, K., Mukherjee, A., & Sil, A. K. (2019). Short-term load forecasting for peak load reduction using artificial neural network technique. In U. Biswas, A. Banerjee, S. Pal, A. Biswas, D. Sarkar & S. Haldar (Eds.), *Advances in computer, communication and control* (pp. 551-559). Springer. https://doi.org/10.1007/978-981-13-3122-0_56
- Ganguly, P., Kalam, A., & Zayegh, A. (2017, May 18-21). Short term load forecasting using fuzzy logic. In *International Conference on Research in Education and Science* (pp. 355-361). Ephesus, Turkey.
- Gohil, P., & Gupta, M. (2014). Short term load forecasting using fuzzy logic 1. *Journal of Engineering Development and Research*, 10(3), 127-130
- Jetcheva, J. G., Majidpour, M., & Chen, W. P. (2014). Neural network model ensembles for building-level electricity load forecasts. *Energy and Buildings*, 84, 214-223. <https://doi.org/10.1016/j.enbuild.2014.08.004>
- Kumar, S., Mishra, S., & Gupta, S. (2016, February). Short term load forecasting using ANN and multiple linear regression. In *2016 Second International Conference on Computational Intelligence & Communication Technology (CICT)* (pp. 184-186). Conference Publishing Services. <https://doi.org/10.1109/CICT.2016.44>
- Kuster, C., Rezgui, Y., & Mourshed, M. (2017). Electrical load forecasting models: A critical systematic review. *Sustainable Cities and Society*, 35, 257-270. <https://doi.org/10.1016/j.scs.2017.08.009>
- Lei, J., Jin, T., Hao, J., & Li, F. (2019). Short-term load forecasting with clustering–regression model in distributed cluster. *Cluster Computing*, 22(4), 10163-10173. <https://doi.org/10.1007/s10586-017-1198-4>

- Mi, J., Fan, L., Duan, X., & Qiu, Y. (2018). Short-term power load forecasting method based on improved exponential smoothing grey model. *Mathematical Problems in Engineering*, 2018, Article 3894723. <https://doi.org/10.1155/2018/3894723>
- Peng, Y., Wang, Y., Lu, X., Li, H., Shi, D., Wang, Z., & Li, J. (2019, May). Short-term load forecasting at different aggregation levels with predictability analysis. In *2019 IEEE Innovative Smart Grid Technologies-Asia (ISGT Asia)* (pp. 3385-3390). Institute of Electrical and Electronics Engineers, Inc. <http://doi.org/10.1109/ISGT-Asia.2019.8881343>
- Razak, I. A. W. A., Majid, S., Aras, M. S. M., & Ahmad, A. (2012). Electricity load forecasting using data mining technique. In A. Karahoca (Ed.), *Advance in data mining knowledge discovery and application* (pp. 235-254). IntechOpen. <http://dx.doi.org/10.5772/48657>
- Rizwan, M., Kumar, D., & Kumar, R. (2012). Fuzzy logic approach for short term electrical load forecasting. *Electrical and Power Engineering Frontier*, 1(1), 8-12.
- Sadaei, H. J., e Silva, P. C. L., Guimarães, F. G., & Lee, M. H. (2019). Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series. *Energy*, 175, 365-377. <https://doi.org/10.1016/j.energy.2019.03.081>
- Silva, G. C., Silva, J. L., Lisboa, A. C., Vieira, D. A., & Saldanha, R. R. (2017, November). Advanced fuzzy time series applied to short term load forecasting. In *2017 IEEE Latin American Conference on Computational Intelligence (LA-CCI)* (pp. 1-6). IEEE Computational Intelligence Society. <http://doi.org/10.1109/LA-CCI.2017.8285726>
- Singhal, R., Choudhary, N. K., & Singh, N. (2020). Short-term load forecasting using hybrid ARIMA and artificial neural network model. In D. Dutta, H. Kar, C. Kumar & V. Bhaduria (Eds.), *Advances in VLSI, Communication, and Signal Processing* (pp. 935-947). Springer. https://doi.org/10.1007/978-981-32-9775-3_83
- Siri, S. T. (2018). *Short-term load forecasting using a hybrid of genetic algorithm (Ga) and particle swarm optimization (Pso) for an optimized neural network* [Doctoral dissertation, University of Nairobi]. University of Nairobi Publications. <http://hdl.handle.net/11295/105048>
- Sivanandam, S. N., Sumathi, S., & Deepa, S. N. (2007). *Introduction to fuzzy logic using MATLAB* (Vol. 1). Springer.
- Srivastava, A. K., Pandey, A. S., & Singh, D. (2016, March). Short-term load forecasting methods: A review. In *2016 International conference on emerging trends in electrical electronics & sustainable energy systems (ICETEESES)* (pp. 130-138). Institute of Electrical and Electronics Engineering, Inc. <http://doi.org/10.1109/ICETEESES.2016.7581373>
- WeatherOnline, W. (2019). *Hourly temperature data of Ilorin, Nigeria*. Retrieved April 24, 2019, from www.worldweatheronline.com.
- Wen, Z., Xie, L., Fan, Q., & Feng, H. (2020). Long term electric load forecasting based on TS-type recurrent fuzzy neural network model. *Electric Power Systems Research*, 179, Article 106106. <https://doi.org/10.1016/j.epsr.2019.106106>
- Yu, F., & Xu, X. (2014). A short-term load forecasting model of natural gas based on optimized genetic algorithm and improved BP neural network. *Applied Energy*, 134, 102-113. <https://doi.org/10.1016/j.apenergy.2014.07.104>

Comparison of Single and MICE Imputation Methods for Missing Values: A Simulation Study

Nurul Azifah Mohd Pauzi¹, Yap Bee Wah^{1,2*}, Sayang Mohd Deni^{1,2}, Siti Khatijah Nor Abdul Rahim³ and Suhartono⁴

¹Centre of Statistical and Decision Science Studies, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

²Advanced Analytics Engineering Centre, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

³Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam, Selangor, Malaysia

⁴Department of Statistics and Data Science, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

ABSTRACT

High quality data is essential in every field of research for valid research findings. The presence of missing data in a dataset is common and occurs for a variety of reasons such as incomplete responses, equipment malfunction and data entry error. Single and multiple data imputation methods have been developed for data imputation of missing values. This study investigated the performance of single imputation using mean and multiple imputation method using Multivariate Imputation by Chained Equations (MICE) via a simulation study. The MCAR which means missing completely at random were generated randomly for ten levels of missing rates (proportion of missing data): 5% to 50% for different sample sizes. Mean Square Error (MSE) was used to evaluate the performance of the imputation methods. Data imputation method depends on data types. Mean imputation is commonly used

to impute missing values for continuous variable while MICE method can handle both continuous and categorical variables. The simulation results indicate that group mean imputation (GMI) performed better compared to overall mean imputation (OMI) and MICE with lowest value of MSE for all sample sizes and missing rates. The MSE of OMI, GMI, and MICE increases when missing rate increases. The MICE method has the lowest performance (i.e. highest

ARTICLE INFO

Article history:

Received: 24 September 2020

Accepted: 30 December 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.15>

E-mail addresses:

azifahpauzi93@yahoo.com (Nurul Azifah Mohd Pauzi)

beewah@tmsk.uitm.edu.my (Yap Bee Wah)

sayang@tmsk.uitm.edu.my (Sayang Mohd Deni)

sitikhathijahnor@tmsk.uitm.edu.my

(Siti Khatijah Nor Abdul Rahim)

suhartono@statistika.its.ac.id (Suhartono)

* Corresponding author

MSE) when percentage of missing rates is more than 15%. Overall, GMI is more superior compared to OMI and MICE for all missing rates and sample size for MCAR mechanism. An application to a real dataset confirmed the findings of the simulation results. The findings of this study can provide knowledge to researchers and practitioners on which imputation method is more suitable when the data involves missing data.

Keywords: MICE, missing data, multiple imputation, simulation, single imputation

INTRODUCTION

High quality data is important to ensure correct information and valid findings for better evidence-based decision-making. One of the key concerns related to data quality is missing data. Missing data can occur for various reasons during the data collection process, such as incomplete responses (Pampaka et al., 2016; Barnett et al., 2017), equipment malfunction (Masconi et al., 2015; Gopal et al., 2019) and manual data entry errors (Bhati & Gupta, 2016). Incomplete data is a serious data quality problem since it leads to a reduction of statistical power, bias in parameter estimates, and loss of efficiency in the analytical process (Kaiser, 2014; Ayilara et al., 2019; Hughes et al., 2019). These problems have led to extensive research on developing methods to treat missing data.

Data imputation is widely used to deal with missing data in many areas such as medical research (Pedersen et al., 2017; Sullivan et al., 2018; Turner et al., 2019; Stavseth et al., 2019), organizational research (Newman, 2003; Fichman & Cummings, 2003; Newman, 2014) and educational research (Grund et al., 2018; Shi et al., 2019). The main objective of data imputation is to replace any missing data with estimated values to obtain a complete dataset. The estimated values could be from the calculated mean, median, mode, predefined value of missing variable or values that are obtained from the predictive model (Malarvizhi & Thanamani, 2012).

One of the important considerations in addressing missing data is identifying the mechanism of missing data. There are three types of missing data mechanisms known as MCAR, missing at random (MAR) or missing not at random (MNAR). The mechanism and pattern of missing data plays an important role during the selection of imputation methods. They have greater impact on research results compared to the proportion of missing data (Tabachnick et al., 2007; Song & Shepperd, 2007). There is a need to address these issues properly to help in selecting appropriate data imputation methods. If a study employs improper imputation methods, it may lead to incorrect data analysis, false conclusions, and erroneous predictions (Chaudhry et al., 2019). Reliable data imputation methods are needed as the outcome of data analysis tasks relies upon efficient handling of missing data.

Several data imputation methods have been developed such as mean imputation, hot deck imputation, regression imputation and Multivariate Imputation by Chained

Equations (MICE). The simplest and commonly used method is by replacing the missing data with the mean of the continuous variable. Various studies have been conducted on the performance of data imputation methods. The study conducted by Le et al. (2018) compared the performance of four data imputation method, Expectation Maximization (EM), Multiple Imputation, K-Nearest Neighbor and Mean Imputation in healthcare dataset. The results showed that imputation using EM outperformed other methods with lowest RMSE value. Jadhav et al. (2019) compared seven imputation methods namely mean imputation, median imputation, k-NN imputation, predictive mean matching, Bayesian Linear Regression (norm), Linear Regression, non-Bayesian (norm.nob). They reported that k-NN imputation method is the more superior imputation method. Recently, Kamatchi and Baranidharan (2019) proposed deep learning imputation method (DNN) and compared it with statistical imputation methods (Mean, imputation with zero or constant, Stochastic Regression imputation, Extrapolation and Interpolation and Hot-Deck imputation, MICE), machine learning methods (K-nearest neighbour) using an Autism dataset. Their results showed that DNN outperformed statistical and machine learning methods. The study by Ochieng'Odhiambo (2020) investigates the three most common conventional methods (Listwise, Mean Imputation, Median Imputation) in handling missing data. Their finding showed that median imputation was a better data imputation method among the conventional methods.

The aim of this study is to evaluate the performance of single and multiple imputation methods for continuous variable via a simulation study. The selected single imputation method is mean imputation while the multiple imputation method is MICE. In this simulation study, MCAR missing data were generated for different sample size and missing rates ranging from 5% to 50%. Then, the performance of these imputation methods was evaluated based on Mean Square Error (MSE). The simulation approach allows the researcher to control the parameter in order to study the pattern and changes in the estimation of each imputation method. The simulation procedures were carried out using R programming language software. The imputation methods were applied to a real dataset taken from the Kaggle website to validate the findings of the simulation study.

MATERIALS AND METHODS

Missing Data Mechanisms

The mechanism of missing data describes the reason the values are missing. Little and Rubin (1987) classified the mechanisms of missing data into two major types known as ignorable and non-ignorable missingness. Non-ignorable means that the missing data mechanism is related to the missing values. It is referred to as non-ignorable since the missing data mechanism itself must be modelled when dealing with missing data. There is a need to include some model for why the data are missing and what the likely values

are. Ignorable missingness however ignores any information about the missing data itself when dealing with missing data. Missing Completely at Random (MCAR) and Missing at Random (MAR) are classified under ignorable missingness, whereas Missing Not at Random (MNAR) is non-ignorable missingness (Aljuaid & Sasi, 2016; Chhabra et al., 2017). The mechanism of missing data is determined by the dependency of K on the variables in the data set where K represents the missing data indicator, Y_{obs} is the observed values, and Y_{miss} is the missing values.

The probability of missingness is related to observed data but not on the missing data itself for the MAR mechanism (Ma & Chen, 2018; van Ginkel et al., 2019; Li et al., 2019). There is a systematic relationship between missing data and observed data in the dataset. The MAR assumption is defined as Equation 1:

$$Pr(K|Y_{obs}, Y_{miss}) = Pr(K|Y_{obs}) \quad [1]$$

while the stronger assumption of MCAR can be written as Equation 2:

$$Pr(K|Y_{obs}, Y_{miss}) = Pr(K) \quad [2]$$

which implies the missing data indicator, K is completely unrelated to both observed and missing data (Aljuaid & Sasi, 2016; Abidin et al., 2018; Madley-Dowd et al., 2019). MCAR is a special case of the MAR mechanism and represents the highest level of randomness. Finally, any missing data that does not satisfy Equation 1 and 2 is classified as MNAR where the missing indicators, K are related to missing data itself (Dettori et al., 2018; Goretzko et al., 2019). The assumption of the MNAR mechanism can be expressed as Equation 3:

$$Pr(K|Y_{obs}, Y_{miss}) = Pr(K|Y_{miss}) \quad [3]$$

Data Imputation Methods

Data imputation is a powerful and widely used method in treating missing data. An imputation method preserves the sample size of data by imputing with estimated values without discarding cases with missing data. There are single or multiple imputation methods (Dettori et al., 2018). Single imputation treats missing data by replacing with a single value for all missing data of that variable (Pederson et al., 2017; Papageorgiou, 2018). As the same values are used to replace each missing data, it ignores the uncertainty of the parameter estimates (Salgado et al., 2016; Yadav & Roychoudhury, 2018). Multiple imputations overcome this problem by considering the uncertainty associated with missing

data, which was unaccounted for in single imputation (Pedersen et al., 2017). However, single imputation methods are popular among the researchers partly due to their simplicity and availability in many statistical software. The most basic single imputation is arithmetic mean imputation proposed by Wilks (1932). Mean imputation can be categorized into overall mean imputation (OMI) and group mean imputation (GMI). Thus, OMI and GMI are categorized under single imputation while Multivariate Imputation by Chained Equations (MICE) falls under multiple imputation methods.

Single Imputation

Overall Mean Imputation (OMI). In overall mean imputation, the mean of the observed values of the variable is used to impute missing data in the same variable. Thus, the mean of the non-missing values of that particular variable will be calculated and then used to replace each missing value. This can be expressed by the mathematical Equation 4 (Sim et al., 2015):

$$X_i^j = \sum_{k \in I(\text{complete})} \frac{X_k^j}{n | I(\text{complete})|} \quad [4]$$

where X_i^j is the i^{th} instance (or case) for the j^{th} variable, while X_k^j is the non-missing value of the j^{th} variable. $I(\text{complete})$ is an index set with 1 if X_i^j is non-missing and 0 otherwise, and $n | I(\text{complete})|$ is the total number of observed values (non-missing cases) in the j^{th} variable.

Group Mean Imputation (GMI). Group mean imputation is also known as similar case imputation. In GMI, the mean of the observed data is calculated based on the underlying group. Thus, the mean of non-missing data of that particular variable will be calculated separately. For instance, mean of group based on male and female and then replace each missing data according to the gender. The group mean can be expressed by the mathematical Equation 5:

$$X_{m,i}^j = \sum_{k \in I(\text{complete})} \frac{X_{m,k}^j}{n | I(\text{complete})|} \quad [5]$$

where $X_{m,i}^j$ is the i^{th} instance (or case) for the j^{th} variable of the m^{th} class. $I(\text{complete})$ is an index set with 1 if $X_{m,i}^j$ is non-missing and 0 otherwise, and $n | I(\text{complete})|$ is the total number of observed values (non-missing cases) in the j^{th} variable of the m^{th} class.

Multiple Imputation

Multivariate Imputation by Chained Equations (MICE). Multivariate Imputation by Chained Equations is a particular multiple imputation technique (Van Buuren, 2007). MICE is also known as fully conditional specification (FCS) that is widely used in handling missing data. It is an extension of single imputation that gives much better results since missing data are estimated multiple times (m times) which reflects the uncertainty of parameter estimates of the imputed variables (Zhang, 2016). In the MICE procedure, imputation is carried out by conducting series of estimations whereby each variable with missing data is modeled according to its distribution. An introduction to the MICE method is given in the paper by Buuren and Groothuis-Oudshoorn (2010) which provided good practical resources to guide researchers in implementing this method. The MICE imputation method can be described in a three-step procedure (Ratolojanahary et al., 2019; Lo et al., 2019):

- (i) Imputation phase
- (ii) Analysis phase
- (iii) Pooling phase

In the imputation phase, missing data are imputed for $m > 1$ times resulting in multiple imputed datasets (m imputed datasets). Then, in the analysis phase, each m imputed dataset is analyzed separately to obtain the parameter estimates and standard errors (m point estimates and standard errors) by using a standard statistical procedure. In the pooling phase, the m point estimates and standard errors for each analysis are then pooled together to get overall estimates and standard errors. The estimated parameter P_i can be denoted by \hat{P}_i from the i^{th} dataset. Then, the pooled estimates of the parameter P can be obtained as Equation 6 (Dong & Peng, 2013):

$$\bar{P} = \frac{1}{m} \sum_{i=1}^m \hat{P}_i \quad [6]$$

while \bar{C} is the within imputation variance (Equation 7)

$$\bar{C} = \frac{1}{m} \sum_{i=1}^m \hat{V}_i \quad [7]$$

which is the average of the estimated variances for i^{th} imputed datasets. The variance estimates from the i^{th} dataset is denoted as \hat{V}_i . The between imputation variance, W is the variability of the imputed values for multiple imputed datasets and can be written as Equation 8:

$$W = \frac{1}{m-1} \sum_{i=1}^m (\hat{P}_i - \bar{P})^2 \quad [8]$$

The variance of the pooled estimate is the weighted sum of two variances (within imputation (\bar{C}), between imputation variance (W) and can be written as Equation 9:

$$\text{Var}(\bar{P}) = \bar{C} + (1 + \frac{1}{m})W \quad [9]$$

where ($\frac{1}{m}$) is an adjustment for the randomness associated with a finite number of imputations.

Performance Evaluation Measure

The evaluation of the imputation methods is based on Mean Square Error (MSE). Based on the past studies (Gad & Abdelkhalek, 2017; Abidin et al., 2018; Nwakuya & Nwabueze, 2018), MSE is the most commonly used metric in *evaluate of the performance* of different *imputation algorithms* for continuous variables. The MSE measures the discrepancy between imputed and actual observed values. The best imputation method will be the one with the lowest MSE which indicates that the predicted values are close to the actual values. MSE is calculated as Equation 10:

$$\text{MSE} = \frac{\sum_{i=1}^n (x_i^{\text{actual}} - x_i^{\text{imputed}})^2}{n} \quad [10]$$

where x^{actual} is the actual observed values and x^{imputed} is the imputed values, and n is the number of cases (observations).

METHODS

Study 1: Simulation Study

The simulation study was carried out to provide empirical evidence on the performance of single and MICE imputation methods in handling missing data. The simulations were conducted using R programming language software. This study adopted the simulation approach by Wah et al. (2018) for generating data with a binary Y and continuous independent variables. As most data imputation methods depend heavily on the assumption of normality (Scheffer, 2002), two continuous independent variables (X_1, X_2) were simulated from standard normal distributions. Data were generated for various sample sizes, $n(30, 50, 100, 200, 500, 1000)$ and ten levels of missing rates, 5% to 50% (with 5% increment) under the MCAR mechanism. The MCAR approach is by removing data randomly from a variable. Little's MCAR Test (Little, 1988) was used to test the assumption of MCAR. The null hypothesis (H_0) is missing data are MCAR. If we reject the null hypothesis under

Little's Test of MCAR, the data may then be assumed to be MAR or MNAR. Some packages in R were used to carry out the simulation study. The 'missForest' (Stekhoven & Bühlmann, 2012) package was used to generate missing data MCAR randomly from the simulated dataset and 'mvnmlc', and 'BayorEdPsych' were used for conducting Little's MCAR Test. The 'mice' (Van Buuren et al., 1999) package was used to impute incomplete multivariate data by chained equations. Although, simulation study was carried out in previous works to compare the performance of imputation method, the findings were based on different parameters such as missing data mechanism, sample size, percentage of missing data. This simulation study, however, dealt with all these parameters simultaneously for continuous missing variables. The simulation involves 1000 replications.

The procedure of the simulation process is as follows:

1. Simulate n data for two continuous independent variables from standard normal distributions $N(0,1)$.
2. Calculate $z = (0.7 + 1.08 * X_1 + 1.69 * X_2)$ and $\pi(x) = \frac{1}{1 + e^{-z}}$
3. Simulate the data u from a random uniform distribution, $U(0,1)$.
4. Generate the Y dependent variable for binary logistic regression by using the rule of $y=1$ if $u \leq \Pr(Y)$, if else assign $y = 0$.
5. Artificially create ten levels of missing rates under the MCAR mechanism: 5% to 50% (with 5% increment)
6. Test the assumption of MCAR using Little's Test of MCAR.
7. Apply data imputation method for each sample size with different missing rates.
8. Obtain MSE.
9. Perform 1000 simulations and average MSE over 1000 simulations.

Steps 1 to 9 were repeated for each sample size, $n(30, 50, 100, 200, 500, 1000)$.

Study 2: Application to a Real Dataset

To validate the findings of the simulation study, the imputation methods were applied to a Human Resource (HR) dataset which was taken from the Kaggle website. The Human Resource dataset consists of 15000 observations with 10 features. The outcome variable is status which represents the status of employee turnover (1 = Left (23.81%) and 0 = Stay (76.19%)). The selected variables are sl and le which represent the values of satisfaction level and last evaluation, respectively. The variables sl and le are selected as these variables are continuous variable while the other variables (number of projects, average monthly hours, time spend at company, accident at work, promotion last 5 Years, position at work, level of salary) were classified as discrete and categorical variables. The summary statistics of satisfaction level (sl) and last evaluation (le) are given in Table 1.

Table 1
Summary Statistics

Variable	Group	Mean	Standard Deviation	Skewness	Kurtosis
Satisfaction Level (<i>sl</i>)	1=Left	0.4402	0.2640	0.291	-1.034
	0= Stay	0.6668	0.2171	-0.605	-0.215
Last Evaluation (<i>le</i>)	1=Left	0.7182	0.1977	-0.014	-1.710
	0= Stay	0.7155	0.1620	-0.039	-1.014

Using the original dataset, sample sizes of 30, 50, 100, 200, 500, and 1000 were randomly selected from the 15000 observations to study the effect of small, medium and large sample size. We follow the work by Cheema (2014), who selected 10 sub-samples size (small to large sample) from ($n = 10000$) simulated dataset to study the effect of sample size on the performance of missing data treatment. Missing data in the HR dataset were generated randomly for 5% to 50% missing rates. Single imputation and MICE methods were then applied and the MSE was obtained to evaluate the performance of each imputation method.

RESULTS AND DISCUSSION

Study 1: Simulation Results

This section presents the simulation results. Table 2 displays the result of Little's Test of MCAR which tests the assumption of MCAR. The result indicates that the assumption of MCAR is satisfied since the p-value for all sample sizes and missing rates is greater than 0.05. This test confirmed that missing data were simulated under the MCAR mechanism.

Table 3 shows the simulation results of each imputation method for different levels of missing rates and sample sizes. The best imputation method is determined based on the lowest MSE.

Based on the simulation results shown in Table 3 and line chart in Figure 1, the MSE of group mean imputation (GMI) is the lowest compared to OMI and MICE regardless of missing rates and small sample size. Imputation using MICE has the highest MSE compared to OMI and GMI regardless of missing rates and small sample size.

Thus, in terms of single imputation methods, GMI is more superior compared to OMI. The line chart in Figure 1 also show that the MSE for all three methods (OMI, GMI, MICE) increases when missing rate increases with MICE having the highest increase in MSE. This simulation results support the findings by Schmitt et al. (2015), whose results showed that the RMSE for mean imputation and MICE increased with increasing missing rates.

Overall, it can be concluded that single imputation method by group mean (GMI) is more superior compared to overall mean (OMI) and MICE in treating missing data for all sample sizes regardless of missing rates. In addition, it is not recommended to perform imputation using these methods if missing rate is more than 15% due to the large MSE.

Table 2
Little's Test for Missing Completely at Random (MCAR) results

Sample Size	p-value									
	5	10	15	20	25	30	35	40	45	50
30	0.514	0.6068	0.567	0.101	0.337	0.2901	0.202	0.152	0.082	0.125
50	0.867	0.2858	0.350	0.928	0.895	0.409	0.340	0.059	0.230	0.651
100	0.398	0.8264	0.409	0.136	0.817	0.077	0.800	0.700	0.766	0.141
200	0.910	0.3677	0.431	0.287	0.141	0.333	0.914	0.689	0.494	0.494
500	0.725	0.9612	0.324	0.568	0.575	0.978	0.879	0.959	0.959	0.782
1000	0.757	0.6522	0.858	0.097	0.057	0.281	0.145	0.237	0.436	0.879

Table 3
Mean Square Error (MSE) of Imputation Methods

Sample Size	Missing Rates									
	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
30	0.02927 ^a	0.08895 ^a	0.11587 ^a	0.17729 ^a	0.20589 ^a	0.26562 ^a	0.29670 ^b	0.35766 ^c	0.38986 ^c	0.45146 ^c
	0.02125^b	0.07179^b	0.10279^b	0.14545^b	0.14761^b	0.19056^b	0.21375^b	0.25976^b	0.28295^b	0.33305^b
	0.03965 ^c	0.12101 ^c	0.16148 ^c	0.24238 ^c	0.28065 ^c	0.35642 ^c	0.40207 ^c	0.48558 ^c	0.52647 ^c	0.61462 ^c
50	0.03387 ^a	0.08573 ^a	0.12013 ^a	0.17208 ^a	0.20512 ^a	0.25671 ^a	0.29322 ^a	0.34744 ^a	0.38039 ^a	0.43655 ^a
	0.02852^b	0.04853^b	0.06673^b	0.09590^b	0.17525^b	0.21805^b	0.25093^b	0.29541^b	0.32792^b	0.37603^b
	0.05551 ^c	0.13501 ^c	0.19408 ^c	0.27735 ^c	0.33524 ^c	0.41746 ^c	0.47461 ^c	0.56411 ^c	0.62280 ^c	0.70729 ^c
100	0.03845 ^a	0.07679 ^a	0.11372 ^a	0.15324 ^a	0.19081 ^a	0.22856 ^a	0.26824 ^a	0.30717 ^a	0.34525 ^a	0.38503 ^a
	0.03170^b	0.06234^b	0.09321^b	0.12628^b	0.15665^b	0.18701^b	0.21976^b	0.25217^b	0.28441^b	0.31652^b
	0.05982 ^c	0.11750 ^c	0.17922 ^c	0.23702 ^c	0.29774 ^c	0.35807 ^c	0.42519 ^c	0.48188 ^c	0.55260 ^c	0.61548 ^c
200	0.03458 ^a	0.06821 ^a	0.10335 ^a	0.13819 ^a	0.17159 ^a	0.20778 ^a	0.24161 ^a	0.27560 ^a	0.31157 ^a	0.34504 ^a
	0.02916^b	0.05704^b	0.08658^b	0.11575^b	0.14402^b	0.17396^b	0.20243^b	0.23167^b	0.26216^b	0.29063^b
	0.05844 ^c	0.11660 ^c	0.17840 ^c	0.23530 ^c	0.29690 ^c	0.35542 ^c	0.41574 ^c	0.47448 ^c	0.53077 ^c	0.58641 ^c
500	0.03350 ^a	0.06679 ^a	0.10000 ^a	0.13384 ^a	0.16717 ^a	0.20040 ^a	0.23398 ^a	0.26686 ^a	0.30102 ^a	0.33327 ^a
	0.02774^b	0.05535^b	0.08313^b	0.11145^b	0.13881^b	0.16665^b	0.19434^b	0.22233^b	0.25044^b	0.27723^b
	0.05276 ^c	0.10646 ^c	0.15951 ^c	0.21291 ^c	0.26779 ^c	0.32188 ^c	0.37583 ^c	0.43280 ^c	0.48800 ^c	0.54432 ^c
1000	0.03371 ^a	0.06703 ^a	0.10093 ^a	0.13446 ^a	0.16781 ^a	0.20181 ^a	0.23529 ^a	0.26870 ^a	0.30302 ^a	0.33634 ^a
	0.02772^b	0.05514^b	0.08292^b	0.11064^b	0.13803^b	0.16581^b	0.19344^b	0.22087^b	0.24896^b	0.27680^b
	0.05449 ^c	0.10949 ^c	0.16325 ^c	0.21909 ^c	0.27310 ^c	0.32938 ^c	0.38511 ^c	0.43943 ^c	0.49507 ^c	0.55291 ^c

Notes: a=OMI; b=GMI; c=MICE

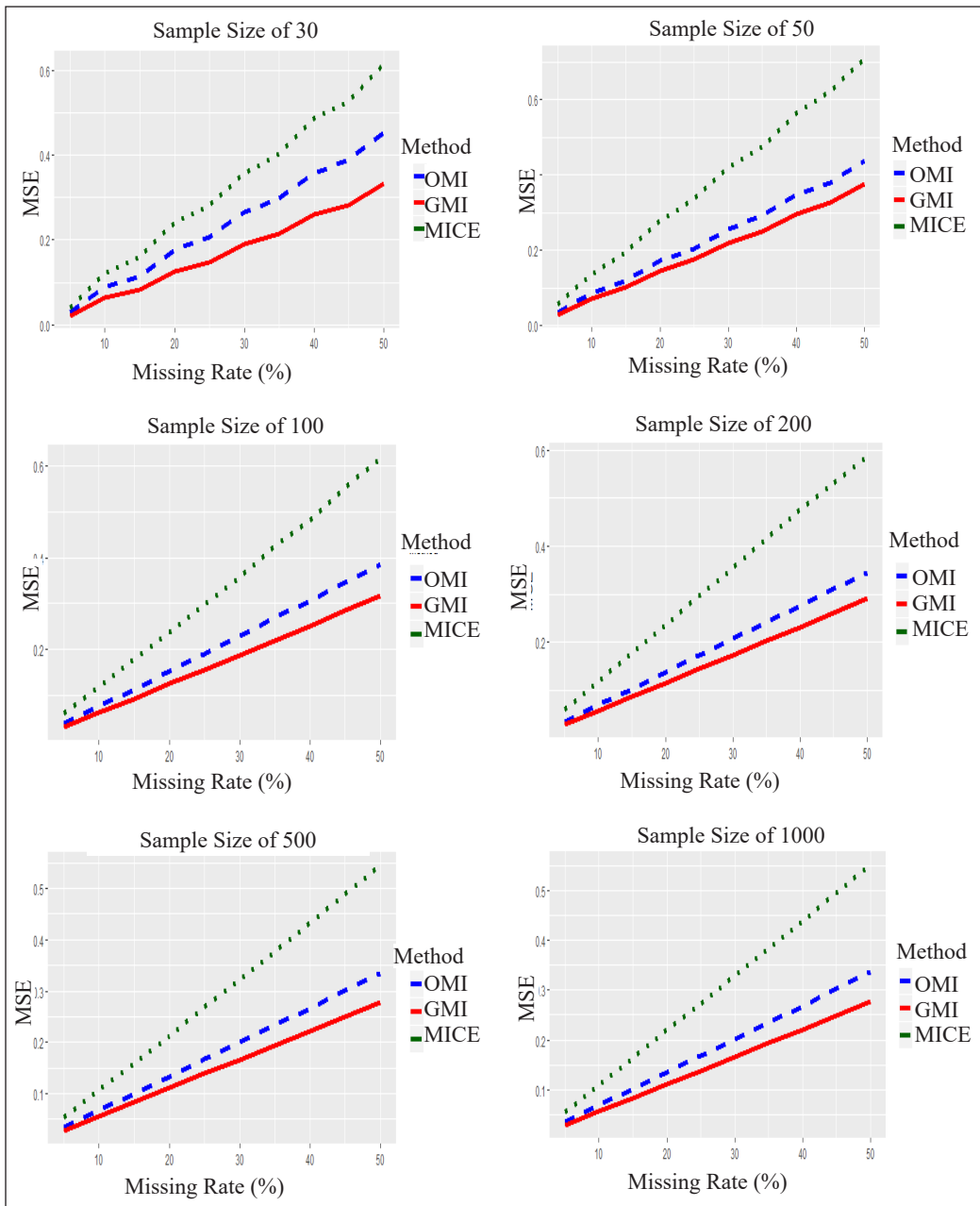


Figure 1. Line chart of MSE for different missing rates

Performance Measures of Different Imputation Methods

The pattern of data imputation methods was observed based on MSE for different missing rate and sample size using clustered boxplots. The clustered boxplots of MSE for the three imputation methods are shown in Figures 2 and 3. Figure 2 shows the pattern of MSE for

different missing rate at a fixed sample size. The clustered boxplots show that MSE of single imputation (OMI and GMI) are quite similar for samples sizes ($n = 30, 50$) and missing rate (5%). For sample sizes 100 and above, GMI has the lowest MSE for all missing rates. In Figure 3, the MSE for different sample size at a fixed missing rate was observed. The clustered boxplots show that the variability of MSE decreases as sample size increases. It can be concluded that the GMI method is more superior compared to OMI and MICE for all sample sizes regardless of missing rates.

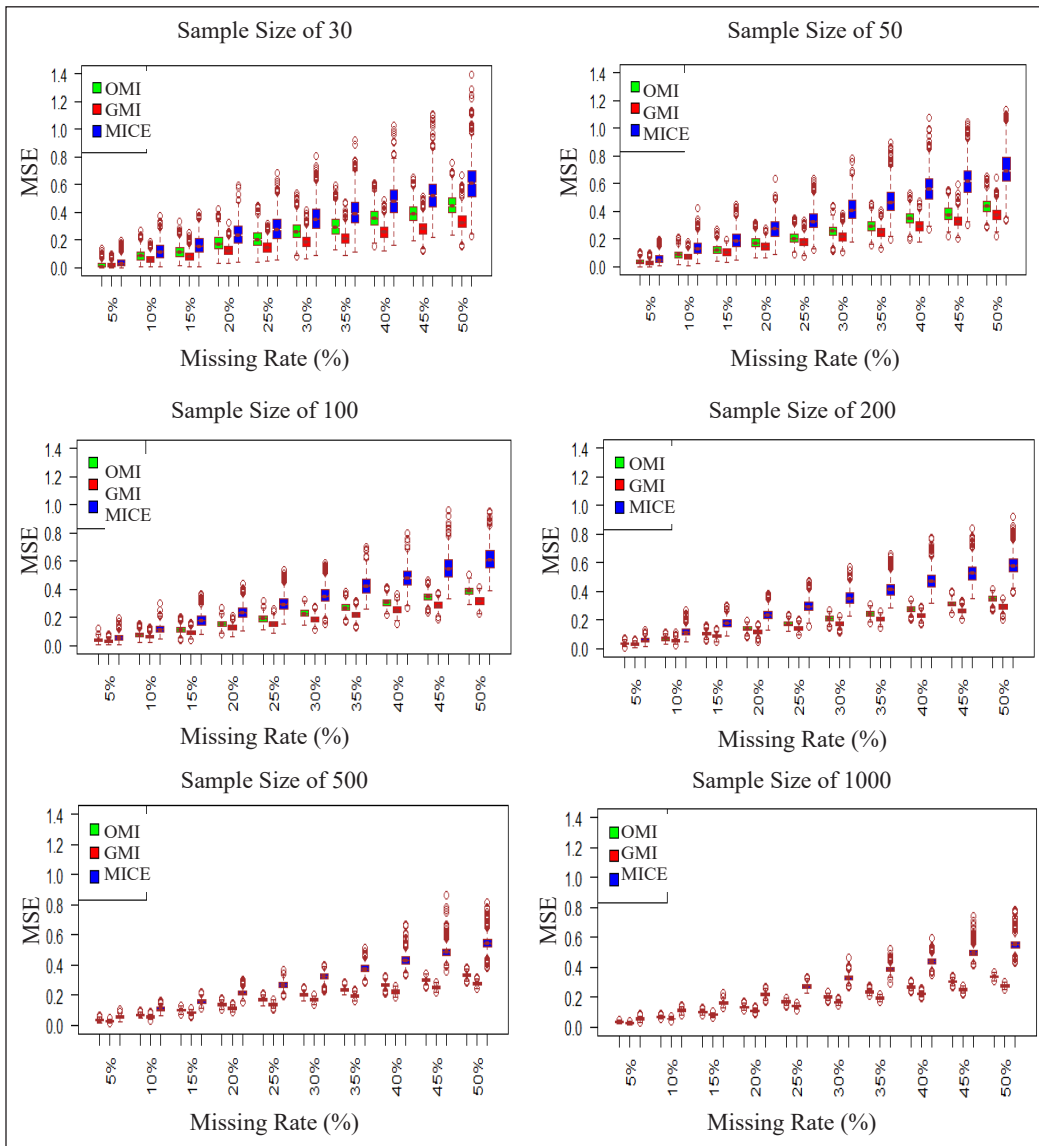


Figure 2. Clustered boxplots of MSE for different missing rates

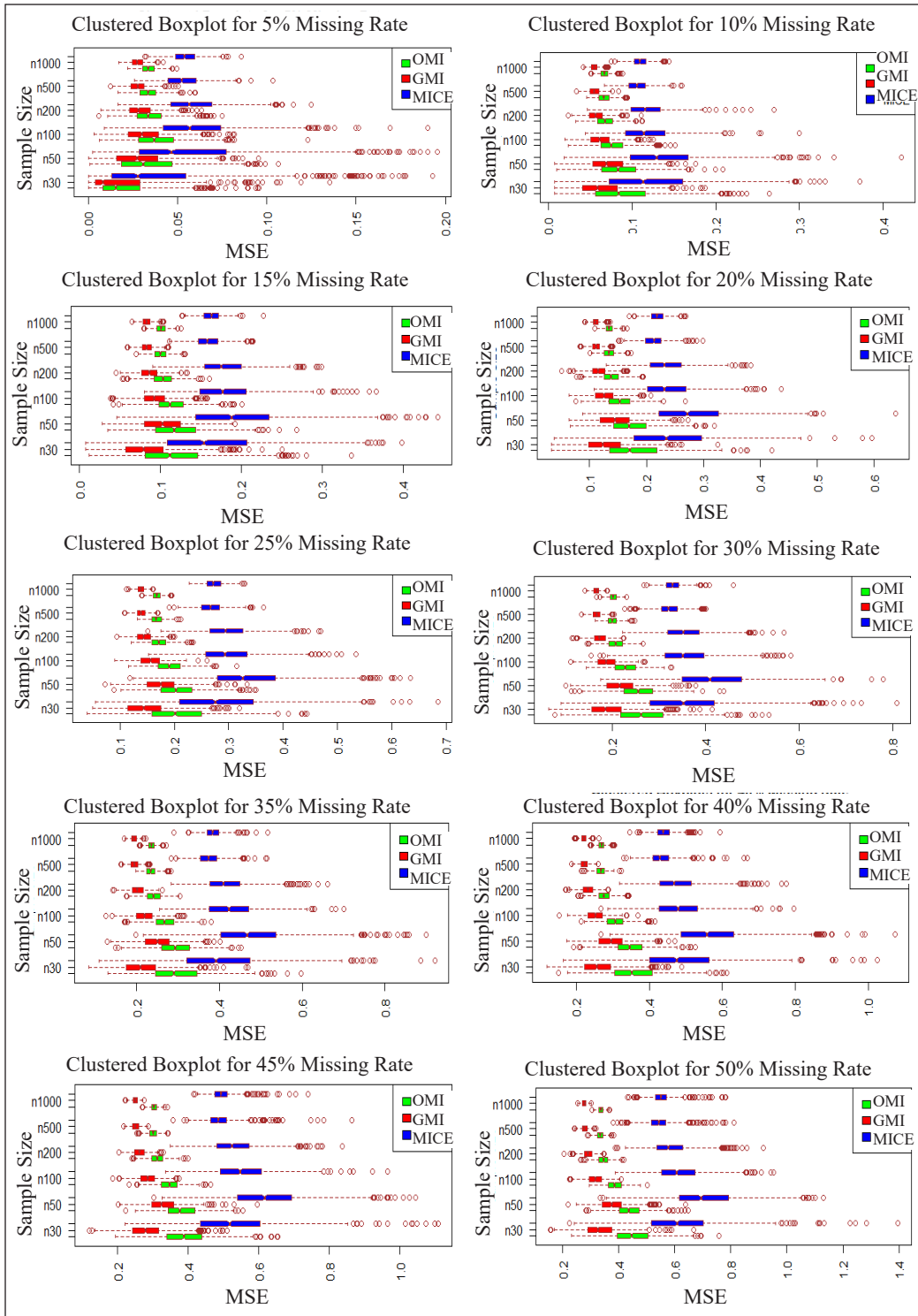


Figure 3. Clustered boxplots of MSE for different sample size

Table 4 shows the recommendation of imputation methods for each level of missing rate and sample size based on the simulation results.

Table 4
Recommended Method for Each Sample Size and Missing Rate

Sample Size	Missing Rates									
	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
30	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI
50	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI
100	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI
200	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI
500	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI
1000	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI	GMI

Study 2: Results of Application to Real Dataset

The imputation methods were then applied to the HR dataset. Table 5 displays the result of Little’s Test of MCAR for the real dataset. This test confirmed that missing data in the real dataset satisfied the assumption of the MCAR mechanism. Table 6 summarizes the results of the experiment. The MSE for GMI were observed to be lowest across different levels of missing rates and sample sizes. This indicates that the single imputation method using GMI outperformed the OMI and MICE imputation methods. The MSE for both single imputation methods and MICE increased with increasing missing rates. These results also confirmed the finding of the simulation results.

CONCLUSION

The study evaluated the performance of single imputation methods using mean with MICE (Multivariate Imputation with Chained Equations). The result of Little’s Test of MCAR for both simulation study and real dataset confirmed that missing data satisfied the assumption of MCAR. The simulation results showed that GMI is superior compared to OMI and MICE for all sample size, regardless of missing rates. However, OMI or GMI can be used when samples size is small ($n = 30, 50$) and missing rate is low (5%). The MSE for OMI, GMI and MICE increased with increasing missing rates. MICE was found not to perform well especially when missing rate was high (more than 15%). The MSE of MICE appeared to be high when missing rates were more than 15% compared to when missing rates were below 15%. An application to a real dataset confirmed the findings of the simulation results that GMI performed well for all sample sizes and missing rates. There are several other advanced imputation methods such as Regression Imputation, Regression Tree Imputation, and KNN Imputation methods. The limitation of this study is that only continuous missing attribute was considered. Future work will include a simulation study involving categorical

Table 5
Little's Test for MCAR: Real Dataset

Sample Size	p-value									
	Missing Rates (%)									
	5	10	15	20	25	30	35	40	45	50
30	0.808	0.853	0.270	0.828	0.829	0.886	0.992	0.482	0.441	0.470
50	0.744	0.259	0.953	0.067	0.760	0.777	0.291	0.673	0.492	0.314
100	0.820	0.296	0.202	0.625	0.901	0.806	0.616	0.163	0.052	0.564
200	0.939	0.664	0.369	0.369	0.564	0.095	0.782	0.846	0.881	0.833
500	0.190	0.083	0.800	0.110	0.411	0.563	0.136	0.723	0.614	0.112
1000	0.724	0.732	0.984	0.472	0.711	0.203	0.695	0.887	0.745	0.615

Table 6
Mean Square Error (MSE) of Imputation Methods (Real Dataset)

Sample Size	Missing rates									
	5%	10%	15%	20%	25%	30%	35%	40%	45%	50%
30	0.00019 ^a	0.00302 ^a	0.00443 ^a	0.00470 ^a	0.00965 ^a	0.01030 ^a	0.01084 ^a	0.01324 ^a	0.01584 ^a	0.01587 ^a
	0.00009^b	0.00230^b	0.00251^b	0.00264^b	0.00376^b	0.00440^b	0.00547^b	0.00590^b	0.00736^b	0.00882^b
	0.00045 ^c	0.00460 ^c	0.00712 ^c	0.01012 ^c	0.01261 ^c	0.01507 ^c	0.01602 ^c	0.01870 ^c	0.01985 ^c	0.01995 ^c
50	0.00056 ^a	0.00391 ^a	0.00439 ^a	0.00603 ^a	0.00654 ^a	0.01034 ^a	0.01166 ^a	0.01312 ^a	0.01315 ^a	0.01513 ^a
	0.00006^b	0.00190^b	0.00242^b	0.00421^b	0.00444^b	0.00748^b	0.00883^b	0.00998^b	0.01020^b	0.01368^b
	0.00096 ^c	0.00475 ^c	0.00486 ^c	0.01011 ^c	0.01192 ^c	0.01494 ^c	0.01790 ^c	0.01894 ^c	0.01963 ^c	0.02066 ^c
100	0.00073 ^a	0.00242 ^a	0.00407 ^a	0.00544 ^a	0.00641 ^a	0.00806 ^a	0.00864 ^a	0.01020 ^a	0.01124 ^a	0.01261 ^a
	0.00071^b	0.00224^b	0.00305^b	0.00494^b	0.00618^b	0.00774^b	0.00810^b	0.00903^b	0.01012^b	0.01109^b
	0.00120 ^c	0.00230 ^c	0.00782 ^c	0.01297 ^c	0.01356 ^c	0.01746 ^c	0.01772 ^c	0.01782 ^c	0.02346 ^c	0.02446 ^c
200	0.00151 ^a	0.00369 ^a	0.00545 ^a	0.00575 ^a	0.00706 ^a	0.00907 ^a	0.01193 ^a	0.01349 ^a	0.01473 ^a	0.01544 ^a
	0.00113^b	0.00294^b	0.00419^b	0.00444^b	0.00589^b	0.00725^b	0.00965^b	0.01058^b	0.01183^b	0.01214^b
	0.00219 ^c	0.00470 ^c	0.00706 ^c	0.01256 ^c	0.01313 ^c	0.01736 ^c	0.01738 ^c	0.02068 ^c	0.02221 ^c	0.02705 ^c
500	0.00145 ^a	0.00289 ^a	0.00473 ^a	0.00629 ^a	0.00757 ^a	0.00908 ^a	0.01036 ^a	0.01161 ^a	0.01275 ^a	0.01476 ^a
	0.00136^b	0.00252^b	0.00410^b	0.00565^b	0.00683^b	0.00816^b	0.00925^b	0.01043^b	0.01141^b	0.01305^b
	0.00205 ^c	0.00486 ^c	0.00736 ^c	0.01142 ^c	0.01309 ^c	0.02197 ^c	0.02297 ^c	0.02383 ^c	0.02438 ^c	0.02799 ^c
1000	0.00155 ^a	0.00296 ^a	0.00427 ^a	0.00598 ^a	0.00750 ^a	0.00918 ^a	0.01073 ^a	0.01262 ^a	0.01397 ^a	0.01527 ^a
	0.00137^b	0.00256^b	0.00368^b	0.00502^b	0.00651^b	0.00790^b	0.00923^b	0.00987^b	0.01214^b	0.01389^b
	0.00242 ^c	0.00483 ^c	0.00858 ^c	0.01132 ^c	0.01286 ^c	0.01676 ^c	0.02029 ^c	0.02207 ^c	0.02429 ^c	0.03114 ^c

Notes: a=OMI; b=GMI; c=MICE

or mixed types of variables. The issue of missing data for auto correlated data or time series data also will be interesting for future research.

The R-syntax for the simulation study and data imputation methods using R programming can be obtained from the first author, Nurul Azifah Mohd Pauzi. The basic data imputation R commands are given in the Appendix.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Dr Hamzah Abdul Hamid from Institute of Engineering Mathematics, Universiti Malaysia Perlis for his permission to use his simulation model as a basis to generate the data for the simulation study. We also thank the Research Management Centre, Universiti Teknologi MARA, Shah Alam, Selangor for the funding of the publication of this paper.

REFERENCES

- Abidin, N. Z., Ismail, A. R., & Emran, N. A. (2018). Performance analysis of machine learning algorithms for missing value imputation. *International Journal of Advanced Computer Science and Applications*, 9(6), 442-447.
- Aljuaid, T., & Sasi, S. (2016). Proper imputation techniques for missing values in data sets. In *International Conference on Data Science and Engineering (ICDSE)* (pp. 1-5). IEEE Conference Publication. <https://doi.org/10.1109/ICDSE.2016.7823957>
- Ayilara, O. F., Zhang, L., Sajobi, T. T., Sawatzky, R., Bohm, E., & Lix, L. M. (2019). Impact of missing data on bias and precision when estimating change in patient-reported outcomes from a clinical registry. *Health and Quality of Life Outcomes*, 17(1), 106. <https://doi.org/10.1186/s12955-019-1181-2>
- Barnett, A. G., McElwee, P., Nathan, A., Burton, N. W., & Turrell, G. (2017). Identifying patterns of item missing survey data using latent groups: An observational study. *BMJ Open*, 7(10), 1-9. <https://doi.org/10.1136/bmjopen-2017-017284>
- Bhati, S., & Gupta, M. K. (2016). Missing data imputation for medical database: Review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 6(4), 754-758.
- Buuren, S. V., & Groothuis-Oudshoorn, K. (2010). Mice: Multivariate imputation by chained equations in R. *Journal of Statistical Software*, 1-68.
- Chaudhry, A., Li, W., Basri, A., & Patenaude, F. (2019). A method for improving imputation and prediction accuracy of highly seasonal univariate data with large periods of missingness. *Wireless Communications and Mobile Computing*, 2019, 1-13. <https://doi.org/10.1155/2019/4039758>
- Cheema, J. R. (2014). Some general guidelines for choosing missing data handling methods in educational research. *Journal of Modern Applied Statistical Methods*, 13(2), 53-75. <https://doi.org/10.22237/jmasm/1414814520>

- Chhabra, G., Vashisht, V., & Ranjan, J. (2017). A comparison of multiple imputation methods for data with missing values. *Indian Journal of Science and Technology*, 10(19), 1-7. <https://doi.org/10.17485/ijst/2017/v10i19/110646>
- Dettoni, J. R., Norvell, D. C., & Chapman, J. R. (2018). The sin of missing data: Is all forgiven by way of imputation? *Global Spine Journal*, 8(8), 892-894. <https://doi.org/10.1177/2192568218811922>
- Dong, Y., & Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2(1), 1-17. <https://doi.org/10.1186/2193-1801-2-222>
- Fichman, M., & Cummings, J. N. (2003). Multiple imputation for missing data: Making the most of what you know. *Organizational Research Methods*, 6(3), 282-308. <https://doi.org/10.1177/1094428103255532>
- Gad, A. M., & Abdelkhalik, R. H. M. (2017). Imputation methods for longitudinal data: A comparative study. *International Journal of Statistical Distributions and Applications*, 3(4), 72. <https://doi.org/10.11648/j.ijstd.20170304.13>
- Gopal, K. M., Durgaprasad, N., Deepa, K. S., Sravan, R. G., & Revanth, R. D. (2019). Comparative analysis of different imputation techniques for handling missing dataset. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(7), 347-351.
- Goretzko, D., Heumann, C., & Bühner, M. (2019). Investigating parallel analysis in the context of missing data: A simulation study comparing six missing data methods. *Educational and Psychological Measurement*, 80(4), 756-774. <https://doi.org/10.1177/0013164419893413>
- Grund, S., Lüdtke, O., & Robitzsch, A. (2018). Multiple imputation of missing data at level 2: A comparison of fully conditional and joint modeling in multilevel designs. *Journal of Educational and Behavioral Statistics*, 43(3), 316-353. <https://doi.org/10.3102/1076998617738087>
- Hughes, R. A., Heron, J., Sterne, J. A., & Tilling, K. (2019). Accounting for missing data in statistical analyses: Multiple imputation is not always the answer. *International Journal of Epidemiology*, 48(4), 1294-1304. <https://doi.org/10.1093/ije/dyz032>
- Jadhav, A., Pramod, D., & Ramanathan, K. (2019). Comparison of performance of data imputation methods for numeric dataset. *Applied Artificial Intelligence*, 33(10), 913-933. <https://doi.org/10.1080/08839514.2019.1637138>
- Kaiser, J. (2014). Dealing with missing values in data. *Journal of Systems Integration*, 5(1), 42- 51. <http://dx.doi.org/10.20470/jsi.v5i1.178>
- Kamatchi P, L., & Baranidharan, C. (2019). Missing data imputation methods for autism prediction. *International Journal of Recent Technology and Engineering*, 8(5), 940-944.
- Le, T. D., Beuran, R., & Tan, Y. (2018). Comparison of the most influential missing data imputation algorithms for healthcare. In *2018 10th International Conference on Knowledge and Systems Engineering (KSE)* (pp. 247-251). IEEE Conference Publication. <http://dx.doi.org/10.1109/KSE.2018.8573344>
- Li, Y., Ji, L., Oravec, Z., Brick, T. R., Hunter, M. D., & Chow, S. M. (2019). dynr. mi: An R program for multiple imputation in dynamic modeling. *World Academy of Science, Engineering and Technology*, 13(5), 302-311. <https://doi.org/10.5281/zenodo.3298841>

- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of The American Statistical Association*, 83(404), 1198-1202.
- Little, R. J., & Rubin, D. B. (1987). *Statistical analysis with missing data*. John Wiley & Sons.
- Lo, A. W., Siah, K. W., & Wong, C. H. (2019). Machine learning with statistical imputation for predicting drug approvals. *Harvard Data Science Review*, 1(1), 1-25. <https://doi.org/10.1162/99608f92.5c5f0525>
- Ma, Z., & Chen, G. (2018). Bayesian methods for dealing with missing data problems. *Journal of The Korean Statistical Society*, 47(3), 297-313. <https://doi.org/10.1016/j.jkss.2018.03.002>
- Madley-Dowd, P., Hughes, R., Tilling, K., & Heron, J. (2019). The proportion of missing data should not be used to guide decisions on multiple imputation. *Journal of Clinical Epidemiology*, 110, 63-73. <https://doi.org/10.1016/j.jclinepi.2019.02.016>
- Malarvizhi, M. R., & Thanamani, A. S. (2012). K-Nearest Neighbor in missing data imputation. *International Journal of Engineering Research and Development*, 5(1), 5-7.
- Masconi, K. L., Matsha, T. E., Echouffo-Tcheugui, J. B., Erasmus, R. T., & Kengne, A. P. (2015). Reporting and handling of missing data in predictive research for prevalent undiagnosed type 2 diabetes mellitus: A systematic review. *The EPMA Journal*, 6(1), 1-11. <https://doi.org/10.1186/s13167-015-0028-0>
- Newman, D. A. (2003). Longitudinal modeling with randomly and systematically missing data: A simulation of ad hoc, maximum likelihood, and multiple imputation techniques. *Organizational Research Methods*, 6(3), 328-362. <https://doi.org/10.1177/1094428103254673>
- Newman, D. A. (2014). Missing data: Five practical guidelines. *Organizational Research Methods*, 17(4), 372-411. <https://doi.org/10.1177/1094428114548590>
- Nwakuya, M. T., & Nwabueze, J. C. (2018). Comparison of shrinkage-based estimators in the presence of missing data: A multiple imputation analysis. *International Journal of Statistics and Applications*, 8(6), 305-308. <https://doi.org/10.5923/j.statistics.20180806.03>
- Ochieng'Odhiambo, F. (2020). Comparative study of various methods of handling missing data. *Mathematical Modelling and Applications*, 5(2), 87.
- Pampaka, M., Hutcheson, G., & Williams, J. (2016). Handling missing data: Analysis of a challenging data set using multiple imputation. *International Journal of Research & Method in Education*, 39(1), 19-37. <https://doi.org/10.1080/1743727X.2014.979146>
- Papageorgiou, G., Grant, S. W., Takkenberg, J. J., & Mokhles, M. M. (2018). Statistical primer: How to deal with missing data in scientific research? *Interactive Cardiovascular and Thoracic Surgery*, 27(2), 153-158. <https://doi.org/10.1093/icvts/ivy102>
- Pedersen, A. B., Mikkelsen, E. M., Cronin-Fenton, D., Kristensen, N. R., Pham, T. M., Pedersen, L., & Petersen, I. (2017). Missing data and multiple imputation in clinical epidemiological research. *Clinical Epidemiology*, 9, 157-166. <https://doi.org/10.2147/CLEPS129785>
- Ratolojanahary, R., Ngouna, R. H., Medjaher, K., Junca-Bouricié, J., Dauriac, F., & Sebilo, M. (2019). Model selection to improve multiple imputation for handling high rate missingness in a water quality dataset. *Expert Systems with Applications*, 131, 299-307. <https://doi.org/10.1016/j.eswa.2019.04.049>

- Salgado C. M., Azevedo C., Proença H., & Vieira S. M. (2016) Missing data. In *Secondary analysis of electronic health records* (pp. 143-162). Springer.
- Scheffer, J. (2002). Dealing with missing data. *Research Letters in the Information and Mathematical Sciences*, 3, 153-160.
- Schmitt, P., Mandel, J., & Guedj, M. (2015). A comparison of six methods for missing data imputation. *Journal of Biometrics & Biostatistics*, 6(1), 1-6. <https://doi.org/10.472/2155-6180.1000224>
- Shi, D., Lee, T., Fairchild, A. J., & Maydeu-Olivares, A. (2019). Fitting ordinal factor analysis models with missing data: A comparison between pairwise deletion and multiple imputation. *Educational and Psychological Measurement*, 80(1), 41-66. <https://doi.org/10.1177/0013164419845039>
- Sim, J., Lee, J. S., & Kwon, O. (2015). Missing values and optimal selection of an imputation method and classification algorithm to improve the accuracy of ubiquitous computing applications. *Mathematical Problems in Engineering*, 2015, 1-14. <https://doi.org/10.1155/2015/538613>
- Song, Q., & Shepperd, M. (2007). Missing data imputation techniques. *International Journal of Business Intelligence and Data Mining*, 2(3), 261-291. <https://doi.org/10.1504/IJBIDM.2007.015485>
- Stavseth, M. R., Clausen, T., & Røislien, J. (2019). How handling missing data may impact conclusions: A comparison of six different imputation methods for categorical questionnaire data. *SAGE Open Medicine*, 7, 1-12. <https://doi.org/10.1177/2050312118822912>
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest - Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112-118. <https://doi.org/10.1093/bioinformatics/btr597>
- Sullivan, T. R., White, I. R., Salter, A. B., Ryan, P., & Lee, K. J. (2018). Should multiple imputation be the method of choice for handling missing data in randomized trials? *Statistical Methods in Medical Research*, 27(9), 2610-2626. <https://doi.org/10.1177/0962280216683570>
- Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Pearson.
- Turner, E. L., Yao, L., Li, F., & Prague, M. (2019). Properties and pitfalls of weighting as an alternative to multilevel multiple imputation in cluster randomized trials with missing binary outcomes under covariate-dependent missingness. *Statistical Methods in Medical Research*, 29(5), 1338-1353. <https://doi.org/10.1177/0962280219859915>
- Van Buuren, S. (2007). Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical Methods in Medical Research*, 16(3), 219-242. <https://doi.org/10.1177/0962280206074463>
- Van Buuren, S., Boshuizen, H. C., & Knook, D. L. (1999). Multiple imputation of missing blood pressure covariates in survival analysis. *Statistics in medicine*, 18(6), 681-694. [https://doi.org/10.1002/\(SICI\)1097-0258\(19990330\)18:6<681::AID-SIM71>3.0.CO;2-R](https://doi.org/10.1002/(SICI)1097-0258(19990330)18:6<681::AID-SIM71>3.0.CO;2-R)
- van Ginkel, J. R., Linting, M., Rippe, R. C., & van der Voort, A. (2019). Rebutting existing misconceptions about multiple imputation as a method for handling missing data. *Journal of Personality Assessment*, 102(3), 297-308. <https://doi.org/10.1080/00223891.2018.1530680>

- Wah, Y. B., Ibrahim, N., Hamid, H. A., Abdul-Rahman, S., & Fong, S. (2018). Feature selection methods: Case of filter and wrapper approaches for maximising classification accuracy. *Pertanika Journal of Science & Technology*, 26, 329-340.
- Wilks, S. S. (1932). Certain generalizations in the analysis of variance. *Biometrika*, 24(3/4), 471-494. <https://doi.org/10.2307/2331979>
- Yadav, M. L., & Roychoudhury, B. (2018). Handling missing values: A study of popular imputation packages in R. *Knowledge-Based Systems*, 160, 104-118. <https://doi.org/10.1016/j.knosys.2018.06.012>
- Zhang, Z. (2016). Missing data imputation: focusing on single imputation. *Annals of Translational Medicine*, 4(1), 1-9. <https://doi.org/10.3978/j.issn.2305-5839.2015.12.38>

Robust Multivariate Correlation Techniques: A Confirmation Analysis using Covid-19 Data Set

Friday Zinzendoff Okwonu^{1,2}, Nor Aishah Ahad^{1*}, Joshua Sarduana Apanapudor² and Festus Irismisose Arunaye²

¹*School of Quantitative Sciences, College of Arts and Sciences, Universiti Utara Malaysia, 06010 Sintok, Kedah, Malaysia*

²*Department of Mathematics, Faculty of Science, Delta State University, P.M.B.1, Abraka, Nigeria*

ABSTRACT

Robust multivariate correlation techniques are proposed to determine the strength of the association between two or more variables of interest since the existing multivariate correlation techniques are susceptible to outliers when the data set contains random outliers. The performances of the proposed techniques were compared with the conventional multivariate correlation techniques. All techniques under study are applied on COVID-19 data sets for Malaysia and Nigeria to determine the level of association between study variables which are confirmed, discharged, and death cases. These techniques' performances are evaluated based on the multivariate correlation (R), multivariate coefficient of determination (R^2), and Adjusted R^2 . The proposed techniques showed $R=0.99$ and the conventional methods showed that R ranges from 0.44 to 0.73. The R^2 and the Adjusted R^2 for proposed methods are 0.98 and 0.97 while the conventional methods showed that R equals 0.53, 0.44, and 0.19 whereas Adjusted R^2 equals 0.52, 0.43, and 0.18, respectively. The proposed techniques strongly affirmed that for any patient to be discharged or die of the Covid-19, the patient must be confirmed Covid-19 positive, whereas the conventional method showed moderate to very weak affirmation. Based on the results, the proposed techniques are robust and show a very strong association between the variables of interest than the conventional techniques.

ARTICLE INFO

Article history:

Received: 13 October 2020

Accepted: 04 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.16>

E-mail addresses:

o.friday.zinzendoff@uum.edu.my (Friday Zinzendoff Okwonu)

aishah@uum.edu.my (Nor Aishah Ahad)

japanpudor@yahoo.com (Joshua Sarduana Apanapudor)

ifaranaye@yahoo.com (Festus Irismisose Arunaye)

* Corresponding author

Keywords: Coefficient of determination, Covid-19, multivariate correlation techniques, robust

INTRODUCTION

Karl Pearson established the product sum formula, widely known as the Pearson product-moment correlation coefficient denoted by “ r ” (Pearson, 1920). The multivariate correlation technique is the generalisation of the conventional Pearson product-moment correlation coefficient. The latter is pairwise involving one independent variable and one dependent variable while the former requires two or more independent variables and one dependent variable. It can be described as the squared correlation between the expected values and the observed values (Abdi, 2007). Historically, Bravais had discussed correlation involving two and three variables (Pearson, 1920). These three variables consideration is the generalisation of the two variable concepts, and by extension is regarded as the multivariate or multiple correlation techniques.

The bivariate correlation analysis has been extensively discussed with numerous applications to different fields of study (Armstrong, 2019; Bareinboim et al., 2014; Brown et al., 2012; Geiss et al., 1991; Mukaka, 2012; Nguyen et al., 2013; Wang et al., 2017). Different techniques describing multivariate correlation analysis and their performance analysis have been discussed (Wang et al., 2017). The emergence of multivariate correlation analysis is due to the ever-increasing data dimension with associated variables and the need to determine associations between the variables. Unlike the conventional correlation technique due to Karl Pearson, multivariate correlation techniques have no basic formulation fundamentals (Wang & Zheng, 2020). Techniques such as the partial correlation and coefficient of determination have been recruited to justify analytically to analyse the association of multivariate independent and dependent variables. Techniques such as unsigned correlation have been proposed to investigate multivariate variables association (Wang & Zheng, 2020).

The multivariate correlation coefficient determines the strength of the association between two or more variables of interest, unlike the univariate correlation coefficient (r) whose values range between ± 1 , the multivariate correlation coefficient values (r_{xyz}) range between 0 and 1. A strong association indicates that r_{xyz} tends to one. On the other hand, if the multivariate correlation coefficient value tends to zero, it implies little or no association between the variables. The uniqueness of r_{xyz} is that it is always positive. The amplifying nature of the multivariate correlation technique is the ability to handle all the independent and dependent variables provide simultaneously (Geiß & Einax, 1996). In most cases, multivariate correlation techniques often reveal a stronger association between variables of interest than the univariate correlation technique (Zhang et al., 2008).

Correlation is applied to determine the level of association between variables of interest (Asuero et al., 2006). It has been investigated that the classical correlation coefficient is not robust when the data set contains influential observations (Abdullah, 1990; Okwonu et al., 2020). Studies have shown that a nonparametric approach based on the Spearman

correlation coefficient is insensitive against the influential observation present in the paired variables (Abdullah, 1990). We observed that when the bivariate correlation method was transformed into a multivariate correlation technique, the issue of underperformance due to the presence of outliers in the data set persisted as such, the existing multivariate correlation techniques were susceptible to outliers when the data set contained random outliers. To solve this problem of outliers, we propose robust and high breakdown multivariate correlation techniques that utilizes all the variables simultaneously and does not require the distributional assumption of the data set to compute the multivariate correlation coefficient value. Therefore, the main contribution to this study is that the proposed methods are insensitive and not susceptible to outliers.

The proposed methods and the conventional methods were applied to the Covid-19 data set. The level of association determined based on the data sets were used to evaluate the health risk of the Covid-19 pandemic. The analysis relying on the level of association could assist the researchers to advise people to abide by safety protocols established by healthcare providers and relevant organs of government.

The rest of this paper is organised as follows: In Section 2 we discussed the conventional methods: paired multivariate correlation; the median-based multivariate correlation technique; and the multivariate correlation by regression technique; followed by the proposed techniques: the multivariate cross-correlation technique, and the multivariate cross-correlation based on deviation method. The second section continued with the multivariate coefficient of determination, adjusted multivariate coefficient of determination, and ended with data collection. Results and discussions are presented in Section 3, and the conclusion follows in Section 4.

MATERIALS AND METHODS

Rodgers and Nicewander (1988) observed that the correlation coefficient (r) could be computed using different mathematical formulations. Different names have been ascribed to the mathematical complement, but the general objective is to determine the degree of associations between interest variables (Rodgers & Nicewander, 1988). This paper complies with the mathematical formulation concept elaborated and enunciated by Rodgers and Nicewander (1988) in developing a new computational framework for the index. Before discussing the new computational framework, we looked at the different multivariate correlation index advanced by different researchers of different mindsets.

Paired Multivariate Correlation

Let $X, Y \in \hat{R}$ be independent random variables and $Z \in \hat{R}$ be the dependent variable. The idea is to investigate whether the independent and dependent variables correlate or not. From observation, the strength of correlation in Pearson's concept is stronger in dual sign

directions than the paired multivariate correlation. The conventional procedures involve pairing the random variables before the correlation values are plugged into the multivariate correlation formula. The following mathematical Equation 1 states the multivariate correlation coefficient method (Geiss et al., 1991; Geiß & Einax, 1996; Wang et al., 2018).

$$R = \frac{\varphi}{\theta}, \tag{1}$$

where $(R = r_{xyz})$, $\varphi = \sqrt{R_{xz}^2 + R_{yz}^2 - 2R_{xz}R_{yz}R_{xy}}$ (Garnett, 1919) and $\theta = \sqrt{1 - R_{xy}^2}$.

Equation 1 is defined by pairing the Pearson's correlation coefficient, where R_{xz} , R_{yz} and R_{xy} are well defined Pearson correlation coefficient. Equation 1 can be simplified as Equation 2-4:

$$D_{xy} = \frac{D_z}{D_{z1}} = \frac{D_{xi} \times D_{yi}}{\sqrt{D_{xi}^2 \times D_{yi}^2}}, \tag{2}$$

$$D_{xz} = \frac{D_k}{D_{z2}} = \frac{D_{xi} \times D_{zi}}{\sqrt{D_{xi}^2 \times D_{zi}^2}}, \tag{3}$$

$$D_{yz} = \frac{D_{kk}}{D_{zk}} = \frac{D_{yi} \times D_{zi}}{\sqrt{D_{yi}^2 \times D_{zi}^2}}, \tag{4}$$

where $D_{xi} = \sum_{i=1}^n (x_i - \bar{x})$, $D_{yi} = \sum_{i=1}^n (y_i - \bar{y})$ and $D_{zi} = \sum_{i=1}^n (z_i - \bar{z})$. Equation 1 can be written based on Equations 2-4 to form the paired multivariate correlation as shown in Equation 5

$$r_{xyz} = \frac{d_{c(xyz)}}{t}, \tag{5}$$

where $d_{c(xyz)} = \sqrt{D_{xz}^2 + D_{yz}^2 - 2D_{xz}D_{yz}D_{xy}}$ and $t = \sqrt{1 - D_{xy}^2}$.

The Median-Based Multivariate Correlation Technique

This procedure involves substituting the mean by the median, that is $\bar{D}_{xi} = \sum_{i=1}^n (x_i - m(x))$, $\bar{D}_{yi} = \sum_{i=1}^n (y_i - m(y))$ and $\bar{D}_{zi} = \sum_{i=1}^n (z_i - m(z))$, where $m(x)$, $m(y)$ and $m(z)$ are the median of the independent and dependent random variables. Substituting these into Equation 2-5, we have the median-based multivariate correlation as presented in Equation 6.

$$r_{dxyz} = \frac{d_{m(xyz)}}{m(t)}, \quad 0 \leq r_{dxyz} \leq 1 \quad (6)$$

where $d_{m(xyz)} = \sqrt{m(\widetilde{D}_{xz}^2) + m(\widetilde{D}_{yz}^2) - 2(m(\widetilde{D}_{xz})m(\widetilde{D}_{yz})m(\widetilde{D}_{xy}))}$ and $m(t) = \sqrt{1 - m(\widetilde{D}_{xy}^2)}$.

Multivariate Correlation by Regression Technique (MCRT)

The concept of regression to compute the correlation coefficient was initiated by Mr. Yule around 1897 (Pearson, 1920; Weida, 1927). Standing on this note, this concept is considered suitable for this paper. The multivariate correlation technique based on regression was coined using the principle of multiple regression procedures (Huberty, 2003; Lewis-Beck et al., 2004). Recall the multiple regression formulation, that is Equation 7,

$$\hat{y} = b_0 + \sum_{k=1}^p b_k x_i, i = 1, 2, \dots, n \quad (7)$$

where x_i are the independent variables (predictors), y denotes a single dependent variable, b_k , where $k = 1, 2, \dots, p$ is the weights and \hat{y} is the estimate of y called the predictor score. In this regard, the computation of b_k strictly relies on the information provided by the independent and dependent variables as such how large or small the correlation between y and \hat{y} depends on the information provided by b_k . This method implies that the correlation between the dependent variable and the linear combination of the independent variables is called the multivariate correlation (Lewis-Beck et al., 2004). Based on Equation 7, the multivariate correlation denoted by $r_{-y\hat{y}}$ (Asuero et al., 2006) can be started by applying the dependent variable and the predictor score, that is y and \hat{y} as given in Equation 8.

$$r_{-y\hat{y}} = \frac{\sum_{i=1}^n (y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2} \sqrt{\sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \quad (8)$$

where $\bar{\hat{y}} = \frac{\sum_{i=1}^n \hat{y}_i}{n}$. An alternative to Equation 8 is to apply the following procedure:

$$\hat{y} = Xb,$$

where $b = (X'X)^{-1}X'y$, $X_{n \times (p+1)}$ denote augmented matrix and $y_{n \times 1}$ vector (Abdi, 2007). Let γ denote the regression sum of square defined by Equation 9.

$$\gamma = b'X'y - \frac{(1'y)^2}{n} \quad (9)$$

where $\mathbf{1}$ is a row vector of 1 that align with the dependent variable y and Y as defined in Equation 9 and the total sum of square θ is defined as Equation 10.

$$\theta = y \cdot y - \frac{(\mathbf{1}y)^2}{n} \tag{10}$$

$$\omega = y \cdot y - b \cdot X \cdot y \tag{11}$$

where ω is the sum of square of the residual (Equation 11). Applying Equation 9-11, we obtain the multivariate correlation as Equation 12.

$$\tau^2 = \frac{\gamma}{\varphi}, \varphi = \gamma + \omega \tag{12}$$

Equation 8 and 12 are based on the regression approach. The proposed methods are described as follows.

Multivariate Cross-Correlation Technique

Let $X, Y \in \hat{R}$ be independent random variables and $Z \in \hat{R}$ be a dependent random variable. These variables or vectors are paired such that each interacts to obtain the level of trend (Wang & Zheng, 2020). Recall the concept of vector, matrix products, and transpose, in this case, when a $(1 \times n)$ row vector is multiplied by an $(n \times 1)$ column vector, the product is a scalar or a real number. The independent and dependent variables may assume $(n \times 1)$ or $(1 \times n)$ row or column depending precisely on the variable transposed from row to column or column to row.

Let $S_{XX} = (XX^T)$ be the autocorrelation and $S_{XY} = (XY^T)$ be the cross-correlation as defined in Geiss et al. (1991) and Geiß and Einax (1996) and invoking the Cauchy-Binet concept we obtain the following Equation 13-15:

$$\Delta_1 = 1 - \left(\frac{X^T Y}{X^T Z} \right), X^T = [x_i]^T, i = 1, 2, \dots, n, X^T Z \neq 0 \tag{13}$$

$$\Delta_2 = 1 - \left(\frac{X^T Z}{Y^T Z} \right), Y^T = [y_i]^T, i = 1, 2, \dots, n, Y^T Z \neq 0 \tag{14}$$

$$\Delta_3 = 1 - \left(\frac{Y^T Z}{X^T Y} \right), Z^T = [z_i]^T, i = 1, 2, \dots, n, X^T Y \neq 0 \tag{15}$$

Where X^T, Y^T, Z^T are the transpose of the data set. Equation 13-15 can be written compactly as Equation 16-19.

$$\alpha_1 = \left[\frac{X^T Y}{X^T Z} \right], \rightarrow \Delta_1 = 1 - \alpha_1 \tag{16}$$

$$\alpha_2 = \left[\frac{X^T Z}{Y^T Z} \right], \rightarrow \Delta_2 = 1 - \alpha_2 \tag{17}$$

$$\alpha_3 = \left[\frac{Y^T Z}{X^T Y} \right], \rightarrow \Delta_3 = 1 - \alpha_3 \tag{18}$$

$$\tau = \sum_{j=1}^k \Delta_j^2, \vartheta = k \prod_{j=1}^{k+1} \Delta_j \text{ and } \phi = \sum_{j=1}^{k+1} \Delta_j^2 \tag{19}$$

The definition of Δ_j mimic Chatillon's correlation $r = \sqrt{1 - \left(\frac{c}{C}\right)^2}$ (Chatillon, 1984). Applying Equation 13-15, this technique produced Multivariate Cross-Correlation and can be stated as in Equation 20 ($sig_del = \delta_\Delta$): :

$$sig_del = \delta_\Delta = \sqrt{\frac{\sum_{j=1}^k \Delta_j^2 - k \left\{ \prod_{j=1}^{k+1} \Delta_j \right\}}{\sum_{j=1}^{k+1} \Delta_j^2}}, 0 \leq \delta_\Delta \leq 1. \tag{20}$$

The values of Δ_j play a vital role in determining whether there is an association between the variables. Equation 20 detects an association if $\Delta_j < 1$ and $\alpha_j \neq 0$. On the other hand, Equation 20 detects no association if $\Delta_j = 1$, $\alpha_j = 0$, and $\Delta_j = 0$ if $\alpha_j = 1$. If the definitions of α_j are achievable for the first instance, that is $\alpha_j \neq 0$, the association of the variables can be determined uniquely. On the other hand, if $\alpha_j = 0, 1$, this implies no association between the variables. It is pertinent to state that the value of α_j is always positive. A similar computational framework has been expressed in Tan et al. (2011). Therefore, the following is true (Equation 21).

$$\delta_\Delta = \begin{cases} \leq 1, & \text{if } \alpha_j \neq 0 \\ 0, & \text{if } \alpha_j = 0, 1 \end{cases} \tag{21}$$

Multivariate Cross-Correlation Based on Deviation Method

Let X, Y and Z be as defined, then the deviation method is stated as Equation 22-24:

$$Z = (X - \bar{X}), \bar{X} = \frac{\sum_{i=1}^l x_i}{l} \tag{22}$$

$$B = (Y - \bar{Y}), \bar{Y} = \frac{\sum_{i=1}^l y_i}{l} \tag{23}$$

$$W = (Z - \bar{Z}), \bar{Z} = \frac{\sum_{i=1}^l z_i}{l} \tag{24}$$

Based on the deviations and invoking the concept defined in Geiss et al. (1991) and Urain and Peters (2019), the following argument applies and mimics Pearson’s procedure. For each $C_j, j = 1,2,3$, we have Equation 25-27.

$$C_1 = \frac{Z^T B}{ZW^T}, ZW^T \neq 0 \tag{25}$$

$$C_2 = \frac{Z^T W}{BW^T}, BW^T \neq 0 \tag{26}$$

$$C_3 = \frac{B^T W}{Z^T B}, Z^T B \neq 0 \tag{27}$$

The values of C_j, j strictly depends on the deviations from the variable of interest. Hence we have Equation 28.

$$\gamma = \sum_{j=1}^k c_j^2 - k(\prod_{j=1}^{k+1} c_j) \text{ and } \delta = \sum_{j=1}^{k+1} c_j^2. \tag{28}$$

Applying the definitions in Equation 28 we obtain Multivariate Cross-Correlation Based on Deviation Method ($\nabla = \nabla$) as stated in Equation 29

$$\nabla = \sqrt{\frac{\sum_{j=1}^k c_j^2 - l(\prod_{j=1}^{k+1} c_j)}{\sum_{j=1}^{k+1} c_j^2}}, 0 \leq \nabla \leq 1 \tag{29}$$

The value of ∇ is always positive. Therefore Equation 29 is true if the following conditions hold as in Equation 30

$$\nabla = \begin{cases} \leq 1, & \text{if } C_j \neq 0 \\ 0, & \text{if } C_j = 1. \end{cases} \tag{30}$$

Therefore, the limits [0,1] for the proposed procedures are established based on the conditions explicitly stated in Equation 21 and 30. The proposed methods are insensitive against influential observations.

Multivariate Coefficient of Determination

In some situations, the correlation value is often questioned as lacking interpretations for its measurement unit as such; the correlation value is squared. This squared correlation

value is called the coefficient of determination. This can be described as the proportion of the variance in one variable explained by the differences in the other variables (Nagelkerke, 1991; Nakagawa et al., 2017; Pyrczak et al., 2018). The strength of the different multivariate correlation techniques discussed in this paper depends on the multivariate coefficient of determination R^2 (Equation 31).

$$R^2 = (r_{dxyz})^2, 0 \leq R^2 \leq 1 \quad (31)$$

Equation 31 denotes the fraction of the sum of interaction explained by the methods discussed in this paper. In this case, R^2 close to one show superior interaction or association between the independent and dependent variables. On the other hand, if R^2 is close to zero it implies that the level of interaction between the variables is not suitably explained. It is vital to note that the univariate coefficient of determination and multivariate coefficient of determination play a similar explanatory role. Explaining Equation 31 further implies that $|R^2| < r_{dxyz}$ may signify a closer association between the independent and dependent variables than R^2 . However, R^2 allows for the comparison of the level of association (Asuero et al., 2006).

In this study, it is proper for us to investigate and compare the Pearson coefficient of determination from the regression approach. The formula in Equation 32 is applied to determine the coefficient of determination for Equation 8.

$$r_{y\hat{y}}^2 = 1 - \frac{\sum_{k=1}^l (y_k - \hat{y}_k)^2}{\sum_{k=1}^l (y_k - \bar{y})^2}, 0 \leq r_{y\hat{y}}^2 \leq 1 \quad (32)$$

Where \hat{y}_k denotes the predicted values of y_k based on the information provided by the regression model, and \bar{y} is the mean of the response variable. Equation 32 was compared with the usual technique of squaring r values. This argument was considered vital because Equation 8 yields the same results as Pearson's technique. Unfortunately, the coefficient of determination computed by applying Equation 32 is equal to the coefficient of determination obtained by squaring the Pearson's r .

Adjusted Multivariate Coefficient of Determination

The adjusted multivariate coefficient of determination is stated as Equation 33.

$$\text{Adjusted } R^2 = 1 - \frac{(l-1)}{l-(k+1)} (1 - R^2) \quad (33)$$

where l denotes the sample size, and k is the number of independent variables, and R^2 denotes the multivariate coefficient of determination. The basic objective of using Equation

33 is to minimise positive estimation bias (Huberty, 2003). Minimisation of bias is assumed large depending on the sample size divided by the independent variables' dimension, that is (n/p) . In general, Equation 33 gives a true estimate and more reliable than the multivariate correlation coefficient of whatever derivation.

Data Collection

The Covid-19 virus is a global problem, and some countries have experience in managing pandemics while others do not. In this study, we used the data set from two countries: Malaysia and Nigeria. These countries are both developing but they have similar weather or climate conditions, different continents, and geographical location but differ in population size. Unfortunately, Covid-19 does not recognise the lifestyle. We use Covid-19 data from these two countries because of easy access to data, similar lockdown measures, and the use of non-pharmaceutical interventions to combat the outbreak of the virus. The uses of the data set from these countries provide more information on the risk associated with the virus and non-compliance with safety protocols.

Case 1. Malaysian Covid-19 data set. The history of Covid-19 in Malaysia started on 25th January 2020 when the Covid-19 case was first reported in Malaysia. The data for this study was collected from 25th January to 31st July 2020 based on recorded data in the portal of the Ministry of Health Malaysia (KKM, 2020). This represents 189 days of the Covid-19 outbreak in Malaysia. As of 31st July 2020, the number of confirmed cases is 8,976, discharged cases are 8,645, representing 96.31% recovering, and 125 death cases representing 1.39% with 2.3% active cases.

Case 2. Nigerian Covid-19 data set: The data set are the daily recording of the Covid-19 outbreak in Nigeria which was first reported on the 27th of February 2020 to 4th August 2020 by the Nigeria centre for disease control (NCDC, 2020). The data consist of 163 days of recording with total confirmed cases (44,433), discharged cases (31,851) representing 71.68%, and death cases (910) representing 2.05%. This implies that as of 4th August 2020, Nigeria has approximately 11,672 (26.27%) active cases. Table 1 consists of variables of interest in this study.

Table 1
Study variables

Independent Variable	Confirmed cases
Dependent variables	Discharged and death cases

RESULTS AND DISCUSSIONS

In this section, we present the values of the multivariate correlation (R), coefficient of determination (R^2), and the adjusted coefficient of determination (Adjusted R^2). The Pearson correlation technique performance is evaluated based on the ± 1 , while the multivariate correlation technique' is evaluated between 0 and +1. The implication of this is that if $R = 0$, it implies no association between the variables, and if $R = 0.1, 0.2, 0.3, 0.4$ it implies weak association, and if $R = 0.5, 0.6, 0.7, 0.8, 0.9, 1.0$ it implies moderate to strong association of the study variables. These performance evaluation procedures were adopted due to consistent and reliability to analyse situational data. However, as the value of R approaches one, the difference between R , R^2 and Adjusted R^2 becomes immaterial. In Figure 1, we used the Covid-19 data set from Malaysia for the period under review. The performance of the conventional multivariate correlation and the proposed techniques are compared.

In Figure 1, the paired multivariate correlation (r_{xyz}), Equation 5 and the median-based paired multivariate correlation (r_{dxyz}), Equation 6 are defined, respectively. Equation 6 is a modification of Equation 5. At the same time, $r_{y\hat{y}}$ and τ^2 are the multivariate regression techniques, that is Equation 8 and Equation 12. The proposed methods are the multivariate cross-correlation technique (sig_del, Equation 20) and the multivariate cross-correlation based on the deviation method (delta, Equation 29). In Figure 1, the proposed methods (Equation 20 & 29) showed that there was a strong association between confirmed, discharged and death cases, the implication is that if a case is confirmed, two reasonable possibilities exist, that is the case may either be discharged or classified as dead. The modified approach (Equation 6) indicates a strong association between the study variables, and Equation 5 and Equation 8 show similar performance, and Equation 12 shows a weak association between the study variables. We observed that the regression techniques (Equations 8 & 12) and the conventional technique (Equation 5) are affected by influential observations (random outliers). This is justified because outliers often affect the mean, and the median is resistant against outliers. The implication is that the mean base techniques are affected when a high number of confirmed or discharge or death cases are reported. In this data set, influential observations or random outliers implies the day with the highest number of confirmed or discharge or death cases against the usual pattern of infection or discharge or death cases. From Figure 1, the comparative analysis indicates that the proposed techniques are robust against random outliers and hence showed a stronger variable association than the conventional techniques. Figure 2 consists of the Covid-19 data set from Nigeria. The performance analyses of the different methods based on the Nigeria Covid-19 data set showed similar comparative performance discussed for the Malaysian Covid-19 data in Figure 1.

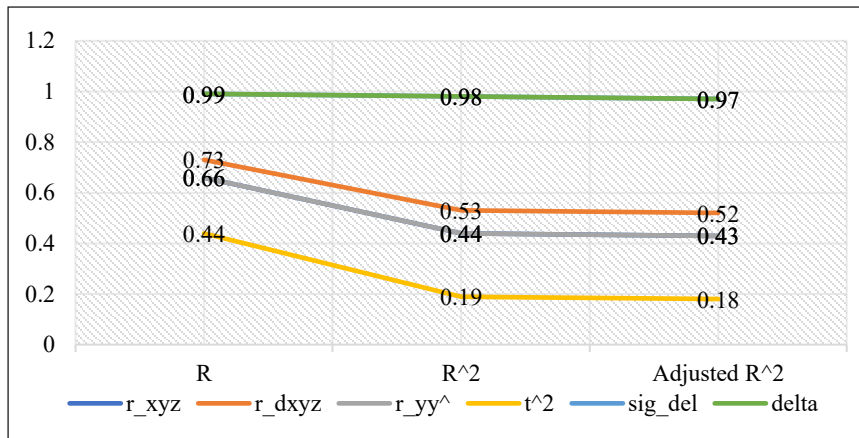


Figure 1. Comparative analysis of different Multivariate correlation techniques (Malaysia)

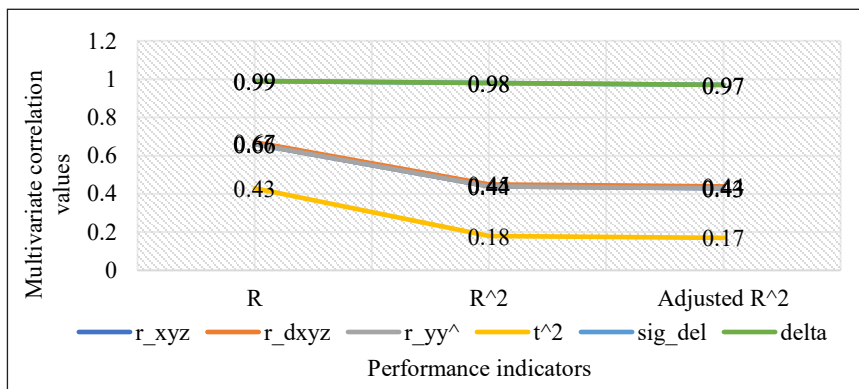


Figure 2. Comparative analysis of different Multivariate correlation techniques (Nigeria)

In Figure 1 and 2 above, τ^2 has the lowest coefficient of determination (R^2) and Adjusted R^2 values. The proposed methods based on R^2 and Adjusted R^2 showed strong association indicating that when somebody was infected with the Covid-19 virus, there was a strong association (relationship) that the person might be discharged or die of the virus. However, the other methods apart from the multivariate regression method τ^2 showed a moderate degree that an infected person might either be discharged or die. This indicates that the Covid-19 data set from the two different continents and countries have similar characteristics. The implication of $R^2 = 0.98$ for the proposed techniques showed that these two methods explained the data set very well.

The conventional methods imply that when someone is confirmed positive, there is a slightly moderate to strong association that the person may be discharged or die of the virus. However, the degree of association to be discharged or dead differ with the different

techniques. The proposed method showed a strong association for the confirmed person to be discharged or confirmed dead. Although, the other methods differ in association because R , R^2 and Adjusted R^2 differ strongly. We observed from Figure 1 and 2 that R^2 values are greater than the values of Adjusted R^2 . From the study, we observed that sig_del and $delta$ performed similarly for both data set. For the Nigeria data set, r_{xyz} and r_{dxyz} performed comparably.

The comparative analysis showed that there were possibilities that any confirmed case might either be discharged or die of the virus. The study had shown that there was a strong to a weak association for the study variables. The study also showed that the confirmed and discharged cases were strongly associated hence there was a minimum number of death cases for both countries. This affirmation concurred with the number of deaths recorded.

In Figure 3, we observed that the discharged cases were 8,770 (97.7%), including death cases for the period under consideration. This implies that about 206 (2.3%) active cases were still receiving treatment in the country’s health facilities. While in Figure 4, about 32,761 (73.7%), including death cases had been discharged, and about 11,672 (26.3%) active cases remained in the country’s health facilities. In Figure 3 and 4, we observed that the number of discharged cases was approaching the number of confirmed cases. The trend in Figures 3 and 4 continues for the two countries.

The total confirmed cases for the two countries for the period under review were 53,105 with 83.2% cases reported in Nigeria and 16.8% cases reported in Malaysia. Malaysia has recorded 96.3% recovering on country-wise analysis than Nigeria with 71.7% analysed

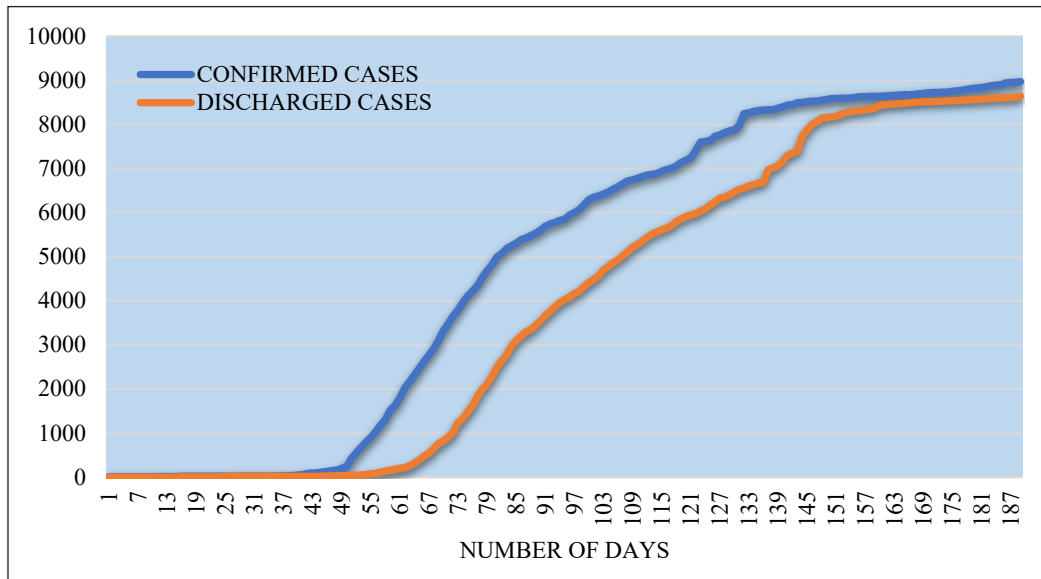


Figure 3. Comparative analysis of confirmed and discharged cases of Covid-19 in Malaysia

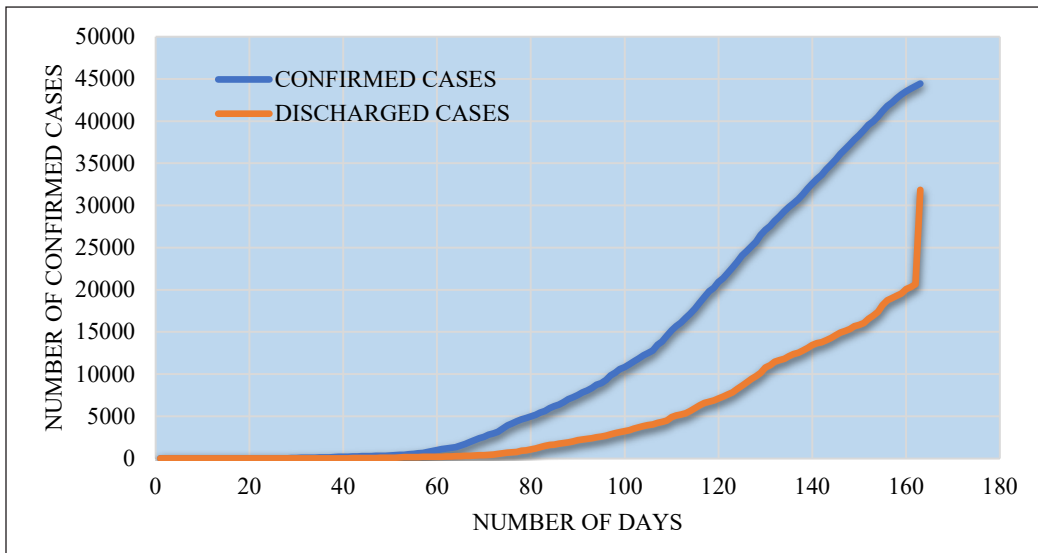


Figure 4. Comparative analysis of confirmed and discharged cases of Covid-19 in Nigeria

using similar conditions for the period under review. As at 31st July 2020, the active cases based on the data source (KKM, 2020), Malaysia had 206 (2.3%) active cases while as at 4th August 2020, Nigeria (NCDC, 2020) had 11,672 (26.3%) active cases (number of deaths inclusive). Based on the results and analysis above, this study confirmed a strong association between confirmed and discharged cases. This implies a fragile association between the confirmed and death cases.

In general, without vaccines and recommended drugs to treat Covid-19 patients, the two countries studied have done remarkably well to fight the pandemic. Based on the current circumstances whereby on availability of drugs and vaccines, it is better to examine the level of association between the confirmed, discharged, and death cases. Therefore, due to the nature of global data on daily confirmed, discharged and death cases, Covid-19 data may not be normally distributed. Using conventional methods to determine the association level may indicate a lower level of association due to the outlier effects on conventional methods. Apart from the proposed method, every other method, especially the multivariate regression technique is heavily affected by the presence of influential observations. Therefore, we may conclude that due to the present scenario surrounding Covid-19 management, robust and precise procedures be applied to determine the level of association to enable a caregiver to infer that there is a strong association that Covid-19 infected person may recover or die. Therefore, this study confirmed that a strong association between the confirmed and recovered cases implied that the number of casualties might be small depending on the infected cases.

CONCLUSIONS

The comparative analysis of the proposed methods and the conventional methods showed that the proposed methods were insensitive to influential observations or random outliers than the conventional methods. The study based on these methods showed moderate to a strong association between confirmed, discharged, and death cases. It implies that someone infected with the virus can only be discharged or die. Although the different methods revealed different degrees of association, the proposed methods showed a robust association indicating that the infected persons could either recover or die. Other techniques showed a different association level based on the multivariate correlation, multivariate coefficient of determination, and adjusted coefficient of determination. Therefore, there is a very weak association between confirmed and death cases and a very strong association between confirmed and discharged cases. The study affirmed that the strong association between confirmed and discharged cases implies a minimal number of deaths recorded in both countries, thereby confirming a high survival rate and minimal death rate. Therefore, it is recommended that the method with strong association be accepted due to the lack of vaccine and globally accepted drugs to treat Covid-19 for effective case management. These strengths of the association will encourage people to adopt non-pharmaceutical interventions such as the “Movement control order” in Malaysia, social distancing, and lockdown in Nigeria. However, these two countries still maintain the same trend of infection, discharge, and low death rate.

ACKNOWLEDGEMENT

We would like to express our gratitude to Universiti Utara Malaysia and Delta State University for the support to carry out this research work. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- Abdi, H. (2007). Multiple correlation coefficient. In N. Salkind (Ed.), *Encyclopedia of Measurement and Statistics* (pp. 648-651). Sage Publication.
- Abdullah, M. B. (1990). On a robust correlation coefficient. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 39(4), 455-460. <https://doi.org/10.2307/2349088>
- Armstrong, R. A. (2019). Should Pearson’s correlation coefficient be avoided? *Ophthalmic and Physiological Optics*, 39(5), 316-327. <https://doi.org/10.1111/opo.12636>
- Asuero, A. G., Sayago, A., & Gonzalez, A. G. (2006). The correlation coefficient: An overview. *Critical Reviews in Analytical Chemistry*, 36(1), 41-59. <https://doi.org/10.1080/10408340500526766>
- Bareinboim, E., Tian, J., & Pearl, J. (2014). Recovering from selection bias in causal and statistical inference. In *Proceedings of the National Conference on Artificial Intelligence* (Vol. 28, No. 1). PKP Publishing Services Network.

- Brown, G., Pocock, A., Zhao, M. J., & Luján, M. (2012). Conditional likelihood maximisation: a unifying framework for information theoretic feature selection. *The Journal of Machine Learning Research*, 13(1), 27-66.
- Châtillon, G. (1984). The balloon rules for a rough estimate of the correlation coefficient. *American Statistician*, 38(1), 58-60. <https://doi.org/10.1080/00031305.1984.10482875>
- Garnett, J. C. (1919). General ability, cleverness and purpose. *British Journal of Psychology*, 9(3), 345-366.
- Geiß, S., & Einax, J. (1996). Multivariate correlation analysis - A method for the analysis of multidimensional time series in environmental studies. *Chemometrics and Intelligent Laboratory Systems*, 32(1), 57-65. [https://doi.org/10.1016/0169-7439\(95\)00067-4](https://doi.org/10.1016/0169-7439(95)00067-4)
- Geiss, S., Einax, J., & Danzer, K. (1991). Multivariate correlation analysis and its application in environmental analysis. *Analytica Chimica Acta*, 242, 5-9. [https://doi.org/10.1016/0003-2670\(91\)87040-E](https://doi.org/10.1016/0003-2670(91)87040-E)
- Huberty, C. J. (2003). Multiple correlations versus multiple regression. *Educational and Psychological Measurement*, 63(2), 271-278. <https://doi.org/10.1177/0013164402250990>
- Lewis-Beck, M. S., Bryman, A., & Futing Liao, T. (2004). *The SAGE Encyclopedia of Social Science Research Methods* (Vols. 1-0). Sage Publications, Inc. <https://doi.org/10.4135/9781412950589>
- KKM. (2020). *Distribution of covid-19 cases according to date of confirmation*. Retrieved October 01, 2020, from <http://covid-19.moh.gov.my/>
- Mukaka M. M. (2012). Statistics corner: A guide to the appropriate use of the Correlation coefficient in medical research. *Malawi Medical Journal*, 24(3), 69-71.
- Nagelkerke, N. J. (1991). A note on a general definition of the coefficient of determination. *Biometrika*, 78(3), 691-692. <https://doi.org/10.1093/biomet/78.3.691>
- Nakagawa, S., Johnson, P. C. D., & Schielzeth, H. (2017). The coefficient of determination R² and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14(134), Article 20170213. <https://doi.org/10.1098/rsif.2017.0213>
- NCDC. (2020). *The official Twitter account of the Nigeria Centre for Disease Control*. Retrieved June 19, 2020, from <https://twitter.com/ncdcgov>
- Nguyen, H. V., Müller, E., Vreeken, J., Keller, F., & Böhm, K. (2013). CMI: An information-theoretic contrast measure for enhancing subspace cluster and outlier detection. In *Proceedings of the 2013 SIAM International Conference on Data Mining* (pp. 198-206). Society for Industrial and Applied Mathematics. <https://doi.org/10.1137/1.9781611972832.22>
- Okwonu, F. Z., Asaju, B. L., & Arunaye, F. I. (2020, September). Breakdown analysis of pearson correlation coefficient and robust correlation methods. In *IOP Conference Series: Materials Science and Engineering* (Vol. 917, No. 1, p. 012065). IOP Publishing. <https://doi.org/10.1088/1757-899X/917/1/012065>
- Pearson, K. (1920). Notes on the history of correlation. *Biometrika*, 13(1), 25-45. <https://doi.org/10.1093/biomet/13.1.25>
- Pyrczak, F., & Oh, D. M. (2018). *Making sense of statistics: A conceptual overview* (7th ed.). Routledge.

- Rodgers, L. J., & Nicewander, W. L. (1988). Thirteen ways to look at the correlation coefficient. *The American Statistician*, 42(1), 59-66. <https://doi.org/10.1080/00031305.1988.10475524>
- Tan, Z., Jamdagni, A., He, X., Nanda, P., & Liu, R. P. (2011). Denial-of-service attack detection based on multivariate correlation analysis. In *International Conference on Neural Information Processing* (pp. 756-765). Springer. https://doi.org/10.1007/978-3-642-24965-5_85
- Urain, J., & Peters, J. (2019). Generalized multiple correlation coefficient as a similarity measurement between trajectories. In *IEEE International Conference on Intelligent Robots and Systems* (pp. 1-7). IEEE Conference Publication. <https://doi.org/10.1109/IROS40897.2019.8967884>
- Wang, J., & Zheng, N. (2020). Correlation with applications (1): Measures of correlation for multiple variables. In *IEEE International Conference on Intelligent Robots and Systems* (pp. 1-18). Cornell University Press.
- Wang, L., Tang, X., Zhang, J., & Guan, D. (2018). Correlation Analysis for Exploring Multivariate Data Sets. *IEEE Access*, 6, 44235-44243. <https://doi.org/10.1109/ACCESS.2018.2864685>
- Wang, Y., Romano, S., Nguyen, V., Bailey, J., Ma, X., & Xia, S. T. (2017). Unbiased multivariate correlation analysis. In *Proceedings of the AAI Conference on Artificial Intelligence* (Vol. 31, No. 1). PKP Publishing Services Network.
- Weida, F. M. (1927). On various conceptions of correlation. *The Annals of Mathematics*, 29(1/4), 276-312. <https://doi.org/10.2307/1968000>
- Zhang, X., Pan, F., Wang, W., & Nobel, A. (2008). Mining non-redundant high order correlations in binary data. In *Proceedings of the VLDB Endowment International Conference on Very Large Data Bases* (Vol. 1, No. 1, p. 1178). NIH Public Access. <https://doi.org/10.14778/1453856.1453981>



Generalized Fibonacci Search Method in One-Dimensional Unconstrained Non-Linear Optimization

Chin Yoon Chong^{1*}, Soo Kar Leow¹ and Hong Seng Sim²

¹*School of Mathematics Sciences, Sunway University, Jalan Universiti, Bandar Sunway, 47500 Selangor, Malaysia*

²*Department of Mathematical and Actuarial Sciences, Universiti Tunku Abdul Rahman, 43000 Kajang, Selangor, Malaysia*

ABSTRACT

In this paper, we develop a generalized Fibonacci search method for one-dimensional unconstrained non-linear optimization of unimodal functions. This method uses the idea of the “ratio length of 1” from the golden section search. Our method takes successive lower Fibonacci numbers as the initial ratio and does not specify beforehand, the number of iterations to be used. We evaluated the method using Microsoft Excel with nine one-dimensional benchmark functions. We found that our generalized Fibonacci search method out-performed the golden section and other Fibonacci-type search methods such as the Fibonacci, Lucas and Pell approaches.

Keywords: Fibonacci-type, Generalized Fibonacci search, golden section

INTRODUCTION

Many techniques have been introduced to solve the problem of the one-dimensional unconstrained non-linear optimization, i.e. to minimize (or maximize) a unimodal function $f(x)$, where x is real.

In this paper, we will focus more on the Fibonacci-type search methods. We first study the importance of the bracketing search method (Bazaraa et al., 2013; Belegundu & Chandrupatla, 2019; Hassin, 1981; Rao, 2019). Subsequently, we proceed to the golden section search (Beamer & Wilde, 1969; Luenberger & Ye, 1984; Solomon, 2015; Rao, 2019), which is a technique to achieve the optimal of a unimodal function

ARTICLE INFO

Article history:

Received: 18 October 2020

Accepted: 14 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.17>

E-mail addresses:

cychong@sunway.edu.my (Chin Yoon Chong)

skleow@sunway.edu.my (Soo Kar Leow)

simhs@utar.edu.my (Hong Seng Sim)

* Corresponding author

within a specified interval by involving a ratio, $r = \frac{1}{2}(\sqrt{5} - 1)$ such that $r = \varphi^{-1}$, where $\varphi = \frac{1}{2}(\sqrt{5} + 1)$, namely, the golden ratio (Bazaraa et al., 2013; Belegundu & Chandrupatla, 2019; Solomon, 2015). Next, we further our research to the well-known Fibonacci search method (Witzgall, 1969; Kiefer, 1953; Avriel & Wilde, 1966; Hassin, 1981).

By applying the idea of “ratio length of 1”, we will introduce a different approach to obtain the ratio (r) of the golden section search in one-dimensional optimization problems, and we also propose a different approach of Fibonacci search by putting the successive lower Fibonacci numbers as the initial ratios. Next, by introducing the *sub-Fibonacci numbers*, we will extend our results to a new search method, called the generalized Fibonacci search. This is the main emphasis in our research. Like the golden section search, both the Fibonacci search and generalized Fibonacci search methods have similar algorithms where their ratios are changed at every iteration.

Based on the necessary and sufficient conditions obtained in the generalized Fibonacci search, together with the bracketing search upon the initial interval with same width, we use Microsoft Excel to test nine one-dimensional benchmark functions and compare our results with the golden section search and other Fibonacci-type search methods such as the Fibonacci, Lucas and Pell (Koshy, 2019). Furthermore, the Fibonacci-like numbers of Mersenne, Jacobsthal (Catarino et al., 2016), arithmetic or geometric (Chong & Ho, 2015) will also be discussed.

Fibonacci-Like Numbers

A famous Italian mathematician, Leonardo Fibonacci (Pisa, c.1170 – Pisa, c.1240) wrote on a variety of mathematical topics and *Fibonacci* is remembered particularly for the sequence numbers as 0, 1, 1, 2, 3, 5, 8, 13, 21, 34, ..., to which his name has been applied. This sequence, even today, is the subject of continuing research, especially by the Fibonacci Association, which publishes *The Fibonacci Quarterly*. The Fibonacci numbers are recursively defined by $F_{n+2} = F_{n+1} + F_n$ for all integers $n \geq 0$, with initial values $F_0 = 0$, $F_1 = 1$ (Koshy, 2019).

Marin Mersenne (Maine, 8 September 1588 – Paris, 1 September 1648), a French mathematician, is perhaps best known today among mathematicians for the Mersenne prime numbers, those which can be written in the form $M_n = 2^n - 1$. This can be described as a sequence of Fibonacci-like numbers namely Mersenne numbers as $M_0 = 0$, $M_1 = 1$, $M_{n+2} = 3M_{n+1} - 2M_n$ (Catarino et al., 2016).

In seventeenth-century, an English Mathematician, John Pell (Southwick, 1 March 1611 – Westminster, 12 December 1685) defined a sequence of numbers 0, 1, 2, 5, 12, 29, ..., also known as Pell numbers, which can be recursively generated as $P_0 = 0$, $P_1 = 1$, $P_{n+2} = 2P_{n+1} + P_n$. Pell numbers can be used to find square triangular numbers, to construct integer approximations

to the right isosceles triangle, and to solve certain combinatorial enumeration problems (Catarino et al., 2016; Kilic, 2007).

Jackques Philippe Marrie Binet (Rennes, 2 February 1786 – Paris, 12 May 1856), a French mathematician developed a Binet formula in 1843 that directly calculate the value of the n th Fibonacci number, that is, $F_n = \frac{1}{\sqrt{5}} \left[\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right]$, where $\frac{1 \pm \sqrt{5}}{2}$ are the roots of $x^2 + x + 1 = 0$. The well-known golden ratio also can be obtained from the limiting value on the ratio of two successive Fibonacci numbers as n approaches infinity, i.e., $\varphi = \lim_{n \rightarrow \infty} \frac{F_n}{F_{n-1}} = \frac{\sqrt{5}+1}{2}$. However, Binet was not the first to find this expression. De Moivre had discovered it much earlier in 1718 (Koshy, 2019).

Another French mathematician, Édouard Anatole Lucas (Amiens, 4 April 1842 – Paris, 3 October 1891) succeeded in branching out to another type of sequence, that is: 2, 1, 3, 4, 7, 11, 18, 29, 47, ..., which is named after him as the Lucas numbers. The Lucas number is denoted recursively as $L_{n+2} = L_{n+1} + L_n$, with the initial values $L_0 = 2$ and $L_1 = 1$. The Lucas numbers are closely related to the Fibonacci numbers (Koshy, 2019).

In 1965, Alwyn Francis Horadam (Hunter Valley, 22 March 1923 – Armidale, July 2016), an Australian mathematician introduced a Horadam-generalized Fibonacci number H_n , which is defined by $H_{n+2} = pH_{n+1} + qH_n$ with initial values $H_0 = a$, $H_1 = b$, where a , b , p and q are real constants (Horadam, 1965; Horzum & Kocer, 2009; Udea, 1996). He developed and obtained the generalized Fibonacci numbers by giving a very useful generalized Binet formula as follows, $H_n = \frac{1}{\alpha - \beta} \left[(b - a\beta)\alpha^n - (b - a\alpha)\beta^n \right]$, where

$$\alpha = \frac{p + \sqrt{p^2 + 4q}}{2} \quad \text{and} \quad \beta = \frac{p - \sqrt{p^2 + 4q}}{2}$$

are two distinct real roots of the quadratic

equation $x^2 - px - q = 0$. α is known as the generalized golden ratio which is also the limit of the ratio of two successive generalized Fibonacci numbers when n approaches

$$\text{infinity, that is, } \varphi = \lim_{n \rightarrow \infty} \frac{H_n}{H_{n-1}} = \frac{p + \sqrt{p^2 + 4q}}{2}.$$

Golden Section Search and Fibonacci Search

An American mathematical statistician, Jack Carl Kiefer (Cincinnati, 25 January 1924 – Berkeley, 10 August 1981) discovered the golden section search and the well-known Fibonacci search (Kiefer, 1953). These are the sequential methods for seeking the optimum of a unimodal function $f(x)$ as the golden section search needs a fixed ratio φ^{-1} , where φ is the golden ratio, while the ratio of the Fibonacci search varies in every stage of the iterations.

The theoretical advantage of the zero-order methods, like the Fibonacci and the golden section search. Thus, these search methods are widely popular as their algorithms do not require differentiability of the unimodal function $f(x)$ as an underlying assumption. In practice, the golden section search is frequently used when the total number of iterations required need not be specified.

Unlike the golden section procedure, the Fibonacci method requires the total number of iterations be chosen beforehand. For n large enough, then $\frac{F_{n-1}}{F_n} = \varphi^{-1}$, and thus the Fibonacci search method and the golden section are identical. It is worth mentioning that among the derivative-free methods that minimize strict unimodal functions over a closed bounded interval, the Fibonacci search method is the most efficient in that it requires the smallest number of iterations for a given reduction in the length of the interval of uncertainty (Bazaraa et al., 2013; Belegundu & Chandrupatla, 2019; Hassin, 1981; Luenberger & Ye, 1984; Solomon, 2015).

Subasi et al. (2004) searched the optimal point by Lucas numbers in the Fibonacci search algorithm, the improved method gave better results, where the optimal values of the objective functions were more sensitive in the improved method when their results were examined by a developed computer program in MAPLE. Demir et al. (2008) presented a mathematical analysis of the Fibonacci search method by k -Lucas numbers. In this study, they developed a new algorithm which determined the maximum point of unimodal functions on closed intervals. As a result, it made the Fibonacci search method more effective.

MATERIALS AND METHODS

In this section, we will focus on the line search algorithm for the one-dimensional optimization problem. We will describe the bracketing, the golden section, the Fibonacci and the generalized Fibonacci search. Further, we will also discuss the conditions to be fulfilled for the generalized Fibonacci search. We will close this section with the comparison of our results with the results from the golden section search and the other Fibonacci-type search methods.

Unimodal Function

A unimodal function is one that has only one peak (maximum) or valley (minimum) in a given domain. For instance, a univariate function $f(x)$ is unimodal on an interval $[x_1, x_2]$ if there exists a minimum $x^* \in (x_1, x_2)$ such that $f(x)$ is monotonically decreasing for $x \in [x_1, x^*]$ and monotonically increasing for $x \in [x^*, x_2]$. In other words, there are no other local minima on the interval (x_1, x_2) (Avriel & Wilde, 1966; Beamer & Wilde, 1969; Hassin, 1981; Kiefer, 1953; Oliver & Wilde, 1964; Witzgall, 1969).

It is perhaps more conventional to define unimodal functions in the opposite sense, such that there is a unique global *maximum* rather than a minimum. However, since any maximization problem can be transformed into a minimization problem by multiplying the function by a negative sign. Moreover, a unimodal function with an optimum $x^* \in (x_1, x_2)$ can be a nondifferentiable or even a discontinuous function.

For the rest of the paper, we will assume $f(x)$ is a unimodal function which has a minimizer x^* in the interval (x_1, x_2) , that is $x_1 < x^* < x_2$. Let x_3 and x_4 be the two points such that $x_1 < x_3 < x_4 < x_2$.

Bracketing Search Method

The search for a local optimum x^* will involve a sequence of function evaluations for some different points of $x \in [x_1, x_2]$. One after the other, using the previous results of the function evaluations to obtain the new interval that bounds the optimum x^* . At each iteration, the closer interval that brackets the optimum x^* is called the *interval of uncertainty*. From the past, bracketing search has been discussed by many researchers (Avriel & Wilde, 1966; Belegundu & Chandrupatla, 2019; Hassin, 1981; Oliver & Wilde, 1964, Solomon, 2015; Witzgall, 1969). We have two possibilities of the interval of uncertainty as follows.

Case 1

If $f(x_3) < f(x_4)$, then $x_1 < x^* < x_3 < x_4 < x_2$ or $x_1 < x_3 < x^* < x_4 < x_2$. Thus, $x^* \in (x'_1, x'_2)$ such that $x'_1 = x_1$ and $x'_2 = x_4$.

Case 2

If $f(x_3) > f(x_4)$, then $x_1 < x_3 < x_4 < x^* < x_2$ or $x_1 < x_3 < x^* < x_4 < x_2$. Thus, $x^* \in (x'_1, x'_2)$ such that $x'_1 = x_3$ and $x'_2 = x_2$.

Remark

$|x'_2 - x'_1| < |x_2 - x_1|$ for the above two cases.

Golden Section Search Method

The ratio r of the golden section search plays a very important role in line search method as described above. This ratio, r acts as a “generator” to generate the values of x_3 and x_4 from the initial interval $[x_1, x_2]$. The following algorithm shows the procedure how can the ratio r be obtained.

Let $x_2 - x_1 = d$ and $x_3 - x_1 = x_2 - x_4$ in the first iteration as shown in Figure 1.

Let r be a ratio such that $\frac{x_4 - x_1}{x_2 - x_1} = \frac{x_2 - x_3}{x_2 - x_1} = r$, where $0 < r < 1$.

$$\text{Then } x_4 - x_1 = x_2 - x_3 = rd \text{ gives } x_3 - x_1 = x_2 - x_4 = (1-r)d \tag{1}$$

$$\text{Simultaneously, let } \frac{x_3 - x_1}{x_4 - x_1} = \frac{x_2 - x_4}{x_2 - x_3} = r. \text{ Then } x_3 - x_1 = x_2 - x_4 = r^2d \tag{2}$$

Equation 1 and 2 will give $r = \frac{1}{2}(\sqrt{5}-1)$. Moreover, by substituting $d = x_2 - x_1$ into Equation 1 yields $x_4 = (1-r)x_1 + rx_2$ and $x_3 = (1-r)x_1 + rx_2$, respectively. Together with the bracketing search, the process will obtain an interval of uncertainty for which $x^* \in (x'_1, x'_2)$. Without loss of generality, at m th iteration, then

$$x_3^{(m-1)} = rx_1^{(m-1)} + (1-r)x_2^{(m-1)} \quad \text{and} \quad x_4^{(m-1)} = (1-r)x_1^{(m-1)} + rx_2^{(m-1)}$$

The similar algorithm has been discussed in some articles (Bazaraa et al., 2013; Solomon, 2015; Belegundu & Chandrupatla, 2019).

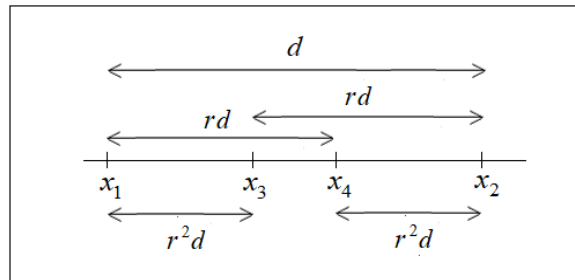


Figure 1. Length measured in terms of r and d

Different Approach to Obtain the Ratio (r) of Golden Section Search

In this section, we introduce a different approach to obtain the ratio r of the golden section search by applying the idea of “ratio length of 1”.

Let $x_2 - x_1 = 1$ and $x_3 - x_1 = x_2 - x_4$, as shown in Figure 2. In other words, we treat the length, $x_2 - x_1$ as a total ratio of 1, and based on the same algorithm as discussed above, we will obtain the same results, where the length from the diagram is measured in terms of r .

This idea of “ratio length of 1” will lead us to a different approach of the Fibonacci search, and also provide us an idea to develop a new search method, that is the generalized Fibonacci search. This will be discussed in the next section of **Different Approach of Fibonacci Search**, and **Generalized Fibonacci Search** as well.

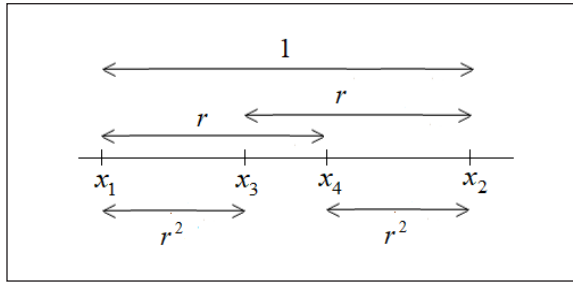


Figure 2. Length measured in terms of r

Fibonacci Search with a Fixed Number of Iterations

In this section, we first describe the Fibonacci search as mentioned by Bazaraa et al. (2013), Belegundu and Chandrupatla (2019), Hassin (1981), Mathews et al. (2004), Subasi et al. (2004), which starts with a ratio with two successive higher Fibonacci numbers, that is $\frac{F_{n-2}}{F_n}$ for m iterations.

F_n Golden section search and Fibonacci search have a very close relationship, and thus we first to investigate the initial ratio of the Fibonacci search by comparing the ratio of the golden section search. Suppose $x_2 - x_1 = d$ and $x_3 - x_1 = x_2 - x_4$, as illustrated in Figure 3. Assume a fixed number of iterations, m is needed to achieve the desired optimal x^* .

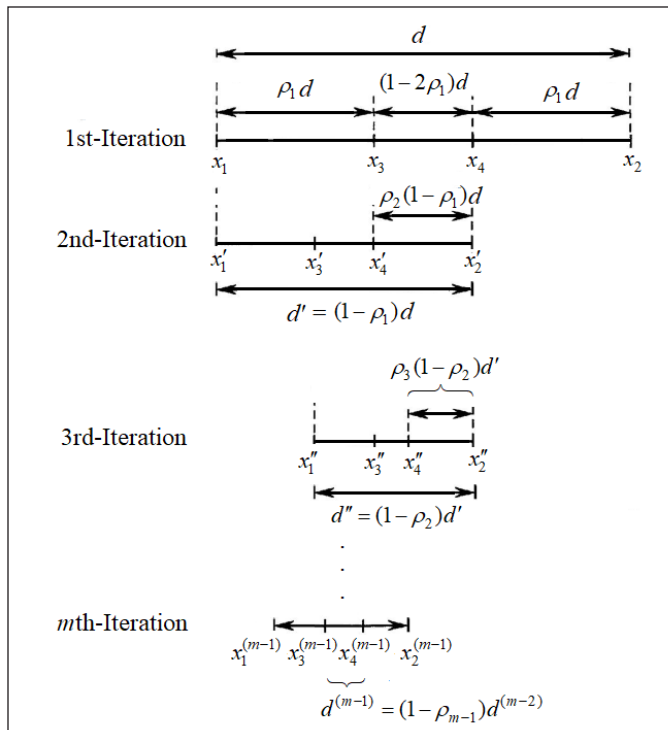


Figure 3. Initial ratio that has higher Fibonacci numbers

In the first iteration, we let $\frac{x_3 - x_1}{x_2 - x_1} = \frac{x_2 - x_4}{x_2 - x_1} = \rho_1$, where $0 < \rho_1 < 1$, then $x_4 - x_1 = (1 - \rho_1)d$. Compared to $x_4 - x_1 = rd$ in Figure 1, therefore we have $1 - \rho_1 = r$. Since all the search methods will give a high efficiency when $n \rightarrow \infty$, therefore, $1 - \rho_1 = r = \frac{1}{\phi} = \lim_{n \rightarrow \infty} \frac{1}{F_n} = \lim_{n \rightarrow \infty} \frac{F_{n-1}}{F_n}$ gives $\rho_1 = 1 - \frac{F_{n-1}}{F_n} = \frac{F_{n-2}}{F_n}$ as n approaches

infinity. Hence, this is proven that the initial ratio of the Fibonacci search will involve the successive higher Fibonacci numbers. Next, we will obtain a new interval of uncertainty, where $x^* \in (x'_1, x'_2)$ after the bracketing search is imposed.

In the second iteration, let $d' = x'_2 - x'_1 = x_4 - x_1 = (1 - \rho_1)d$ such that $x'_2 - x'_4 = x_4 - x_3$ and exists a ratio ρ_2 for which $\frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{x'_2 - x'_4}{x'_2 - x'_1} = \rho_2$. Hence, $\rho_2(1 - \rho_1)d = (1 - 2\rho_1)d$ gives

$$\rho_2 = \frac{1 - 2\rho_1}{1 - \rho_1} = \frac{1 - \frac{2F_{n-2}}{F_n}}{1 - \frac{F_{n-2}}{F_n}} = \frac{F_n - 2F_{n-2}}{F_n - F_{n-2}} = \frac{F_{n-3}}{F_{n-1}}$$

Next, the bracketing search gives a new interval of uncertainty such that $d'' = (1 - \rho_2)d'$. Repeat the above process, then $d'' = (1 - \rho_2)d'$ and $\rho_3 = \frac{F_{n-4}}{F_{n-2}}$ in the third iteration.

Generally, $d^{(m-1)} = (1 - \rho_{m-1})d^{(m-2)}$ and $\rho_m = \frac{F_{n-m-1}}{F_{n-m+1}}$ in m th iteration.

Thus, when $m = n - 2$, the results obtained:

$$\begin{aligned} d^{(n-3)} &= (1 - \rho_{n-3})d^{(n-4)} = (1 - \rho_{n-3})(1 - \rho_{n-4}) \cdots (1 - \rho_3)(1 - \rho_2)(1 - \rho_1)d \\ &= \left(1 - \frac{F_2}{F_4}\right) \left(1 - \frac{F_3}{F_5}\right) \cdots \left(1 - \frac{F_{n-4}}{F_{n-2}}\right) \left(1 - \frac{F_{n-3}}{F_{n-1}}\right) \left(1 - \frac{F_{n-2}}{F_n}\right) d \\ &= \frac{F_3}{F_4} \frac{F_4}{F_5} \cdots \frac{F_{n-2}}{F_{n-1}} \frac{F_{n-1}}{F_n} d \\ &= \frac{F_3 d}{F_n} \end{aligned}$$

Fibonacci Search Without a Fixed Number of Iterations

What if we begin the ratio with the two successive lower Fibonacci numbers and without a fixed number of iterations? We will propose an algorithm procedure that begins with the

first ratio with lower Fibonacci numbers and the results will show that both procedures are equivalent. This result brews an idea and lead us to develop the new approach of search method in next sections.

Let $x_2 - x_1 = d$ and $x_3 - x_1 = x_2 - x_4$, as shown in Figure 4. Next, we proposed another method by considering the assumption of first ratio with the lower Fibonacci numbers as $\lambda_1 = \frac{F_2}{F_4}$ in the first iteration such that $\frac{x_3 - x_1}{x_2 - x_1} = \frac{x_2 - x_4}{x_2 - x_1} = \lambda_1$, and thus $x_4 - x_3 = (1 - 2\lambda_1)d$. Subsequently, we will obtain a new interval of uncertainty for which $x^* \in (x'_1, x'_2)$ upon the bracketing search.

In the second iteration with the assumption by taking another ratio with one-step ahead of two successive lower Fibonacci numbers as $\lambda_2 = \frac{F_3}{F_5}$, and let

$$D' = x'_2 - x'_1 = x_4 - x_1 = (1 - \lambda_1)d \text{ such that } \frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{x'_2 - x'_4}{x'_2 - x'_1} = \lambda_2. \text{ Take note that } x'_2 - x'_4$$

and $x_4 - x_3$ are not necessarily in equal length, whereby $x'_2 - x'_4 = x_4 - x_3$ in Figure 3. Next, the bracketing search gives a new interval of uncertainty such that $x^* \in (x''_1, x''_2)$.

As such, without loss of generality, we will have $D'' = (1 - \lambda_2)D'$ and $\lambda_3 = \frac{F_4}{F_6}$ in the third iteration, and $x^* \in (x''_1, x''_2)$. Generally, $D^{(m-1)} = (1 - \lambda_{m-1})D^{(m-2)}$ and $\lambda_m = \frac{F_{m+1}}{F_{m+3}}$

in m th iteration, where $D^{(0)} = d$. Thus, when $m = n - 2$, we get

$$\begin{aligned} D^{(n-3)} &= (1 - \lambda_{n-3})D^{(n-4)} = (1 - \lambda_{n-3})(1 - \lambda_{n-4}) \cdots (1 - \lambda_3)(1 - \lambda_2)(1 - \lambda_1)d \\ &= \left(1 - \frac{F_{n-2}}{F_n}\right) \left(1 - \frac{F_{n-3}}{F_{n-1}}\right) \cdots \left(1 - \frac{F_3}{F_5}\right) \left(1 - \frac{F_2}{F_4}\right) d \\ &= \frac{F_{n-1}}{F_n} \frac{F_{n-2}}{F_{n-1}} \cdots \frac{F_4}{F_5} \frac{F_3}{F_4} d \\ &= \frac{F_3 d}{F_n} \end{aligned}$$

which is equal to the result of $d^{(n-3)}$ in Figure 3. This is proven that the Fibonacci search can be iterated with the first ratio of successive lower Fibonacci numbers. Hence, compared to the previous method, this will advantage us to run the algorithm with the successive lower Fibonacci numbers and without a fixed number of iterations at the beginning. In other words, we could see the limitation in fixing m iterations at the start of the algorithm.

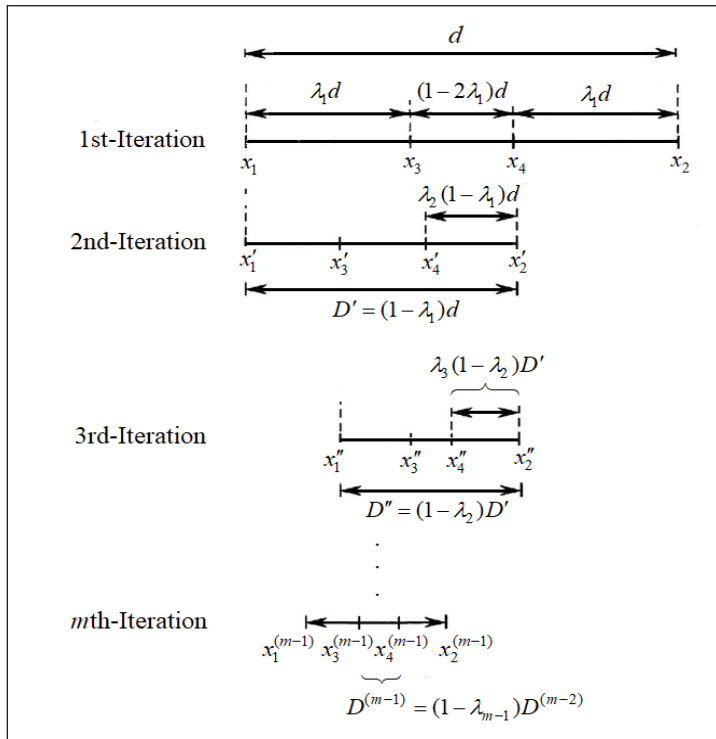


Figure 4. Initial ratio that has lower Fibonacci numbers

Different Approach of Fibonacci Search

In the following section, we proposed a different approach of the Fibonacci search method compared to the Fibonacci search as discussed by Bazaraa et al. (2013), Belegundu and Chandrupatla (2019), Hassin (1981), and Subasi et al. (2004).

Let F_n be the Fibonacci number which is defined recursively by $F_0 = 0$, $F_1 = 1$ and $F_{n+2} = F_{n+1} + F_n$ for all integers $n \geq 0$. Based on the idea of “ratio length of 1” as discussed in the golden section search above, we will consider the successive lower Fibonacci numbers,

i.e., $\frac{F_2}{F_4}$ and $\frac{F_3}{F_4}$ as the initial ratios in the first iteration. From Figure 5, we have Equation 3

$$\frac{x_2 - x_3}{x_2 - x_1} = \frac{x_4 - x_1}{x_2 - x_1} = \frac{F_3}{F_4} \text{ and } \frac{x_3 - x_1}{x_2 - x_1} = \frac{x_2 - x_4}{x_2 - x_1} = \frac{F_2}{F_4} \tag{3}$$

Then the following equalities hold:

$$\frac{x_2 - x_3}{x_2 - x_1} + \frac{x_3 - x_1}{x_2 - x_1} = \frac{F_3}{F_4} + \frac{F_2}{F_4} = 1 \text{ and } \frac{x_4 - x_1}{x_2 - x_1} + \frac{x_2 - x_4}{x_2 - x_1} = \frac{F_3}{F_4} + \frac{F_2}{F_4} = 1.$$

Hence, from Equation 3, we obtained the following results, $x_3 = \frac{F_3}{F_4}x_1 + \frac{F_2}{F_4}x_2$ and $x_4 = \frac{F_2}{F_4}x_1 + \frac{F_3}{F_4}x_2$.

Using the bracketing search, we will find the new interval of uncertainty for $x^* \in (x'_1, x'_2)$, where $|x'_2 - x'_1| < |x_2 - x_1|$.

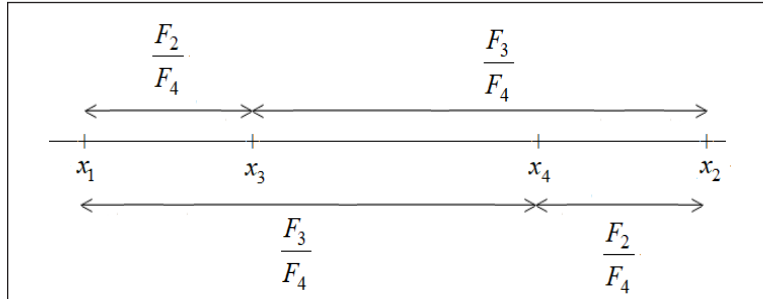


Figure 5. Ratios with lower Fibonacci numbers at 1st iteration

Subsequently, with the same process at the second iteration, we consider different ratios of one-step ahead of two successive Fibonacci numbers $\frac{F_3}{F_5}$ and $\frac{F_4}{F_5}$, and we let x'_3 and x'_4 be two points for which $x'_1 < x'_3 < x'_4 < x'_2$, such that

$$\frac{x'_2 - x'_3}{x'_2 - x'_1} = \frac{x'_4 - x'_1}{x'_2 - x'_1} = \frac{F_4}{F_5} \text{ and } \frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{x'_2 - x'_4}{x'_2 - x'_1} = \frac{F_3}{F_5} \tag{4}$$

subject to $\frac{x'_2 - x'_3}{x'_2 - x'_1} + \frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{F_4}{F_5} + \frac{F_3}{F_5} = 1$ and $\frac{x'_4 - x'_1}{x'_2 - x'_1} + \frac{x'_2 - x'_4}{x'_2 - x'_1} = \frac{F_4}{F_5} + \frac{F_3}{F_5} = 1$.

Then, Equation 4, we will obtain the following results, $x'_3 = \frac{F_4}{F_5}x'_1 + \frac{F_3}{F_5}x'_2$ and $x'_4 = \frac{F_3}{F_5}x'_1 + \frac{F_4}{F_5}x'_2$.

Next, we will obtain a new interval of uncertainty for $x^* \in (x''_1, x''_2)$ after the bracketing search is imposed, where $|x''_2 - x''_1| < |x'_2 - x'_1|$.

Repeat the above process until m th iteration as shown in Figure 6 so that

$$\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+2}}{F_{m+3}} \text{ and } \frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+1}}{F_{m+3}} \tag{5}$$

subject to
$$\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} + \frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+2}}{F_{m+3}} + \frac{F_{m+1}}{F_{m+3}} = 1$$

and
$$\frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} + \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+2}}{F_{m+3}} + \frac{F_{m+1}}{F_{m+3}} = 1.$$

Thus, Equation 5 gives the following results,

$$x_3^{(m-1)} = \frac{F_{m+2}}{F_{m+3}} x_1^{(m-1)} + \frac{F_{m+1}}{F_{m+3}} x_2^{(m-1)} \text{ and } x_4^{(m-1)} = \frac{F_{m+1}}{F_{m+3}} x_1^{(m-1)} + \frac{F_{m+2}}{F_{m+3}} x_2^{(m-1)}$$

such that $x_1^{(m-1)} < x_3^{(m-1)} < x_4^{(m-1)} < x_2^{(m-1)}$, where $\frac{F_{m+1}}{F_{m+3}} < \frac{F_{m+2}}{F_{m+3}}$ for all $m \geq 1$.

Next, the bracketing search yields a closer interval of uncertainty for which $x^* \in (x_1^{(m)}, x_2^{(m)})$ such that $|x_2^{(m)} - x_1^{(m)}| < |x_2^{(m-1)} - x_1^{(m-1)}|$, where $|x_2^{(m)} - x_1^{(m)}| \approx 0$ as $m \rightarrow \infty$.

The proof of the results of $x_3^{(m-1)}$ and $x_4^{(m-1)}$ is as follows:

At m th iteration, since
$$\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+2}}{F_{m+3}}, \text{ therefore } x_2^{(m-1)} - x_3^{(m-1)} = \frac{F_{m+2}}{F_{m+3}} [x_2^{(m-1)} - x_1^{(m-1)}]$$

gives
$$x_3^{(m-1)} = \frac{F_{m+2}}{F_{m+3}} x_1^{(m-1)} + \left(1 - \frac{F_{m+2}}{F_{m+3}}\right) x_2^{(m-1)} = \frac{F_{m+2}}{F_{m+3}} x_1^{(m-1)} + \frac{F_{m+1}}{F_{m+3}} x_2^{(m-1)}.$$

Similarly,
$$\frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+2}}{F_{m+3}} \text{ gives } x_4^{(m-1)} = \frac{F_{m+1}}{F_{m+3}} x_1^{(m-1)} + \frac{F_{m+2}}{F_{m+3}} x_2^{(m-1)}.$$

Alternatively,
$$\frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{F_{m+1}}{F_{m+3}}$$
 also gives the same results of $x_4^{(m-1)}$

and $x_4^{(m-1)}$.

Take note that the above results of $x_3^{(m-1)}$ and $x_4^{(m-1)}$ matched the results of golden section search as m approaches infinity. Since
$$\lim_{m \rightarrow \infty} \frac{F_{m+2}}{F_{m+3}} = \lim_{m \rightarrow \infty} \frac{1}{\frac{F_{m+3}}{F_{m+2}}} = \frac{1}{\phi} = r,$$
 therefore the above

results will give
$$\begin{aligned} x_3^{(m-1)} &= \frac{F_{m+2}}{F_{m+3}} x_1^{(m-1)} + \frac{F_{m+1}}{F_{m+3}} x_2^{(m-1)} = \frac{F_{m+2}}{F_{m+3}} x_1^{(m-1)} + \left(1 - \frac{F_{m+2}}{F_{m+3}}\right) x_2^{(m-1)} \\ &= r x_1^{(m-1)} + (1-r) x_2^{(m-1)} \end{aligned}$$

as $m \rightarrow \infty$. Thus, similarly for $x_4^{(m-1)}$, which will equal to $(1-r)x_1^{(m-1)} + r x_2^{(m-1)}$.

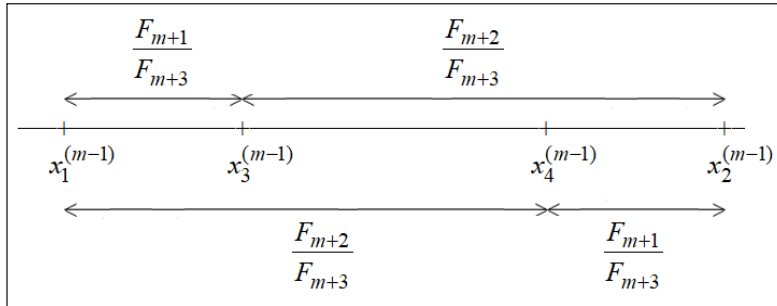


Figure 6. Ratios with higher Fibonacci numbers at m th iteration

Generalized Fibonacci Search

The generalized Fibonacci search method is the main focus of our research. It may have more complexities compared to the Fibonacci search. Unlike the Fibonacci search discussed previously, not all the Fibonacci-like numbers can be applied in the algorithm of the generalized Fibonacci search method. Thus, to satisfy the algorithm, we have to restrict the conditions of the combination of a , b , p and q from the generalized Fibonacci number (Horadam, 1965), $H_n = pH_{n-1} + qH_{n-2}$, $H_0 = a$, $H_1 = b$ for all real values of a , b , p and q . Thus, we will introduce a new term, *sub-Fibonacci number* before proceeding to discussion of the generalized Fibonacci search.

Definition 1. Let a and b be two nonnegative integers such that $a + b > 0$. Let p and q be two positive integers. Then $S_n(a, b, p, q)$ denote the n th *sub-Fibonacci number* for all integers $n \geq 0$, where $S_0 = a$, $S_1 = b$ and $S_{n+2} = pS_{n+1} + qS_n$.

Lemma 1. If $S_n(a, b, p, q)$ is an n th sub-Fibonacci number, then $S_n > 0$ for all $n \geq 2$.

Proof. Since $S_n(a, b, p, q)$ gives $S_0 = a \geq 0$, $S_1 = b \geq 0$ and $p, q > 0$, where a and b never take zero simultaneously, therefore $S_n = pS_{n-1} + qS_{n-2} > 0$ for all integers $n \geq 2$. We complete the proof.

Algorithm of Generalized Fibonacci Search

Let $S_n(a, b, p, q)$ be the n th sub-Fibonacci number for all integers $n \geq 0$. We begin the algorithm with the initial ratios of the successive lower Fibonacci numbers for $n \geq 2$,

i.e., we use $\frac{qS_2}{S_4}$ and $\frac{pS_3}{S_4}$ in the first iteration as shown in Figure 7, we have Equation 6

$$\frac{x_3 - x_1}{x_2 - x_1} = \frac{x_2 - x_4}{x_2 - x_1} = \frac{qS_2}{S_4} \text{ and } \frac{x_3 - x_1}{x_2 - x_1} = \frac{x_2 - x_4}{x_2 - x_1} = \frac{pS_3}{S_4} \tag{6}$$

subject to $\frac{x_2 - x_3}{x_2 - x_1} + \frac{x_3 - x_1}{x_2 - x_1} = \frac{pS_3}{S_4} + \frac{qS_2}{S_4} = 1$ and $\frac{x_4 - x_1}{x_2 - x_1} + \frac{x_2 - x_4}{x_2 - x_1} = \frac{pS_3}{S_4} + \frac{qS_2}{S_4} = 1$.

From Equation 6, we obtained the following results, $x_3 = \frac{pS_3}{S_4}x_1 + \frac{qS_2}{S_4}x_2$ and $x_4 = \frac{qS_2}{S_4}x_1 + \frac{pS_3}{S_4}x_2$.

We proceed to the bracketing search to find the new interval of uncertainty such that $x^* \in (x'_1, x'_2)$, where $|x'_2 - x'_1| < |x_2 - x_1|$.

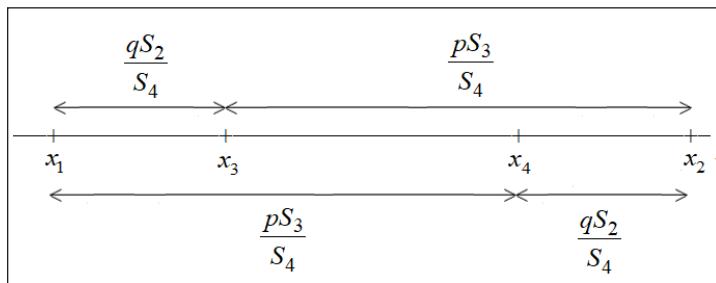


Figure 7. Ratios with lower generalized Fibonacci numbers at 1st iteration

Next, we repeat the above process for the second iteration with different ratios of one-step ahead of two successive sub-Fibonacci numbers, i.e., $\frac{qS_3}{S_5}$ and $\frac{pS_4}{S_5}$. Let x'_3 and x'_4 be another two points for which $x'_1 < x'_3 < x'_4 < x'_2$, such that:

$$\frac{x'_2 - x'_3}{x'_2 - x'_1} = \frac{x'_4 - x'_1}{x'_2 - x'_1} = \frac{pS_4}{S_5} \text{ and } \frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{x'_2 - x'_4}{x'_2 - x'_1} = \frac{qS_3}{S_5} \tag{7}$$

subject to $\frac{x'_2 - x'_3}{x'_2 - x'_1} + \frac{x'_3 - x'_1}{x'_2 - x'_1} = \frac{pS_4}{S_5} + \frac{qS_3}{S_5} = 1$ and $\frac{x'_4 - x'_1}{x'_2 - x'_1} + \frac{x'_2 - x'_4}{x'_2 - x'_1} = \frac{pS_4}{S_5} + \frac{qS_3}{S_5} = 1$.

Then, from Equation 7, we will obtain the following results,

$$x'_3 = \frac{pS_4}{S_5}x'_1 + \frac{qS_3}{S_5}x'_2 \text{ and } x'_4 = \frac{qS_3}{S_5}x'_1 + \frac{pS_4}{S_5}x'_2.$$

Continuing with the bracketing search, yields another new interval of uncertainty for $x^* \in (x''_1, x''_2)$, where $|x''_2 - x''_1| < |x'_2 - x'_1|$. Subsequently, after the third and fourth iteration, we will have the respective results:

$$x_3'' = \frac{pS_5}{S_6}x_1'' + \frac{qS_4}{S_6}x_2'', \quad x_4'' = \frac{qS_4}{S_6}x_1'' + \frac{pS_5}{S_6}x_2'', \text{ where } x_1'' < x_3'' < x_4'' < x_2'' \text{ and give } x^* \in (x_1'', x_2''),$$

$$x_3''' = \frac{pS_6}{S_7}x_1''' + \frac{qS_5}{S_7}x_2''', \quad x_4''' = \frac{qS_5}{S_7}x_1''' + \frac{pS_6}{S_7}x_2''', \text{ where } x_1''' < x_3''' < x_4''' < x_2''' \text{ and give } x^* \in (x_1^{(4)}, x_2^{(4)}).$$

Thus, generally at m th iteration, suppose $x_3^{(m-1)}$ and $x_4^{(m-1)}$ are the two new points after the bracketing search is imposed, where $x_1^{(m-1)} < x_3^{(m-1)} < x_4^{(m-1)} < x_2^{(m-1)}$, as shown in Figure 8, such that

$$\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{pS_{m+2}}{S_{m+3}} \text{ and } \frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{qS_{m+1}}{S_{m+3}} \quad [8]$$

subject to $\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} + \frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{pS_{m+2}}{S_{m+3}} + \frac{qS_{m+1}}{S_{m+3}} = 1$ and

$$\frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} + \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{pS_{m+2}}{S_{m+3}} + \frac{qS_{m+1}}{S_{m+3}} = 1.$$

Then, Equation 8 gives the results of $x_3^{(m-1)}$ and $x_4^{(m-1)}$ as

$$x_3^{(m-1)} = \frac{pS_{m+2}}{S_{m+3}}x_1^{(m-1)} + \frac{qS_{m+1}}{S_{m+3}}x_2^{(m-1)} \text{ and } x_4^{(m-1)} = \frac{qS_{m+1}}{S_{m+3}}x_1^{(m-1)} + \frac{pS_{m+2}}{S_{m+3}}x_2^{(m-1)}.$$

Next, the bracketing search yields the closer interval of uncertainty for $x^* \in (x_1^{(m)}, x_2^{(m)})$

such that $|x_2^{(m)} - x_1^{(m)}| < |x_2^{(m-1)} - x_1^{(m-1)}|$, where $|x_2^{(m)} - x_1^{(m)}| \approx 0$ as $m \rightarrow \infty$.

The following is the proof for the results of $x_3^{(m-1)}$ and $x_4^{(m-1)}$.

At m th iteration, since $\frac{x_2^{(m-1)} - x_3^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{pS_{m+2}}{S_{m+3}}$, then $x_2^{(m-1)} - x_3^{(m-1)} = \frac{pS_{m+2}}{S_{m+3}}[x_2^{(m-1)} - x_1^{(m-1)}]$

gives $x_3^{(m-1)} = \frac{pS_{m+2}}{S_{m+3}}x_1^{(m-1)} + \left(1 - \frac{pS_{m+2}}{S_{m+3}}\right)x_2^{(m-1)} = \frac{pS_{m+2}}{S_{m+3}}x_1^{(m-1)} + \frac{qS_{m+1}}{S_{m+3}}x_2^{(m-1)}.$

Similarly, $\frac{x_4^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{pS_{m+2}}{S_{m+3}}$ gives $x_4^{(m-1)} = \frac{qS_{m+1}}{S_{m+3}}x_1^{(m-1)} + \frac{pS_{m+2}}{S_{m+3}}x_2^{(m-1)}.$

Alternatively, $\frac{x_3^{(m-1)} - x_1^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{x_2^{(m-1)} - x_4^{(m-1)}}{x_2^{(m-1)} - x_1^{(m-1)}} = \frac{qS_{m+1}}{S_{m+3}}$ also give the same results $x_3^{(m-1)}$ and $x_4^{(m-1)}$.

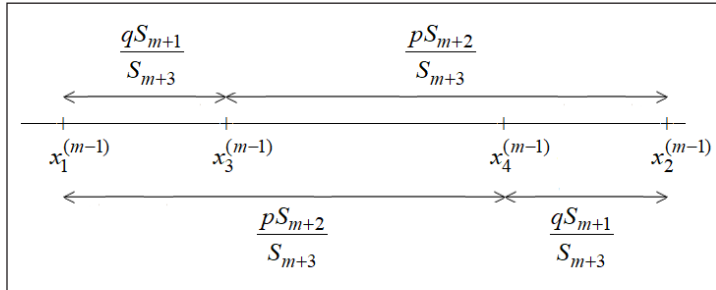


Figure 8. Ratios with higher generalized Fibonacci numbers at m th iteration

Not all the sub-Fibonacci numbers will fulfill the algorithm with different sets of combinations of a, b, p and q . Next, we will present the necessary and sufficient conditions of a, b, p and q to ensure the convergence of the optimal solution as follows.

Lemma 2. Consider an n th sub-Fibonacci number, $S_n(a, b, p, q)$ for all integers $n = k \geq 0$. If $qS_k < pS_{k+1}$ and $qS_{k+1} < pS_{k+2}$, then $qS_i < pS_{i+1}$ for all $i \geq k + 2$.

Proof. We shall prove this by mathematical induction for $S_n(a, b, p, q)$, where $a + b > 0$ and $n = k \geq 0$. Suppose the results of the following Equation 9 and 10 are true for $n = k \geq 0$. i.e.,

$$qS_k < pS_{k+1} \text{ gives } \frac{q^2}{p} S_k < qS_{k+1} \tag{9}$$

$$\text{and } qS_{k+1} < pS_{k+2} \tag{10}$$

$$\text{Now } qS_{k+2} = q(pS_{k+1} + qS_k) = p \left(qS_{k+1} + \frac{q^2}{p} S_k \right) < p(pS_{k+2} + qS_{k+1}) = pS_{k+3} \tag{11}$$

Thus, Equation 11 show that the result is also true for $n = k + 1$. Similarly, Equation 10 and 11 will lead to $qS_{k+3} < pS_{k+4}$. We complete the proof.

Definition 2. An optimal x^* of a unimodal function $f(x)$ is achievable if and only if x^* can be approximated by any line search method such that x^* is bounded within all the intervals of uncertainty.

Theorem 1. Let $S_n(a, b, p, q)$ denote the n th sub-Fibonacci number for all for all $n \geq 0$. Then the following statements are equivalent such that an optimal x^* of a unimodal function $f(x)$ is achievable by the generalized Fibonacci search for all $n \geq 2$ if and only if

- (a) $qS_n < pS_{n+1}$.
- (b) $0 < \frac{qS_n}{S_{n+2}} < \frac{1}{2} < \frac{pS_{n+1}}{S_{n+2}} < 1$.
- (c) $q(q-p^2)a < p^3b$ and $q(q-p^2)b < p^3(qa+pb)$.

Proof. Consider $S_n(a, b, p, q) > 0$ for all $n \geq 2$ such that $a + b > 0$ and $p, q > 0$.

(a) For all $m \geq 1$, the generalized Fibonacci search at m th iteration gives

$$x_3^{(m-1)} = \frac{pS_{m+2}}{S_{m+3}}x_1^{(m-1)} + \frac{qS_{m+1}}{S_{m+3}}x_2^{(m-1)} \quad \text{and} \quad x_4^{(m-1)} = \frac{qS_{m+1}}{S_{m+3}}x_1^{(m-1)} + \frac{pS_{m+2}}{S_{m+3}}x_2^{(m-1)}.$$

So, x^* is achievable if and only if $x_1^{(m-1)} < x_3^{(m-1)} < x_4^{(m-1)} < x_2^{(m-1)}$.

Thus, $x_3^{(m-1)} < x_4^{(m-1)}$ gives $\frac{pS_{m+2}}{S_{m+3}}x_1^{(m-1)} + \frac{qS_{m+1}}{S_{m+3}}x_2^{(m-1)} < \frac{qS_{m+1}}{S_{m+3}}x_1^{(m-1)} + \frac{pS_{m+2}}{S_{m+3}}x_2^{(m-1)}$

$$\frac{qS_{m+1}}{S_{m+3}}(x_2^{(m-1)} - x_1^{(m-1)}) < \frac{pS_{m+2}}{S_{m+3}}(x_2^{(m-1)} - x_1^{(m-1)})$$

$$qS_{m+1} < pS_{m+2}$$

Obviously, $qS_n < pS_{n+1}$ for all $n \geq 2$. Thus, we complete the proof.

(b) From (a), $qS_n < pS_{n+1}$ gives $0 < \frac{qS_n}{S_{n+2}} < \frac{pS_{n+1}}{S_{n+2}}$. Now, since $\frac{pS_{n+1}}{S_{n+2}} + \frac{qS_n}{S_{n+2}} = 1$,

therefore $\frac{qS_n}{S_{n+2}} < \frac{1}{2}$ and $\frac{pS_{n+1}}{S_{n+2}} > \frac{1}{2}$. Also, $S_{n+2} = pS_{n+1} + qS_n$ gives $S_{n+2} > pS_{n+1}$, i.e.,

$\frac{pS_{n+1}}{S_{n+2}} < 1$. Hence, the proof is completed.

(c) Since $qS_2 < pS_3$ gives $q(q-p^2)a < p^3b$ and $qS_3 < pS_4$ gives $q(q-p^2)b < p^3(qa+pb)$, thus, Lemma 2 gives $qS_i < pS_{i+1}$ for all $i \geq 4$. Next, Theorem 1(a) completes the proof.

Theorem 2. Let $S_n(a, b, p, q)$ denote the n th sub-Fibonacci number for all $n \geq 0$. An optimal x^* of a unimodal function $f(x)$ is achievable by generalized Fibonacci search for all $n \geq 2$ if either one of the following is true.

- (a) $q < p^2$
- (b) $q = p^2$ and $b > 0$.

Proof. Consider an n th sub-Fibonacci number, $S_n(a, b, p, q)$ such that $a + b > 0$ and $p, q > 0$.

- (a) Notice that $S_0 = a \geq 0$ and $S_1 = b \geq 0$, where a and b never take zero simultaneously. If $q < p^2$, then

$$qS_1 - pS_2 = qS_1 - p(pS_1 + qS_0) = (q - p^2)b - pqa < 0 \tag{12}$$

$$\text{And } qS_2 - pS_3 = qS_2 - p(pS_2 + qS_1) = q(q - p^2)a - p^3b < 0 \tag{13}$$

Lemma 2 leads Equation 12 and 13 to generate $qS_i < pS_{i+1}$ for all $i \geq 3$. Next, Theorem 1(a) completes the proof.

- (b) Since $p^2 = q$, $a \geq 0$ and $b > 0$, therefore

$$qS_2 - pS_3 = qS_2 - p(pS_2 + qS_1) = q(q - p^2)a - p^3b = -p^3b < 0 \tag{14}$$

and

$$qS_3 - pS_4 = qS_3 - p(pS_3 + qS_2) = [q(q - p^2) - p^4]b - p^3qa = -p^3(qa + pb) < 0 \tag{15}$$

Lemma 2 leads Equation 14 and 15 to produce $qS_i < pS_{i+1}$ for all $i \geq 4$. Then, Theorem 1 (a) completes the proof.

RESULTS AND DISCUSSIONS

Table 1 shows the benchmark functions of two variables obtained from *Test Functions for Optimization* at the following link: https://en.wikipedia.org/wiki/Test_functions_for_optimization.

We converted the two-dimensional benchmark functions to a one-dimensional function by fixing the values of y in $f(x, y)$ with the optimal y -coordinate, y^* , or vice-versa the values of x in $f(x, y)$ with the optimal x^* . Then we searched the cross section along the

lines $y = y^*$ or $x = x^*$ using the line search method for one variable since $f(x, y^*)$ and $f(x^*, y)$ have only one degree of freedom.

We implemented the generalized Fibonacci search with $a = 2, b = 7, p = 2$ and $q = 7$, namely Fibo2727, which is involved the sub-Fibonacci numbers, $S_n(2, 7, 2, 7)$. Based on the same initial intervals of $[x_1, x_2]$ and compared to the golden section search and other Fibonacci-type search methods such as Fibonacci, Lucas, or Pell, we found that experimentally tested with a self-developed function in Microsoft Excel, Fibo2727 could give the least number of iterations to achieve the optimization. Pell is the weakest, as shown in Table 2, where the optimal value has rounded up to 7 decimal places.

Figure 9 shows the flow chart of Fibonacci-like numbers. Not all the Fibonacci-like numbers will fulfil the generalized Fibonacci search method. For instance, the Mersenne

Table 1
Benchmark functions of two variables

No.	Name of function	$f(x, y)$
1	Easom	$-e^{-[(x-\pi)^2-(y-\pi)^2]} \cos x \cos y$
2	Holder Table	$- \left e^{\left 1 - \frac{\sqrt{x^2+y^2}}{\pi} \right } \sin x \cos y \right $
3	Rosenbrock	$\sum_{i=1}^{n-1} [100(x_{i+1} - x_i^2)^2 - (1 - x_i)^2]$
4	Eggholder	$-(y+47) \sin \sqrt{\left \frac{x}{2} + (y+47) \right } - x \sin \sqrt{ x - (y+47) }$
5	Schaffer N.4	$0.5 + \frac{\cos^2 \left(\sin x^2 - y^2 \right) - 0.5}{[1 + 0.001(x^2 + y^2)]^2}$
6	Ackley	$-20e^{-0.2\sqrt{0.5(x^2+y^2)}} - e^{0.5(\cos 2\pi x + \cos 2\pi y)} + e + 20$
7	Rastrigin	$10n + \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i)]$
8	Bukin	$100\sqrt{ y - 0.01x^2 } + 0.01 x + 10 $
9	Cross-in-tray	$-\frac{1}{10000} \left[\left (\sin x \sin y) \exp \left(\left 100 - \frac{\sqrt{x^2 + y^2}}{\pi} \right \right) \right + 1 \right]^{0.1}$

Table 2

Comparison of the performance of golden section search and other Fibonacci-type search methods on benchmark functions

No.	Function	Initial interval (x_1, x_2)	Optimal (min)	Fibonacci-Type (a, b, p, q)				
				Golden Section	Fibonacci (0,1,1,1)	Lucas (2,1,1,1)	Pell (0,1,2,1)	Fibo2727 (2,7,2,7)
				Number of Iterations				
1	Easom at $y = \pi$	(2, 4)	$x^* =$ 3.1415927	35	35	35	88	26
2	Holder Table at $y = 9.66459$	(6, 9)	$x^* =$ 8.0550236	35	37	36	90	26
3	Rosenbrock at $y = 1, n = 2$	(0, 3)	$x^* = 1$	38	38	36	93	28
4	Eggholder at $x = 512$	(350, 450)	$x^* =$ 404.2318503	40	41	41	105	30
5	Schaffer N.4 at $x = 0$	(0, 3)	$x^* =$ 1.2531318	38	38	37	90	28
6	Ackley at $x = 0$	(-2, 1.5)	$y^* = 0$	38	38	38	83	28
7	Rastrigin at $x = 0, n = 2$	(-1, 3)	$y^* = 0$	39	38	39	97	29
8	Bukin at $y = 1$	(-20, 0)	$x^* = -10$	42	42	42	104	31
9	Cross-in-tray at $y = 1.3494064$	(0, 2)	$x^* =$ 1.3494064	34	36	34	87	24

(Catarino et al., 2016), arithmetic or geometric numbers (Chong & Ho, 2015) failed in generalized Fibonacci search method since they were not the members of sub-Fibonacci number as their q 's values were negative. However, although the Jacobsthal number is a sub-Fibonacci number (Catarino et al., 2016), but it also failed as well as it does not fulfil Theorem 1.

The results significantly show that the algorithm of both the Fibonacci search and generalized Fibonacci search methods can be expressed with the initial ratio of two successive lower Fibonacci numbers in the first iteration. This will lead us to investigate the combination of the values of a, b, p, q to further improve our search algorithm in future. We will also attempt to extend the generalized Fibonacci search method to two-dimensional or higher dimensional optimization problems of unimodal functions.

CONCLUSIONS

In this paper, we found another approach to obtain the ratio of the golden section search, which led us to an idea of the "ratio length of 1". This idea opened up a big step for us to explore deeper research in Fibonacci search, because $F_{n+2} = F_{n+1} + F_n$ always give

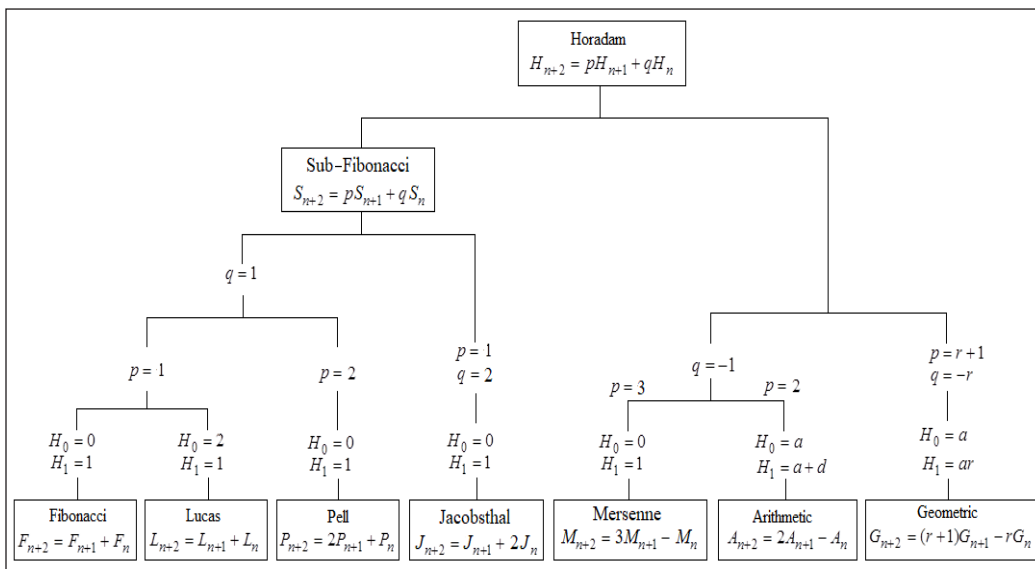


Figure 9. Flow chart of Fibonacci-like numbers

$\frac{F_{n+1}}{F_{n+2}} + \frac{F_n}{F_{n+2}} = 1$. Additionally, we proved that both the algorithms of the Fibonacci and the generalized Fibonacci search could be iterated with the successive lower Fibonacci numbers. As such, from the generalized Fibonacci number $S_4 = pS_3 + qS_2$, where $\frac{pS_3}{S_4} + \frac{qS_2}{S_4} = 1$, we considered $\frac{qS_2}{S_4}$ and $\frac{pS_3}{S_4}$ as the initial ratios for the first iteration. Hence this would advantage us to run the algorithm without fixing the number of iterations.

Unfortunately, certain sub-Fibonacci numbers failed the algorithm of the generalized Fibonacci search. However, we found the necessary and sufficient conditions of a, b, p and q to ensure the convergence of the optimal solution.

In the last section of this paper, we compared the performance of the generalized Fibonacci search with the golden section and other line Fibonacci-type search methods such as Fibonacci, Lucas and Pell approaches using nine one-dimensional benchmark functions. Based on the same width of the initial interval for each search method, we used a self-developed function in Microsoft Excel to test a number of different combination values of a, b, p and q , the generalized Fibonacci search performed more effectively and successfully gave significant results, correct to 7 decimal places, when $a = 2, b = 7, p = 2$ and $q = 7$. In other words, the generalized Fibonacci search with $S_n(2, 7, 2, 7)$ required less iterations to achieve the optimal when compared to others.

We strongly believe that there is no best combination of a, b, p and q to achieve the optimum by the generalized Fibonacci search. We shall attempt to prove this statement in future.

ACKNOWLEDGEMENT

The authors wish to express their gratitude to the School of Mathematical Sciences, Sunway University for the financial support in the conduct of this research.

REFERENCES

- Avriel, M., & Wilde, D. J. (1966). Optimally proof for the symmetric Fibonacci search technique. *Fibonacci Quarterly Journal*, 265-269.
- Bazaraa, M. S., Sherali, H. D., & Shetty, C. M. (2013). *Nonlinear programming: Theory and algorithms*. John Wiley & Sons.
- Beamer, J. H., & Wilde, D. J. (1971). Minimax optimization of a unimodal function by variable block derivative search with time delay. *Journal of Combinatorial Theory, Series A*, 10(2), 160-173. [https://doi.org/10.1016/0097-3165\(71\)90019-7](https://doi.org/10.1016/0097-3165(71)90019-7)
- Belegundu, A. D., & Chandrupatla, T. R. (2019). *Optimization concepts and applications in engineering*. Cambridge University Press.
- Catarino, P., Campos, H., & Vasco, P. (2016). On the Mersenne sequence. *CM-Centro de Matemática*, 46, 37-53.
- Chong, C. Y., & Ho, C. K. (2015, December). Some Properties of generalized Fibonacci sequence. In *AIP Conference Proceedings* (Vol. 1691, No. 1, p. 040004). AIP Publishing LLC. <https://doi.org/10.1063/1.4937054>
- Demir, A., Omur, N., & Ulutas, Y. T. (2008). Optimization by k-Lucas numbers. *Applied mathematics and computation*, 197(1), 366-371. <https://doi.org/10.1016/j.amc.2007.07.045>
- Hassin, R. (1981). On maximizing functions by Fibonacci search. *Fibonacci Quarterly*, 19, 347-351.
- Horadam, A. F. (1965). Basic properties of a certain generalized sequence of numbers. *The Fibonacci Quarterly*, 3(3), 161-176.
- Horzum, T., & Kocer, E. G. (2009). On some properties of Horadam polynomials. *International Mathematical Forum*, 4(25), 1243-1252.
- Kiefer, J. (1953). Sequential minimax search for a maximum. *Proceedings of the American Mathematical Society*, 4(3), 502-506. <https://doi.org/10.2307/2032161>
- Kilic, E. (2007). The generalized order-k Fibonacci–Pell sequence by matrix methods. *Journal of Computational and Applied Mathematics*, 209(2), 133-145. <https://doi.org/10.1016/j.cam.2006.10.071>
- Koshy, T. (2019). *Fibonacci and Lucas numbers with applications* (2nd Ed.). John Wiley & Sons.
- Luenberger, D. G., & Ye, Y. (1984). *Linear and nonlinear programming* (Vol. 2). Addison-wesley.
- Mathews, J. H., & Fink, K. D. (2004). *Numerical methods using MATLAB* (Vol. 4). Pearson prentice hall.
- Oliver, L. T., & Wilde, D. J. (1964). Symmetric sequential minimax search for a maximum. *Fibonacci Quarterly*, 3, 169-176.
- Rao, S. S. (2019). *Engineering optimization: Theory and practice*. John Wiley & Sons.

- Solomon, J. (2015). *Numerical algorithms: Methods for computer vision, machine learning, and graphics*. CRC press.
- Subasi, M., Yildirim, N., & Yildiz, B. (2004). An improvement on Fibonacci search method in optimization theory. *Applied mathematics and computation*, 147(3), 893-901. [https://doi.org/10.1016/S0096-3003\(02\)00828-7](https://doi.org/10.1016/S0096-3003(02)00828-7)
- Udrea, G. (1996). A note on the sequence $(W_n)_{n \geq 0}$ of AF Horadam. *Portugaliae Mathematica*, 53(2), 143-156.
- Witzgall, C. (1969). *Fibonacci search with arbitrary first evaluation*. Boeing Scientific Research Labs Seattle Wa Mathematics Research Lab.



Mathematical Modeling and Availability Analysis of Leaf Spring Manufacturing Plant

Sohan Lal Tyagi*, Shikha Bansal, Priyanka Agarwal and Ajay Singh Yadav

Department of Mathematics, SRM Institute of Science and Technology, Delhi-NCR Campus, Modinagar, 201204, India

ABSTRACT

In this manuscript, we portray the composition and availability analysis of leaf spring. In wheeled vehicles for the suspension, an elementary form of spring i.e. leaf spring is usually utilized. Particularly in industrial vehicles, Leaf springs are one of the widely recognized suspension segments they are frequently used. A method is presented that tests the availability of each part involved in the system as well as the system as a whole. Besides, the analysis demonstrates an optimization model that helps to boost the availability of the leaf spring production plant system. The system is divided into different subsystems considering various phases in the production of leaf spring. The framework of this system consists of four components i.e. shearing, punching, heating, and assembling. Using the Markov birth-death strategy we established the mathematical model of the plant. The matrix method is applied to simplify the differential equations and C programming survey the fluctuation of availability related to time. The numerical results for the different framework are given which are effective to enhance the maintenance policy of the system.

Keywords: Availability, failure rate, leaf spring, matrix method

ARTICLE INFO

Article history:

Received: 12 November 2020

Accepted: 25 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.18>

E-mail addresses:

drsohantya@gmail.com (Sohan Lal Tyagi)

srbansal2008@gmail.com (Shikha Bansal)

priyanka1354@gmail.com (Priyanka Agarwal)

ajay29011984@gmail.com (Ajay Singh Yadav)

* Corresponding author

INTRODUCTION

The present organizations require the potential to tackle the latest proficiency that is accessible with the present innovation. In the era of competition among the manufacturing organizations, the appeal of quality production inside the expressed time span increases for assuring the requirement of patrons. The availability of production systems plays a significant role

in accomplishing this necessity. The production system's availability can be improved by maximizing the efficiency of its sub-systems, which can be possible only through proper maintenance of the systems.

Kumar et al. (1989) used the probabilistic approach in the paper industry for the availability of the washing system. Singh and Mahajan (1999) examined the long-run availability and reliability of a utensils manufacturing plant. Garg et al. (2010) proposed a new technique in the plywood industry to find the reliability model of a block-board manufacturing system. Aggarwal et al. (2015) reported a mathematical model of the fertilizer plant of the urea synthesis system to find the mean time between failure and reliability analysis. Mangey and Manglik (2016) considered numerous aspects of failure and the reliability parameters of the dynamic repairable manufacturing system had been assessed.

In the field of reliability theory, mechanical systems have drawn the attention of researchers. Venkatesan and Devaraj (2012) defined a composite leaf spring concept and experimental study. Kumar and Reddy (2017) presented a finite element method to study the functionality of Leaf springs, design parameters that affected the suspension system on an automobile vehicle. So far, the leaf spring plant, which also plays an important role in our everyday lives, has not been addressed.

Jain et al. (2004) dealt with the degraded machine repair problem for the finite population Bernoulli feedback model. Yadav et al. (2008) developed a technique in the absence of data concerning failures. Agarwal and Bansal (2009) studied the head of line repair strategy to evaluate the reliability of a dynamic standby redundancy system with evolving environmental conditions. Shakuntala et al. (2011) applied a supplementary variable technique applied to compute the reliability of polytube manufacturing plant with variable rates of failure and repair. Mehta et al. (2017) emphasized on steel plant for High productivity. Kumar and Kadyan (2018) evaluated the profit and availability analysis of distillery plant through supplementary variable techniques. Bansal and Tyagi (2018) had applied the orthogonal matrix method to obtain reliability analysis of screw manufacturing plant. Gupta et al. (2019) introduced a neural network approach to predict different parameters of reliability. Jain et al. (2020) obtained the performance of redundant machining systems having imperfect fault detection and reboot delay.

Various techniques i.e., Runge-Kutta method, Laplace transform method, or Lagrange's method has been used to solve differential equations which include laborious estimations. Subsequently, it has been seen that in complex frameworks the computation of availability is extremely troublesome. So, we have built up the matrix method to illuminate differential equations and a computer program is created to ascertain the time-dependent availability. Table 1 to 7 for the estimation of time-dependent availability are prepared by C Programming. The findings of the analysis depict which subsystems and hardware

are critical from an availability point of view. Hence, weak spots of the industry could be distinguished and maintained.

MATERIAL AND METHODS

In this paper, we consider a leaf spring manufacturing plant. The leaf spring plant consists of four components and detailed descriptions are as follows:

Shearing Component

In the leaf spring manufacturing plant, the shearing section is the most essential usual and inevitable tactics. For the cutting of the metal sheet, shearing is used normally to produce the primary workpiece, referred to as a blank, and for other sheet metal processes. It employs a single edge machine, the failure of which causes the leaf plant to fully fail.

Punching Component

For punching, the workpiece is set out in the appropriate orientation. To keep all the plates jointly a ditch at the center of each leaf is maintained. For this, a punching operation is carried out. It employs two units. This component can work in reduced capacity with one unit.

Heating Component

The heating component is achieved most effectively at the ends to be smooth to bend the ends. That is performed to fix the leaves to the frame of the vehicle. It consists of two units with minor and major failures.

Assembly Component

To check the quality examination and trial are carried out in the assembly component of the leaf springs. After inspection, the leaf springs with precise excellent and overall performance are allowed to stock storage. This component can work in reduced capacity with one unit.

Assumptions

- (a) Originally all components are functioning.
- (b) The rates of repair and failure are constant and statistically independent.
- (c) Switch-over devices are perfect.
- (d) Performance-wise, a repaired unit is in good condition as new.
- (e) The capacity of the working units are similar to the standby units and are of the same nature

Notations

- U, V, W, X : Good states condition of the leaf plant
- \bar{V}, \bar{X} : Reduced capacity states
- u, v, w, x : Represents the respective failed states of the system
- η_i : Represents failure rates of i^{th} unit of the system where $i= u, v_1, v_2, w, x_1, x_2$
- γ_i : Represents repair, rates of i^{th} unit of the system where $i= u, v_1, v_2, w, x_1, x_2$
- $P_i(t)$: Probability of i^{th} state at time t
- $P_i'(t)$: Rate of change of probability function $P_i(t)$
- $Av(t)$: Availability of the system

The transition illustration of the leaf spring manufacturing plant is developed and shown in Figure 1.

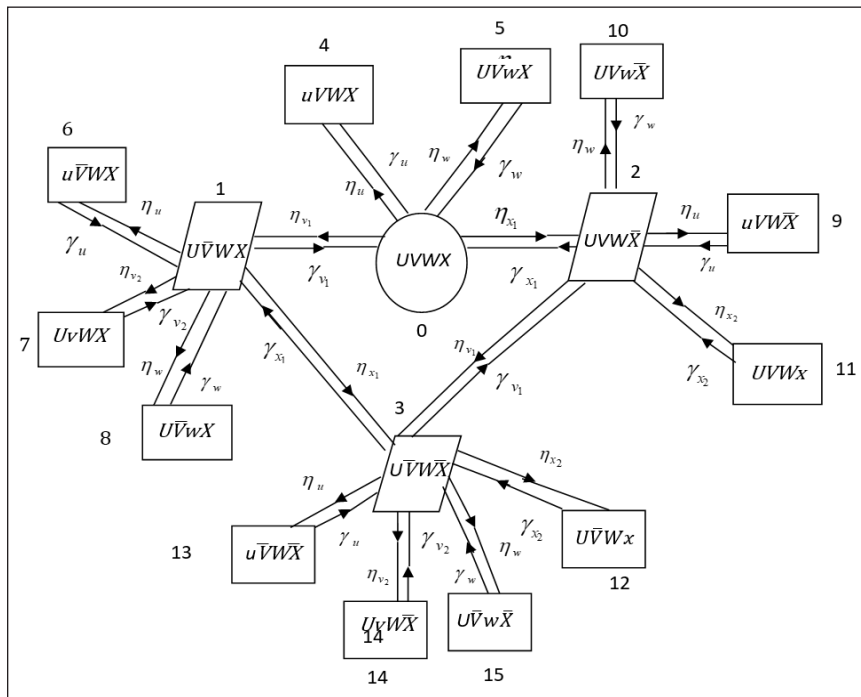


Figure 1. State transition illustration of leaf spring manufacturing plant

Mathematical Modeling

Using Markov birth-death process the following differential equations, correlated with the transition states are given by Figure 1.

$$P_0'(t) + (\eta_u + \eta_w + \eta_{x_1} + \eta_{v_1})P_0(t) = \gamma_u P_4(t) + \gamma_w P_5(t) + \gamma_{v_1} P_1(t) + \gamma_{x_1} P_2(t) \tag{1}$$

$$P_1'(t) + (\eta_u + \eta_w + \eta_{x_1} + \eta_{v_2} + \gamma_{v_1})P_1(t) = \gamma_{v_2}P_7(t) + \eta_{v_1}P_0(t) + \gamma_uP_6(t) + \gamma_wP_8(t) + \gamma_{x_1}P_3(t) \quad (2)$$

$$P_2'(t) + (\eta_u + \eta_w + \eta_{x_2} + \eta_{v_1} + \gamma_{x_1})P_2(t) = \eta_{x_1}P_0(t) + \gamma_uP_9(t) + \gamma_wP_{10}(t) + \gamma_{x_2}P_{11}(t) + \gamma_{v_1}P_3(t) \quad (3)$$

$$P_3'(t) + (\eta_u + \eta_w + \eta_{x_2} + \eta_{v_2} + \gamma_{x_1} + \gamma_{v_1})P_3(t) = \eta_{x_1}P_1(t) + \eta_{v_1}P_2(t) + \gamma_{x_2}P_{12}(t) + \gamma_uP_{13}(t) + \gamma_{v_2}P_{14}(t) + \gamma_wP_{15}(t) \quad (4)$$

$$P_i'(t) + \gamma_jP_i(t) = \mu_jP_k(t) \quad (5)$$

Where,

$i = 4, 5, j = u, w$ when $k=0$

$i = 6, 7, 8, j = u, v_2, w$ when $k=1$

$i = 9, 10, 11, j = u, w, x_2$ when $k=2$

$i = 12, 13, 14, 15, j = u, v_2, w, x_2$ when $k=3$

With initial conditions at $t = 0$,

$$P_i(t) = \begin{cases} 1, & \text{for } i=0, \\ 0, & \text{Otherwise} \end{cases} \quad (6)$$

Equation 1-6 are simplified using matrix method and C programming.

Since '16' is the number of all the possible transition states of the complex structure.

So, the system of differential equations can be written for the above equations as:

$$(\theta I - Q) \bar{P}_i(t) = O$$

Where $\theta = d/dt$, O is the null matrix, Q is the matrix of $P_i(t)$ coefficients in the equation of differential differences. $\bar{P}_i(t) = [P_1(t)P_2(t)P_3(t).....P_{16}(t)]^T$

The equations reduce to $P^{-1}(\theta I - E)\bar{P}_i(t) = O$, where P is the matrix such that $P^{-1}QP = E$, and $E = (e_1, e_2, \dots, e_n)$ is the matrix of Eigenvalues of the matrix Q .

The availability of the system is the sum of the availabilities of working subsystems.

$$Av(t) = P_0(t) + P_1(t) + P_2(t) + P_3(t) \\ = 1 + (a_{11} + a_{21} + a_{31} + a_{41} + \dots) t + (b_{11} + b_{21} + b_{31} + b_{41} + \dots) t^2 / 2! + \dots$$

The entries a_{ij} , and b_{ij} , are respective entries in Q , $R = Q\bar{P}_i(0)$, $S = QR = Q^2\bar{P}_i(0)$ so on.

$$\text{Matrix } Q = \begin{bmatrix} -a_1 & \gamma_{v_1} & \gamma_{x_1} & 0 & \gamma_u & \gamma_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \eta_{v_1} & -a_2 & 0 & \gamma_{x_1} & 0 & 0 & \gamma_u & \gamma_{v_2} & \gamma_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \eta_{x_1} & 0 & -a_3 & \gamma_{v_1} & 0 & 0 & 0 & 0 & 0 & \gamma_u & \gamma_w & \gamma_{x_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_{x_1} & \eta_{v_1} & -a_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \gamma_{x_2} & \gamma_u & \gamma_{v_2} & \gamma_w & 0 \\ \eta_u & 0 & 0 & 0 & -\gamma_u & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \eta_w & 0 & 0 & 0 & 0 & -\gamma_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_u & 0 & 0 & 0 & 0 & -\gamma_u & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_{v_2} & 0 & 0 & 0 & 0 & 0 & -\gamma_{v_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \eta_w & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta_u & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_u & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_w & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \eta_{x_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_{x_2} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta_{x_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_{x_2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta_u & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_u & 0 & 0 & 0 \\ 0 & 0 & 0 & \eta_{v_2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_{v_2} & 0 & 0 \\ 0 & 0 & 0 & \eta_w & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -\gamma_w & 0 \end{bmatrix}$$

Where

$$a_1 = \eta_u + \eta_w + \eta_{x_1} + \eta_{v_1}$$

$$a_2 = \eta_u + \eta_w + \eta_{x_1} + \eta_{v_2} + \gamma_{v_1}$$

$$a_3 = \eta_u + \eta_w + \eta_{x_2} + \eta_{v_1} + \gamma_{x_1}$$

$$a_4 = \eta_u + \eta_w + \eta_{x_2} + \eta_{v_2} + \gamma_{x_1} + \gamma_{v_1}$$

RESULTS AND DISCUSSION

Availability Analysis

For distinctive values of the failure and repair rates, the availability of the leaf plant is determined. The failure values and rates of repair are taken as constant:

$$\eta_u = 0.002, \eta_{v_1} = 0.001, \eta_{v_2} = 0.004, \gamma_u = 0.02, \gamma_{v_1} = 0.01, \gamma_{v_2} = 0.04,$$

$$\eta_w = 0.0025, \eta_{x_1} = 0.0025, \eta_{x_2} = 0.005, \gamma_w = 0.02, \gamma_{x_1} = 0.02, \gamma_{x_2} = 0.05$$

Table 1 is shown in the support of the above data. The analog graph is shown in Figure 2.

Numerical Analysis

With the help of the C program, we calculated the consequence of parameters on the system’s availability and display in Table 2 to 7 in the time horizon of 10 months to 50 months. From the Table 1, it has been noticed that as the failure rate rises, the availability of the system reduces.

Table 1
Availability corresponding to the time

Time	Availability	Time	Availability
0	1	40	0.880433
5	0.978644	45	0.871865
10	0.959463	50	0.864330
15	0.942279	55	0.857839
20	0.926918	60	0.852500
25	0.913207	65	0.848621
30	0.900986	70	0.846812
35	0.890105	75	0.848161

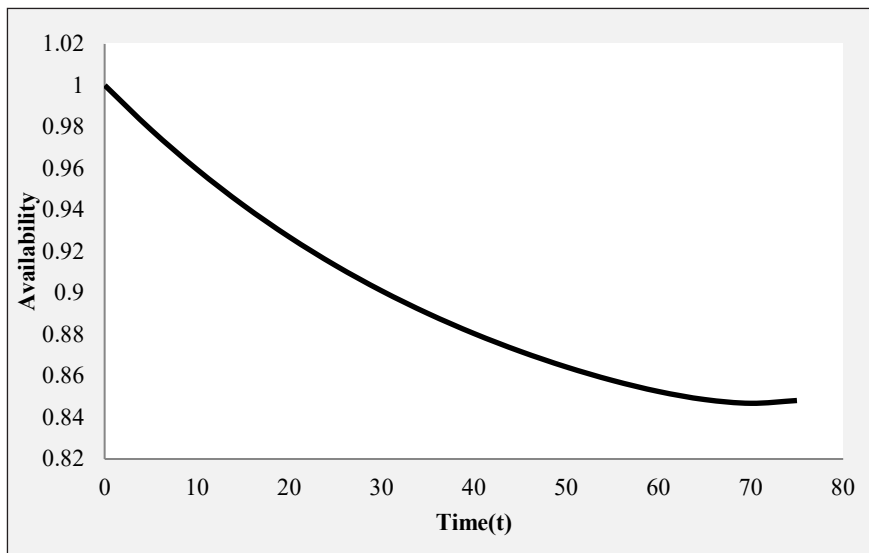


Figure 2. Time-dependent availability graph

Table 2
Impact of the shearing component failure rate on system availability

Months/ η_u	0.002	0.004	0.006	0.008
10	0.95946	0.94228	0.92543	0.90891
20	0.92691	0.89723	0.86863	0.84107
30	0.90098	0.86225	0.82554	0.70733
40	0.88043	0.83521	0.79298	0.75352
50	0.86433	0.81456	0.76873	0.72653

Table 3
Impact of the first punching component failure rate on system availability

Months/ η_{v_1}	0.001	0.002	0.003	0.004
10	0.95946	0.95930	0.95914	0.95898
20	0.92691	0.92639	0.92588	0.92538
30	0.90098	0.90001	0.89906	0.89813
40	0.88043	0.87898	0.87758	0.87621
50	0.86433	0.86244	0.86063	0.85890

Table 4
Impact of the second punching component failure rate on system availability

Months/ η_{v_2}	0.004	0.005	0.006	0.007
10	0.95946	0.95942	0.95938	0.95934
20	0.92691	0.92679	0.92666	0.92654
30	0.90098	0.90075	0.90052	0.90030
40	0.88043	0.88009	0.87976	0.87943
50	0.86433	0.86389	0.86347	0.86306

Table 5
Impact of the heating component failure rate on system availability

Months/ η_w	0.0025	0.0035	0.0045	0.0055
10	0.95946	0.95083	0.94228	0.93382
20	0.92691	0.91193	0.89723	0.88279
30	0.90098	0.88136	0.86225	0.84365
40	0.88043	0.85743	0.83521	0.81373
50	0.86433	0.83893	0.81456	0.79117

Table 6
Impact of the first assembling component failure rate on system availability

Months/ η_{x_1}	0.0025	0.0035	0.0045	0.0055
10	0.95946	0.95927	0.95909	0.95891
20	0.92691	0.92636	0.92582	0.92528
30	0.90098	0.90003	0.89910	0.89819
40	0.88043	0.87912	0.87785	0.87662
50	0.86433	0.86280	0.86133	0.85993

Table 7
Impact of the second assembling component failure rate on system availability

Months/ η_{x_2}	0.005	0.006	0.007	0.008
10	0.95946	0.95937	0.95928	0.95916
20	0.92691	0.92664	0.92637	0.92610
30	0.90098	0.90051	0.90005	0.89960
40	0.88043	0.87979	0.87916	0.87854
50	0.86433	0.86359	0.86288	0.86219

The result shown in Table 2 reveals the impact of the shearing component on the system's availability. By differing the failure rate from 0.002 to 0.008 and fixing other parameters, around 24.27% reduces the availability of the leaf plant in the time from 10 months to 50 months. Table 3 demonstrates the effect of η_{v_1} which decreases the availability about 10.48% at the same time while varying the failure rate from 0.001 to 0.004. Similarly, Table 4 shows that as the failure rate of η_{v_2} increases from 0.004 to 0.007, the availability also decreases 10.04% of the leaf plant. Table 5 presents the system's availability of heating component. By varying the failure rate from 0.0025 to 0.0055 and fixing other parameters around 17.53% reduces the availability. Table 6 and 7 reveals the impact of η_{x_1} and η_{x_2} on the system's availability, around 10.37% and 10.13% reduces the availability of the leaf plant in the time from 10 months to 50 months. By differing the failure rate from 0.0025 to 0.0055 and 0.005 to 0.008, respectively.

CONCLUSIONS

The target of this paper is to recognize the critical points of the system so that appropriate maintenance strategies can be applied. The approach proposed is simple to use with a large number of differential equations in the complex system and helps to measure the availability of the leaf spring manufacturing plant. The variation in availability concerning time is shown in Table 1 and Figure 2. Availability initially gradually decreases with respect to time and becomes almost steady after a long period of time. From Table 2 to 7, we can easily study the effects of failure rates of different subsystems on the availability of the complete system. A comparative report from Table 1 to 7 shows that the shearing component should pay more attention to improve the availability of the system because it influences the availability of the entire system than any other. The heating component also has a minor influence on system availability. The other subsystems almost have the same effect on the performance of the system. Such findings are extremely useful in enhancing the efficiency of the system by implementing effective maintenance policies.

ACKNOWLEDGEMENT

The reviewers were extremely helpful in improving the article, which the authors greatly appreciate.

REFERENCES

- Agarwal, S. C., & Bansal, S. (2009). Reliability analysis of a standby redundant complex system with changing environment under head of line repair discipline. *Bulletin of pure and applied sciences*, 28(1), 165-173.
- Aggarwal, A. K., Kumar, S., Singh, V., & Garg, T. K. (2015). Markov modeling and reliability analysis of urea synthesis system of a fertilizer plant. *Journal of Industrial Engineering International*, 11, 1-14. <https://doi.org/10.1007/s40092-014-0091-5>
- Bansal, S., & Tyagi, S. (2018). Reliability analysis of screw manufacturing plant using orthogonal matrix method. *Pertanika Journal of Science & Technology*, 26(4), 1789-1800.
- Garg, S., Singh, J., & Singh, D. V. (2010). Availability and maintenance scheduling of a repairable block-board manufacturing system. *International Journal of Reliability and Safety*, 4(1), 104-118. <https://doi.org/10.1504/IJRS.2010.029567>
- Gupta, R., Ekata, & Batra, C. M. (2019). Estimate reliability parameters in bio fuel plant using neural network architecture. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 8(11), 440-445. <https://doi.org/10.35940/ijitee.K1392.0981119>
- Jain, M., Kumar, R. M., & Kumar, P. (2020). Maintainability of redundant machining system with vacation, imperfect recovery and reboot delay. *Arabian Journal for Science and Engineering*, 45, 2145-2161. <https://doi.org/10.1007/s13369-019-04060-w>
- Jain, M., Sharma, G. C., & Baghel, K. P. S. (2004). N policy for M/G/1 machine repair model with mixed standby components, degraded failure and Bernoulli feedback. *International Journal of Engineering*, 17(3), 279-288.
- Kumar, D., Singh J., & Pandey, P. C. (1989). Availability of a washing system in the paper industry. *Microelectronics Reliability*, 29(5), 775-778. [https://doi.org/10.1016/0026-2714\(89\)90177-7](https://doi.org/10.1016/0026-2714(89)90177-7)
- Kumar, R., & Kadyan, M. S. (2018). Reliability modelling and study of failure mechanism of distillery plant using supplementary variable technique. *Life Cycle Reliability and Safety Engineering*, 7, 137-146. <https://doi.org/10.1007/s41872-018-0053-9>
- Kumar, S., & Reddy, B. C. (2017). Modelling and structural analysis of leaf spring using finite element method. *International Research Journal of Engineering and Technology*, 4(12), 1155-1161.
- Mangey, R., & Manglik, M. (2016). An analysis to multi-state manufacturing system with common cause failure and waiting repair strategy. *Cogent Engineering*, 3, 1-20. <https://doi.org/10.1080/23311916.2016.1266185>
- Mehta, M., Singh, J., & Singh, M. (2017). Reliability and availability evaluation of a series-parallel system subject to random failure. *Indian Journal of Science and Technology*, 10(31), 1-11. <https://doi.org/10.17485/ijst/2017/v10i31/116145>

- Shakuntla, S., Lal, A. K., Bhatia, S. S., & Singh, J. (2011). Reliability analysis of polytube industry using supplementary variable technique. *Applied Mathematics and Computation*, 218(8), 3981-3992. <https://doi.org/10.1016/j.amc.2011.10.016>
- Singh, J., & Mahajan, P. (1999). Reliability of utensils manufacturing plant - A case study. *Opsearch*, 36(3), 260-269. <https://doi.org/10.1007/BF03398580>
- Venkatesan, M., & Devaraj, H. D. (2012). Design and analysis of composite leaf spring in light vehicle. *International Journal of Modern Engineering Research*, 2(1), 213-218.
- Yadav, O P., Choudhary, N., & Bilen, N. (2008). Complex system reliability estimation methodology in the absence of failure data. *Quality and Reliability Engineering International*, 24, 745-764. <https://doi.org/10.1002/qre.920>



Penalized LAD-SCAD Estimator Based on Robust Wrapped Correlation Screening Method for High Dimensional Models

Ishaq Abdullahi Baba^{1,3}, Habshah Midi^{1,2*}, Leong Wah June^{1,2} and Gafurjan Ibragimov^{1,2}

¹ Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

² Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor

³ Department of Mathematical Sciences, Faculty of Science, Taraba State University Jalingo, Taraba State Nigeria

ABSTRACT

The widely used least absolute deviation (LAD) estimator with the smoothly clipped absolute deviation (SCAD) penalty function (abbreviated as LAD-SCAD) is known to produce corrupt estimates in the presence of outlying observations. The problem becomes more complicated when the number of predictors diverges. To overcome these problems, the LAD-SCAD based on sure independence screening (SIS) technique is put forward. The SIS method uses the rank correlation screening (RCS) algorithm in the pre-screening step and the traditional Pathwise coordinate descent algorithm for computing the sequence of the regularization parameters in the post screening step for onward model selection. It is now evident that the rank correlation is less robust against outliers. Motivated by these inadequacies, we propose to improvise the LAD-SCAD estimator using robust wrapped

correlation screening (WCS) method by replacing the rank correlation in the SIS method with robust wrapped correlation. The proposed estimator is denoted as WCS+LAD-SCAD and will be employed for variable selection. The simulation study and real-life data examples show that the proposed procedure produces more efficient results compared to the existing methods.

ARTICLE INFO

Article history:

Received: 19 August 2020

Accepted: 6 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.19>

E-mail addresses:

ishaqbaba@yahoo.com; ishaqabdullahibaba@gmail.com (Ishaq Abdullahi Baba)

habshah@upm.edu.my (Habshah Midi)

leongwj@upm.edu.my (Leong Wah June)

ibragimov@gmail.com (Gafurjan Ibragimov)

*Corresponding author

Keywords: LAD-SCAD estimators, robust screening, ultrahigh dimensional data, variable selection

INTRODUCTION

The variable selection has become the key ingredient in building reliable and reproducible prediction models and hence, is fundamental to scientific studies in areas such as text processing, gene expression analysis, epidemiology, and combinatorial chemistry, among others. The main objective of variable selection is to find a subset of the predictor variables with the highest predictive power and better interpretability. In practice, most of the real-life dataset contains a huge number of the predictor variables, though some may not be significant to the target response. As a result, the accuracy of the model selection poses a major challenge due to the fact that redundant variables may affect the selection frequency of the authentic predictor variables (George, 2000; Heinze et al., 2018; Uraibi et al., 2017). Stepwise selection procedures are often used to find the most relevant variables which influence the value of the dependent variable. However, this method is known to produce inconsistent and biased estimates in addition to poor prediction and therefore, it is considered impractical for variables selection (Whittingham et al., 2006; Desboulets, 2018). In line with this, penalized least square estimators such as the least absolute shrinkage and selection operator are developed to remedy these shortcomings (Tibshirani, 1996). The attractive features of Lasso are that they can simultaneously perform both estimation and variable selection, and they can also be applied to the high dimensional dataset. Furthermore, Leng et al. (2006) and Meinshausen and Bühlmann (2006) exemplified the inconsistency of the Lasso condition. Fan and Li (2001) and Fan and Peng (2004) noted that the Lasso did not enjoy the oracle property, which was the ability to correctly estimate the insignificant coefficient with probability converging to one. This led to the development of the Adaptive Lasso estimator which had been proven to enjoy the oracle properties under regularity conditions (Zou, 2006). It is noted that the two objective functions of Lasso and Adaptive Lasso are convex; therefore, they do not achieve the closed form numerical solutions. Conversely, smoothly clipped absolute deviation (SCAD) penalized least squared objective function is concave (Xie & Huang, 2009) and note that SCAD penalized estimators can achieve sparse estimates and unbiased solution for the large coefficients. There are several interesting penalized estimators in the literature which include the bridge regression (Frank & Friedman, 1993), Dantzig selector (Candes & Tao, 2007), and the Elastic Net (Zou & Hastie, 2005). The difficulty of using penalized methods becomes obvious when the dimensionality is ultrahigh. To address this problem, Fan and Lv (2008) introduced the concept of sure independence screening (SIS) to reduce dimensionality from ultrahigh scale to moderate which was below the sample size and then selected the most significant variables into the linear model. As mentioned in Fan and Song (2010), it is explicit that SIS is computed based on the ordinary least squares (OLS) and is heavily dependent on the joint normality assumption between predictors and the response. Consequently, compared with Lasso and smoothly clipped absolute deviation

(SCAD) method, SIS proves to produce accurate prediction and interpretable model. Li et al. (2012) and Fan et al. (2009) presented feature screening-based distance correlations. The extended ultrahigh dimensional sure independence screening for the generalized linear model was discussed in Saldana and Feng (2018) and Ahmed and Bajwa (2019). However, since the preceding studies created a synergy between the Lasso type estimators and the sure screening methods for ultrahigh dimensional data, the Lasso type based sure screening methods had been studied by Ghaoui et al. (2010), Tibshirani et al. (2012), Xiang and Ramadge (2012) as cited in Ahmed and Bajwa (2019).

All the aforesaid estimators are known to be inefficient when the errors come from a heavy tailed distribution and/or in the presence of outliers in the response, as all of them technically utilize either the OLS loss function and/or the classical correlation learning algorithm. To address this issue, Wang et al. (2007) and Wu and Liu (2009) had put forward the least absolute deviation lasso (LAD lasso) and the LAD-SCAD estimators, respectively. Fan and Li (2001) observed that the penalized SCAD satisfied oracle only when the numbers of parameters were finite. Fan et al. (2009) noted that using a combination of the penalized estimators with the SIS procedure substantially improved the performance of the penalized estimators. The SIS is computed based on the Pearson correlation between the dependent and independent variables; however, it is known to be sensitive to outliers or heavy tailed errors (Li et al., 2011; Li et al., 2012). However, the authors suggested the use of rank correlation screening (RCS) method based on Kendall instead of the Pearson correlation in the pre-screening step of the SIS method. Li et al. (2011) combined the RCS with the penalized LAD-SCAD to achieve robust variable selection and parameter estimation when the number of predictor variables diverged.

Li et al. (2011) and Li et al. (2012) innovative approach and outstanding results are challenged at both the pre-screening and post screening steps as their screening algorithms and computation of the regularization parameter are based on Kendall rank correlation and the traditional path descent algorithm, which are both not robust against outlying observations (Wang et al., 2015). As an alternative to the existing correlation learning algorithm, Raymaekers & Rousseeuw (2019) advocated a robust and fast wrapped correlation algorithm which was based on the concept of g-product moment transformation. A comparison between this approach and other correlation algorithms such as the Kendall tau, Gnanadesikan Kettenring (GK) and scale estimator, among others, had clearly shown that the wrapped correlation algorithm was more robust against outliers than other existing methods (Raymaekers & Rousseeuw, 2019). Inspired by these, we proposed incorporating the wrapped correlation learning algorithm in the LAD-SCAD method to serve as a screening algorithm to reduce dimensionality from high to below sample size. The proposed method is expected to be more efficient than the existing methods in this study.

MATERIALS AND METHODS

Consider the linear regression Equation 1

$$y_i = X_i^T \beta_0 + e_i, \quad (1)$$

where β_0 is the vector of the regression parameters, X_i represents p_n the dimensional vectors of the predictors, y_i is the vector of the response variable, and e_i is independent identically distributed vector of the random errors with mean 0 and the constant variance σ^2 . Under these conditions, the classical linear regression estimators are applied. In the presence of anomalous observations in a dataset, and/or when the distribution of the errors is not normal, the robust regression loss function is applied to shrink the effect of outlying observations on the estimates of the regression (Bai & Wu, 1997; Maronna et al., 2019). In the application of this concept, the relationship between the vector of the response variable y_i and the set of predictors $X_i = (x_1, \dots, x_p)$ can be modelled by minimizing the following objective function (Equation 2):

$$\sum_{i=1}^n \rho(y_i - X_i^T \beta_0), \quad (2)$$

where ρ is a continuous symmetric function called the objective function with unique minimum at 0 (Rousseeuw & Leroy, 1987). Various choices of ρ functions have been suggested in the robust works (Bai & Wu, 1997; Stuart, 2011) for linear regression; for example, the quantile regression with $\rho_\alpha(x) = \alpha x^+ + (1 - \alpha)(-x)^+$, $0 < \alpha < 1$, where $x^+ = \max(x, 0)$ and the L_q regression estimates with $\rho(x) = |x|^q$, $1 \leq q \leq 2$. If $q = 1$, the minimizer in Equation 2 is generally named the least absolute deviation (LAD) estimator. Conversely, for sparse regression model, namely $\beta_0 = (\beta_a^T, \beta_b^T)^T$ where β_a is a $k_n \times 1$ vector of the significant parameters ($\beta_a \neq 0$) and β_b is an $m_n \times 1$ vector of the insignificant parameters ($\beta_b = 0$) such that $k_n + m_n = p_n$, we define correctly fitted model as a model which has k_n significant and m_n insignificant coefficients. Throughout this paper we used λ_n instead of λ to accentuate the dependency of λ on the sample size n (Fan & Li, 2001). Following Wang et al. (2015) the Huber penalized loss function can be expressed as Equation 3:

$$\sum_{i=1}^n \rho(y_i - X_i^T \beta_n) + n \sum_{j=1}^{p_n} p_\lambda(|\beta_{nj}|), \quad (3)$$

where p_n is the dimension of β_n , $p_\lambda(\cdot)$ is the SCAD penalty function which depend on the regularization parameter λ as defined in Fan and Li (2001) and Wang et al. (2015). There are various versions of regularization functions in the literature, but the Lasso and SCAD penalties have been cited to be more efficient. Comparison of the two functions by different scholars (Fan & Li, 2001; Xie & Huang, 2009) has shown that SCAD penalty enjoys all the three desirable properties of good penalty function: continuity, sparsity, and biasness while the Lasso penalty generates estimation bias but enjoys sparsity and continuity properties (Fan & Li, 2001; Li et al., 2011). Motivated by these properties,

Huang & Xie (2007) and Li et al. (2011) proposed the penalized least squares SCAD and SCAD penalized M estimators for estimation and variable selection. More recently, Wang et al. (2015) suggested the penalized LAD-SCAD estimator with divergence number of predictors. By following Wang et al. (2015), Equation 3 can be written as Equation 4:

$$\sum_{i=1}^n |y_i - X_i^T \beta_n| + n \sum_{j=1}^{p_n} p'_{\lambda_n}(|\beta_{nj}^0|) |\beta_{nj}|, \tag{4}$$

Where $p'_{\lambda_n}(\cdot)$ presents a vector of $p_n \times 1$ dimension whose j^{th} elements are the first derivative of the SCAD penalty function defined by Equation 5:

$$p'_{\lambda_n}(\beta_{nj}^0, a) = \begin{cases} \text{sgn}(\beta_{nj}^0)\lambda, & |\beta_{nj}^0| \leq \lambda, \\ \text{sgn}(\beta_{nj}^0)(a\lambda - |\beta_{nj}^0|)/(a-1), & \lambda < |\beta_{nj}^0| \leq a\lambda, \\ 0, & |\beta_{nj}^0| > a\lambda. \end{cases} \tag{5}$$

and $\beta_{nj}^0 = (\beta_{n1}^0, \dots, \beta_{np_n}^0)^T$ represents a vector of an initial estimates, which is usually obtained by minimizing the unpenalized objective function in Equation 2 when $q = 1$, and as a result, Equation 4 followed. Hence, the objective function in Equation 4 is continuous and differentiable which can achieve global minimum. To simplify the selection of regularization parameter, Fan and Li (2001) recommended using $a = 3.7$ and $\lambda > 0$ as the tuning parameter in Equation 5 for the SCAD regularization function. The minimization function in Equation 4 is called penalized LAD-SCAD estimator.

RESULTS AND DISCUSSION

Asymptotic Properties of the LAD-SCAD Estimator

In this section, we discuss some important asymptotic properties of the LAD-SCAD estimator when the number of predictor variables diverges with increasing number of samples. Pursuing Wang et al. (2015), the following definitions and assumptions are adapted:

Assumption 1. The errors are continuous and has a positive density at origin with median 0.

Assumption 2. There exists a positive fixed value $M < \infty : \max_{1 \leq i \leq n, 1 \leq j \leq p_n} |X_{ij}| \leq M$.

Assumption 3. $p_n^3/n \rightarrow 0$ as $n \rightarrow \infty$

Assumption 4. There exists fixed values $0 < \rho_1 < \rho_2 < \infty$ and $0 < \tau_1 < \tau_2 < \infty : \rho_1 \leq \rho_{n1} \leq \rho_{n2} \leq \rho_2$ and $\tau_1 \leq \tau_{n1} \leq \tau_{n2} \leq \tau_2$.

Assumption 5. $\sqrt{n/p_n} \lambda_n \rightarrow \infty$ and $\lambda_n \rightarrow 0$

The preceding notations and assumptions have been used in the literature to study the asymptotic normality, consistency, and sparsity properties of the penalized LAD-SCAD estimators (Li et al., 2011; Wang et al., 2015). To establish asymptotic properties of the LAD-SCAD, the following important notations and definitions are required.

Partitioning the $p_n \times 1$ vector of parameters β_n as $\beta_n = (\beta_{1n}^T, \beta_{2n}^T)^T$ in the same way as β_0 , where $\beta_{1n} = (\beta_1, \dots, \beta_{p_0})^T$ represents the vector corresponding to all significant coefficients and $\beta_{2n} = (\beta_{p_0+1}, \dots, \beta_p)^T$ represents a vector corresponding to all the insignificant coefficients. Let $\hat{\beta}_S = (\hat{\beta}_{S1n}^T, \hat{\beta}_{S2n}^T)$ be the corresponding LAD-SCAD estimator. We also divided the predictors x_i for $i = 1, \dots, n$ as $x_i = (x_{i1n}^T, x_{i2n}^T)$ with $x_{i1n} = (x_{i1}, \dots, x_{ip_0})^T$ and $x_{i2n} = (x_{ip_0+1}, \dots, x_{ip})^T$. Furthermore, define $a_{1n} = \max\{p'_{\lambda_n}(\beta_{nj}^0) : 1 \leq j \leq p_0\}$ and $b_{2n} = \min\{p'_{\lambda_n}(\beta_{nj}^0) : p_0 < j \leq p\}$, where $p'_{\lambda_n}(\cdot)$ is a function of sample size n as mentioned in (Fan & Peng, 2004; Li et al., 2011). Based on the preceding assumptions and notations, the following theorems can be established for the modified penalized LAD-SCAD estimator with divergence number of predictors.

Theorem 1. (Consistency). Assuming that (y_i, x_i) , $i = 1, \dots, n$ are i.i.d. and the conditions given in Assumption 1-5 are satisfied, then there exists a penalized LAD-SCAD estimator $\hat{\beta}_n : \|\hat{\beta}_n - \beta_0\| = Op(\sqrt{p_n/n})$

Theorem 2. (Oracle property). Assume that $k_n^3 p_n^3/n \rightarrow 0, E_n(r(\hat{\theta}_n)) = op(1/\sqrt{p_n})$, and the assumptions 1-5 are satisfied, then the penalized LAD-SCAD estimator $\hat{\beta}_n = (\hat{\beta}_{1n}^T, \hat{\beta}_{2n}^T)^T$ satisfies the following:

- i. (Sparsity) $\hat{\beta}_{2n} = 0$ with $\Pr(\hat{\beta}_{2n} = 0) \rightarrow 1$ as $n \rightarrow \infty$.
- ii. (Asymptotic normality)

$$n^{1/2} \alpha^T \Sigma_{n,11}^{-1/2} (\hat{\beta}_{1n} - \beta_{10}) = \frac{-1}{2f(0)\sqrt{n}} \alpha^T \Sigma_{n,11}^{-1/2} \sum_{i=1}^n (I(e_i < 0) - I(e_i > 0)) w_i \xrightarrow{D} N(0, 0.25(f(0))^{-2}),$$

Here α is a random $k_n \times 1$ vector with $\|\alpha\| = 1$, and the notation \xrightarrow{D} means the convergence in distribution. The proofs of the foregoing theorems follow from the adaptation of the proofs in Wang et al. (2015).

Wrapping Sure Independent Screening for Ultrahigh Dimensional Data

In the preceding section, we discussed the desirable properties of the penalized LAD-SCAD estimator when the number of predictors is less than the number observations. In the present section, we will give special attention on the ultrahigh dimensional scenario which is $p_n > n$. Fan and Lv (2008) noted that the traditional penalized selection procedure like Lasso or SCAD or Dantzing selector tended to misbehave when the number of the predictor variables P_n was too large compare to the number of observations. This had motivated Fan et al. (2009) and Wang et al. (2015) to combine the penalized and the SIS procedures to have a two steps procedure called penalized sure screening for high dimensional linear regression models. In the first step, SIS is implemented to reduce the dimension from the ultrahigh to below sample size and then used a suitable penalized estimator to accomplish the final variable selection and parameters estimation concurrently. The SIS is computed according to the Pearson correlation between the dependent variable and the predictors.

Li et al. (2012) noted that Pearson correlation was affected by outliers and or heavy tailed errors. This led to the procedure of LAD-SCAD given in Wang et al. (2015) where it employed rank correlation screening method based on rank correlation, denoted as RCS+LAD-SCAD. The appealing features of the combined procedure includes: (i) It is easy to apply because the computational burden for huge or large scale problems can easily be handled by any well-known correlation algorithm for example Pearson correlation and (ii) it enjoys the oracle properties (Fan & Li, 2001). Despite these striking properties, the existing combined procedure that is RCS+LAD-SCAD still have two shortfalls: (i) it uses rank based correlation which may or may not be preferable in a given application as they measure monotonicity instead of the linear relationship in addition to being too sensitive to multiple outliers (Croux & Dehhoon, 2010; Raymaekers & Rousseeuw, 2019) (ii) the RCS combined with LAD-SCAD estimator, does not take into account the effect of outlier in the computation of the sequence of regularization parameters which is considered as the key determinant factor in selecting predictable and interpretable model (Zhang et al., 2010; Shevlykov & Simironov, 2011). Recently, Raymaekers and Rousseeuw (2019) had shown that the rank correlation based on Kendall was liable in the presence of contamination and or heavy tailed errors. Comparisons between the robust and the non-robust correlation methods can be found in Shevlykov and Simironov (2011). The latter works motivate us to improvise the RCS+LAD-SCAD by employing the wrapped correlation algorithm to achieve robust sure independent screening and compute the set of the regularization parameters for model selection for onward model selection and prediction. We call this method WCS+LAD-SCAD estimator. Following Fan and Lv (2008), let $S_* = \{1 \leq p : \beta_j \neq 0\}$ be the true model parameter values with non-sparsity size $s \ll n$. Under the assumption of sparsity, the other $p - s$ variables can also be associated with the response variable by linkage to the predictor variables that are enclosed in the model. Let $w = (w_1, \dots, w_p)$ be a $p_n \times 1$ vector computed based on the component wise regression that is Equation 6.

$$w = X^T y \tag{6}$$

For any given $\gamma \in (0,1)$, the component wise magnitude of the vector w is sorted in decreasing order and a submodel can be defined as:

$S_\gamma = \{1 \leq j \leq p : |w_j| \text{ is among the first } [\mathcal{m}] \text{ largest of all}\}$, where $[\mathcal{m}]$ represents the integer part of \mathcal{m} . This is the commonly used procedure to reduce the full model from high down to a submodel S_γ with a size $d_n = [\mathcal{m}] < n$. In similar style, Wang et al. (2015), replaced Equation 6 with Kendall correlation learning algorithm while maintaining the same selection criteria to better estimates since the latter is proved to be non-robust estimator. The Kendall correlation values can be obtained using the Equation 7:

$$w_k = \frac{1}{n(n-1)} \sum I(x_{ij} < x_{ij})I(y_i < y_i) - \frac{1}{4}, \quad j = 1, \dots, p_n \tag{7}$$

Using the same approach as in Fan and Lv (1996) and Wang et al. (2015), but with fast and more robust correlation estimator built upon the wrapped correlation algorithm of Raymaekers and Rousseeuw (2019), we will replace the correlation formula in Equation 6 and 7 with the wrapped correlation algorithm introduced by Raymaekers and Rousseeuw (2019) while maintaining the selection procedure in Fan and Lv (1996) and Wang et al. (2015). The wrapping correlation procedure can be defined as follows: let the correlation values of the wrapped variables be defined by the entries as Equation 8.

$$w_j = cor(X^*, y^*) = cor\left(\psi_{b,c}\left(\frac{x_{ij} - \hat{\mu}_{xj}}{\hat{\sigma}_{xj}}\right), \left(\frac{y_i - \hat{\mu}_{yj}}{\hat{\sigma}_y}\right)\right) \tag{8}$$

where X^* and y^* are the transformed variables, $\hat{\mu}_{xj}$ and $\hat{\mu}_y$ are the estimates of location computed based on the one step M estimator of location with the wrapping function $\psi_{b,c}$, with $b = 1.5$ and $c = 4$, $\hat{\sigma}_{xj}$ and $\hat{\sigma}_y$ are computed using the MAD estimator (Raymaekers & Rousseeuw, 2019). We will present the broader steps involved for the implementation of the proposed WCS+LAD-LASSO in the next section.

Computational Procedures

The LAD-SCAD estimator can be computed easily by augmenting a dataset and using any suitable existing software; for example, Matlab or Phytion or R for quantile regression. Define an augmented dataset $\{(\tilde{y}_i, \tilde{X}_i), i = 1, \dots, n, n+1, \dots, n+p_n\}$ where $\tilde{y}_i = y_i/n, \tilde{X}_i = X_i/n$ for $i = 1, \dots, n$ and $\tilde{y}_i = 0, \tilde{X}_i = p'_{\lambda_n}(\beta_{nj}^0)\tilde{e}_i$ for $i = n+1, \dots, n+p_n$ and \tilde{e}_i is a $p_n \times 1$ dimensional vector with i^{th} term component equal to 1 such that Equation 4 can be expressed as Equation 9.

$$\sum_{i=1}^n |y_i - X_i^T \beta_n| + n \sum_{j=1}^{p_n} p'_{\lambda_n}(\beta_{nj}^0) |\beta_{nj}| = \sum_{i=1}^{n+p_n} |\tilde{y}_i - \tilde{X}_i^T \beta_n| \tag{9}$$

Equation 9 is equivalent to the traditional LAD objective function for the dataset $(\tilde{y}_i, \tilde{X}_i)$ for $i = 1, 2, \dots, n, n+1, \dots, n+p_n$. In this paper, an R software package `quantreg` is used for solving Equation 9. Furthermore, in the robust Lasso type regression, the selection of appropriate regularization parameter controls the complexity and improves the prediction accuracy of the selected model (Wang & Zhu, 2011; Friedman et al., 2010; Gao & Huang, 2010). Although there are several methods for selecting the best regularization parameter in the literature, the issue of robustness and computational efficiency desire special attention as well. Friedman et al. (2010) suggested a coordinate descent algorithm for computing the regularization parameters which is adapted in Wang et al. (2015). The computation of the sequence of K values of λ is a function of λ_{\min} and λ_{\max} where $\lambda_{\min} = \epsilon \lambda_{\max}$, $\lambda_{\max} = \max\{|X_i, y|\}/n, \epsilon = 0.001$ and $K = 100$ (Friedman et al., 2010). Following this concept, we replace $\{|X_i, y|\}$ with the wrapped correlation to robustly estimate the maximal correlation between the columns of X matrix and the vector y based on the idea of the product moment

transformation of the dataset as backed in Raymaekers and Rousseeuw (2019). Our reason for this is to reduce the effect of outliers before applying the selection criteria for onward model selection. This is followed by selecting a suitable model selection criterion such as the type considered in Wang et al. (2007) and Chang et al. (2018). In this paper, we consider the BIC type criteria used in Gao and Huang (2010) as cited in Wang et al. (2015). This is defined by Equation 10:

$$BIC = \log\left(\frac{1}{n} \sum_{i=1}^n |y_i - X_i^T \hat{\beta}_n|\right) + d_\lambda \frac{\log(n)}{n}, \tag{10}$$

where d_λ is the number of significant regression coefficients. The computation procedure for the proposed WCS+LAD-SCAD estimator can be summarized as follows:

Step 1. For a given dataset (X, y) where $X \in \mathfrak{R}^{n \times p}$ is the design matrix and $y \in \mathfrak{R}^n$ is the response variable and $d_n \in \mathbb{Z}_+$

1. Compute the robust initial scale $\hat{\sigma}_{xy}$ for $1, 2, \dots, p$ based on the median absolute deviation (MAD) for the matrix $X \in \mathfrak{R}^{n \times p}$ and $\hat{\sigma}_y$ for the vector of response $y_i \in \mathfrak{R}^{n \times p}$.
2. Compute a one-step M location estimator $\hat{\mu}_{xy}$ for the matrix $X \in \mathfrak{R}^{n \times p}$ and $\hat{\mu}_y$ for the vector of response $y_i \in \mathfrak{R}^{n \times p}$ with wrapping function $\psi_{b,c}$ where $b = 1.5, c = 4$
3. Let the variables X^* and y^* denote the transformed of the original variables X and y , then we can compute w_j for $j = 1, \dots, p$ based on Equation 8 and select the subset of predictor variables $d_n < n$ such that $|w_j|$ is among the first largest ones. Then apply the LAD-SCAD penalized estimator to Equation 1 by employing the following steps.

Step 2. Compute the sequence of regularization parameter λ defined on the interval $[\lambda_{\max}, \lambda_{\min}]$, where $\lambda_{\max} = \max\langle g_x(X) g_y(y) \rangle$ and $\lambda_{\min} = \epsilon \lambda_{\max}$ with $\epsilon = 0.001$

Step 3. Compute the initial estimates $\hat{\beta}$ by minimizing the unpenalized lad objective function $\sum_{i=1}^n |y - x\beta|$. Using the initial estimates in step 3 and the set of regularization parameter λ in step2, compute the SCAD penalty function as in Equation 5.

Step 4. Form the augmented dataset (\tilde{X}, \tilde{y}) for $i = 1, 2, \dots, n, n+1, \dots, n+p$, where $(\tilde{y}, \tilde{X}) = (\frac{y}{n}, \frac{X}{n})$, $i = 1, \dots, n$ and $(\tilde{y}, \tilde{X}) = (0, p'_\lambda(\beta_{nj}^0) \tilde{e}_i)$ for $i = 1, 2, \dots, n, n+1, \dots, n+p$, and \tilde{e}_i is a vector of p by 1 with all the i^{th} component equal to 1.

Step 5. Use any of the LAD regression procedure, for example quantreg package in R to compute the lad regression estimators for the augmented dataset (\tilde{X}, \tilde{y}) for $i = 1, 2, \dots, n, n+1, \dots, n+p$,

Step 6. Select the best model by computing the formula in Equation 9 for each value of λ . The model that corresponds to the minimum value of λ is considered as the best.

Numerical Evaluation

To assess the performance of the proposed WCS+LAD-SCAD estimator as explained in the proceeding sections, we carried a simulation study and analysis three real datasets namely NIR, octane and cookie dataset.

Simulation Study

A simulation study is carried out to compare the variable selection properties and prediction accuracy of the proposed method and the existing RCS+LAD-SCAD estimator with divergent number of predictors. As per Wang et al. (2015) the following three cases are considered:

Case 1: $y = X_i^T \beta + e_i$, with $e_i \sim N(0,1), i = 1, \dots, n$

Case 2: $y = X_i^T \beta + e_i$, with $e_i \sim t(3), i = 1, \dots, n$

Case 3: $y = X_i^T \beta + e_i$, with $e_i \sim 0.9N(0,1) + 0.1Cauchy(3), i = 1, \dots, n$

Here, for each case, we set the vector of coefficients β such that $\beta_1 = 3, \beta_2 = 1.5, \beta_5 = 2$ and $\beta_j = 0$ for $j \notin \{1, 2, 5\}$. The vector of predictors $X_i = (x_1, \dots, x_p)^T$ are generated from the multivariate normal distribution with mean 0 and covariance matrix $\Sigma = \sigma_j$ with $\sigma_{ij} = 0.5^{|i-j|}$. The number of predictors p_n and the number of sample size n are set to $p_n = (500, 1000)$ and $n = (100, 200)$ which are repeated 200 times in each case. We used the same seed number, 1234 throughout this paper both in the simulation and real date examples. The threshold $d_n = 4n/\log(n)$ is used with wrapped correlation screening WCS method to reduce the dimension from p_n to d_n . Table 1 to 3 exhibit the results based on the average number of zero coefficients correctly estimated as zero (NC), the average number of non-zero coefficients incorrectly estimated to zero (NIC). Following Aslan (2012) and Wang et al. (2015), the average median estimation error (MEE) defined as $\|\hat{\beta} - \beta\|$ is also used as performance measures of the various estimators. The values in parenthesis are for $n = 200$ and $p_n = 1000$

Table 1

Results for normal distributed errors with $n = 100, p = 500 ; n = 200, p = 1000$ in parenthesis

Method	NC	NIC	MEE	%Efficiency
RCS+LAD-Lasso	495.050(995.435)	0.000(0.000)	0.585(0.467)	37.43(31.69)
RSC+LAD-SCAD	496.570(996.975)	0.003(0.000)	0.281(0.160)	77.94(92.50)
WSC+LAD-Lasso	495.275(995.510)	0.000(0.000)	0.553 (0.398)	39.60(37.19)
WSC+LAD-SCAD	496.615(996.980)	0.000(0.000)	0.261(0.150)	83.91(98.66)
Oracle	497.000(997.000)	0.000(0.000)	0.219(0.148)	100(100)

Table 2

Results for t_3 distributed errors with $n = 100$, $p = 500$; $n = 200$, $p = 1000$ in parenthesis

Method	NC	NIC	MEE	%Efficiency
RCS+LAD-Lasso	495.175(995.985)	0.000(0.000)	0.720(0.580)	31.81(31.55)
RSC+LAD-SCAD	495.775(996.815)	0.004(0.000)	0.520(0.214)	44.04(85.51)
WSC+LAD-Lasso	495.525(996.705)	0.000(0.000)	0.654(0.465)	35.02(39.35)
WSC+LAD-SCAD	496.061(997.000)	0.003(0.000)	0.482(0.183)	47.51(100.00)
Oracle	497.000(997.000)	0.000(0.000)	0.229(0.183)	100(100)

Table 3

Results for normal errors with 10% contaminated observations, $n = 100$, $p = 500$; $n = 200$, $p = 1000$ in parenthesis

Method	NC	NIC	MEE	%Efficiency
RCS+LAD-Lasso	495.110(996.175)	0.006(0.001)	0.623(0.521)	36.44(29.37)
RSC+LAD-SCAD	496.310(996.970)	0.051(0.002)	0.381(0.163)	59.58(93.87)
WSC+LAD-Lasso	495.465(995.895)	0.005(0.000)	0.597(0.424)	38.02(36.08)
WSC+LAD-SCAD	496.430(997.000)	0.030(0.000)	0.360(0.156)	63.06(98.08)
Oracle	497.000(997.00)	0.000(0.000)	0.227(0.153)	100(100)

It can be observed From Table 1 to 3 that the RCS+LAD-Lasso and WSC+LAD-Lasso provide the worst results in terms of having NC and MEE values far away from the oracle values. Although their results are close, the values of MEE for WSC+LAD-Lasso are consistently smaller than those of the RCS+LAD-Lasso which indicate that WSC screening algorithm tends to increase the estimation accuracy. On the other hand, the values of the NC, NIC and MEE of WCS+LAD-SCAD estimator tend to be closer to the Oracle estimator. Nevertheless, the WSC+LAD-SCAD estimator provides NC, NIC and MEE values which are in best agreement with the oracle values especially when the $n = 200$ and $p = 1000$. The performance of RSC-LAD-SCAD is quite good both in terms of NC and MEE values; however, its accomplishment cannot outperform the WSC+LAD-SCAD. The performances of WSC-LAD-SCAD compared to other estimators is further assessed based on efficiency criterion (Dhhan et al., 2017) defined as follows:

$$\text{Efficiency} = (\text{MEE of the (oracle estimator)}/\text{MEE of the (oracle competitors)}) \times 100\%.$$

It can be clearly seen from Table 1 to 3 that our proposed WSC-LAD-SCAD has the highest value of efficiency, followed by RSC+LAD-SCAD, WSC+LAD-Lasso and RCS+LAD-Lasso. The results seem to be consistent for each sample $n = (100,200)$, and the number of predictors, $p_n = (500,1000)$, respectively. Hence, the WSC+LAD-SCAD may give a better alternative estimator for handling ultrahigh dimensional data in the presence of outliers. Furthermore, the results incline to suggest that by using the WSC screening algorithm, the efficiency of the LAD-SCAD estimator can substantially be improved.

Real Life Application

In this section, several real-life examples are presented to show the applicability and merit of our proposed method. These datasets include near infrared spectroscopy (NIR), cookies, and octane. The dataset can be obtained from the chemometrics, pppls, and rrcov packages in R, which have been formerly analysed in Brown et al. (2001), Liebmann et al. (2009), and Hubert et al. (2005), respectively. The octane data set consists of 39 samples and 226 predictors. After some preliminary investigation of the response observations based on Hubert and Van der Veecken (2008) as used in <https://github.com/marcellodo/univOutl>, we detected observation 25, 26, 36-39 as outliers. This procedure was repeated on the responses of the cookies and NIR data set, but all detected zero outlying points. The dimension for NIR data is 166 by 235 and 72 by 700 for the cookies data which represent number of sample and predictors, respectively. Motivated by the skewness property of the octane data plus our motive to assess the vigour of the proposed procedure, we contaminated the responses of NIR and cookies data by multiplying 0.05 percent of randomly selected observation by 100 after splitting the data into training and test set (the percentage are given in parenthesis) as given in Table 4. Applying the computational procedure described in the previous sections produced the results as shown in Table 4 for test data set, where the second column before the last column, represents the number of variables selected (NVS), the column before the last represents the average median absolute error, and the last represents the average robust R^2 statistics over 100 repeated simulations, as used in (Liu et al, 2016). Here, we employed the robust R^2 which is adapted from Wang et al. (2007) as defined Equation 11:

$$R^2 = 1 - \left(\frac{\text{med}(|y_i - \hat{y}_i|)}{\text{mad}(y_i)} \right)^2 \quad [11]$$

Generally speaking, the range of R^2 value start from 0 to 1 for a reasonable fit, $R^2=1$ signifies perfect fit and $R^2 < 0$ corresponds to bad model fit. From Table 4, it is interesting to observe that the proposed technique outperformed its competitors on all the data set, and is considered as our method to select a reasonable number of variables with minimum median absolute error (MAE) and better R^2 statistics. The threshold value of $d=n/\log(n)$ is used in all the data sets. A boxplot of the median absolute error and R^2 statistics for the

original data (raw values) as shown in Figure 1 to 2 explicitly demonstrate the effectiveness of our method against its competitors.

Table 4
Real data application

No.	Dataset	Method	#train (%)	#test (%)	NVS	MAE	R^2
1	NIR (166 × 235)	RCS+LAD-Lasso	120(72)	46(28)	1	0.659	-0.8597
		RSC+LAD-SCAD			3	0.253	0.2222
		WSC+LAD-Lasso			1	0.368	-0.1848
		WSC+LAD-SCAD			8	0.212	0.3472
2	Octane (39 × 226)	RCS+LAD-Lasso	25(64)	14(36)	3	0.008	0.9994
		RSC+LAD-SCAD			4	0.007	0.9999
		WSC+LAD-Lasso			6	0.005	0.9995
		WSC+LAD-SCAD			6	0.004	1.0000
3	Cookie (72 × 700)	RCS+LAD-Lasso	40(56)	32(44)	1	0.433	0.7589
		RSC+LAD-SCAD			8	0.237	0.9281
		WSC+LAD-Lasso			1	0.432	0.7559
		WSC+LAD-SCAD			10	0.143	0.9697

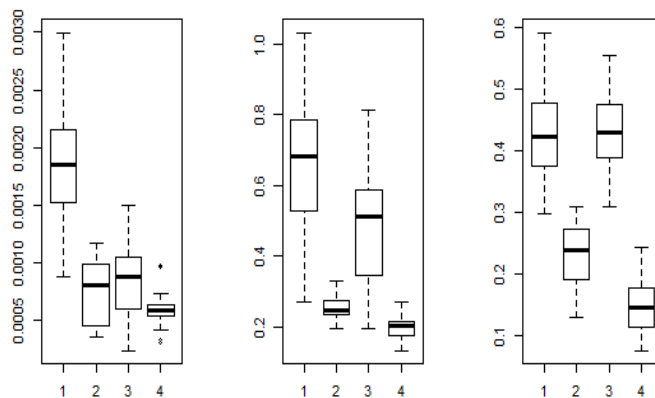


Figure 1. Boxplot of the median absolute error for the octane, NIR and cookies data set with RCS+LAD-Lasso=1, RCS+LAD-SCAD=2, WCS+LAD-Lasso=3 and WCS+LAD-SCAD=4 respectively.

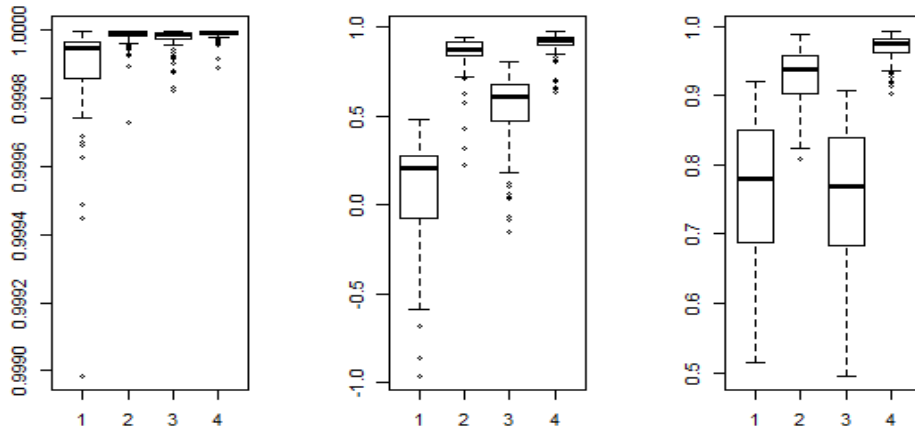


Figure 2. Boxplot of the R^2 statistics for the octane, NIR and cookies data set with RCS+LAD-Lasso=1, RCS+LAD-SCAD=2, WCS+LAD-Lasso=3 and WCS+LAD-SCAD =4, respectively.

CONCLUSION

Inspired by the robust rank correlation sure screening-based LAD-SCAD, we proposed a wrapped based sure screening LAD-SCAD to achieve better robust estimates. The main advantage of our method is that it deals with outliers in both the pre-screening and post-screening step by using the robust wrapped transformation in the computation of best regularization parameter. The proposed procedure shows more success as it appears to be more robust and efficient than the existing RCS+LAD-SCAD method for solving linear regression in the presence of outlying observations. Therefore, the proposed procedure can be used by practitioners for parameter estimations and variable selection when the response observation contains some outliers. Future work will consider the impact of both vertical and horizontal outliers.

ACKNOWLEDGEMENT

This research was partially supported by the Fundamental Research Grant Scheme (FRGS) under Ministry of Education with project number FRGS/1/2019/STG06/UPM/01/1 and Petroleum Technology Development Fund of Nigeria.

REFERENCE

Ahmed, T., & Bajwa, W. U. (2019). ExSIS: Extended sure independence screening for ultrahigh-dimensional linear models. *Signal Processing*, 159, 33-48. <https://doi.org/10.1016/j.sigpro.2019.01.018>

- Arslan, O. (2012). Weighted LAD-LASSO method for robust parameter estimation and variable selection in regression. *Computational Statistics & Data Analysis*, 56(6), 1952-1965. <https://doi.org/10.1016/j.csda.2011.11.022>
- Bai, Z. D., & Wu, Y. (1997). General M-estimation. *Journal of Multivariate Analysis*, 63(1), 119-135. <https://doi.org/10.1006/jmva.1997.1694>
- Brown, P. J., Fearn, T., & Vannucci, M. (2001). Bayesian wavelet regression on curves with application to a spectroscopic calibration problem. *Journal of the American Statistical Association*, 96(454), 398-408. <https://doi.org/10.1198/016214501753168118>
- Candes, E., & Tao, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n . *The Annals of Statistics*, 35(6), 2313-2351. <https://doi.org/10.1214/009053606000001523>
- Chang, L., Roberts, S., & Welsh, A. (2018). Robust Lasso Regression Using Tukey's Biweight Criterion. *Technometrics*, 60(1), 36-47. <https://doi.org/10.1080/00401706.2017.1305299>
- Croux, C., & Dehon, C. (2010). Influence functions of the Spearman and Kendall correlation measures. *Statistical Methods & Applications*, 19(4), 497-515. <https://doi.org/10.1007/s10260-010-0142-z>
- Desboulets, L. D. D. (2018). A review on variable selection in regression analysis. *Econometrics*, 6(4), Article 45. <https://doi.org/10.3390/econometrics6040045>
- Dhhan, W., Rana, S., & Midi, H. (2017). A high breakdown, high efficiency and bounded influence modified GM estimator based on support vector regression. *Journal of Applied Statistics*, 44(4), 700-714. <https://doi.org/10.1080/02664763.2016.1182133>
- Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 96(456), 1348-1360. <https://doi.org/10.1198/016214501753382273>
- Fan, J., & Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 70(5), 849-911. <https://doi.org/10.1111/j.1467-9868.2008.00674.x>
- Fan, J., & Peng, H. (2004). Nonconcave penalized likelihood with a diverging number of parameters. *The Annals of Statistics*, 32(3), 928-961. <https://doi.org/10.1214/009053604000000256>
- Fan, J., & Song, R. (2010). Sure independence screening in generalized linear models with NP-dimensionality. *The Annals of Statistics*, 38(6), 3567-3604.
- Fan, J., Samworth, R., & Wu, Y. (2009). Ultrahigh dimensional feature selection: beyond the linear model. *The Journal of Machine Learning Research*, 10, 2013-2038.
- Frank, L. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics*, 35(2), 109-135.
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1.

- Gao, X., & Huang, J. (2010). Asymptotic analysis of high-dimensional LAD regression with LASSO. *Statistica Sinica*, 1485-1506.
- George, E. I. (2000). The variable selection problem. *Journal of the American Statistical Association*, 95(452), 1304-1308.
- Ghaoui, L. E., Viallon, V., & Rabbani, T. (2010). Safe feature elimination for the lasso and sparse supervised learning problems. *Machine Learning*, 2000, 1-31.
- Heinze, G., Wallisch, C., & Dunkler, D. (2018). Variable selection—a review and recommendations for the practicing statistician. *Biometrical Journal*, 60(3), 431-449. <https://doi.org/10.1002/bimj.201700067>
- Huang, J., & Xie, H. (2007). Asymptotic oracle properties of SCAD-penalized least squares estimators. In *Asymptotics: Particles, Processes and Inverse Problems* (pp. 149-166). Institute of Mathematical Statistics. <https://doi.org/10.1214/074921707000000337>
- Hubert, M., & Van der Veeken, S. (2008). Outlier detection for skewed data. *Journal of Chemometrics: A Journal of the Chemometrics Society*, 22(3-4), 235-246. <https://doi.org/10.1002/cem.1123>
- Hubert, M., Rousseeuw, P. J., & Branden, K. V. (2005). ROBPCA: a new approach to robust principal component analysis. *Technometrics*, 47(1), 64-79. <https://doi.org/10.1198/004017004000000563>
- Leng, C., Lin, Y., & Wahba, G. (2006). A note on the lasso and related procedures in model selection. *Statistica Sinica*, 1273-1284.
- Li, G., Peng, H., & Zhu, L. (2011). Nonconcave penalized M-estimation with a diverging number of parameters. *Statistica Sinica*, 391-419.
- Li, R., Zhong, W., & Zhu, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association*, 107(499), 1129-1139. <https://doi.org/10.1080/01621459.2012.695654>
- Liebmann, B., Friedl, A., & Varmuza, K. (2009). Determination of glucose and ethanol in bioethanol production by near infrared spectroscopy and chemometrics. *Analytica Chimica Acta*, 642(1-2), 171-178. <https://doi.org/10.1016/j.aca.2008.10.069>
- Liu, J., Wang, Y., Fu, C., Guo, J., & Yu, Q. (2016). A robust regression based on weighted LSSVM and penalized trimmed squares. *Chaos, Solitons & Fractals*, 89, 328-334. <https://doi.org/10.1016/j.chaos.2015.12.012>
- Maronna, R. A., Martin, R. D., & Yohai, V. J. (2006). *Robust statistics: Theory and methods (with R)*. John Wiley & Sons.
- Meinshausen, N., & Bühlmann, P. (2006). High-dimensional graphs and variable selection with the lasso. *The Annals of Statistics*, 34(3), 1436-1462. <https://doi.org/10.1214/009053606000000281>
- Raymaekers, J., & Rousseeuw, P. J. (2019). Fast robust correlation for high-dimensional data. *Technometrics*, 1-15. <https://doi.org/10.1080/00401706.2019.1677270>
- Rousseeuw, P. J., & Leroy, A. M. (1987). *Robust regression and outlier detection*. Wiley.

- Saldana, D. F., & Feng, Y. (2018). SIS: An R package for sure independence screening in ultrahigh dimensional statistical models. *Journal of Statistical Software*, 83(2), 1-25. <https://doi.org/10.18637/jss.v083.i02>
- Shevlyakov, G., & Smirnov, P. (2011). Robust estimation of the correlation coefficient: An attempt of survey. *Austrian Journal of Statistics*, 40(1&2), 147-156. <https://doi.org/10.17713/ajs.v40i1&2.206>
- Stuart, C. (2011). *Robust regression*. Durham University.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Tibshirani, R., Bien, J., Friedman, J., Hastie, T., Simon, N., Taylor, J., & Tibshirani, R. J. (2012). Strong rules for discarding predictors in lasso-type problems. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(2), 245-266. <https://doi.org/10.1111/j.1467-9868.2011.01004.x>
- Uraibi, H. S., Midi, H., & Rana, S. (2017). Selective overview of forward selection in terms of robust correlations. *Communications in Statistics: Simulation and Computation*, 46(7), 5479-5503. <https://doi.org/10.1080/03610918.2016.1164862>
- Wang, H., Li, G., & Jiang, G. (2007). Robust regression shrinkage and consistent variable selection through the LAD-Lasso. *Journal of Business & Economic Statistics*, 25(3), 347-355. <https://doi.org/10.1198/073500106000000251>
- Wang, M., Song, L., & Tian, G. L. (2015). SCAD-penalized least absolute deviation regression in high-dimensional models. *Communications in Statistics-Theory and Methods*, 44(12), 2452-2472. <https://doi.org/10.1080/03610926.2013.781643>
- Wang, T., & Zhu, L. (2011). Consistent tuning parameter selection in high dimensional sparse linear regression. *Journal of Multivariate Analysis*, 102(7), 1141-1151. <https://doi.org/10.1016/j.jmva.2011.03.007>
- Whittingham, M. J., Stephens, P. A., Bradbury, R. B., & Freckleton, R. P. (2006). Why do we still use stepwise modelling in ecology and behaviour? *Journal of Animal Ecology*, 75(5), 1182-1189. <https://doi.org/10.1111/j.1365-2656.2006.01141.x>
- Wu, Y., & Liu, Y. (2009). Variable selection in quantile regression. *Statistica Sinica*, 19(2), 801-817.
- Xiang, Z. J., & Ramadge, P. J. (2012). Fast lasso screening tests based on correlations. In *2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 2137-2140). IEEE Conference Publication. <https://doi.org/10.1109/ICASSP.2012.6288334>
- Xie, H., & Huang, J. (2009). SCAD-penalized regression in high-dimensional partially linear models. *The Annals of Statistics*, 37(2), 673-696. <https://doi.org/10.1214/07-AOS580>
- Zhang, Y., Li, R., & Tsai, C. L. (2010). Regularization parameter selections via generalized information criterion. *Journal of the American Statistical Association*, 105(489), 312-323. <https://doi.org/10.1198/jasa.2009.tm08013>
- Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association*, 101(476), 1418-1429. <https://doi.org/10.1198/016214506000000735>

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (statistical methodology)*, 67(2), 301-320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

Relative Risk Estimation for Human Leptospirosis Disease in Malaysia Based on Existing Models and Discrete Space-Time Stochastic Sir Model

Sufi Hafawati Ideris*, Muhammad Rozi Malim and Norshahida Shaadan

Faculty of Computer and Mathematical Sciences, MARA University of Technology, Shah Alam, 40450 Selangor, Malaysia

ABSTRACT

The disease leptospirosis is known to be endemic in Malaysia, and it significantly impacts human wellbeing and the national economy. Current surveillance systems are based on morbidity and mortality leptospirosis national data from the Ministry of Health and remain inadequate due to the number of unreported and misdiagnosed cases. A robust surveillance system is needed to monitor temporal and spatial changes which yield improvements in terms of identifying high-risk areas and disease behaviour. The objective of this study is to identify high-risk areas by estimating relative risk using existing models which are the Standardized Morbidity Ratio (SMR), Poisson-gamma, log-normal, Besag, York and Mollié (BYM) and mixture models. An alternative model is also proposed which involves transmission systems and stochastic elements, namely the stochastic Susceptible-Infected-Removed (SIR) transmission model. This estimation of risk is expected to assist in the early detection of high-risk areas which can be applied as a strategy for preventive and control

measures. The methodology in this paper applies relative risk estimates to determine the infection risk for all states in Malaysia based on monthly data from 2011 to 2018 using WinBUGS 1.4 software. The results of relative risks are discussed and presented in tables and graphs for each model to disclose high-risk areas across the country. Based on the risk estimates, different models used have different risk interpretations and drawbacks which make each model different in its use depending on the objectives of the

ARTICLE INFO

Article history:

Received: 9 September 2020

Accepted: 25 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.20>

E-mail addresses:

sufi.hafawati@gmail.com (Sufi Hafawati Ideris)

rozi@tmsk.uitm.edu.my (Muhammad Rozi Malim)

shahida@tmsk.uitm.edu.my (Norshahida Shaadan)

*Corresponding author

study. As a result, the deviance information criteria (DIC) values obtained do not differ greatly from each expected risk which was estimated.

Keywords: BYM, DIC, leptospirosis, log-normal, mixture model, poisson-gamma, relative risk, SIR transmission model, SMR

INTRODUCTION

Leptospirosis is a primarily zoonotic disease originating from the pathogenic leptospire species and living widely throughout nature. Small mammals like rodents are the reservoir hosts in which pathogenic leptospire are carried and maintained, reproducing itself in their hosts' kidneys. Excretion of leptospire through the urine of those infected animals causes acquisition in new hosts, whether other animals or humans (Haake & Lavett, 2014). Transmission may occur via two major routes: direct contact with urine or other bodily fluids of infected animals through the skin and intact mucus membranes such as the eyes, nose or mouth; and indirect contact with broken skin or through being scratched with soil, food or water contaminated with the urine of infected animals (Wahab, 2015). Leptospirosis is widespread and can be a serious public health issue especially in temperate or tropical climates (Haake & Levett, 2014). Malaysia is known to be a location in which leptospirosis is endemic among the south-east Asia region and significantly impacts human wellbeing and the national economy (Garba et al., 2017a). This is because the humid weather and climate conditions in Malaysia are conducive for the growth of the leptospire bacteria. The abundance of leptospire has therefore caused death and suffering in Malaysia since it was first diagnosed by Fletcher (1928). Based on Figure 1, the frequency of incidence and the mortality rate due to leptospirosis increased between 2004 and 2015.

Despite the dramatic increase in the number of cases and deaths, leptospirosis is a seriously neglected tropical disease. Factors like misdiagnosis of clinical symptoms which resemble other similar diseases (such as dengue and hepatitis) and the limited or lack of rapid diagnosing facilities causes the disease to continually be under-reported. This is also may create barriers to government efforts to determine the true incidence rate for the human population. The elimination of leptospirosis is therefore crucial to minimise the risk of infection. This prevention action can be performed through appropriate controls, including a robust and recognised vaccination programme, promotion of the use of hygiene and personal protective equipment as well as awareness of high incidence areas by the authorities and society. This is crucial for the risk estimates in geographical areas and revealing the hotspot areas of leptospirosis infection.

In Malaysia, the surveillance systems currently and typically rely on available leptospirosis national databases, which are the morbidity and mortality reports from

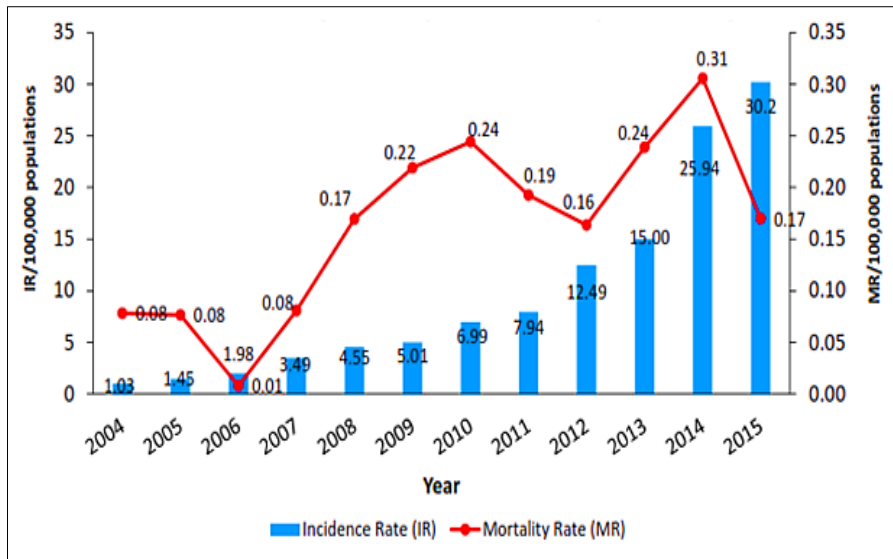


Figure 1. Incidence and mortality rate of leptospirosis in Malaysia, 2004 to 2015 (Wahab, 2015)

Ministry of Health Hospitals. However, the existing surveillance system is still inadequate to combat this disease due to the number of unreported and underestimated cases (Benacer et al., 2016). A robust surveillance system which can monitor temporal and spatial changes is needed to yield improvements in identifying at-risk areas and disease behaviour. This at the same time can assist the early detection of high-risk areas which can be applied as a strategy for preventive measures in Malaysia and other countries.

Estimation of relative risk is very common in the study of epidemiology and is defined as the estimation of the probability of infection after exposure to the risk variable compared with the likelihood of disease occurrence in a sample study group or population (Andrade, 2015). Relative risk (also known as risk ratio) calculation is usually expressed as a ratio to a denominator instead of a percentage. Relative risk in this study is defined as the probability that a person within a specific region would contract the disease against the likelihood that a person in the general population would contract the disease. To determine the disease risk status of an area, the first action is to assess the local supposed 'expected' incidence of the disease, followed by a comparison with local observations of the incidence value of the disease. Several common methods are used in the epidemiology field to estimate the ratio of risk by using data for counts within tracts. These are the SMR model, the Poisson-gamma model, the Log-normal model, the Besag, York and Mollié (BYM) model and the mixture model. These approaches are known as classic methods of estimating disease risk which are commonly used in disease mapping. However, this paper will not discuss in detail

the representatives of the disease risk map. The focus is on the estimation of relative risk approaches to determine the risk of leptospirosis disease across the Malaysian region. The methodology is discussed from a basic Poisson likelihood to more complex risk structure models which include spatial correlation and spatiotemporal.

The goal of epidemiology research is to gain an insight into the nature or history of transmission model development (Brauer, 2017). Biologically, leptospirosis disease involves direct and indirect processes of transmission from infectious to susceptible individuals. Therefore, a model that can incorporate this transmission system and an alternative method for estimating the relative risk of the leptospirosis SIR transmission model are required in this paper for the human population using monthly data from Malaysia. The results of relative risk estimated from both the existing and proposed transmission models are discussed and the performance is evaluated using deviance information criteria (DIC) to assess model complexity and best fitting model to the data.

METHODOLOGY

Data and Study Location

The relative risk estimation in this paper used secondary data for observed cases of leptospirosis infection in the form of count cases for all 13 states and 3 federal territories in Malaysia. These states and federal territories are Perlis, Kedah, Penang, Perak, Selangor, Federal Territory of Kuala Lumpur, Federal Territory of Putrajaya, Negeri Sembilan, Malacca, Johor, Pahang, Terengganu, Kelantan, Sarawak, Sabah, and Federal Territory of Labuan. In Malaysia, leptospirosis was gazetted as a notifiable disease in December 2010 (Garba et al., 2017a; Sulong et al., 2011). Thus, this study was designed to accommodate the monthly observed cases starting from January 2011 up to December 2018 so cumulative data could be taken at annual intervals (starting in January and ending in December) for all states and federal territories in Malaysia.

This secondary data was obtained from the Ministry of Health with ethic approval from the National Medical Research Registration (NMRR): NMRR-18-3006-43002 (IIR). Analyses are based on space and time dimensions since data from the public health authorities (MOH) is often available for time windows for several years. Here, the observed counts for the new infected human cases are based on the time $(j - 1, j)$ for $j = 1, 2, \dots, 96$ months per calendar month and study region i for all 13 states and three federal territories in Malaysia. A one-month observation time was chosen based on the development of the leptospira bacteria in the human body once a person is infected (Triampo et al., 2008). In this study, the observed number for humans infected with leptospirosis was categorised as the dependent variable and time (monthly) was the independent variable.

Analyses Framework

This paper involves several analytical steps such as data collection, calculation, estimation, presentation of results as well as comparison. This progress flow is explained in more detail in the form of a flow chart in Figure 2.

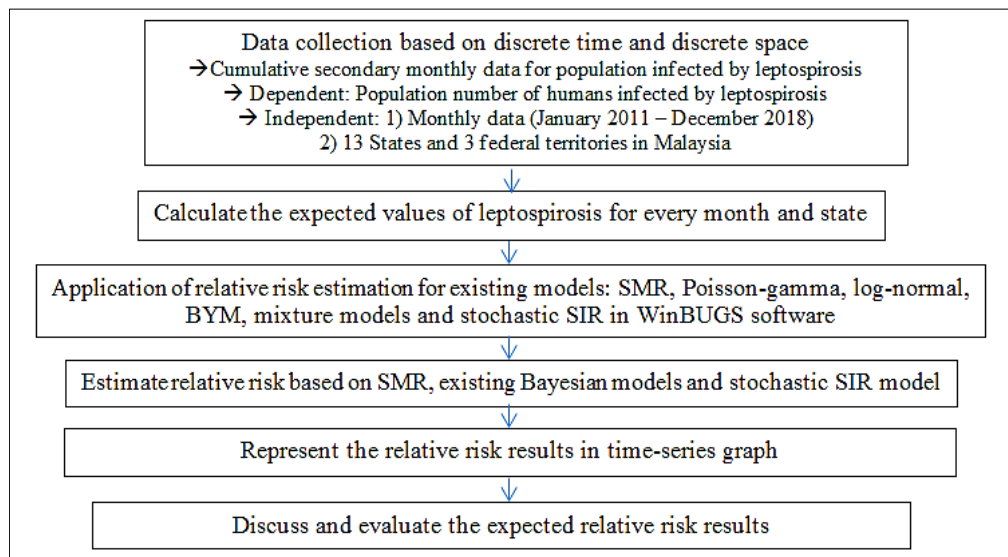


Figure 2. Flow of the process

Relative Risk Estimator for Leptospirosis using Existing Models

In statistical analysis, inferences are made by selecting a value or pattern for a certain phenomenon from empirical data, a procedure known as an estimator (Long, 2012). An estimator aims to measure or derive such risk estimates on a statistical spatial method. A basic summary measurement for an estimator often computes a local relative risk which is taken to measure excess risk within a population (Lawson, 2006). This risk estimate has been commonly used in spatial epidemiology and is calculated by comparing the observed incidence to that expected from the background population. This assessment of risk status of the ratio of observed to expected counts is called standardised mortality or morbidity ratio (SMR), a ratio which is an estimate of relative risk within each area or region. Usually, confounding, or unobserved variables can be found within observed data in spatial units affecting individuals or associated tracts or covariates, these are known as random effects. Due to its simple estimates, this type of primary descriptive method often ignores such variables. Thus, more substantial models are used to include these effects in the analysis, which is discussed further in this study.

Standardized Morbidity Ratio (SMR). In broad epidemiological terms, SMR can be defined as either Standardised Mortality Ratio or Standardised Morbidity Ratio. The term mortality refers to death and the mortality rate is the number of deaths divided by the size of the composite population. Meanwhile, morbidity means disease or illness. It can range from mild illnesses such as common flu to chronic diseases such as cancer or traumatic brain injury. In addition, a study of nonfatal disease is sometimes known as the Standardised Morbidity Ratio (Richardson et al., 2004) where the status burden of a disease in an area respective to the disease is the ratio of observed to expected counts.

Suppose that the study area to be mapped is divided into M mutually exclusive states ($i = 1, 2, \dots, M$) and T time ($j = 1, 2, \dots, T$). Each state has its own observed number of cases O_i and expected number of cases E_i , as shown in Equation 1 and 2.

$$E_i = \text{Population of state } i \text{ in time } j \times \frac{\sum \text{number of observed cases for all states in time } j}{\sum \text{number of population in time } j} \quad [1]$$

$$E_i = N_{i,j} \left(\frac{\sum_{i=1}^M O_j}{\sum_{i=1}^M N_j} \right) \quad [2]$$

Where N_i is the population of state i and the summations are for $i = 1, 2, \dots, M$. Here the standardisation is performed based on the total population at risk, assuming everybody is equally at risk. Using O_i and E_i as obtained from the available data, the derivation of relative risk value, which is denoted as $\hat{\theta}_i$, can be calculated using Equation 3. The relative risk, $\hat{\theta}_{i,j}$ for state i at time j is defined as the probability that a person within the state i contracts the disease divided by the probability that a person in the population contracts the disease (Samat & Percy, 2012).

$$\text{Relative risk} = \frac{\text{Observed number of cases in state } i \text{ at time } j}{\text{Expected number of cases in state } i \text{ at time } j}$$

$$\begin{aligned} \hat{\theta}_{i,j} &= \frac{O_{i,j}}{E_{i,j}} \\ \hat{\theta}_{i,j} &= \frac{O_{i,j}}{N_{i,j} \left(\frac{\sum_{i=1}^M O_j}{\sum_{i=1}^M N_j} \right)} \\ \hat{\theta}_{i,j} &= \frac{\left(\frac{O_{i,j}}{N_{i,j}} \right)}{\left(\frac{\sum_{i=1}^M O_j}{\sum_{i=1}^M N_j} \right)} \end{aligned} \quad [3]$$

However, this method has several drawbacks that make interpretation of the SMR difficult. According to Kismiantini (2009), the variance is large in a region that has small

expected values, especially an area with a small population size or region with rare diseases cases. This extra variation in a small region with a small population size can cause artefacts or errors in the estimated risk (Lawson & Williams, 2003). The SMR is measured by observed over expected incident numbers in the form of a ratio. If the ratio is more than 1 it may suggest excess in risk, but the SMR will become more than 1 or even large if the expected number of incident cases is too small. For instance, the observed count number for Labuan in April 2016 was 11 with the expected number counted as 1, giving an SMR of 11. Otherwise, if the expected value is too large or there is no observed value of cases in a region, then the SMR is very small or equal to zero (Awang & Samat, 2017). In this case, the zero value of relative risk indicates that there is no risk for the people in the study region infected by the disease even though everybody has the possibility of catching the disease. In addition, if the expected number is zero in a count tract, then the estimation ratio of the SMR is infinite. All these types of artefacts for the SMR often occur in a region with a small population size and area and with a small count of cases (such as Labuan, Perlis, and Putrajaya). Thus, the SMR is a reliable measure of relative risk for large geographical regions such as countries but may be unreliable for small areas such as counties (Meza, 2003). To overcome these drawbacks, many researchers have sought alternative solutions by introducing better alternative methods for relative risk estimation (Kismiantini, 2009). These include the use of Bayesian methods such as the Poisson-gamma model.

Poisson-Gamma Model. The Poisson-gamma model is one of the earliest examples of a Bayesian approach used to estimate disease relative risk especially when the data is typically small (Richardson et al., 2004). By using Bayes' rule, a combination of observed data and some prior knowledge would produce the posterior distribution. In Bayes' method, relative risk estimation is based on the posterior distribution which is defined as the product of likelihood function and prior distribution in Equation 4.

$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior} \quad [4]$$

For $i = 1, 2, \dots, M$ study states and $j = 1, 2, \dots, T$ time, $y_{i,j}$ is the observed count of cases in state i and at time j . Likelihood is defined as the probability distribution for data that conditionally acts upon unknown parameter, $\theta_{i,j}$. This probability distribution for likelihood is often assumed to follow a Poisson distribution within a given period, with mean and variance $e_{i,j}\theta_{i,j}$ as follows:

$$y_{i,j} | e_{i,j}, \theta_{i,j} \sim \text{Poisson}(e_{i,j}\theta_{i,j}) \quad [5]$$

Based on Equation 5, $e_{i,j}$ is the expected number of cases and the parameter of interest, $\theta_{i,j}$ is assumed to be a parameter of relative risk for state i and treated as a random variable. The random variable distribution of parameter $\theta_{i,j}$ has a gamma prior distribution with parameters α and β as written in Equation 6.

$$\theta_{i,j} \sim \text{Gamma}(\alpha, \beta) \quad [6]$$

This gamma prior provides for a measure of extra variation affected by random effects within the estimation of risk (Lawson, 2006). Hence, the final output of the Bayes estimate includes the posterior expected of relative risk for all states and time periods. Previous studies by Awang and Samat (2017) and Ideris and Samat (2015) suggested that Poisson-gamma gave a better estimation compared to the SMR when dealing with zero or clustered counts data. This is due to the modelling of random effects applied within the Bayesian framework in the Poisson-gamma model.

Log-Normal Model. Although Poisson-gamma is convenient, it may be too complicated to allow spatial correlation between risk in nearby areas, while the covariate adjustment is also difficult when heterogeneity is involved (Lawson & Williams, 2003; Nurmalasari & Pramana, 2014). Spatial random effects may arise when the regions are believed to have site-dependent covariates. Thus, the log-normal model is introduced which allows spatial correlation between risk in nearby areas. A log-normal model can be written as Equation 7:

$$\begin{aligned} y_i &\sim \text{Poisson}(e_i \theta_i) \\ \log \theta_i &= \alpha + v_i \\ v_i &\sim N(0, \tau_v^2) \end{aligned} \quad [7]$$

For a log-normal model, y_i are assumed to be Poisson distributed with parameter $e_i \theta_i$. The observed number of cases θ_i is described in logarithmic terms, with α as a semiparametric model of improper prior distribution for an intercept mean of relative risk in the study region (Lawson, 2006), while v_i represents the correlation parameter between the risk of neighbouring areas. This correlation is assumed to be normally distributed with mean 0 and variance τ_v^2 .

Besag, York and Mollié (BYM) Model. In the BYM model, random effects are specified into the above components for the model of relative risk estimation that considers effects that vary within a cluster or correlated heterogeneity, as well as effects that vary in uncorrelated heterogeneity. The relative risk for the BYM model is formulated as Equation 8 where u_i is the correlated heterogeneity while v_i is the uncorrelated heterogeneity. Meanwhile α is assigned as an improper uniform prior for the conditional autoregressive model (CAR) distribution to ensure that the model is identifiable.

$$\begin{aligned} y_i &\sim \text{Poisson}(e_i \theta_i) \\ \log \theta_i &= \alpha + u_i + v_i \\ v_i &\sim N(0, \tau_v^2) \end{aligned} \quad [8]$$

In the form of extra variation, there is a spatial autocorrelation between spatial units due to the unobserved environmental effect, and a naturally clustered scale observed for the disease for its distribution as discussed before, where estimation of risk in any area depends on neighbouring regions. Therefore, the conditional autoregressive (CAR) model is used to unveil and describe the quantities for spatial relations between the data and spatial variation in the form of sub-regions aggregation as well as to detect the cluster areas. The CAR model is defined as Equation 9:

$$[u_i | u_k, i \neq k, \tau_u^2] \sim N(\bar{u}_i, \tau_i^2) \quad [9]$$

where,

$$\bar{u}_i = \frac{1}{\sum_j \omega_{ik}} \sum_k u_k \omega_{ik} \quad [10]$$

$$\tau_1^2 = \frac{\tau_u^2}{\sum_k \omega_{ik}}$$

$\omega_{ik} = 1$ if i, k are adjacent (or 0 otherwise)

By referring to Equation 10, ω_{ik} represents the weightage for neighbouring areas. Summation of weightage ω_{ik} is equal to 1 if i, k is adjacent and 0 when i, k are not adjacent. Meanwhile, variables v and u are controlled by parameters τ_v^2 and τ_u^2 representing the inverse-variance of random effects. These hyperprior distributions are assumed to follow gamma distributions.

Mixture Model. To estimate true relative risk, an appropriate model should be able to detect and maintain discontinuities and variations between neighbouring areas. This means that an area should not be too dependent on its nearby areas, which causes the estimated risk to become almost identical to the neighbouring areas. To overcome this problem of the BYM model, the mixture model was proposed by contributing discrete jumps components to maintain the variation risk for each tract. The relative risk model is composed of three components: v as unstructured heterogeneity and two mixing components; and (u, φ) denoted as a combination of a component for spatial correlation and a component that models the discrete jumps between neighbouring areas.

$$\log \theta_i = \alpha + v_i + p_i u_i + (1 - p_i) \varphi_i \quad [11]$$

Equation 11 depends on the value of p_i as a cluster assignment which subdivides the data points into clusters (Long, 2012). When $p_i = 1$ the standard BYM model arises for the smoothing, and when $p_i = 0$, a pure jump model arises which is distributed as Equation 12.

$$p_i \sim \text{Beta}(0.5, 0.5) \quad \forall i \quad [12]$$

When the discrete jump model is considered (for $p_i = 0$), the spatial correlation follows a chosen prior which examines the total difference of two independent random variables between neighbours. The results of Lawson and Williams (2003) show that estimated risk through the mixture model performs better than the BYM by maintaining the areas with jumps, and at the same time displaying smoothing in the other areas.

Based on the discussion of all the risk estimation methods above, the output risk varies according to the approaches used. Each method has different interpretations and elements used in estimating the spatial random variables. A more effective model is often assumed to produce true risk estimation which takes random effects into account (Lawson, 2006). In the next subtopic, a goodness-of-fit method for the model comparison for the Bayesian model, which is a package from WinBUGS software, is discussed.

Deviance Information Criterion (DIC). Model comparisons for complexity and fit of the data to the model in Bayesian measurements can be conducted using deviance information criterion (DIC) monitored in WinBUGS, given as Equation 13.

$$\text{DIC} = \text{Dbar} + \text{pD} \quad [13]$$

Based on Equation 13 of DIC above, Dbar is used to calculate the goodness of fit for the model while component pD calculates the complexity of the model by estimating the effective number of parameters used in the model (Spiegelhalter et al., 2003). A model with a small DIC number is estimated to be the best model to fit with the data. In statistical models of analysis, there is no perfect or good model, and no mathematical model can accurately estimate detailed and real situations. DIC is used to compare the models which best fit the data. However, some studies disputed the capacity of DIC to compare hierarchical or multi-level models when re-parameterisation occurs (Pooley & Marion, 2018; Dey et al., 2019). On the other hand, it is very elusive to find a study which compares DICs between different forms of mathematical models. This is because models with different operator scopes have different mathematical functions and operations. Therefore, comparing different types of mathematical models with different mathematical functions can cause misleading interpretation of DICs.

Expected Relative Risk for Leptospirosis using Discrete Space-Time Stochastic SIR Model for Human Transmission

The focus of this paper is the application of mathematical and statistical approaches that quantify transmission models and data uncertainty for the estimation of relative risk.

Descriptions of mathematical transmission models are often based on the understanding of a disease spread between susceptible and infectious individuals (Brauer, 2017). This study led to the development of a compartmental model which has long been introduced and applied by mathematicians and physicians. The basic development of a compartmental model is based on the ideas behind the SIR model proposed by Kermack and McKendrick (1927). The simple compartmental model assumes that the mixing population in the compartments is homogeneous. However, a small number of individuals are infected at the beginning of the disease outbreak and the transmission is a probabilistic event based on the pattern of contact between susceptible and infective individuals.

Therefore, in this paper a basic compartmental stochastic SIR model is proposed specifically for humans infected by leptospirosis disease. This model involves a transmission system (known as a compartmental model) and a stochastic process which involves the probabilistic process of infection adapted from an early study by Reed-Frost in the 1920s. The population was divided into compartments with a time rate of transfer from one compartment to another. The terminology SIR is used to describe individuals from susceptible state S to infected state I and eventually to recovered state R . Susceptible is classified for individuals who will be the subject of the disease but are not yet infected. Individuals classified as infected are the ones who had been infected with the disease from contact with infected rats' urine. Once the infected individuals recuperated from the disease, they moved to the recovered state.

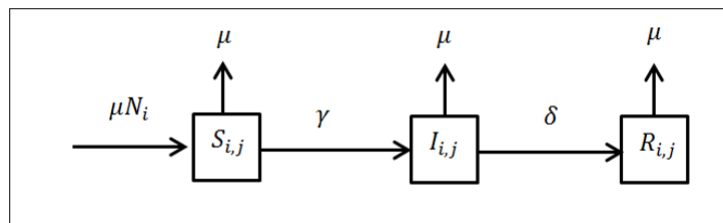


Figure 3. A compartmental SIR model for human leptospirosis transmission

Let ' S ' be the susceptible, ' I ' is the infected and ' R ' is recovered for $i=1, 2, \dots, M$ study regions (for all states and federal territories in Malaysia) and $j=1, 2, \dots, T$ time periods (which is monthly data from January to December per calendar month). From Figure 3, the compartment $S_{i,j}$ represents the total number of susceptible persons for leptospirosis for state i at j^{th} month, $I_{i,j}$ refers to the total number of persons infected with leptospirosis at state i for month j , and $R_{i,j}$ is the total number of recovered persons from leptospirosis disease for state i at month j .

The arrow from susceptible to infected represents the new infections occurring at rate γ . While δ denotes the recovery rate for an infected person. Furthermore, μ represents

the rate for persons who leave the group because of natural death such as from an illness unrelated to leptospirosis, like diabetic complications, myocardial infarct, malignancy, and motor-vehicle accidents among the human population. For human population, N_i is assumed to have closed population comprises of susceptible, infected, and recovered individuals written as Equation 14.

$$N_i = S_i + I_i + R_i \tag{14}$$

The total susceptible human population comprises the existing susceptible population which increases with the rate of birth, μ , of which every person born is assumed to be susceptible to leptospirosis. At the same time, the susceptible population decreases following effective contact with the infectious individuals at rate γ . The susceptible human population is further decreased by natural death at rate μ . In this model, birth and death rates are assumed to be equal for simplicity and the rate of death μ is constant for all population subgroups of susceptible, infected, and recovered individuals. Thus, the deterministic difference model for susceptible humans can be written as Equation 15.

$$S_{i,j} = S_{i,j-1} + \mu N_i - \gamma I_{i,j} S_{i,j-1} - \mu S_{i,j-1} \tag{15}$$

Next, the population of infected humans is generated from the existing previously infected individuals and increase by newly infective individuals by new infections of susceptible humans' rate γ . The population of infectious humans is decreased by natural death at rate μ , as well as by recovery from infection at rate δ . The infectives human formula can be written as Equation 16.

$$I_{i,j} = I_{i,j-1} + \gamma I_{i,j} S_{i,j-1} - \mu I_{i,j-1} - \delta I_{i,j-1} \tag{16}$$

The population of individuals in the recovered compartment is formulated by the existing number of previously recovered persons and increased by the number of newly recovered persons from the infected class at rate δ while it is decreased by natural death μ which is constant for all compartments as written in Equation 17.

$$R_{i,j} = R_{i,j-1} + \delta I_{i,j-1} - \mu R_{i,j-1} \tag{17}$$

The system of transmission behaviour for human leptospirosis disease (Figure 3) can be written as a linear system of difference deterministic Equation 18.

$$\begin{aligned} S_{i,j} &= \mu N_i + \{1 - \mu - \gamma I_{i,j-1}\} S_{i,j-1} \\ I_{i,j} &= (1 - \mu - \delta) I_{i,j-1} + \gamma I_{i,j-1} S_{i,j-1} \\ R_{i,j} &= \delta I_{i,j-1} + (1 - \mu) R_{i,j-1} \end{aligned} \tag{18}$$

Based on this proposed compartmental model, the deterministic Equation 16 was derived to develop the corresponding discrete-time discrete-space stochastic SIR model specifically for leptospirosis disease transmission for human population. Initially there must be an infected human in the population; $I_{i,0} = 1$. Meanwhile the initial number of susceptible humans is $S_{i,0} = N - 1$. Both $I_{i,j}$ and $S_{i,j}$ denote the number of infected and susceptible humans for state i at time j , respectively. The probability of contact between a susceptible human with infectious individuals is b . The infection probability for a susceptible human to become infected is β . The transmission coefficient of infectious individuals to the human population is denoted as $\frac{I}{N_i}$. Thus, the transmission rate or probability of susceptible individuals being infected is summarised as Equation 19.

$$\gamma = \frac{\beta b I}{N_i} \tag{19}$$

This stochastic component has been adapted from the original model designed by Reed-Frost in the 1920s for leptospirosis infection. It is reasonable to represent uncertainty or randomness in the leptospirosis infection model. This stochastic model includes a new term $I_{i,j}$ representing the number of newly infected leptospirosis cases in the human population. The time interval is Poisson stationary and the independent increment properties ($j-1$ and j) in the study region are represented by i . A stochastic process is a Poisson distribution process with random likelihood that may be analysed statistically but may not be predicted precisely. In this analysis, a random Poisson distribution will be used to generate the probabilities of outcomes of a system called the Monte Carlo simulation which includes a probability distribution to reflect the inherent randomness within the data.

For human leptospirosis, $I_{i,j}$ = the total number of newly infected human leptospirosis cases,

$$\begin{aligned} S_{i,j} &= \mu N_i + S_{i,j-1} - \mu S_{i,j-1} - \iota_{i,j} \\ \iota_{i,j} &\sim \text{Poisson}(\lambda_{i,j}) \\ \lambda_{i,j} &= [\exp(\beta_o + c_i)] \left(\frac{\beta b}{N_i} I_{i,j-1} S_{i,j-1} \right) \\ I_{i,j} &= (1 - \mu - \delta) I_{i,j-1} + \iota_{i,j} \\ R_{i,j} &= \delta I_{i,j-1} + (1 - \mu) R_{i,j-1} \end{aligned} \tag{20}$$

Based on Equation 20, β_o represents the constant term to describe the overall rates of the process. The intrinsic conditional autoregressive (CAR) priors are applied to fit the model for random effects c_i presented among the observed data and random effects for the unobserved variables in nearby areas. The discrete time discrete space stochastic SIR

model for human leptospirosis proposed in this study thereby used the estimation of relative risk, the results of which are discussed in the next section.

In this study, the intrinsic conditional autoregressive (CAR) priors are applied to fit the model for random effect c_i . The CAR-Normal distribution is used as the prior to allow for spatial dependence between the random effects c_i in nearby areas. These CAR priors were proposed by Besag et al. (1991), where the probability densities of values at any given location are conditional on the neighbouring areas. The CAR priors model used is as Equation 21.

$$[c_i | c_j, i \neq j, \sigma_c^2] \sim N(\bar{c}_i, \sigma_i^2) \tag{21}$$

where the mean

$$\bar{c}_i = \frac{1}{\sum_j \kappa_{i,j}} \sum_j c_j \kappa_{i,j} \tag{22}$$

and variance

$$\sigma_i^2 = \frac{\sigma_c^2}{\sum_j \kappa_{i,j}} \tag{23}$$

In this context, the κ_{ij} in Equation 22 and 23 are equal to 1 if i, j are adjacent and 0 if they are not adjacent. For a full Bayesian analysis, prior distributions must be specified for this hyperparameter. Here, this analysis considers gamma distribution for parameter σ_1^2 that controls the variability of c . The discrete time discrete space stochastic SIR model proposed in this study is used as an estimate of relative risk specifically for leptospirosis disease.

Generally, for $i = 1, 2, \dots, M$ study region and $j = 1, 2, \dots, T$ time periods, the observation sample used is the pseudo-random sample where λ_{ijk} for $k = 1, 2, \dots, n$ is generated from a posterior distribution for the mean number of leptospirosis infections, $\lambda_{i,j}$ in Equation 19. From this sample of the posterior mean number, the posterior expected mean is approximated using the Equation 24.

$$\tilde{\lambda}_{i,j} = \frac{1}{n} \sum_{k=1}^n \lambda_{ijk} \tag{24}$$

Meanwhile, the respective relative risk parameter is $\theta_{i,j}$ is defined as Equation 25.

$$Relative\ risk = \theta_{i,j} = \frac{Posterior\ mean\ number, \lambda_{i,j}}{expected\ number\ of\ new\ infectives, e_{i,j}} \tag{25}$$

Based on Equation 25, the pseudo-random sample generated for relative risk parameter is approximated as Equation 26:

$$\tilde{\theta}_{i,j} = \frac{1}{n} \sum_{k=1}^n \theta_{ijk} = \frac{1}{n} \sum_{k=1}^n \frac{\lambda_{ijk}}{e_{ijk}} = \frac{\tilde{\lambda}_{i,j}}{e_{i,j}} \tag{26}$$

Therefore, the posterior expected relative risk for leptospirosis, $\tilde{\theta}_{i,j}$ can be written as the posterior expected mean number of infections, $\tilde{\lambda}_{i,j}$ over the expected number of new infections, $e_{i,j}$ based on the human population of the study regions.

In this study, the value of relative risk is defined based on research by Samat and Percy (2012). A value of relative risk that is equal to zero means people within the study region have no relative risk or infection risk for the disease compared to the people in the population. However, the value of relative risk not ought to be zero since everybody in fact has the possibility or chance of contracting the disease. This is one major drawback of the SMR method that this present study aims to overcome. Next, when the value of relative risk is close to 1 this means that there is no significant difference between the conditional probability that a person in the study region contracts the disease and the conditional probability that a person in the general population contracts the disease.

On the other hand, if the value of relative risk is greater than 1, it indicates that the people within the study region are more likely to contract the disease compared to the people in the population. Conversely, a value of relative risk below 1 indicates a decrease in the likelihood of contracting the disease which means that the people in the study region are less likely to contract the disease compared to people in the population. All the traditional and proposed methods discussed in this study will be applied using WinBUGS software to process leptospirosis data from Malaysia. The results are presented in graphs and tables in the next section.

APPLICATION OF RELATIVE RISK ESTIMATION FOR LEPTOSPIROSIS DISEASE IN MALAYSIA

Application of Relative Risk Estimates based on Existing Models

This section discusses the results of the analyses using the SMR, Poisson-Gamma, Log-normal, BYM and mixture models and utilising leptospirosis data for all states in Malaysia from January 2011 to December 2018. Figure 4 depicts the time series plots of human leptospirosis cases reported for epidemiology months from January 2011 to December 2018 in four selected states with the highest cases throughout the period. It can be clearly seen that the state of Selangor had the highest number of leptospirosis cases in the early years. However, the trend changed in early 2015 as Kelantan recorded more cases than Selangor, after which point the numbers generally decreased in both states from 2016 to 2018. The average number of cases was between zero and 80 from 2011 until the end of 2014 and then rose sharply to 150 at the beginning of 2015, with Kelantan showing the highest number, 278 total cases, followed by Selangor with 143 cases in March 2015. Based on the frequency of reporting, it can be concluded that Selangor and Kelantan had the highest number of human leptospirosis cases. However, from a particular point of view, the higher number of cases reported in Figure 4 is most likely due to the greater

population density and the large size of the two states, which make the data inappropriate for representing true risk.

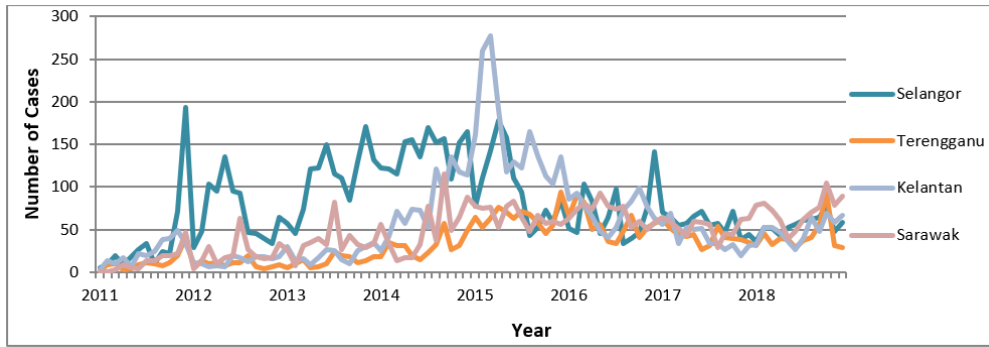


Figure 4. Numbers of human leptospirosis cases in Malaysia from 2011 to 2018

The results of relative risk for the selected highest-risk states using the SMR method from 2011 to 2018 are displayed in Figure 5. Most states have risk between zero to 2.5 except for Melaka (from the end of 2011 to early 2013), Kelantan (from 2014 to 2016) and Terengganu (from 2016 to 2018). This leads to a general conclusion that people in the state of Melaka, Kelantan, Terengganu, and Labuan are more likely to contract leptospirosis compared to the entire population. The risk of Labuan fluctuates over the years due to the high number of observed cases to expected value. Figure 4 shows Selangor frequently has a high number of human leptospirosis cases. However, by using the SMR model, Melaka has shown the highest risk, observed between early 2012 and early 2013 compared to the other states. This is followed by Kelantan and Terengganu in which states the risk began to increase in 2015 until 2018. Thus, the higher number of cases shown in Figure 4 does not reflect the risk of disease infection.

The time series pattern for posterior relative risk plotted for the Poisson-gamma model in Figure 6 is reasonably like the SMR in Figure 5. However, the SMR has a slightly higher range value of risk compared to Poisson-gamma risk. The range risk, which is the largest value minus the lowest value of the risk, in Figure 5 for the SMR is 9 (9 – 0). Meanwhile the range value of risk estimated from Poisson-gamma is 6.629 (6.664 – 0.035). The range is used to compare the spread of risk estimated from different models using a similar dataset. A large range of risk is assumed to have larger variability with more dispersion in the data and vice versa. However, a large range of a dataset may result from outliers. These outliers are unusual values of risk output probably caused by data entry errors, measurement errors or sampling problems. Outliers may also occur because of natural variation especially

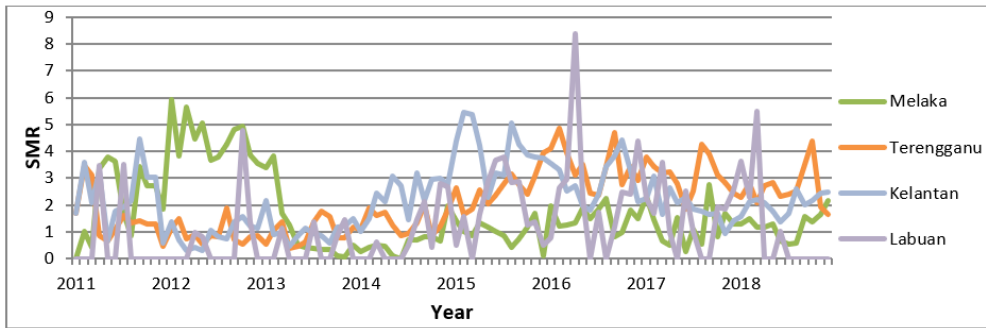


Figure 5. Estimates of relative risk for selected highest risk states based on the SMR of human leptospirosis, 2011 to 2018

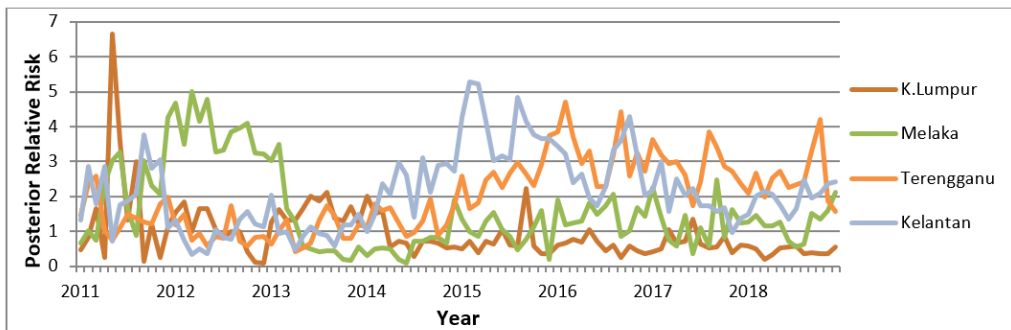


Figure 6. Estimates of relative risk for selected highest risk states based on the Poisson-gamma model of human leptospirosis, 2011 to 2018

when big data is involved. Outliers are useful for providing information about the highest and lowest variability risks inherent in the study areas. In this case, the estimated risk does not necessarily need to be precise towards a mean point. Risk is a variable that occurs naturally and removing outliers sometimes causes the subject areas to appear less variable than they are in reality.

It can be concluded that the use of the SMR offers a slightly higher variability than Poisson-gamma. Besides, the SMR has a drawback when dealing with zero cases. The risk estimated from a zero-case area will have zero risk estimated, as shown in Figure 5, which means that the area has no possibility for anyone there to catch the disease. But everyone has the possibility of being infected with the disease either low or high. Hence, Poisson-gamma in Figure 6 is used to overcome this shortcoming by estimating risk based on gamma posterior distribution that gives weightage for small areas with rare diseases (Lawson & Williams, 2003). The estimated risk shown in the time series graph is more realistic as there are no states with zero potential to contract the disease.

The time series pattern of posterior relative risk for the log-normal model in Figure 7 is quite like the Poisson-gamma model in Figure 6, with a slightly higher range value of risk. Melaka was still recorded as high risk from the end of 2011 to early 2013, while the risk for Kelantan and Terengganu rose gradually and consistently starting from 2014 up to 2018. This slight change of risk values was due to the spatial correlation in which the risk of an area will be affected by the risk from neighbouring areas such as between Kelantan and Terengganu. Furthermore, changes in range of risk may exist once the spatial random effects for correlated heterogeneity in the population are considered.

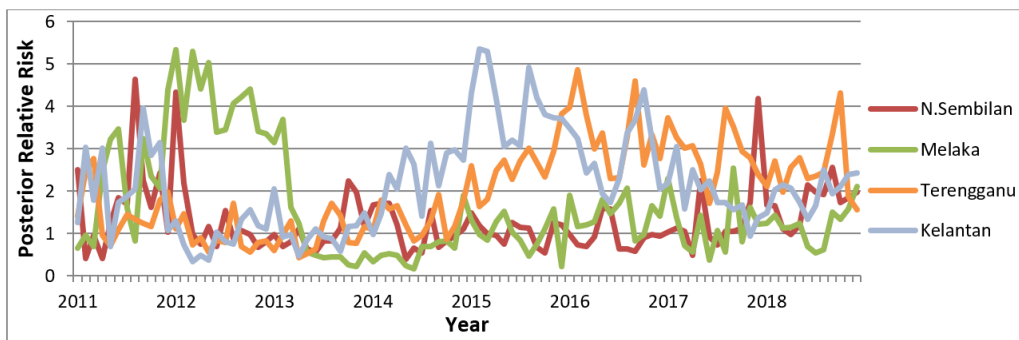


Figure 7. Estimates of relative risk for selected highest risk states based on the Log-normal model of human leptospirosis 2011 to 2018

Based on Figure 8, the time series plots for the BYM model are too smooth and the risk shows less variation than the observed risks for the SMR, Poisson-gamma and log-normal models. Some of the extreme values estimated from the SMR have disappeared. This is because the components in the model consider correlated and uncorrelated heterogeneity of area-specific random effects. The variability of risk estimated depends on uncorrelated heterogeneity. However, in this case variable u controlled by parameter τ_u^2 is small which causes lower variability of risk. Thus, the risk of a particular area is slightly like that of neighbouring areas. Melaka had the highest risk, which gradually decreased from 2011 to 2014, while the risk of Kelantan increased from 2012 to 2016. As the risk in Kelantan decreased, the risk in Terengganu increased from the end of 2016 until 2018.

As discussed before, extreme risk values estimated by the SMR often causes outliers. Without considering the natural variation, these outliers are believed to disfigure statistical analysis and perhaps contravene the assumptions. Subsequently, they produce unclear spatial patterns of risk if any mapping analysis is involved (Lawson & Williams, 2003). Therefore, many researchers have sought appropriate models (known as the smoothing approach) to overcome this problem by estimating the random effects. As a result, and using the same dataset, models such as Poisson-gamma, log-normal and BYM can overcome

the outliers and zero case problems of the SMR. Nevertheless, the BYM model shown in Figure 8 is smoothing over large discontinuities for risk estimated, which may be crucial to maintain. This means that some risk surfaces undergo over-smoothing and the risks produced might not depict the true risk. According to Lawson (2006), the BYM model is a smoothing model which is not designed specifically for cluster detection. Therefore, a mixture model was used to overcome this problem and simultaneously allow for smoothness and discontinuities.

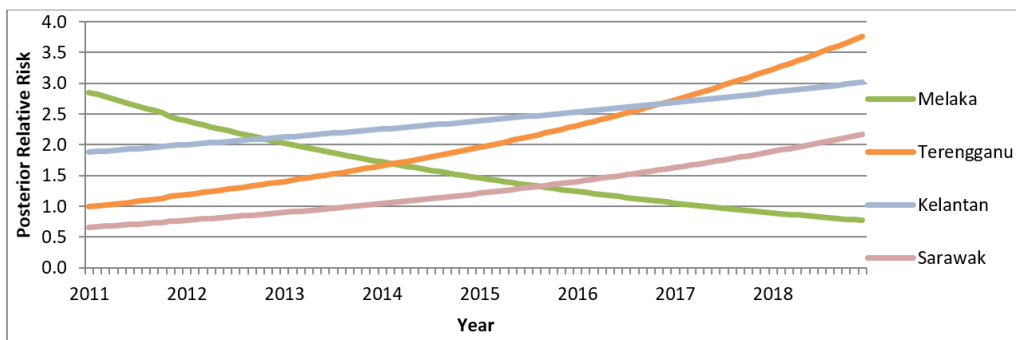


Figure 8. Estimates of Relative Risk for selected highest risk states based on the BYM model of Human Leptospirosis, 2011 to 2018

According to Figure 9, the state of Melaka recorded the highest risk from 2011 to early 2013 and was followed by Kelantan from 2014 to 2016 and Terengganu from 2015 to 2018. The pattern is still like the other observed method in which most states have a risk less than 2 with a small range of risk. For this method, the discontinuity of risks estimation is maintained, and the smoothness is employed at the same time.

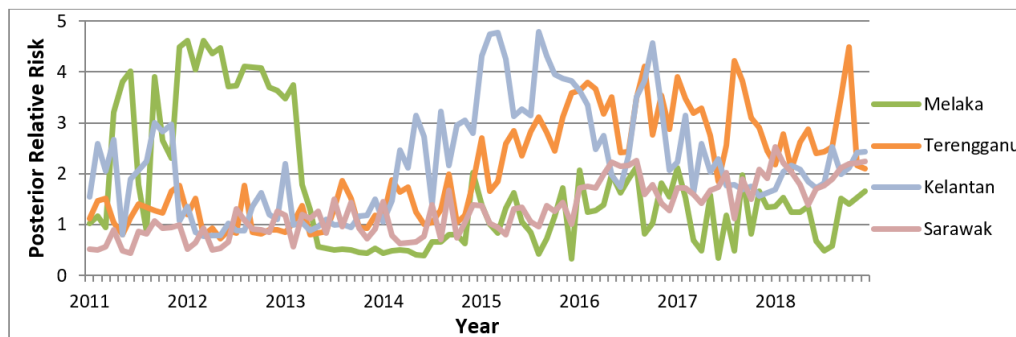


Figure 9. Estimates of relative risk for selected highest risk states based on the mixture model of human leptospirosis, 2011 to 2018

Application of Relative Risk based on Discrete Space-Time Stochastic SIR Transmission Model

The SIR transmission model proposed in this paper utilises the stochastic element in relative risk estimation for leptospirosis in Malaysia. The result of expected relative risk for posterior mean is plotted in the time-series graph in Figure 10.

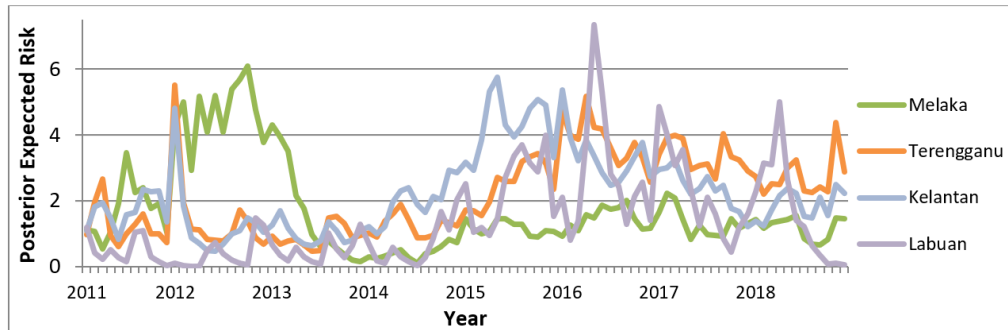


Figure 10. Expected relative risk for selected highest risk states based on the stochastic SIR model of human leptospirosis, 2011 to 2018

The average expected relative risk for most states based on the stochastic SIR transmission model for human leptospirosis is less than 2. It can be assumed that states with an average risk of under 2 are less likely to contract the disease compared to people in the overall population of Malaysia. According to the time-series graph plotted in Figure 10, Melaka, Kelantan, Labuan, and Terengganu have a higher risk compared to the other states. This indicates that these areas are more likely to contract leptospirosis compared to the total population of Malaysia. Besides that, in early 2011 and 2012, Terengganu and Kelantan had relatively high risk which gradually decreased over time. After that, starting in 2015 the risk increased again until 2018. An early sign, like this one from 2011 to 2012, should alert the authorities and communities to any potential increase in the risk of the disease occurring again in these areas.

Based on the expected relative risk between the existing and stochastic SIR models in this paper, it can be concluded that utilising different models produces different patterns of time-series for disease risk plotted. In line with the objective of this study, the proposed model is designed to yield information about areas of excess risk by using the transmission and stochastic approach naturally along with the uncertainty of leptospirosis infection behaviour.

Expected Relative Risk Results of Leptospirosis based on Existing Models and Discrete Space-Time Stochastic SIR Model for Human Leptospirosis Disease

Table 1 shows the comparison of estimated expected relative risk results for the epidemiology month of May 2014, based on the SMR, Poisson-gamma, log-normal, BYM, mixture models and the discrete time-space stochastic SIR model for human leptospirosis. This epidemiology month was chosen to clearly demonstrate the result of posterior expected mean of relative risk for the existing and stochastic SIR models of leptospirosis transmissions.

Table 1

Estimates relative risk for the epidemiology month of May 2014

No.	States	Number of Cases	SMR	Poisson-Gamma	Log-normal	BYM	Mixture	SIR
1	Perlis	0	0	0.3613	0.4097	0.4513	0.3352	0.3736
2	Kedah	1	0	0.0866	0.1303	1.1550	0.5979	0.6504
3	Pulau Pinang	10	0	0.3817	0.3831	0.2892	0.3589	0.1786
4	Perak	59	2	1.7080	1.7040	0.9622	1.4970	1.9580
5	Selangor	155	2	1.8320	1.8330	1.0580	1.5600	1.6380
6	Kuala Lumpur	17	1	0.7317	0.7148	0.9621	0.6821	1.2600
7	Putrajaya	0	0	0.6574	0.6385	0.6288	0.2548	0.0021
8	Negeri Sembilan	5	0	0.4025	0.4041	1.2220	0.7628	1.7090
9	Melaka	1	0	0.1954	0.2420	1.6290	0.4119	0.2621
10	Johor	20	0	0.4199	0.4201	0.2945	0.3993	0.3088
11	Pahang	10	0	0.4951	0.4882	0.8887	0.6389	0.9885
12	Terengganu	20	1	1.2360	1.2110	1.7550	1.2380	1.8860
13	Kelantan	74	3	2.9650	3.0060	2.2990	3.1500	2.3080
14	Sarawak	17	0	0.4831	0.4800	1.0970	0.6566	0.6304
15	Sabah	39	1	0.7736	0.7683	0.6225	0.8922	1.0860
16	Labuan	0	0	0.6573	0.6498	1.4050	1.3820	0.3078
	DIC	-	-	55.264	55.745	56.325	54.197	62.524

Table 1 indicates that Selangor had the highest number of leptospirosis cases in May 2014, 155, followed by Kelantan with 74 leptospirosis cases and Perak with 59 cases. However, expected relative risks estimated by all the existing and proposed models show that the state with the highest risk of catching the disease was Kelantan instead of Selangor. Intrinsically, the higher number of cases in the state of Selangor is because of the large population size and area, and not because of the high-risk occurrence. Meanwhile, the high risk for Kelantan is due to the flood season that affects the north-eastern states of Malaysia (Garba et al., 2017b).

From Table 1, the expected relative risk for the mixture model shows the state with the highest risk is Kelantan with relative risk value 3.15 followed by Selangor, 1.56 and Perak, 1.497. Since Kelantan is the highest risk, the nearest area; Terengganu is expected to have relatively high risk. However, in this case, Terengganu shows low risks of infection which is 1.238. Similarly, with the BYM model, Kelantan is at the highest risk of leptospirosis infection followed by Terengganu and Melaka. In this case, Kelantan neighbours Terengganu, while Melaka is located further away and does not have spatial correlation between those high-risk states. This indicates that these models do not appear to be robust in terms of the covariate element proposed when neighbouring areas have poor relatively influenced by the highest risk adjacent areas. The posterior expected relative risk for the stochastic SIR, on the other hand, shows Kelantan as at higher risk followed by Perak and Terengganu. Geographically, these three states abut against each other and have relatively correlated heterogeneity.

The smallest DIC value in Table 1 is the mixture model with a difference of approximately 1 to 2 with other models. However, it is difficult to conclude that the mixture model is the model most fitted to the data since the DIC value is not greatly dissimilar to the other existing Bayesian models. Next, the DIC value for the stochastic SIR model is 62.524 which is not too far from the others. However, as discussed before, comparing different mathematical models may not be appropriate and the interpretation of DIC values may be misleading. The DIC value comparison for the stochastic SIR model is more suitable compared to the same properties of transmission models. Therefore, regardless of the DIC value, the selection of the model is dependent on the objectives of the study and the availability of data.

CONCLUSION AND FUTURE WORK

In this study, the expected relative risk estimates from existing models have demonstrated high-risk areas contracting leptospirosis across Malaysia from 2011 to 2018. Each model gives different estimates of relative risk based on the components used for estimates of random variables in spatial units, such as an effort to overcome the drawbacks of the SMR. Next, in line with the objective of this study, expected relative risk is derived from

the transmission stochastic SIR model as an estimate of risk. The results disclosed areas at high risk of contracting leptospirosis disease for states in Malaysia which is significant for surveillance systems and control measures.

Compared with existing methods of relative risk estimates, the stochastic SIR model is more relevant and in line with infections due to leptospirosis transmission. Infectious disease is naturally transmitted to susceptible individuals from infectious individuals which finally recovered from the disease. In epidemiology studies, having this transition enables researchers to determine the target population for control measures. For example, susceptible individuals which comprise different age-groups will have different rates of infection. The high infection age-group/s can be easily targeted to stop the transmission. In another case, for different groups of the sub-population, only certain groups are vaccinated to prevent disease outbreak. This indicates the importance of transmission systems in the study of infectious disease. Next, the expected relative risk estimates from newly infective persons are based on the stochastic process. It is important to consider the probability or uncertainty of disease infection so the risk estimates can be more flexible depending on the behaviour of the disease and infection.

However, in this study, the stochastic SIR model only involves human transmission. Since leptospirosis disease originates from rodents, especially rats, it will be more specific and realistic if transmission from the rat population is considered in the model. An extension to this work consisting of a rat population in the transmission model should be the subject of future study. Furthermore, the application of risk would be more effective with the presentation of a risk map.

ACKNOWLEDGEMENT

The authors would like to thank and acknowledge the Institute of Research Management & Innovation (IRMI), Universiti Teknologi MARA (UiTM) for providing FRGS research fund (600-IRMI/FRGS 5/3 (388/2019)) and Ministry of Health Malaysia for the provision of human data infected by leptospirosis disease.

REFERENCES

- Andrade, C. (2015). Understanding relative risk, odds ratio, and related terms: as simple as it can get. *The Journal of clinical psychiatry*, 76(7), 857-861. <https://doi.org/10.4088/JCP.15f10150>
- Awang, A. C., & Samat, N. A. (2017). Standardized morbidity ratio for leptospirosis mapping in Malaysia. In *AIP Conference Proceedings* (Vol. 1847, No. 1, p. 020006). AIP Publishing LLC. <https://doi.org/10.1063/1.4983861>
- Benacer, D., Thong, K. L., Verasahib, K. B., Galloway, R. L., Hartskeerl, R. A., Lewis, J. W., & Mohd Zain, S. N. (2016). Human leptospirosis in Malaysia: reviewing the challenges after 8 decades (1925-2012). *Asia Pacific Journal of Public Health*, 28(4), 290-302. <https://doi.org/10.1177/1010539516640350>

- Besag, J., York, J., & Molli, A. (1991). Bayesian image restoration, with two applications in spatial statistics. *Annals of the Institute of Statistical Mathematics*, 43(1), 1-20. <https://doi.org/10.1007/BF00116466>
- Brauer, F. (2017, February 4). *Mathematical epidemiology: Past, present, and future*. Retrieved August 29, 2020, from <https://www.ncbi.nlm.nih.gov/pubmed/29928732>
- Dey, S., Delampady, M., & Gopalaswamy, A. M. (2019). Bayesian model selection for spatial capture–recapture models. *Ecology and evolution*, 9(20), 11569-11583. <https://doi.org/10.1002/ece3.5551>
- Fletcher, W. (1928). Recent work on leptospirosis, tsutsugamushi disease, and tropical typhus in the federated Malay States. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 21(4), 265-288. [https://doi.org/10.1016/s0035-9203\(28\)90019-x](https://doi.org/10.1016/s0035-9203(28)90019-x)
- Garba, B., Bahaman, A. R., Bejo, S. K., Zakaria, Z., Mutalib, A. R., & Bande, F. (2017a). Major epidemiological factors associated with leptospirosis in Malaysia. *Acta Tropica*, 178, 242-247. <https://doi.org/10.1016/j.actatropica.2017.12.010>.
- Garba, B., Bahaman, A. R., Khairani-Bejo, S., Zakaria, Z., & Mutalib, A. R. (2017b). Retrospective Study of Leptospirosis in Malaysia. *EcoHealth*, 14(2), 389-398. <https://doi.org/10.1007/s10393-017-1234-0>
- Haake, D. A., & Levett, P. N. (2014). Leptospirosis in humans. *Current Topics in Microbiology and Immunology*, (387), 65–97. https://doi.org/10.1007/978-3-662-45059-8_5
- Ideris, S. H., & Samat, N. A. (2015). Comparison of HIV and AIDS diseases mapping in Malaysia based on standardized morbidity ratio and poisson-gamma model. *EDUCATUM Journal of Science, Mathematics and Technology*, 2(1), 69-81.
- Kermack, W. O., & McKendrick, A. G. (1927). A contribution to the mathematical theory of epidemics. Proceedings of the royal society of london. *Series A, Containing Papers of a Mathematical and Physical Character*, 115(772), 700-721. <https://doi.org/10.1098/rspa.1927.0118>
- Kismiantini. (2009, January 25). *Relative risk of disease using linear generalized mixed model*. Staff Site Universitas. Retrieved August 29, 2020, from <http://staff.uny.ac.id/sites/defaultfiles/penelitian/Kismiantini,S.Si.,M.Si/A2009B1.pdf>
- Lawson, A. B. (2006). *Statistical methods in spatial epidemiology*. John Wiley & Sons.
- Lawson, A. B., & Williams, F. L. (2003). *An introductory guide to disease mapping*. John Wiley.
- Long, N. X. (2012). *Clustering problems, mixture models and Bayesian nonparametrics*. Lecture presented at Vietnam Institute of Advanced Studies of Mathematics (VIASM), Hanoi. Retrieved December 05, 2020, from http://dept.stat.lsa.umich.edu/~xuanlong/Talks/nguyen_viasm12_part1.pdf
- Meza, J. L. (2003). Empirical Bayes estimation smoothing of relative risks in disease mapping. *Journal of Statistical Planning and Inference*, 112(1-2), 43-62. [https://doi.org/10.1016/S0378-3758\(02\)00322-1](https://doi.org/10.1016/S0378-3758(02)00322-1)
- Nurmalasari, M., & Pramana, S. (2014). Analysis of appendectomy in Belgium using disease mapping techniques. *Proceedings of the 13th Islamic Countries Conference on Statistical Sciences*, 27, 417-430.
- Pooley, C. M., & Marion, G. (2018). Bayesian model evidence as a practical alternative to deviance information criterion. *Royal Society Open Science*, 5(3), Article 171519. <https://doi.org/10.1098/rsos.171519>

- Richardson, S., Thomson, A., Best, N., & Elliott, P. (2004). Interpreting posterior relative risk estimates in disease-mapping studies. *Environmental Health Perspectives*, *112*(9), 1016-1025. <https://doi.org/10.1289/ehp.6740>
- Samat, N. A., & Percy, D. F. (2012). Vector-borne infectious disease mapping with stochastic difference equations: An analysis of dengue disease in Malaysia. *Journal of Applied Statistics*, *39*(9), 2029-2046. <https://doi.org/10.1080/02664763.2012.700450>
- Spiegelhalter, D., Thomas, A., Best, N., & Luun, D. (2003). *WinBUGS user manual version 1.4*. MRC Biostatistics Unit.
- Sulong, M. R., Shafei, M. N., Yaacob, N. A., Hasan, H., Daud, A., Mohamad, W. M. Z. W., Zaliha, I., & Mohamed, R. A. (2011). Risk factors associated with leptospirosis among town service workers. *International Medical Journal*, *18*(2), 83-88.
- Triampo, W., Baowan, D., Tang, I. M., Nuttavut, N., Wong-Ekkabut, J., & Coungchawee, G. (2008). A simple deterministic model for the spread of leptospirosis in Thailand. *International Journal of Medical and Health Sciences*, *2*(1), 22-26.
- Wahab, Z. A. (2015, September 8-9). Epidemiology and current situation of leptospirosis in Malaysia. In *Persidangan Kesihatan Persekitaran Pihak Berkuasa Tempatan 2015*. WP



Factors in Ergonomic Design of 6-to-18-month Baby Carriers for Elderly People

Ariya Atthawuttikul^{1*} and Sorat Khongkharat²

¹*Department of Product Design, Faculty of Decorative Arts, Silpakorn University (Sanam Chandra Palace Campus), 6 Rajamankha Nai Road, Muang District, Nakhon Pathom 73000, Thailand*

²*Research Center of Trend and Design, Baramizi Lab, Baramizi Co., Ltd., Sukhumvit 63, Khlong Tan Nuea, Watthana, Bangkok 10110, Thailand*

ABSTRACT

The main objective of this study was to formulate an ergonomic design factor for 6-to-18-month baby carrier from simulated loading point data and physical parameters of an average Thai elderly person. The research steps were as follows: (1) reviewing the literature on 6-to-18-month baby carrier designs; (2) simulating the forces acting on a few selected baby carriers and on the body of an average Thai elderly person carrying them; and (3) synthesizing ergonomic design factors based on load-bearing data as well as the physical parameters of an average Thai elderly person. Four ergonomic design factors were synthesized: 1) their shoulder and hip posture while they are using the carrier, 2) the seat and backrest of the baby carrier, 3) the load-bearing points on the body of the carrying person, and 4) the type of baby carrier. Our findings may benefit designers of cushion seat and baby stroller.

Keywords: 6-to-18-month, baby carrier, ergonomic design

ARTICLE INFO

Article history:

Received: 4 April 2020

Accepted: 27 July 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.21>

E-mail addresses:

atthawuttikulariya@gmail.com (Ariya Atthawuttikul)

sorat.khong@gmail.com (Sorat Khongkharat)

*Corresponding author

INTRODUCTION

Traditional lifestyle of people in Thailand has been drastically changed, socially and economically, by modern advances in science, technology, and innovation. The changes have exerted tremendous impacts on families, the smallest unit of the society, in terms of family structure, size, and lifestyle as well as the relationships among family members (Photisita, 2009),

especially on big families where grandparents and young children live together (Husserl, 1965). Boonkwang and Ayuwat (2017) stated that the movement of laborers from the country to big cities to work had resulted in isolation of members of the older generation. Many workers have also put their children into the sole care of their grandparents. This trend is expected to be ever increasing in the rural areas of Thailand.

Temporary childcare strongly affects the lifestyle of elderly people. Komjakkraphan and Chansawang (2015) found that currently, for many families in Thailand, childcare had been relegated solely to elderly members of the family. Moreover, they may also have to earn some auxiliary income for the family in addition to taking care of their own affairs. Jendreck (1993) stated that as elderly people took on the role of temporary parents; they were affected in four ways: (1) they became exhausted; (2) they lost their relationship with their friends; (3) they might have less time to spend with their family members; and (4) misunderstanding might occur between elderly people and their mates. To conclude, Thai elderly people among elderly populations of countries around the world, have a distinctive lifestyle of obligated to raise their grandchildren in a modern rural society. Thus, they need good tools that will help perform this task effectively and conveniently. An ergonomic baby carrier is one such tool for Thai elderly people.

World Health Organization has released the following statistics: one-third of the elderly population slips and falls at least once in a year, and the trend is that 28-35% of elderlies over 65 years of age will slip and fall in a year. This percentage increases to 32-42% for elderlies 70 years of age. As people grow old, they suffer calcium loss from bones, making them fragile. In combination with weaker muscles, their bones are easily broken from a physical (Atthawuttikul & Chavalkul, 2018). It is expected that Thailand will be faced with an aged society in 2021. This means that Thailand has a very short time to prepare a response to demographic change. Therefore, lifestyle and social restructuring, including collaboration of all sectors, will need to support the elderly society and relationship to a family this situation (Punyakaew et al., 2019).

A good relationship between parents and children, from infant to three years old, helps stimulate the children's development. It helps promote development of their physical abilities such as maintaining a good balance, making precise movement, practicing recognition by the sense of touch, coordinating the eyes and the hands to do certain activities, and developing gross motor skills (Elaine et al., 2017), as well as their mental abilities such as recognition, observation, and memorization of objects and their surroundings, analytical ability, and communication by physical actions (Thanatchang, 2007). Figure 1 shows three types of baby carriers for 0-to-36-month-old babies (Pollack-Nelson, 2000; Cirelli, 2014). Parents are instinctively attached to infants by nature, whereas elderly people taking the role of the parents are not by nature. Nevertheless, they provide love and care for the children as surrogate parents.

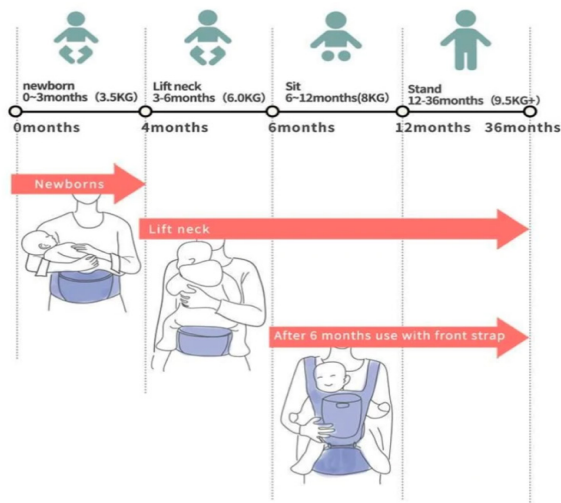


Figure 1. Three different types of baby carrier for 0-to-36-month-old babies. (Infantino Baby Carriers, 2020)

In taking care of small children and doing some activities together, oftentimes, elderly people need to carry the child or support its weight. A baby carrier is divided into 5 types: (1) stretchy wrap; (2) ring sling; (3) woven wrap; (4) Mei-tai; and (5) structured carrier (Dr. Sears Wellness Institute, 2019), as shown in Figure 2 and Table 1. The authors were interested in investigating the structured carrier which should be suitable for babies at 6 to 18 months of age and older till they reached a weight of 18-20 kg. For elderly people, this kind of carrier has a wide and padded hip belt that can distribute the weight of the baby to the hip and not to their shoulders. Andersson (2019) stated that a child should sit in an ergonomic posture, a frog posture, for the parent. Its weight should be supported by the pelvis and thighs of the parent. The back of the parent also supports the baby weight. A parent can place a baby in a tummy-to-tummy position or back-to-tummy position. Some carriers also let the parent place a baby on his or her hip. Structured carriers are among the most comfortable carriers, especially if one still intends to carry a child when it gets bigger and heavier.

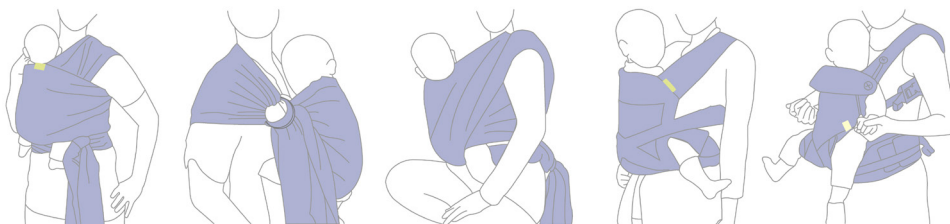


Figure 2. Five common types of baby carrier. (Andersson, 2019)

The physiological postures of child carrying are of 4 types; (1) front-carry (facing-in), (2) front-carry (facing-out), (3) back-carry, and (4) hip-carry, as shown in Figure 2. Baby carriers for these different postures are of different designs. Since a 6-18-month-old baby should have opportunities to be stimulated by its parent and the environment, the design of a carrier should allow a baby those opportunities, i.e., the carrier should have a padded frame that can support the baby for a long time without tiring it, and the carrier should let the baby face out toward the environment and away from the torso of the parent.

Table 1

References for types of baby carrier research

Type of baby carriers	References
1. Stretchy wrap	Fista et al., 2019
2. Ring sling	Fista et al., 2019
3. Woven wrap	Fista et al., 2019; Brown et al., 2018; Jones, 2017
4. Mei tai	Kim & Yun, 2013
5. Structured carrier	Bleah & Ellett, 2010; Fista et al., 2019; Brown et al., 2018; Lee & Hong, 2018

The structure of a baby carrier today has not been designed with the ergonomic of elderly people in mind because the shape of the carrier does not fit their body well (Bleah & Ellett, 2010). In addition, there is a relatively wide gap between the carrier and the person carrying the baby, hence it is possible that the baby may wriggle out of the cradle of the carrier and fall. This is especially dangerous for a small baby because it still cannot hold its posture tightly enough, and when it must put its weight on the seat of the carrier for a long time, it may wriggle or limp and fall. Today's carrier is also not so ergonomic for the elderly person carrying the baby because the weight of the baby is supported only by the torso of the elderly person. Moreover, for a young baby, its neck cannot be held rigidly by itself, so the elderly person carrying the baby may have some difficulties at handling the neck of the baby so that it does not grate with the side bands of the carrier, i.e., the baby may get injured, and it is difficult for an elderly person carrying it to prevent that.

Therefore, good design factors for a 6-18-month-old baby carrier for an elderly person play an important role in preventing baby's neck grating injury and breathing difficulty. The weight distribution of the carrier on the body of the elderly person carrying the baby is also needed to be aligned according to the ergonomic of the elderly person and anyone who may carry the baby. Finally, good design factors should enable a baby to fully develop, both physically and mentally.

The first main objective of this study was to determine, by computer simulation, high loading points of 4 baby carrier designs and loading points on a person carrying each one

of them. The second was to use the obtained loading points and physical parameters of elderly people to synthesize ergonomic design factors of baby carriers for elderly people.

METHODS

Theoretical Consideration

Four types of 6-to-18-month baby carriers for elderly people were gleaned from Andersson, 2019 and considered in this study. Front-carry (facing-out) type was considered the best type because it allows the baby to observe and learn the world in front of it. Hence, the authors focused on the design of this type of baby carriers shown in Figure 3.

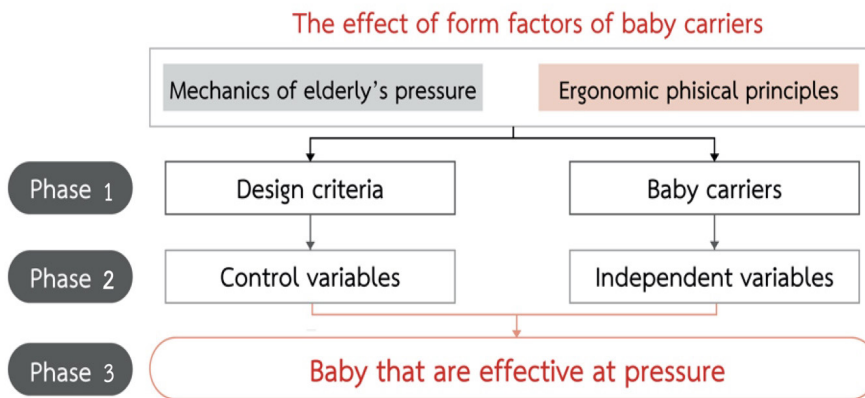


Figure 3. Methods theoretical framework.

Baby Carriers. Three front-carry baby carriers were purchased from a shop in Bangkok, Thailand. The three baby carriers: D1, D2, and D3 shown in Figure 4, were of different designs. All three were imported from England and rather quite expensive compared to other baby carriers that the shop was offering.



Figure 4. Infantino brand 6-to-8-month baby carrier. (Baby Best Buy in Thailand. 2020)

Sampling Method. A questionnaire survey was conducted on samples of a population of prospective consumers of baby carriers, a Thai border senior population-population of seniors with good physical conditions but still at risk of some disability, as defined by Okochi et al., (2005) in his classification of seniors. Samples were purposely selected from a senior care center in Ladkrabang district, Thailand. Living in a suburb of Bangkok Metropolitan, their viewpoints and attitudes on living and lifestyle were various. They were life-embraced elders as defined by Komjakkraphan & Chansawang (2015). The items in the questionnaire were questions on the participant's demographic information, on their use of baby carrier in everyday life, and on baby carrier usage issues that they encountered. Their responses to the questionnaire were crucial information for our synthesis of ergonomic design factors of baby carriers for elderly people.

Simulation. Three-dimensional models of baby carriers were constructed with 3D Studio Max software. The models were then inputted into ANSYS Engineering Simulation & 3D Design Software. ANSYS determined the load of a baby's weight on any point of the three-dimensional model of each carrier. The main purpose was to find the point on each carrier that bore the maximum load, which would be key data for synthesizing factors in ergonomic design. ANSYS is a force-simulation program used by many big corporations such as Toyota, Honda, and Hyundai to simulate the designs of their products.

Procedure. (i) The first step of this work was to consider all available choices of 6-18 months-old baby carriers reported in the research literature and commercial market to find several suitable carriers to test. In the end, three facing-out carriers were selected. They were the top three models, most popularly picked by Thai consumers according to a survey by a website called 'Baby Best Buy in Thailand (2020)'. They were also the three top-selling items.

(ii) Next is the construction of 3D models of the carriers. The models were constructed with 3D Studio Max software. Images of the models—D1, D2, and D3—are shown in Figure 5. They differed in the shape of the and the attachment locations of the shoulder straps. The dimensions and geometry of the models were identical to the real carriers, and all physical differences between the carriers were accurately incorporated into the models.

(iii) The three 3D models were imported into ANSYS force simulation software along with all necessary physical parameters such as earth gravity constant = $\sim 9.600 \text{ N.m/s}$ (Brown et al. 2018), stiffness = $\sim 0.40 \text{ MPa.}$, flexibility = $\sim 0.26 \text{ MPa.}$, and durability = 1.6 MPa. (Michael & Ash, 2007) against impact force.

(iv) The simulation software was run on a PC on Windows platform with an i5 CPU and 8 gigabytes of RAM. High and low loading points were represented by different colors

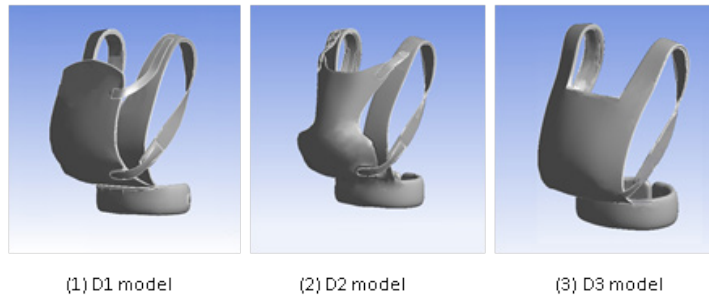


Figure 5. Dimensions and geometry of D1, D2, and D3 simulated baby carrier models.

in the simulated images of the models. The highest loading points are represented by red color, and the lowest loading points are represented by blue color.

(v) The nature of the loading-points of all 6-to-18-month carriers' baby from ANSYS force simulation results was analyzed, and the components of the carrier associated with low loading-points were included in the ergonomic design factors.

RESULTS AND DISCUSSION

Table 2 shows the results of our simulation experiment—simulated images of three baby carriers and their load-bearing points along with the materials and physical dimensions of their parts. It can be observed that D3 was able to support the highest load, while D1 supported the lowest load. The reason that D3 performed better than D1 and D2 is that it did not have as many weak points (areas supporting a high load) as the other two (Fista et al., 2019). Incidentally, the high load areas on the person wearing a carrier found from the simulation output were quite consistent with the responses to an item in the questionnaire that the participants have experienced pains in their shoulders, waist, and torso. Taking this simulation output into account, we synthesized factors in ergonomic design for Thai elderly people.

The four main factors in ergonomic design for elderly people were the following: (1) their shoulder and hip posture while they were using the carrier; (2) the seat and backrest areas discussed by Moore et al. (2012) as load-bearing points of baby carriers; (3) the load-bearing points on their body as investigated by Glover (2012); and (4) the type of baby carrier.

Regarding (1), from a survey by Punyakaew et al. (2019), eight elderly people 62 to 85 years of age (a mean of 70) offered their opinions that the carrier did not fit the baby well—the baby's head could not be positioned above the carrier's cover, and it might get suffocated. Another survey by Brown et al. (2018) reported an opinion that the shoulder strap of a baby carrier was not strong enough or did not fit the physique of elderly people well which made the elderly people fear that the baby might fall from the baby carrier.

Table 2
 Simulated images and load-bearing points of baby carriers of 3 types (areas color-coded by load magnitude)

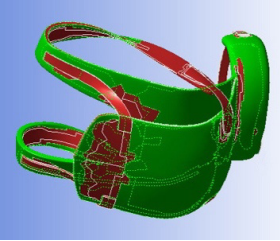
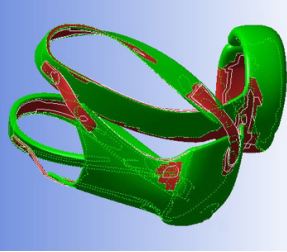
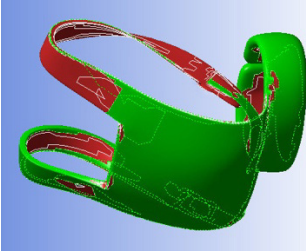
Item	Model	Material	Size (inch.)	Weight (g)	Load bearing-point
D1		<p>Main materials: 60% cotton and 40% polyester Lining: 100% cotton Waist belt: 100% polyester</p>	<ul style="list-style-type: none"> - Waistband thickness 4" - Width of shoulder straps 3" - Main Body including waistband 18" tall, 13.5 wide - Hip Strap 28" to 53" with 27.5" of padding - Shortest possible length of shoulder straps 17" - Fit baby from 8-40 mm. 	995	Shoulder, back, and pelvis belt
D2		<p>Main material: 50% cotton and 50% polyester Lining: 100% cotton Waist belt: 100% polyester</p>	<ul style="list-style-type: none"> - Waistband thickness 4" - Width of shoulder straps 3" - Main body including waistband 21" tall, 19" wide - Hip Strap 28" to 32" with 12.5" of padding - Shortest possible length of shoulder strap 15" - Fit baby from 15-45 mm. 	993	Shoulder, chest, and pelvis belt

Table 2 (Continued)

Item	Model	Material	Size (inch.)	Weight (g)	Load bearing-point
D3		<p>Main material: 40% cotton and 60% polyester Lining: 100% cotton Waist belt: 100% polyester</p>	<ul style="list-style-type: none"> - Waistband thickness 4" - Width of shoulder straps 3" - Main body including waistband 15" tall, 18" wide - Hip Strap 28" to 32" with 12.5" of padding - Shortest possible length of shoulder strap 15" - Fit baby from 20-60 mm. 	997	Shoulder and back belt

Green areas are areas of low load; red areas are areas of high load.

Additionally, another survey by Jones (2017) reported an opinion that the shoulder straps were too long and awkward so that the baby carrier was not convenient to use.

Regarding (2), seat and backrest were discussed by Moore et al. (2012) as load-bearing points of baby carriers. From our simulation, the D3 carrier (H-type baby) was the best at supporting a load. In our opinion, this was because the back-cushion seat of D3 was very large and wide compared to those of D1 and D2 (H-hip type). This finding is not in accordance with a report by Lee & Hong (2018) that X-type baby carriers were able to support a load well, especially the shoulder straps.

Regarding (3), Glover (2012) investigated load-bearing points on the body of elderly people and reported that the dimensions of the soft slings used for shoulder attachment were not ergonomic for elderly people carrying a baby. From our simulation, the high loading points (the red areas in the simulated image in Table. 1) on an average Thai elderly person were at his or her shoulder, back, and hips which might cause the person shoulder, back, and hip pain.

Finally, regarding (4), Bleah and Ellett (2010) reported that a structured carrier compressed the chest area of the baby, making it difficult for it to breathe, but helps it to learn and remember the environment better than other types of baby carriers. They went on and stated that the main issue with an elderly person carrying a baby was that he or she was likely to feel pain at the shoulder and waist. Some may even get cuts at their forearm and wrist from scraping the edge of the shoulder straps as they are trying to maneuver a baby into the carrier.

We had expected that the findings from this study would help considerably in the design of these types of baby carriers. However, the findings did not give us conclusive evidence for either positive or negative effect of the design of the current face-out-baby-carriers on the ergonomic for elderly people or their posture. The evidence for a positive effect of a type of baby carrier was not definite, only suggestive, and the lack of reproducibility of the results prevented a firm conclusion from being drawn. Similarly, the results of the three simulation studies investigating the factors in ergonomic design of 6-18-month baby carriers for elderly people were inconclusive. No design provided decisively better ergonomic factors than the other two. Nonetheless, the investigation provided an interesting basis for further study. The simulations pointed out causal relationships between the types of baby carrier, ergonomic design principles, and elderly people's loading points at their chest, shoulder tip, back, knee, and ankle in their standing posture.

CONCLUSION

Our main objective was to investigate by simulation the load-bearing points of several types of baby carriers and the load-bearing points on the body of an average Thai elderly people. Mainly based on the simulation results, we synthesized four factors in ergonomic design of

baby carriers for Thai elderly people. We obtained valuable information on loading points of the baby carriers and on the body of an average Thai elderly person, which to a certain extent, helped us to synthesize these factors. According to these factors, out of the three baby carrier designs that we investigated, the facing-out, H-type baby carrier design was the best ergonomic design for an average Thai elderly people. A good baby carrier design should have straps that fit a baby snugly, a large, but with good fit, back-cushion seat, and not-too-long shoulder straps. Our findings may directly benefit designers of baby carrier, car seat, and baby stroller in their effort to make their design ergonomic. Future studies should investigate the ergonomic design of cushion seat of baby carriers for elderly people as well as the of back-cushion seat, waist belt, and shoulder straps of baby carriers.

ACKNOWLEDGEMENTS

This research work was supported by the National of Research Council of Thailand (NRCT) as a scholarship for the first author as a lecturer at Rajamangala University of Technology Tawan-Ok (Uthenthawai Campus), specifically the industrial Design Department, Faculty of Engineering and Architecture. The authors would also like to thank Mr. Pratana Kangsadal for his help in revising the English of this paper.

REFERENCES

- Andersson, V. (2019). *5 types of baby carriers – which one is right for your baby?* Retrieved May 6, 2020, from http://whatbabieslove.com/types-of-baby-carriers/#1_Stretchy_Wrap.
- Atthawuttikul, A., & Chavalkul, Y. (2018). The effect of form factors of shock absorption Mat's subunits on attenuation of force from elderly's fall. *The Open Construction and Building Technology Journal*, 2018(12), 350-361.
- Baby Best by in Thailand. (2020). *Top three models for baby carriers*. Retrieved May 1, 2020, from <https://www.babybestbuy.in.th/shop/>
- Bleah, D. A., & Ellett, M. L. (2010). Infant crying among recent African immigrants. *Health Care for Women International*, 31(7), 652-663.
- Boonkwang, P., & Ayuwat, D. (2017). Household security of Isan migrant workers working abroad. *Nakhon Phanom University Journal*, 7(2), 108-115.
- Brown, M. B., Digby-Bowl, C. J., & Todd, S. D. (2018). Assessing infant carriage systems: ground reaction force implications for gait of the caregiver. *Human factors*, 60(2), 160-171.
- Cirelli, L. K., Wan, S. J., & Trainor, L. J. (2014). Fourteen-month-old infants use interpersonal synchrony as a cue to direct helpfulness. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), 1-8.
- Dr. Sears Wellness Institute. (2019). *Benefits of baby wearing*. Retrieved June 19, 2020, from <https://www.askdrsears.com/topics/health-concerns/fussy-baby/baby-wearing/benefits-babywearing>

- Elaine, K. H., Tham., Nora Schneider., & Birit, FP. Broekman. (2017). Infant sleep and its relation with cognition and growth: A narrative review. *Nature and Science of Sleep*, 9, 135-149.
- Fista, B., Widyanti, A., Muslim, K., & Salma, S. A. (2019, August). Evaluation of Baby Carriers in Indonesia: Physiological and Biomechanical Approach. *IOP Conference Series: Materials Science and Engineering*, 598(1), 1-6.
- Glover, R. (2012). Research overview: Is there evidence to support the use of soft slings? *Perspective-NCT's Journal on Preparing Parents for Birth and Early Parenthood*, 16, 18-20.
- Husserl, E. (1965). *Phenomenology and crisis of philosophy*. New York, NY: Harper Torchbook.
- Infantino Baby Carriers. (2020). *Three different types of baby carrier for 0-to-36-month-old babies*. Retrieved May 6, 2020, from <https://www.martpanda.co.uk/products/infantino-baby-carrier>
- Jendreck, M. P. (1993). Grandparents who parent their grandchildren: Effects of lifestyle. *Journal of Marriage and Family*, 55, 609-621.
- Jones, M. (2017). *Baby carriers—and ancient tool adapted to modern times—offer parents a way to bond with their child and express their individuality*. Retrieved May 1, 2020, from <https://www.sapiens.org/culture/babywearing-culture-mainstream/>
- Komjakkraphan, P., & Chansawang, W. (2015). The lived experience of living in skipped-generation household among Southern Thai elders. *Songlanagarind Journal of Nursing*, 35(30), 35-56.
- Kim, K., & Yun, K. H. (2013). The effects of body posture by using Baby Carrier in different ways. *Journal of Korean Society of Physical Medicine*, 8(2), 193-200.
- Lee, H., & Hong, K. H. (2018). Type and wearing method-dependent COP and muscle fatigue measurement of baby carriers for the development of smart baby carriers. *International Journal of Clothing Science and Technology*, 30(5), 610-627.
- Michael, & Ash, I. (2007). *Handbook of fillers, extenders, and diluents* (2nd Ed.). New York, USA: Synapse Information Resources.
- Moore, E. R., Anderson, G. C., Bergman, N., & Dowswell, T. (2012). Early skin-to-skin contact for mothers and their healthy newborn infants. *Cochrane Database of Systematic*, 5(5), 1-75.
- Okochi, J., Utsunomiya, S., & Takahashi, T. (2005). Health measurement using the ICF: test-retest reliability study of ICF codes and qualifiers in geriatric care. *Health and Quality of Life Outcomes*, 3(1), 1-13.
- Photisita, C. (2009). Prelude: What happened to the family. *Population and Society 2009*, 1-15.
- Pollack-Nelson, C. (2000). Fall and suffocation injuries associated with in-home use of car seats and baby carriers. *Pediatric Emergency Care*, 16(2), 77-79.
- Punyakaew, A., Lersilp, S., & Puttinoui, S. (2019). Active ageing level and time use of elderly person in a Thai Suburban Community. *Occupational Therapy International*, 2019, 1-8.
- Thanatchang, T. (2007). *The first for 3 years on the province of development*. Bangkok, Thailand: Sirirat Hospital.

Numerical Analysis of the Crack Inspections Using Hybrid Approach for the Application the Circular Cantilever Rods

Saddam Hussein Raheemah¹, Kareem Idan Fadheel¹, Qais Hussein Hassan², Ashham Mohammed Aned¹, Alaa Abdulazeez Turki Al-Taie^{3,4} and Hussein Kadhim Sharaf^{5*}

¹Department of Mechanical Technical/Production, Al kut Technical Institution, Middle Technical University, Baghdad, 10001 Iraq

²Department of Power, Al kut Technical Institution, Middle Technical University, Baghdad, 10001 Iraq

³SOMO Oil Marketing Company, Ministry of Oil, Baghdad, 10001 Iraq

⁴Department of Chemical engineering, Faculty of Engineering, Al-Nahrain University, Baghdad, 10005 Iraq

⁵Department of Aerospace Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

The present study aims to investigate crack presence in a rigid steel beam so that it can be considered in structural design. A finite element method (FEM) had been used with the Ansys 16.1 software to simulate the whole steel body with three different forces and moments with a magnitude force subjected at the free end of the beam. The steel rod had been considered as simple cantilever to be modelled by the software. Von Mises stress had been considered in the simulation process where the maximum value of stress due to applied load and moment was 1.9 MPa. Total deformation of the whole body had also

been considered to instigate the maximum deformation (4.3mm) due to applied loads and moments. Furthermore, MATLAB and through fuzzy logic had been used to assist in the investigation of cracks. Both approaches had been governed by the Euler-Bernoulli theory for free vibration of motion equations. The other aim of this study is to evaluate results received from the Ansys with MATLAB for the same boundary conditions as the case.

Keywords: ANSYS; Euler- Bernoulli; fem; fuzzy logic; MATLAB

ARTICLE INFO

Article history:

Received: 28 July 2020

Accepted: 13 October 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.22>

E-mail addresses:

raheemh.h@gmail.com (Saddam Hussein Raheemah)

kareem.f33@gmail.com (Kareem Idan Fadheel)

qaishussein123@gmail.com (Qais Hussein Hassan)

ashham.m@mtu.edu.iq (Ashham Mohammed Aned)

taei.turk@gmail.com (Alaa Abdulazeez Turki Al-Taie)

hk.sharaf92@gmail.com (Hussein Kadhim Sharaf)

*Corresponding author

INTRODUCTION

Steel beams are widely used in the machinery and construction industries and need safety monitoring. Thus, damage analysis in beam structures is of critical importance. Many solutions have been extended in this respect, and there are several ongoing attempts to find a workable solution for crack detection in beam structures (Sapuan, 2005). A method was submitted to evaluate the difference between the two modal properties, i.e., mode beam shapes and their natural frequencies, leading to the identification of the crack on a damaged beam (Meade & Presley, 2002). This method simulated the crack with a similar spring at the crack spot. It is obvious that cracks in the turbocharging shaft and vibration parameters are related to each other. The purpose of this study is to show that the proposed model has a non-linear effect.

Zavadskas et al. (2014) and Kremer et al. (2012) have documented the introduction of local flexibility at the site of a vibrationally cracked structure and studied the response of its vibration. Due to the high cost of vibration-based maintenance of these machines, the dynamics of rotors have been very significant in the design of rotary machines in recent years. Due to its extensive usage in manufacturing and power plants, the rotor is the most important component of rotary machines. This provides the key purpose of the unit which focuses on the study of an uncracked and broken rotor (Sharaf et al., 2020a). There are several origins of excitation forces that can be found in the rotary machines. The most recognized cause of excitation in rotor-bearing structures is mass eccentricity, which creates unbalance power. Several structures with a single transverse surface crack are taken based on previous studies (Utili, 2013). When more than one crack occurs in a system, based on the relative locations and depths of such cracks, the hierarchical solution becomes more nuanced. In addition, different combinations of position and depth can contribute to the same changes in the natural frequencies. Several investigators recently examined the impact of two surface cracks on structural vibration reaction. Sander et al. (2013) analyzed the effect on the normal velocity of cantilever beams of the locations and depths of two-cracks.

Sheu and Lee (2011) had developed a method for investigating the crack position and depth of a vibrational beam structure without its release from a system, assuming that the local spring was at a crack position. Ertay et al. (2005) had developed an adequate method for damage detection in a transverse beam with a dynamic vibration status, which was supposed to be a cross-sectional open crack. Nepal et al. (2010) and Al-Harbi (2001) had shown a co-relation between the location of a crack, the crack depth and the difference in mode. The method used specific modes of vibration beam structure. The objective of the current study is to investigate the crack location numerically by using the results of the fuzzy logic approach along with a FEM.

General Theorem

For MATLAB analyzes and the FEM used by the ANSYS, the Euler-Bernoulli approach is assumed. The crack is called an open crack in this particular case and no damping has been called. For the formulation, a single cross crack is considered (Srinivasan & Kraslawski, 2006).

Fuzzy logic theory of analysis fuzzy theory means a computational approach that considers several truths, without numerically identifying them as true or false, i.e. (1,0). For its linguistic variables, this approach uses specific functions (Al-Harbi, 2001). Fuzzy 's logic extends from control theory to artificial intelligence. Modern computation allows use of accurate data with confidence, but soft computing can use inaccurate data and calculate precisely to produce output. For describing such mapping laws, fuzzy logic uses terms instead of numbers (Chang et al., 2007).

Governor Equations

By following the difference equation, a free bending vibration of an Euler-Bernoulli strap of a constant rectangular cross part is given as Equation 1:

$$EI \frac{d^4y}{dx^4} - m\omega^2 = 0 \tag{1}$$

Where 'm' is the mass of the beam per unit length (kg/m), 'ωi' is the natural force of the mode (rad/s), E is the modulus of elasticity (N/m²) and I is the moment of inertia (m⁴). By defining $\lambda^4 = m \omega^2 y / EI$ equation is rearranged as a fourth-order differential equation (Imihezri et al., 2005). Now applying moment of inertia for shape with and without crack as in Equation 2:

$$I = \frac{BW^3}{12} \tag{2}$$

In order to locate position (γ) of crack, following Equation 3 of moment of inertia should be used:

$$I = \frac{B(W - a)^3}{12} \tag{3}$$

The Equation 3 in characterization form is a function of position of the crack (γ), stiffness matrix (K), circular force (ω) and r crack depth (φ). The pictorial view of a free end rod geometry with dimensions (a1) as crack depth, (L1) as crack location and (BxW) as cross-sectional area which is under the influence of axial force (P1) that subject the beam to a single effect (Su & Lin, 2008).

MATERIAL AND METHOD

Boundary Condition Set Up

MATLAB (Coding) Set Up. MATLAB had been conducted to analysis input data by using fuzzy logic tool and input data as follow: First force = “FF”, Second force = “SF”, and Third force = “TF”. For the output of fuzzy logic controller process, following terms had been used: Crack depth = “CD” and Crack length= “CL” (Sharaf et al., 2021a). According to the fuzzy logic subset, the rules fuzzy are known with a general code as follows: “If (FF is FFi and SF is SFj and TF is TFk) then (CD is CDijk and CL is CLijk), where i= 1 to 9, j=1 to 9, and k=1 to 9” Because of “FF”, “SF”, and “TF” there are nine functions. Two kinds of rules can be written “If (FF is FFi and SF is SFj and TF is TFk) then CD is CDijk (Sharaf et al., 2021b) If (FF is FFi and SF is SFj and TF is TFk) then CL is CLijk”.

ANSYS Software Set Up. In ANSYS applications the FEM was used to calculate the position of the crack in steel bar depending on the natural force. As the governor equation for the formulation of the finite element, the Euler-Bernoulli Governor Equation was considered. The circular cross-sectional rod with a cross-sectional crack is set on the right-hand end, free of height, and continuously cut with a triangular cross. Drawing was done using the Model Modular in ANSYS, with segment 1000 mm and diameter 50 mm. The built model used Tri pave components in difficult areas and Quad paves in most of the areas. ANSYS offers full mesh flexibility with amorphous meshing solution and, once the grid is read in ANSYS, it can be polished or roughened in a solution (Kim et al., 2009).

Material Properties

Based on previous research by Sharaf et al. (2020b), the mechanical properties of the steel bar are shown in Table 1. The main mechanical properties were conducted in the simulation process and the parameters are Young Modulus (E), density kg/m³ and Poisson’s Ratio (ϑ).

Table 1

Mechanical properties of steel rod

Young Modulus (E),	density kg/m ³	Poisson's Ratio (ϑ).
2.05× 10 ¹¹	7850	0.29

Geometry and Meshing

The geometry of the current research has been modeled by using Design Modeler (DM) and it is a tool of the ANSYS 16.1 software. Figure 1 shows the meshed model uses Tri type pave components in difficult areas and paves of Quad type in the rest of the areas based on these types of mesh give the best solution. ANSYS allows the solution a full

mesh stability for amorphous meshes and the grid can be polished or roughened based on the solution if the grid was read into ANSYS.

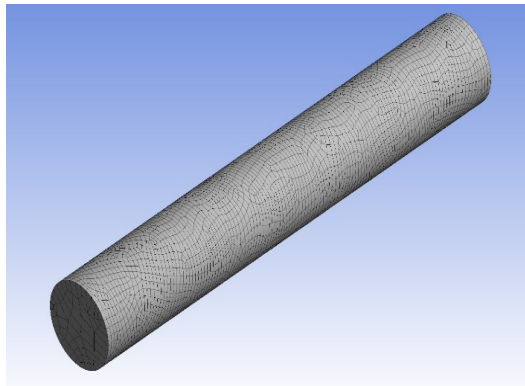


Figure 1. Meshed model of the steel rod

RESULTS AND DISCUSSION

Grid Independent Study

Based on the test, the incidence is shown to be commensurate with the number of components. Crack was set at 5 mm for four different frequencies; the first attempt was set at 10 KN when there were 261794 elements. Furthermore, no change in force occurred when the number of elements was considered to increase to 282893 at both 20 KN,30 KN.

Investigation-based on Normal Force

Based on the outcomes, the simulation processes of the circular cantilever beam to predict the location of the possibility of the presence of the cracks of the supported beam is shown in Figure 2 (total deformation) and Figure 3 (equivalent stresses).

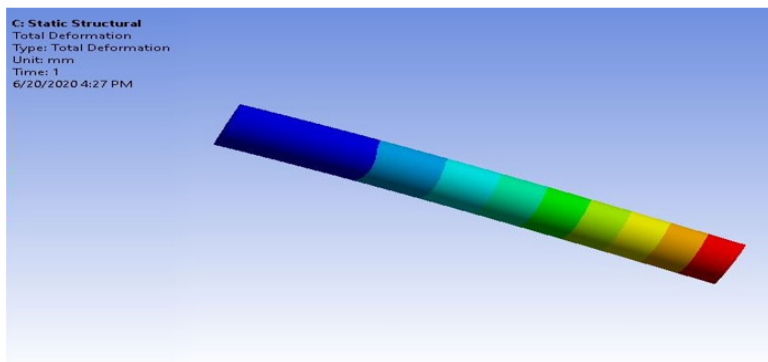


Figure 2. Total deformation due to the normal force

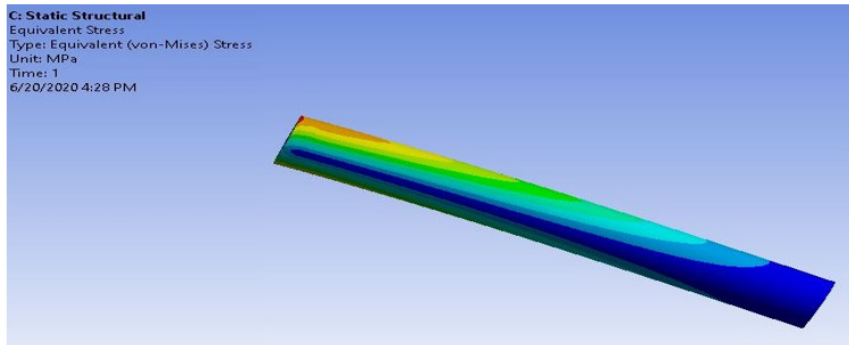


Figure 3. Equivalent stress due to the normal force

Figure 4 shows the maximum relationship between the forces to the total deformation of the amin body. It shows that the maximum deformation occurred at 10KN and deflection was at 4.3 mm.

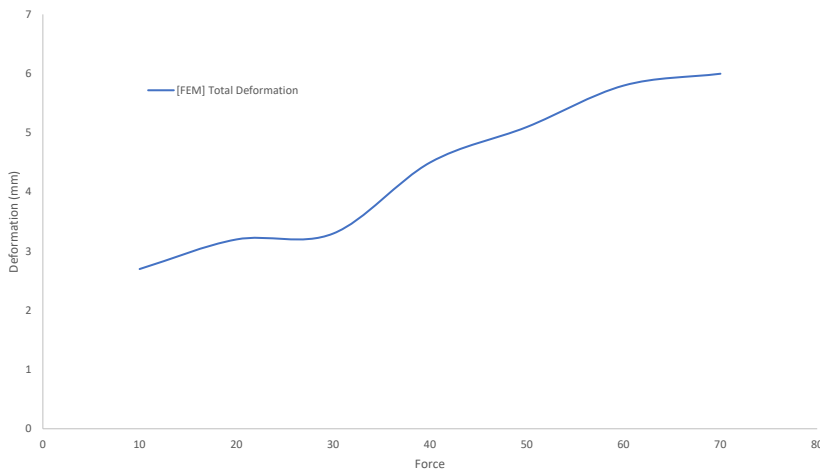


Figure 4. Deformation due to applied forces

Investigation-based on Normal Moment

Normal moment has been considered for further validation of the prediction of possibility and presence of internal damage in the supported beams. Based on the outcomes, the simulation processes of the circular cantilever beam to predict the location of the possibility of the presence of the cracks of the supported beam is shown in Figure 5 (total deformation) and Figure 6 (equivalent stresses).

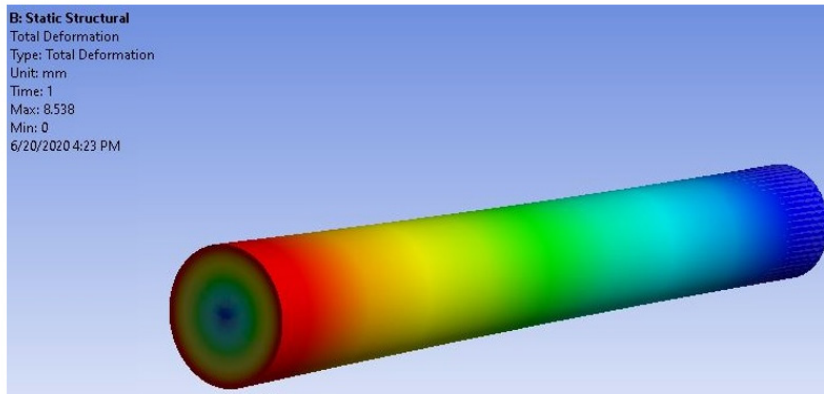


Figure 5. Total deformation due to the normal force

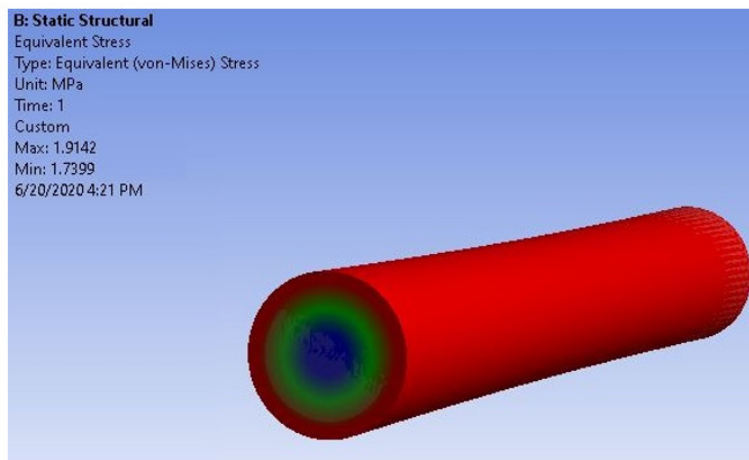


Figure 6. Equivalent stress due to the normal moment

Crack investigation of two main phases, force and moment has been considered to simulate in ANSYS structural. Assumption of crack depth is 5mm and located at distance of L1 from fixed end. Figure 6 shows the simulation result for those different forces. Structural analysis has been done in ANSYS for total deformation which was done by force.

CONCLUSION

The principal objective of the current study was to confirm the outcome of the MATLAB analysis using ANSYS with the FEM simulation. The FEM had been used in ANSYS 16.1 to simulate the whole steel body with three different forces and moments with a magnitude force subjected at the free end of the beam. The steel rod had been considered

a simple cantilever to be modelled by the software. Von Mises stress had been considered in the simulation process where the maximum value of stress due to the applied load and moment was 1.9 MPa. Total deformation for the whole body had also been considered to instigate the maximum deformation due to the applied loads and moments. Furthermore, MATLAB had been conducted with fuzzy logic to assist in the investigation of cracks. Both approaches had been governed by the Euler- Bernoulli theory for free vibration of motion equations. The other main aim of this study was to evaluate results received from the Ansys with MATLAB for the same boundary conditions as the case.

ACKNOWLEDGEMENT

Authors would like to thank Middle Technical University for their support.

REFERENCES

- Al-Harbi, K. M. A. S. (2001). Application of the AHP in project management. *International Journal of Project Management*, 19(1), 19-27. [https://doi.org/10.1016/S0263-7863\(99\)00038-1](https://doi.org/10.1016/S0263-7863(99)00038-1)
- Chang, C. W., Wu, C. R., Lin, C. T., & Chen, H. C. (2007). An application of AHP and sensitivity analysis for selecting the best slicing machine. *Computers and Industrial Engineering*, 52(2), 296-307. <https://doi.org/10.1016/j.cie.2006.11.006>
- Ertay, T., Büyüközkan, G., Kahraman, C., & Ruan, D. (2005). Quality function deployment implementation based on analytic network process with linguistic data: An application in automotive industry. *Journal of Intelligent and Fuzzy Systems*, 16(3), 221-232.
- Imihezri, S. S. S., Sapuan, S. M., Ahmad, M. M. H. M., & Sulaiman, S. (2005). A study of the comparison of “V” and “X” ribbing in a composite pedal using mold flow analysis software. *Materials and Design*, 26(2), 157-166. <https://doi.org/10.1016/j.matdes.2004.05.011>
- Kim, J., Kim, J., Lee, Y., Lim, W., & Moon, I. (2009). Application of TRIZ creativity intensification approach to chemical process safety. *Journal of Loss Prevention in the Process Industries*, 22(6), 1039-1043. <https://doi.org/10.1016/j.jlp.2009.06.015>
- Kremer, G. O., Chiu, M. C., Lin, C. Y., Gupta, S., Claudio, D., & Thevenot, H. (2012). Application of axiomatic design, TRIZ, and mixed integer programming to develop innovative designs: A locomotive ballast arrangement case study. *International Journal of Advanced Manufacturing Technology*, 61(5-8), 827-842. <https://doi.org/10.1007/s00170-011-3752-1>
- Meade, L. M., & Presley, A. (2002). R&D project selection using the analytic network process. *IEEE Transactions on Engineering Management*, 49(1), 59-66. <https://doi.org/10.1109/17.985748>
- Nepal, B., Yadav, O. P., & Murat, A. (2010). A fuzzy-AHP approach to prioritization of CS attributes in target planning for automotive product development. *Expert Systems with Applications*, 37(10), 6775-6786. <https://doi.org/10.1016/j.eswa.2010.03.048>

- Sapuan, S. M. (2005). A conceptual design of the concurrent engineering design system for polymeric-based composite automotive pedals. *American Journal of Applied Sciences*, 2(2), 514-525. <https://doi.org/10.3844/ajassp.2005.514.525>
- Sheu, D. D., & Lee, H. K. (2011). A proposed process for systematic innovation. *International Journal of Production Research*, 49(3), 847-868. <https://doi.org/10.1080/00207540903280549>
- Srinivasan, R., & Kraslawski, A. (2006). Application of the TRIZ creativity enhancement approach to design of inherently safer chemical processes. *Chemical Engineering and Processing: Process Intensification*, 45(6), 507-514. <https://doi.org/10.1016/j.cep.2005.11.009>
- Su, C. T., & Lin, C. S. (2008). A case study on the application of fuzzy QFD in TRIZ for service quality improvement. *Quality and Quantity*, 42(5), 563-578. <https://doi.org/10.1007/s11135-006-9058-y>
- Sharaf, H. K., Ishak, M. R., Sapuan, S. M., Yidris, N., & Fattahi, A. (2020a). Experimental and numerical investigation of the mechanical behavior of full-scale wooden cross arm in the transmission towers in terms of load-deflection test. *Journal of Materials Research and Technology*, 9(4), 7937-7946. <https://doi.org/10.1016/j.jmrt.2020.04.069>
- Sharaf, H. K., Ishak, M. R., Sapuan, S. M., & Yidris, N. (2020b). Conceptual design of the cross-arm for the application in the transmission towers by using TRIZ–morphological chart–ANP methods. *Journal of Materials Research and Technology*, 9(4), 9182-9188. <https://doi.org/10.1016/j.jmrt.2020.05.129>
- Sharaf, H. K., Salman, S., Dindarloo, M. H., Kondrashchenko, V. I., Davidyants, A. A., & Kuznetsov, S. V. (2021a). The effects of the viscosity and density on the natural frequency of the cylindrical nanoshells conveying viscous fluid. *The European Physical Journal Plus*, 136(1), 1-19. <https://doi.org/10.1140/epjp/s13360-020-01026-y>
- Sharaf, H. K., Salman, S., Abdulateef, M. H., Magizov, R. R., Troitskii, V. I., Mahmoud, Z. H., ... & Mohanty, H. (2021b). Role of initial stored energy on hydrogen microalloying of ZrCoAl (Nb) bulk metallic glasses. *Applied Physics A*, 127(1), 1-7. <https://doi.org/10.1007/s00339-020-04191-0>
- Sander, M., Dietrich, S., Pander, M., Ebert, M., & Bagdahn, J. (2013). Systematic investigation of cracks in encapsulated solar cells after mechanical loading. *Solar Energy Materials and Solar Cells*, 111, 82-89. <https://doi.org/10.1016/j.solmat.2012.12.031>
- Utili, S. (2013). Investigation by limit analysis on the stability of slopes with cracks. *Geotechnique*, 63(2), 140-154. <https://doi.org/10.1680/geot.11.P.068>
- Zavadskas, E. K., Turskis, Z., & Kildiene, S. (2014). State of art surveys of overviews on MCDM/MADM methods. *Technological and Economic Development of Economy*, 20(1), 165-179. <https://doi.org/10.3846/20294913.2014.892037>



Radiation Modified CS/PVA Film with PVP Coating as Cu Adsorbent Material

Norhashidah Talip^{1,2*}, Amalina Muhammad Afifi^{1,3}, Mohd Yusof Hamzah²,
Maznah Mahmud² and Sarada Idris²

¹Department of Mechanical Engineering, Faculty of Engineering, University of Malaya, 50603 UM, Kuala Lumpur, Malaysia

²Radiation Processing Technology Division, Malaysian Nuclear Technology, 43000 Kajang, Selangor, Malaysia

³Centre of Advanced Materials (CAM), Faculty of Engineering, University of Malaya, 50603 UM, Kuala Lumpur, Malaysia

ABSTRACT

Chitosan/Polyvinyl alcohol (CS/PVA) films were coated with layer of polyvinyl pyrrolidone (PVP) to improve its mechanical strength. In this study, chitosan/PVA mixture was first cast into film. The films were later dipped into different concentrations of PVP solutions at 1%, 3%, and 5%. They were then exposed to gamma (γ) radiation at 5,10,20 and 30 kGy. The tensile strength of chitosan/PVA films coated with 1% PVP was found to increase up to 31%. Elongation at break was improved by 1% PVP coating. The CS/PVA Removal Efficiency (RE) with or without PVP coatings were not affected by γ irradiation. Result shows that the best settings to enhance the tensile strength of CS/PVA films for Cu adsorption application was 1% PVP coating at 5 kGy γ radiation.

ARTICLE INFO

Article history:

Received: 4 August 2020

Accepted: 10 February 2020

Published: 30 April 2021

Keywords: Chitosan, copper removal, polyvinyl alcohol, polyvinyl pyrrolidone, radiation

DOI: <https://doi.org/10.47836/pjst.29.2.23>

E-mail addresses:

nora_talip@yahoo.com (Norhashidah Talip)

amalina@um.edu.my (Amalina Muhammad Afifi)

m_yusof@nuclearmalaysia.gov.my (Mohd Yusof Hamzah)

maznah@nuclearmalaysia.gov.my (Maznah Mahmud)

sarada@nuclearmalaysia.gov.my (Sarada Idris)

*Corresponding author

INTRODUCTION

Rapid industrial growth has immensely help to drive the country economy, but it comes at a cost where it causes massive environmental pollution. Water is among one of the most critical part of pollution that

requires utmost attention. Typical industrial wastewater pollutants include heavy metals such as copper (Cu), cadmium (Cd), zinc (Zn), and lead (Pb). These are toxic and harmful to the environment as they cause adverse effects to the health of human beings and other organisms due to their presence in drinking water and aquatic habitat (Barakat, 2011). The World Health Organization (WHO) (as cited in Y. Zhang et al., 2014) recommended that 1.5 mg/L was the maximum accepted value of Cu ion concentration in drinking water.

At present, several methods are employed for wastewater treatment and Cu ions removal such as membrane filtration, chemical precipitation, reverse osmosis, evaporation, oxidation, and adsorption (Chen et al., 2010; Sarioglu et al., 2009; Šćiban et al., 2008). Among these, adsorption method is the most common due to its economic factor, efficiency, simplicity, and environmental safety (Tripathi et al., 2009; Senthilkumar et al., 2011).

Chemical adsorption method depends on metal chelators, often they are prepared from commercial polymers such as poly(ethyleneimine) and poly(acrylic acid). Lately, naturally occurring polysaccharides such as chitosan (CS) is gaining attention due to its abundance and renewable sources (Kurita, 1998; Varma et al., 2004). CS possesses a few favorable properties (Xiao, 2007). Some of the more important properties are its capability to be shaped into films, chelate metal ions, and basic optical features (Kalantari et al., 2017; Kumar, 2000). On the polymeric backbone of chitosan consist of $-NH_2$ and $-OH$ groups which acts as active sites and responsible for the excellent performance as natural adsorbent for metal ions (Ngah & Fatinathan, 2008; Alver et al., 2014). Despite all these useful properties, CS has strong intra- and inter-molecular hydrogen bonding along its polymeric chains. This makes CS partly crystalline and thus too brittle and has weak strain at break for adsorption films and castings applications (Huang et al., 2012).

To overcome the deficit in mechanical strength of CS films, it is commonly blended with other polymers (Tripathi et al., 2009; Ngah et al., 2004; J. M. Yang et al. 2004). Poly(vinyl alcohol) (PVA) is one candidate that has been extensively used to blend with chitosan. However, according to Huang et al. (2012), CS/PVA blend needs to be improvised in order to enhance the mechanical properties and water resistance of the film for practical applications. The objective of the present work is to further improvise CS/PVA film mechanical strength and properties with polyvinyl pyrrolidone (PVP) coating and γ irradiation to introduce cross-linkages on the coated film. Interactions between polymer materials and radiation from sources like γ ray will produce free radicals. This is due to the energy of irradiation which has been absorbed by the backbone of polymer materials. This reaction can occur without the presence of catalyst, initiator, monomers, or other cross-linkers (Bhattacharya, 2000; Idris, 2008). Radiation technique was chosen due to the its effective, clean, and versatile process which can be done in convenient temperature, in any physical form and no initiator needed in this process (Al-Assaf et al., 2016).

PVP is one of the polymers that will undergo cross-linking with radiation technique (Rosiak et al., 1990). PVP solutions when exposed to γ ray irradiation, it will induce the formation of free radicals from both carbon atoms in PVP and hydrogen atoms from water that promotes cross-linking by forming C-C bonding (Demeter et al., 2017; Jeong et al., 2020). Besides, PVP has excellent properties of bonding to various metal ions such as Cu(II), Pb(II), and Cd(II), low cytotoxicity, and biocompatibility (Lu et al., 2011). Cu ions adsorption capability of this film is also discussed to evaluate the possibility/effectiveness of PVP coated chitosan as wastewater treatment membrane.

MATERIALS AND METHODS

Chitosan, (Molecular weight 190,000 – 310,000 Da) and degree of deacetylation (DD) of 75-85% was purchased from Sigma Aldrich. PVA, 96% hydrolyzed with approximation of molecular weight 146,000 – 186,000 Da and PVP with molecular weight of 58,000 Da were purchased from Acros Organics. Acetic acid (glacial) and sodium hydroxide (analytical reagent) were purchased from HmbG Chemicals and Fisher Scientific, respectively. Materials were used without further purification.

Preparation of CS/PVA Film Modified with PVP and Gamma Irradiation

Chitosan (2%, w/v) was dissolved in 1% (v/v) aqueous acetic acid and stirred overnight at room temperature. The solution was then filtered to remove dust and other impurities. PVA solution with a concentration of 5% (w/v) was prepared by dissolving PVA powder in ultra-pure water and autoclaved at 121°C for 15 minutes. CS and PVA solutions were mixed at weight ratio of 70:30. The mixed solutions were stirred at room temperature for 30 minutes. Films were obtained by casting method, where 7 ml of the solutions were poured onto polystyrene petri dishes followed by drying at room temperature for 48 h. The dried films were then soaked in 1N of NaOH for 1 h, then, washed four times with 250 ml of Reverse Osmosis (RO) water to remove the excess of acid and NaOH, respectively. The films were then soaked in 1, 3 and 5% (w/v) PVP solutions for 15 – 20 minutes and immediately kept in plastic bag prior to irradiation. The films were irradiated at 5, 10, 20, and 30 kGy from Co-60 source facility at Sinagama, Malaysian Nuclear Agency. Dose rate of the irradiation was 1.62kGy/h. Upon completion of irradiation process, the films were dried again at room temperature for 24 h. Dried films were kept in desiccator until further use. All films obtained were transparent, uniform, and free of air bubbles.

Characterization of CS/PVA Films

The Fourier Transform Infrared Spectroscopy (FTIR) analysis of blended film and modified film were performed using Bruker TENSOR FTIR. The spectrums were obtained at a frequency range of 3600 cm^{-1} to 500 cm^{-1} with a resolution of 4 cm^{-1} and 16 times scanning.

Mechanical properties of films were tested using Instron Universal machine model 4201 according to ASTM Standard Method D 882 (ASTM D882-12, 2012). Testing specimens were cut into strips with 40 mm length and 10 mm width. Thickness of the films was measured at seven various positions using Digital Thickness Gauge model Mitutoyo Crop (Japan). Gauge length and crosshead speed were set at 20 mm and 5 mm/min, respectively with load cell of 1kN. Minimum of five specimens were tested for each sample and average were calculated.

Gel fractions of irradiated films were carried out. The films were cut into small pieces and placed in the tea bags. The tea bags were soaked in distilled water and autoclave for 15 minutes at 121°C. The samples were then dried in oven at 60°C until reached constant weight (Idris, 2008). The estimation of gel fraction was measured by the insoluble part after immersion in water. A minimum of five specimens were tested for each sample and average were calculated. Gel fraction was calculated using Equation 1.

$$\% \text{ Gel fraction} = \frac{W_2}{W_1} \times 100 \quad [1]$$

where W_1 is the initial weight of film and W_2 is the weight of insoluble part after immersion.

Surface roughness and morphology of the films were studied using Atomic Force Microscope (AFM) from JPK Instruments. Samples were cut into small size and located on glass slide. Measurement was done by using contact mode cantilever.

Adsorption Study of the Film

Removal of copper ions was performed via batch adsorption analysis. This experiment was carried out by soaking 0.1 g of film in 50 ml of copper (II) nitrate solution with initial concentration of 20 mg/L for 24 h at room temperature with 120 rpm. pH value of the solution was adjusted to 5.75 by adding 1% of ammonium hydroxide or 1% of HCl. Atomic Absorption Spectrophotometer (Perkin Elmer 800 Analyst) was employed for determination of copper ions. The amount of copper adsorbed onto the film was calculated by using Equation 2.

$$q = \frac{(C_o - C_t)}{m} \times V \quad [2]$$

where, C_o and C_t are initial copper concentrations and at t time in (mg/L), V was the volume of adsorbate (L) and m was the weight of the film (g) (Liu et al., 2009).

Removal efficiency (RE) of metal ion in percentage was calculated by using Equation 3 as describe by Lee et al. (2015).

$$RE (\%) = \frac{(C_o - C_e)}{C_o} \times 100 \quad [3]$$

where, C_0 and C_e represent the initial metal ion concentration and concentration at equilibrium, respectively.

RESULTS AND DISCUSSION

Figure 1 shows the FTIR spectra of (a) CS film (b) CS/PVA film, (c) PVP, modified CS/PVA film with 1% of PVP and irradiated at (d) 0 kGy and (e) 5 kGy. The broad band for all films observed at 3450 - 3200 cm^{-1} are attributable to hydroxyl -OH stretching vibration as well as -NH stretching vibration (Ghobadi et al., 2017). Introduction of PVA into chitosan is made evident by a peak at 1446 cm^{-1} which was caused by -CH-OH bending vibration of PVA (Zhuang et al., 2012). The shifts of absorption peaks to 3357 and 3283 cm^{-1} indicate introduction of -OH group from PVA. Increment in peak intensity at 1650 cm^{-1} indicates the interaction between amine from CS and -C-OH from PVA. A reduction in peak intensity at 1576 cm^{-1} indicates a reduction in -NH₂ group where, some degradation took place. Figure 1(b) shows PVP spectra, strong band at 1645 cm^{-1} , 1421 cm^{-1} , 1281 cm^{-1} and weak band at 2917 cm^{-1} are attributed to C=O stretching, C-N stretching, O-H bending and C-H stretching vibrations, respectively (Can, 2005; Singh & Pal, 2011; C. Yang et al., 2010).

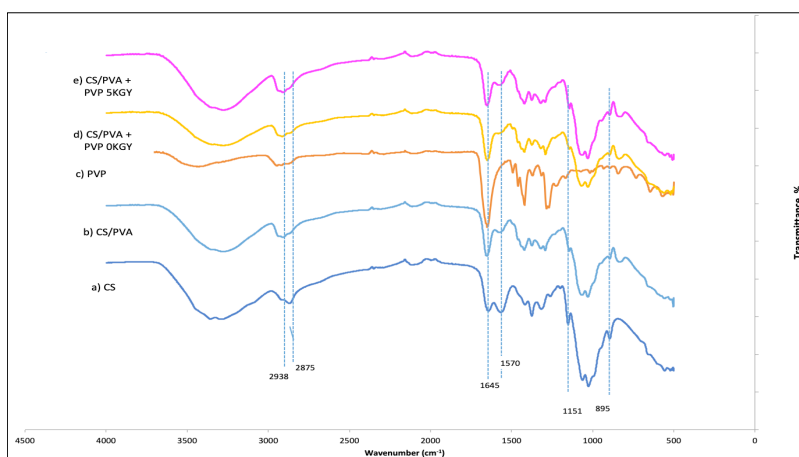


Figure 1. FTIR spectra of (a) CS, (b) CS/PVA, (c) PVP, (d) CS/PVA + PVP 0 kGy, (e) CS/PVA + PVP 5 kGy

As shown in Figure 1(c), 1(d) and 1(e), carbonyl group from PVP at peak 1645 cm^{-1} is very intense. However, this reduced after the coated films were exposed to γ ray at doses mentioned above. The reduction in the intensity suggests an interaction of C=O groups with H bending either from -OH groups from CS or PVA. This is further supported by the increase in the intensity of -OH peak in CS/PVA films upon introduction of PVP. When the films were exposed to γ ray at doses mentioned above, the peaks at around 2938 cm^{-1} , 2909 cm^{-1} , and 2875 cm^{-1} become more intense and shifted to lower wavenumbers. These

indicate formation of C-C linkages that may be formed from chemical cross-linking by γ irradiation.

Chitosan in the mixture of CS/PVA films when exposed to γ irradiation will undergo degradation. It can be shown in decreasing of peaks intensity at 895 and 1151 cm^{-1} as suggested by Casimiro et al. (2016) that some degradation by cleavage of β -glucosidic linkages due to reaction by OH radicals.

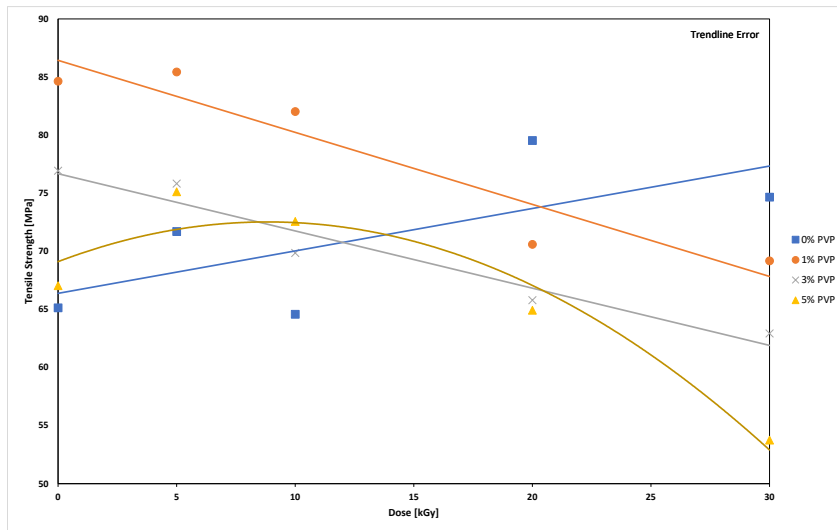


Figure 2. The effect of concentrations of PVP used for coating on tensile strength of CS/PVA films as a function of gamma irradiation dose.

Table 1

Thickness of CS/PVA film and CS/PVA films with PVP coating

Concentration of PVP (%)	Thickness (mm)
0	0.015
1	0.016
3	0.017
5	0.018

Concentrations of PVP used for coatings up to 5% showed no significant impact on the final thickness of the films where, bare CS/PVA film at 0.015 mm while, coated with 5% at 0.018 mm as can be seen in Table 1. Figure 2 shows the tensile strength of bare CS/PVA films and PVP coated CS/PVA film, irradiated with γ irradiation at 5, 10, 20, and 30 kGy. Generally, for non-irradiated samples, addition of PVP enhanced the CS/PVA film

tensile strength. This is expected as coating of PVP provides physical support on CS/PVA films. Bare CS/PVA films generally improved upon irradiation, suggesting that the PVA chains in the CS/PVA mixture have formed cross-linkages that can support the structure. However, this is still lower than the tensile strength provided by PVP coatings especially below 10 kGy. Coating of 1% PVP at 5 kGy, improved the tensile strength as much as 31% compared to the bare CS/PVA film. This can be attributed to PVP-PVP chain cross-linking that provide support on the coated CS/PVA film. However, a slight decrease could be observed at 10 kGy, where the tensile strength improved at only 26% compared to the bare CS/PVA film. This improvement, although lower than that of 5 kGy, was caused by the competing nature between cross-linking and degradation when polymers were subjected to ionizing radiation (Rosiak, 1998). At this dose, the chains degradation effect was more predominant than that of cross-linking compared to the one at 5 kGy as shown in Figure 3.

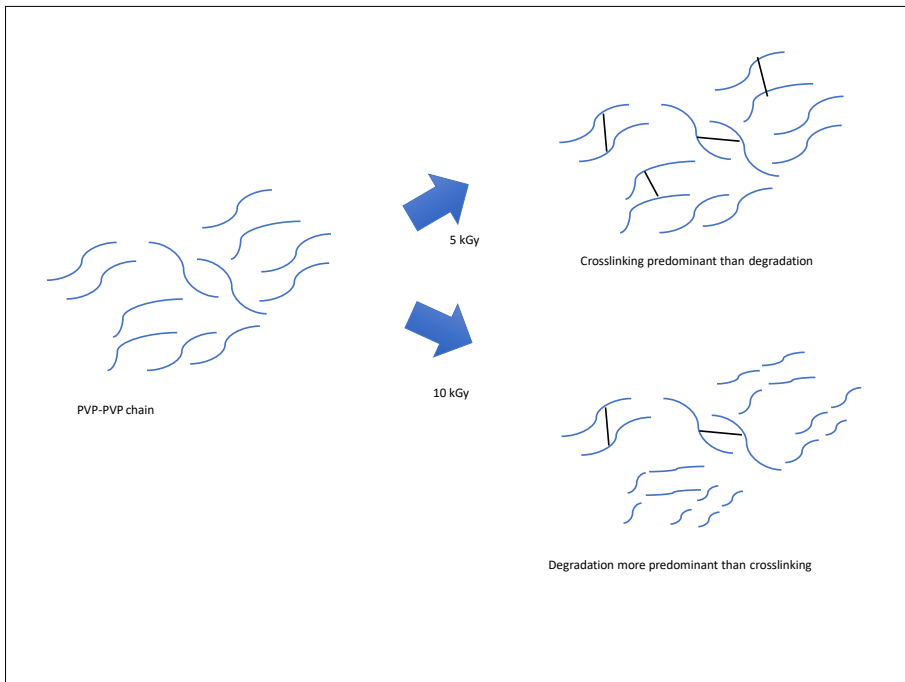


Figure 3. Schematic diagram of competing nature between crosslinking and degradation.

The same trend was observed for CS/PVA films coated with 3% and 5% at greater degradation rates. The stronger decline in the tensile strength at higher PVP concentration coatings can be discussed in term of number of radicals generated by water radiolysis when the coated films are subjected to ionizing radiation. At higher solution concentration

regime, the required dose to form gel phase is proportionate to the concentration of polymer (Rosiak & Ulanski, 1999). Therefore, higher dose is required to produce enough radicals for polymers at higher concentration to produce gel compared to a more dilute system. In this case, from a covalent cross-linking point of view, lower numbers of cross-linking networks are formed on CS/PVA film coated at higher concentration of PVP at the same irradiation dose. The lower amount of cross-linkages means lower strengthening effect from PVP, hence the lower tensile strength values.

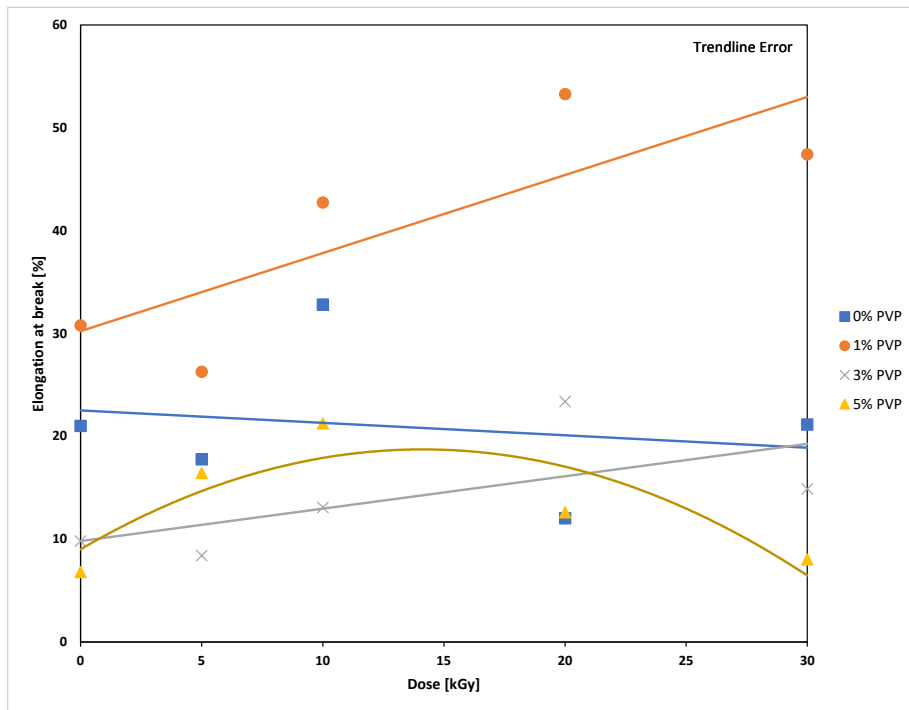


Figure 4. The effect of PVP coated concentration and radiation dose on elongation at break of CS/PVA films

Figure 4 shows the elongation at break data for bare CS/PVA films and its PVP coated form, irradiated with gamma irradiation at 5, 10, 20, and 30 kGy. Unlike its improvement in tensile strength when subjected to γ radiation, at around 20% elongation, bare CS/PVA film lacks elasticity throughout the dose range. This is expected due to the partial crystallinity of CS. Coating of 1% PVP improves the elongation at break throughout the irradiation doses up to 53%. This suggests that PVP-PVP cross-linking provides a useful flexible support to CS/PVA films. PVP coatings at higher concentration do not provide such support. This agrees with the higher gelation dose for higher samples concentration from tensile strength studies. A plateau as seen in uncoated CS/PVA films suggests that dose increase will not

improve the elongation at break for samples with higher concentration PVP coatings. At this point, it was decided that 1% was the best PVP coating concentration. It gave the best tensile strength and elongation at break, thus further characterization and testing were directed on this particular concentration.

Gel fraction was used to estimate the degree of cross-linking of the polymers. Gel fraction of CS/PVA film coated with 1% of PVP and irradiated at various γ doses (5, 10, 20, and 30 kGy) is shown in Figure 5.

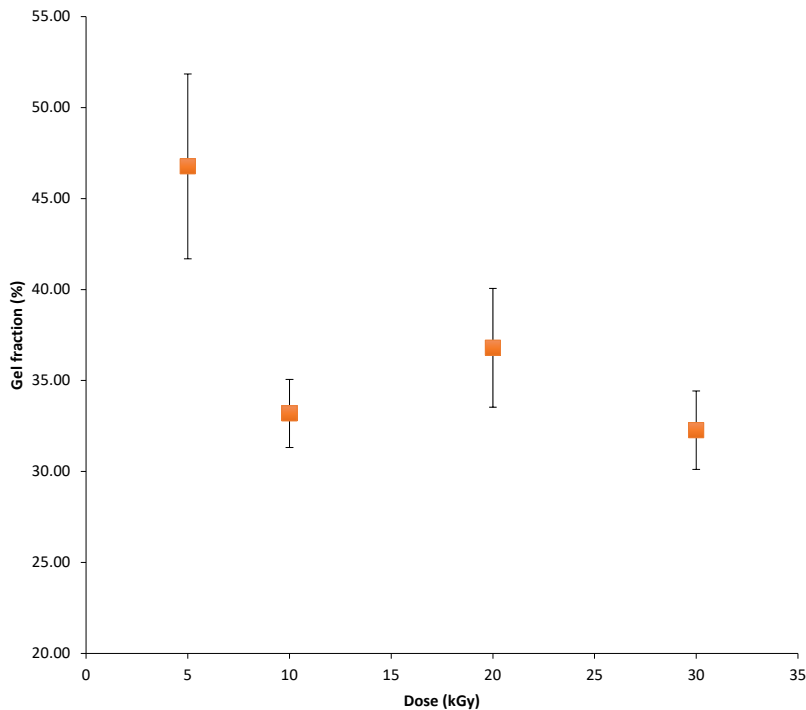


Figure 5. Gel fraction of CS/PVA film coated with 1% PVP and irradiated at various dose.

At 5 kGy of gamma irradiation dose, the value of gel fraction was the highest when compared to other irradiation doses. At this particular dose, it was the highest cross-linking formation between PVP-PVP chains, and this can support the results obtained from the tensile strength where, CS/PVA film with 1% PVP coating irradiated at 5 kGy dose gave the highest value in strength. The mechanism of PVP crosslink by γ irradiation is that, water molecules have absorbed the ionizing radiation and formed free radicals such as hydroxyl radicals and hydrogen atoms which were responsible for macroradicals formation (Demeter et al., 2017). In general, macroradicals may involve in the inter- and intramolecular reactions (Rosiak et al., 1990).

However, when the irradiation dose increasing, there was a drastic decreased. For the creation of the film by using radiation technique to form crosslinking, this is a common situation. It was called critical absorbed dose or gelation dose where, there was a formation of three-dimensional (3D) network. This 3D formation of network is occurred due to the recombination of polymer chains either with the same polymer chains (intra molecules) or with different polymer chains (inter molecules) (C. Yang et al., 2010). Further exposure of the film at the higher irradiation dose, the degradation take place where, termination of network dominate over recombination of the inter or intra molecular radicals (Qiu et al., 2007; Zhao et al., 2006). Therefore, CS/PVA coated with 1% PVP and irradiated at 5 kGy successfully improved the mechanical strength of the film. This can support the statement by Q. G. Zhang et al. (2013) and Can (2005) where, crosslinking between PVP chains can be produced through irradiation and it is one of the convenient methods to improve mechanical strength.

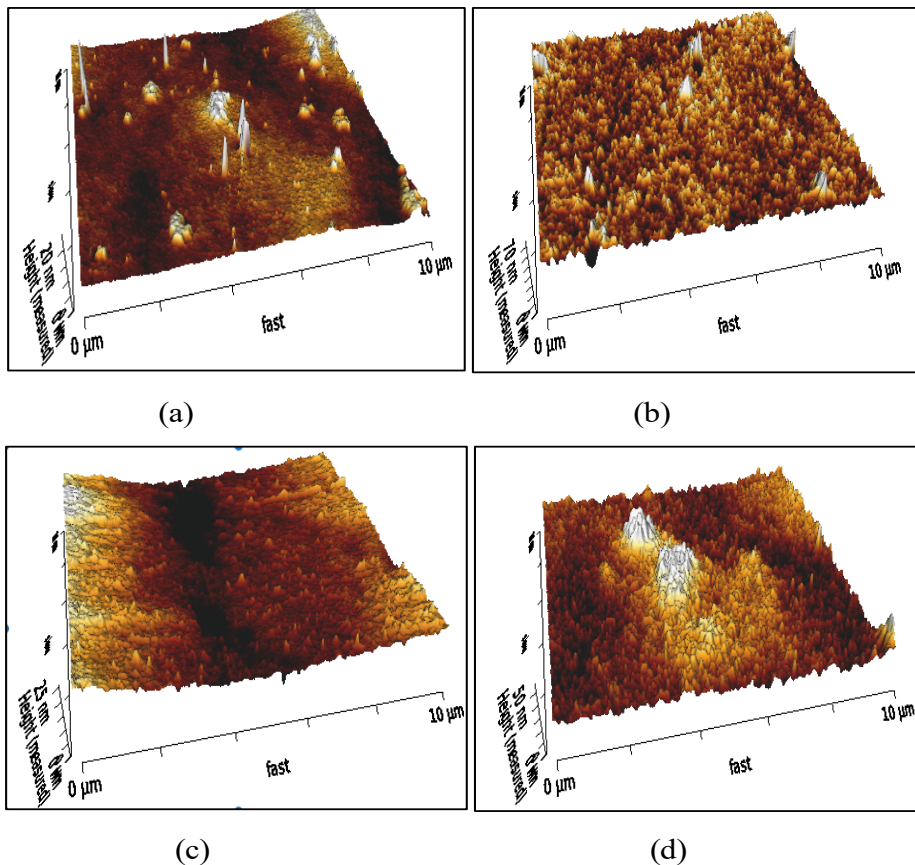


Figure 6. AFM images of (a) CS/PVA/1%PVP – 0 kGy, (b) CS/PVA/1%PVP – 5 kGy, (c) CS/PVA/1%PVP – 10 kGy and (d) CS/PVA/1%PVP – 30 kGy

Figure 6 shows the surface condition of the bare and PVP coated films and the effect of irradiation on its surface roughness. PVP coatings and γ irradiation increases the surface roughness of the films. According to Salehi et al. (2013), surface roughness enhances the reactivity for ions adsorption, however, as seen in Figure 7, this is not true in this study. Although the roughness of the 1% PVP coated CS/PVA films increase with increasing irradiation doses, the adsorption values remain almost unchanged at around 9.6 mg/g. This shows that the adsorption performance does not depend on the surface roughness of the coated films.

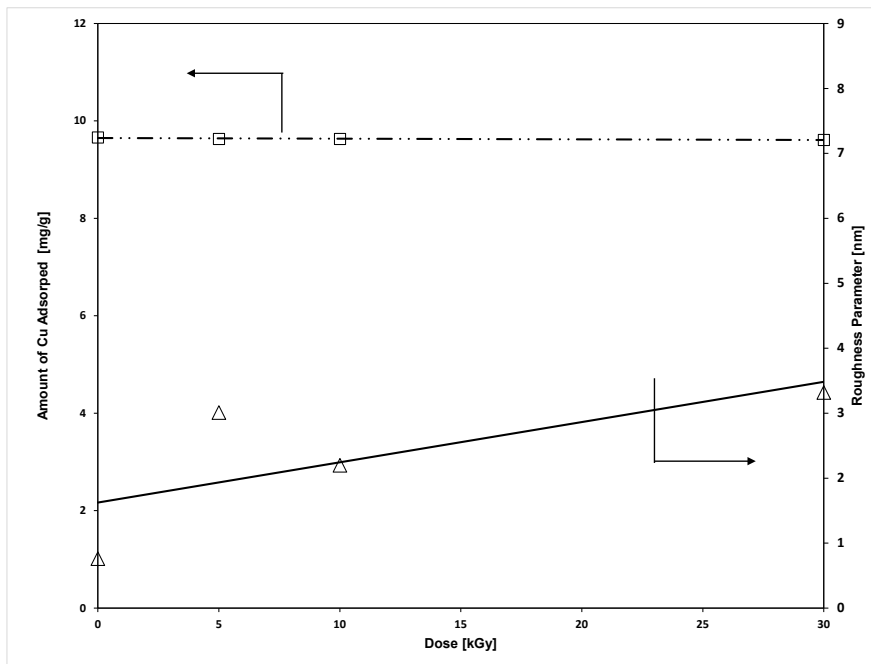


Figure 7. The relation of amount of Cu adsorption and surface roughness of CS/PVA films coated with 1%PVP at different gamma irradiation doses.

Figure 8 shows the removal efficiency (RE) of copper by CS/PVA films coated with 1% PVP and the effect of irradiation. The RE values remain high, in excess of 95% for PVP coated films at doses from 5 to 30 kGy. Note that the RE value for bare CS/PVA has a value of 98%. Heavy metals adsorption ability of chitosan is due to the existence of amine and hydroxyl functional groups in the structure. Although through γ irradiation process chitosan undergo some degradation, but in this study, the RE has no significant change up to 30kGy of irradiation dose when compared with un-modified CS/PVA film with γ irradiation and PVP coatings.

PVP coated on the film might also contribute to this minor change of RE because PVP has the characteristic of reversibly coordinating to various metal ions and has been used in wastewater treatment (Lu et al., 2011). This may be caused by the presence of oxygen in the carbonyl group in the PVP structure that attributes to the attraction of the metal ions with the film (Ali et al., 2003). According to Çaykara and İnam, (2002) there are few non-specifically ways of metals ions that can be adsorbed either by physical or chemical adsorption. In the case of film containing PVP, the adsorption may be attributed by chemical adsorption, which related to exchange of electrons and formation of chemical bond between metal ions and most likely with carbonyl groups of the PVP.

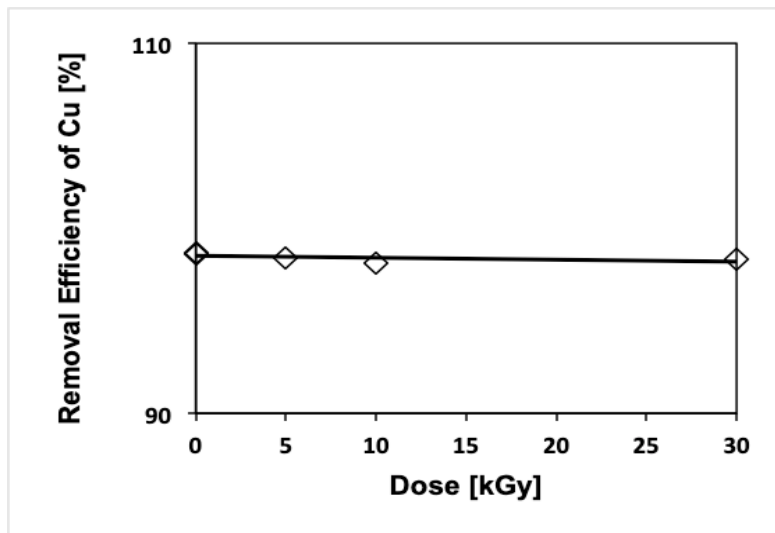


Figure 8. Removal efficiency of 1% PVP coated CS/PVA films copper as a function of gamma irradiation dose.

CONCLUSION

PVP coatings and irradiation method successfully improved the deficit in mechanical strength of CS/PVA films. Coating of PVP even at 0 kGy, enhanced the CS/PVA film tensile strength. Although, with irradiation only (without PVP coating) may help strengthen bare CS/PVA films, the effect is lower compared to the ones with PVP coatings, which is evident for dose below 10 kGy. Elongation at break, however, was only improved by 1% PVP coating. Surface roughness was improved by irradiation; however, it did not improve the Cu ion adsorption. This is because Cu adsorption is very much dependent on the CS portion. RE values were similar for both bare and coated CS/PVA films and were not dependent on irradiation doses. This proves that radiation method can be used to improve the strength

of CS/PVA films without compromising its RE. The result shows that the best settings to improve the mechanical strength of CS/PVA films for Cu adsorption application are 1% PVP coating at 5 kGy.

ACKNOWLEDGEMENTS

The first author would like to thank Assoc. Prof. Dr. Amalina M. Afifi for supervising and assisting this research work. Also, big appreciation to colleagues and supporting staff of Malaysian Nuclear Agency for their expert advices, guidance, and valuable knowledge while conducting this research work.

REFERENCES

- Ali, A. E. H., Shawky, H. A., Rehim, H. A. A. E., & Hegazy, E. A. (2003). Synthesis and characterization of PVP/AAc copolymer hydrogel and its applications in the removal of heavy metals from aqueous solution. *European Polymer Journal*, 39(12), 2337-2344. [https://doi.org/10.1016/s0014-3057\(03\)00150-2](https://doi.org/10.1016/s0014-3057(03)00150-2)
- Al-Assaf, S., Conqueret, X., Dahlan, K. Z. H. M., Sen, M., & Ulanski, P. (Eds.). (2016). *The Radiation Chemistry of Polysaccharides*. International Atomic Energy Agency.
- Alver, E., Metin, A. Ü., & Çiftçi, H. (2014). Synthesis and characterization of Chitosan/Polyvinylpyrrolidone/Zeolite composite by solution blending method. *Journal of Inorganic and Organometallic Polymers and Materials*, 24(6), 1048-1054. <https://doi.org/10.1007/s10904-014-0087-z>
- ASTM D882-12. (2012). *Standard test method for tensile properties of thin plastic sheeting*. ASTM International, West Conshohocken, PA. <https://www.astm.org/DATABASE.CART/HISTORICAL/D882-12.htm>
- Barakat, M. A. (2011). New trends in removing heavy metals from industrial wastewater. *Arabian Journal of Chemistry*, 4(4), 361-377. <https://doi.org/10.1016/j.arabjc.2010.07.019>
- Bhattacharya, A. (2000). Radiation and industrial polymers. *Progress in Polymer Science*, 25(3), 371-401. [https://doi.org/10.1016/s0079-6700\(00\)00009-5](https://doi.org/10.1016/s0079-6700(00)00009-5)
- Can, H. K. (2005). Synthesis of persulfate containing poly (N-vinyl-2-pyrrolidone) (PVP) hydrogels in aqueous solutions by γ -induced radiation. *Radiation Physics and Chemistry*, 72(6), 703-710. <https://doi.org/10.1016/j.radphyschem.2004.04.028>
- Casimiro, M. H., Lancastre, J. J. H., Rodrigues, A. P., Gomes, S. R., Rodrigues, G., & Ferreira, L. M. (2016). Chitosan-Based Matrices Prepared by Gamma Irradiation for Tissue Regeneration: Structural Properties vs. Preparation Method. *Topics in Current Chemistry*, 375(1). <https://doi.org/10.1007/s41061-016-0092-5>
- Çaykara, T., & İnam, R. (2002). Determination of the competitive adsorption of heavy metal ions on poly(n-vinyl-2-pyrrolidone:acrylic acid) hydrogels by differential pulse polarography. *Journal of Applied Polymer Science*, 8(89), 2013-2018. <https://doi.org/10.1002/app.12328>
- Chen, H., Dai, G., Zhao, J., Zhong, A., Wu, J., & Yan, H. (2010). Removal of copper(II) ions by a biosorbent-Cinnamomum camphora leaves powder. *Journal of Hazardous Materials*, 177(1-3), 228-236. <https://doi.org/10.1016/j.jhazmat.2009.12.022>

- Demeter, M., Virgolici, M., Vancea, C., Scarisoreanu, A., Kaya, M. G. A., & Meltzer, V. (2017). Network structure studies on γ -irradiated collagen-PVP superabsorbent hydrogels. *Radiation Physics and Chemistry*, *131*, 51-59. <https://doi.org/10.1016/j.radphyschem.2016.09.029>
- Ghobadi, N., Mohammadi, T., Kasiri, N., & Kazemimoghadam, M. (2017). Modified poly(vinyl alcohol)/chitosan blended membranes for isopropanol dehydration via pervaporation: Synthesis optimization and modeling by response surface methodology. *Journal of Applied Polymer Science*, *134*(11). <https://doi.org/10.1002/app.44587>
- Huang, D., Wang, W., Xua, J., & Wang, A. (2012). Mechanical and water resistance properties of chitosan/poly(vinyl alcohol) films reinforced with attapulgite dispersed by high-pressure homogenization. *Chemical Engineering Journal*, *210*, 166-172. <https://doi.org/10.1016/j.cej.2012.08.096>
- Idris, S. (2008). *Charcterization of electron beam-irradiated sago starch-polyvinyl alcohol blend films* [Unpublished Master Degree]. University Putra Malaysia.
- Jeong, J. O., Park, J. S., Kim, Y. A., Yang, S. J., Jeong, S. I., Lee, J. Y., & Lim, Y. M. (2020). Gamma ray-induced polymerization and cross-linking for optimization of PPy/PVP hydrogel as biomaterial. *Polymers*, *12*(1). <https://doi.org/10.3390/polym12010111>
- Kalantari, K., Afiffi, A. M., Salleh, A., Mohamad, E. N., & Izadiyan, Z. (2017). Evaluation of heavy metals removal by cross-linked (polyvinyl alcohol/chitosan/magnetic) nanofibrous membrane prepared by electrospinning technique. *Desalination and Water Treatment*, *98*, 266-275. <https://doi.org/10.5004/dwt.2017.21741>
- Kumar, M. N. V. R. (2000). A review of chitin and chitosan applications. *Reactive & Functional Polymers*, *46*(1), 1-27. [https://doi.org/10.1016/S1381-5148\(00\)00038-9](https://doi.org/10.1016/S1381-5148(00)00038-9)
- Kurita, K. (1998). Chemistry and application of chitin and chitosan. *Polymer Degradation and Stability*, *59*, 117-120.
- Lee, D., Chen, D. W. C., Chiu, S. F., & Liu, S. J. (2015). Electrospun nanofibrous polylactide/chitosan mats for the filtration of silver ions. *Textile Research Journal*, *85*(4), 346-355. <https://doi.org/10.1177/0040517514548752>
- Liu, X., Cheng, Z., & Ma, W. (2009). Removal of copper by modified chitosan adsorptive membrane. *Frontiers of Chemical Engineering in China*, *3*(1), 102-106. <https://doi.org/10.1007/s11705-009-0123-7>
- Lu, Q., Yu, J., Gao, J., Yang, W., & Li, Y. (2011). Glow-discharge electrolysis plasma induced synthesis of polyvinylpyrrolidone/acrylic acid hydrogel and its adsorption properties for heavy-metal ions. *Plasma Processes and Polymers*, *8*(9), 803-814. <https://doi.org/10.1002/ppap.201000144>
- Ngah, W. S. W., & Fatinathan, S. (2008). Adsorption of Cu(II) ions in aqueous solution using chitosan beads, chitosan-GLA beads and chitosan-alginate beads. *Chemical Engineering Journal*, *143*(1-3), 62-72. <https://doi.org/10.1016/j.cej.2007.12.006>
- Ngah, W. W., Kamari, A., & Koay, Y. J. (2004). Equilibrium and kinetics studies of adsorption of copper (II) on chitosan and chitosan/PVA beads. *International Journal of Biological Macromolecules*, *34*(3), 155-161. <https://doi.org/10.1016/j.ijbiomac.2004.03.001>

- Qiu, J., Xu, L., Peng, J., Zhai, M., Zhao, L., Li, J., & Wei, G. (2007). Effect of activated carbon on the properties of carboxymethylcellulose/activated carbon hybrid hydrogels synthesized by γ -radiation technique. *Carbohydrate Polymers*, 70(2), 236-242. <https://doi.org/10.1016/j.carbpol.2007.04.001>
- Rosiak, J. M. (1998). Gel/Sol analysis of irradiated polymers. *Radiation Physics and Chemistry*, 51(1), 13-17. [https://doi.org/10.1016/S0969-806X\(97\)00254-5](https://doi.org/10.1016/S0969-806X(97)00254-5)
- Rosiak, J. M., Olejniczak, J., & Pekala, W. (1990). Fast reaction of irradiated polymers - I. Crosslinking and degradation of polyvinylpyrrolidone. *Radiation Physics and Chemistry*, 36(6), 747-755. [https://doi.org/10.1016/1359-0197\(90\)90173-F](https://doi.org/10.1016/1359-0197(90)90173-F)
- Rosiak, J. M., & Ulański, P. (1999). Synthesis of hydrogels by irradiation of polymers in aqueous solution. *Radiation Physics and Chemistry*, 55(2), 139-151. [https://doi.org/10.1016/S0969-806X\(98\)00319-3](https://doi.org/10.1016/S0969-806X(98)00319-3)
- Salehi, E., Madaeni, S. S., Rajabi, L., Derakhshan, A. A., Daraei, S., & Vatanpour, V. (2013). Static and dynamic adsorption of copper ions on chitosan/polyvinyl alcohol thin adsorptive membranes: Combined effect of polyethylene glycol and aminated multi-walled carbon nanotubes. *Chemical Engineering Journal*, 215-216, 791-801. <https://doi.org/10.1016/j.cej.2012.11.071>
- Sarioglu, M., Güler, U. A., & Beyazit, N. (2009). Removal of copper from aqueous solutions using biosolids. *Desalination*, 239(1-3), 167-174. <https://doi.org/10.1016/j.desal.2007.03.020>
- Šćiban, M., Klačnja, M., & Škrbić, B. (2008). Adsorption of copper ions from water by modified agricultural by-products. *Desalination*, 229(1-3), 170-180. <https://doi.org/10.1016/j.desal.2007.08.017>
- SenthilKumar, P., Ramalingam, S., Sathyaselvabala, V., Kirupha, S. D., & Sivanesan, S. (2011). Removal of copper(II) ions from aqueous solution by adsorption using cashew nut shell. *Desalination*, 266(1-3), 63-71. <https://doi.org/10.1016/j.desal.2010.08.003>
- Singh, B., & Pal, L. (2011). Radiation crosslinking polymerization of sterculia polysaccharide-PVA-PVP for making hydrogel wound dressings. *International Journal of Biological Macromolecules*, 48(3), 501-510. <https://doi.org/10.1016/j.ijbiomac.2011.01.013>
- Tripathi, S., Mehrotra, G. K., & Dutta, P. K. (2009). Physicochemical and bioactivity of cross-linked chitosan-PVA film for food packaging applications. *International Journal of Biological Macromolecules*, 45(4), 372-376. <https://doi.org/10.1016/j.ijbiomac.2009.07.006>
- Varma, A. J., Deshpande, S. V., & Kennedy, J. F. (2004). Metal complexation by chitosan and its derivatives: A review. *Carbohydrate Polymers*, 55(1), 77-93. <https://doi.org/10.1016/j.carbpol.2003.08.005>
- Xiao, J. B. (2007). Adsorption of guanosine, cytidine, and uridine on a β -cyclodextrin derivative grafted chitosan. *Journal of Applied Polymer Science*, 103(5), 3050-3055. <https://doi.org/10.1002/app.25465>
- Yang, C., Xu, L., Zhou, Y., Zhang, X., Huang, X., Wang, M., Han, Y., Zhai, M., Wei, S., & Li, J. (2010). A green fabrication approach of gelatin/CM-chitosan hybrid hydrogel for wound healing. *Carbohydrate Polymers*, 82(4), 1297-1305. <https://doi.org/10.1016/j.carbpol.2010.07.013>
- Yang, J. M., Su, W. Y., Leu, T. L., & Yang, M. C. (2004). Evaluation of chitosan/PVA blended hydrogel membranes. *Journal of Membrane Science*, 236(1-2), 39-51. <https://doi.org/10.1016/j.memsci.2004.02.005>
- Zhang, Q. G., Hu, W. W., Zhu, A. M., & Liu, Q. L. (2013). UV-crosslinked chitosan/polyvinylpyrrolidone blended membranes for pervaporation. *RSC Advances*, 3, 1855-1861. <https://doi.org/10.1039/c2ra21827e>

- Zhang, Y., Chi, H., Zhang, W., Sun, Y., Liang, Q., Gu, Y., & Jing, R. (2014). Highly efficient adsorption of copper ions by a PVP-reduced graphene oxide based on a new adsorptions mechanism. *Nano-Micro Letters*, 6(1), 80-87. <https://doi.org/10.1007/bf03353772>
- Zhao, L., Xu, L., Mitomo, H., & Yoshii, F. (2006). Synthesis of pH-sensitive PVP/CM-chitosan hydrogels with improved surface property by irradiation. *Carbohydrate Polymers*, 64(3), 473-480. <https://doi.org/10.1016/j.carbpol.2005.12.014>
- Zhuang, P. Y., Li, Y. L., Fan, L., Lin, J., & Hu, Q. L. (2012). Modification of chitosan membrane with poly(vinyl alcohol) and biocompatibility evaluation. *International Journal of Biological Macromolecules*, 50(3), 658-663. <https://doi.org/10.1016/j.ijbiomac.2012.01.026>

An Investigation on Indoor Temperature of Modern Double Storey House with Adapted Common Passive Design Strategies of Malay Traditional House

Maryam Qays Oleiwi* and Mohd Farid Mohamed

Department of Architecture, Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM)

ABSTRACT

Past years have witnessed the popularity of traditional Malay house as a common housing type in Malaysia. However, double-storey house has become one of the common types of low-rise housing in Malaysia. Several passive cooling strategies have been adopted to cope with the hot-humid climate of Malaysia. In this study, the thermal comfort of a double-storey house was examined when different passive cooling strategies that were adopted from traditional Malay houses were applied using IES-VE 2019 building simulation software. The simulation was conducted for various design strategies such as changing concrete roof tiles to clay roof tiles, adding two small openings to the attic, removing the ceiling between the upper floor and the attic, and extending the overhang by 50% of its length for all the four facades. All these strategies were tested and compared between full-day natural ventilation

and without any ventilation. The thermal comfort of these strategies was graphically defined based on the operative temperature. These analyses revealed that protecting the building envelope by extending the overhang by 50% of its length for all the four facades could ensure the best thermal comfort is achieved compared to other selected strategies. Recommendations for further studies are also outlined in this paper.

ARTICLE INFO

Article history:

Received: 21 September 2020

Accepted: 14 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.24>

E-mail addresses:

mar_mka@yahoo.com (Maryam Qays Oleiwi)

faridmohamed@ukm.edu.my (Mohd Farid Mohamed)

*Corresponding author

Keywords: Double storey house, IES-VE software, indoor temperature, simulation, thermal comfort, traditional Malay house

INTRODUCTION

Traditional Malay House

For hundreds of years, the traditional Malay house concept has been developed and modified by locals to suit their immediate environmental and cultural needs. Various design strategies can be observed from traditional Malay houses from different states in Malaysia. Although the form and architectural design of traditional Malay houses in all states in Peninsular Malaysia may look different, their response to the local environment and culture does not differ much due to similar local climatic condition. According to the Malaysian Meteorological Department (MMD), Malaysia is an equatorial country characterized by uniform temperature (21-35°C), high humidity (42 - 100%), high solar radiation, and profuse precipitation (1000 mm per year) (MMD, 2017). This tropical climate is designated by the Köppen climate classification (Yau & Hasbi, 2017).

The design strategies of the traditional Malay house are seen as sustainable solutions to the current modern housing development in Malaysia. Two important strategies are the optimization of natural lighting and ventilation through many openings that lead to good indoor environmental quality as well as reduced energy consumption of the house. However, the recent housing development in Malaysia such as double-storey terrace house does not adopt these green strategies and leads to a poor indoor environment (Al-Obaidi & Woods, 2006; Ibrahim et al., 2014; Nugroho et al., 2007; Sadafi et al., 2012; Tinker et al., 2004). Thus, this can lead to possible dependence on air conditioning system or other mechanical systems to ensure indoor thermal comfort which subsequently leads to high energy consumption. Hence, it is crucial for the community, especially designers or developers, to understand and adopt the green design approach that exists in the traditional Malay house design and appropriately applied into modern house design to ensure sustainable development in Malaysia.

Figure 1 and 2 show two traditional Malay houses from two different states, Terengganu and Kelantan, respectively. Both states are located in the Northeast of Malaysia. Both houses share common green strategies such as a raised floor system, use of timber as the main material, ventilation through walls including just below the roof, and many more. The houses demonstrated similarities with the common double-storey terrace or detached houses which is around two-storey height.

The Green/Sustainable Strategies of the Traditional Malay House

Numerous green design approaches can be studied from the traditional Malay house. A review of the literature and fieldwork were completed to list the common green or sustainable design strategies that were used for traditional Malay houses.

Fieldwork visits were completed to at least one traditional house from each state in Peninsular Malaysia to understand the basis of their designs. These include traditional



Figure 1. Original condition of Traditional Malay House of Terengganu at Terengganu Museum, Kuala Terengganu, Terengganu, Malaysia.



Figure 2. Renovated and extended (including building rooms with brick walls under the original on stilts traditional house) Traditional Malay House of Kelantan at Kota Bharu, Kelantan, Malaysia.

houses in Universiti Putra Malaysia (4 units from Perak, Negeri Sembilan, Terengganu, and Pahang), Universiti Kebangsaan Malaysia (1 unit from Kelantan), Politeknik Port Dickson (1 unit of Negeri Sembilan Traditional houses), Lembaga Muzium Negeri Sembilan (2 units

from Negeri Sembilan), Terengganu State Museum (5 units of Terengganu's Traditional houses), Kedah State Museum (2 units Kedah's House), and many more. This fieldwork is completed within 3 years. Literature review was also completed from relevant publications (Lim, 1987; Mohamed, 2018a, Mohamed, 2018b; Nasir, 1985; Nasir & Teh, 2011; Nasir & Teh, 1997; Surat, 2018, Surat, 2016; Surat et al., 2012; Yusoff & Mohamed, 2017). Authors' observation during the fieldwork and literature review revealed 23 significant green or sustainable approaches (Table 1) that commonly existed in various types of traditional Malay houses. However, the list is not limited to this.

Table 1

Selected significant green or sustainable approaches that commonly exist in various traditional Malay houses.

No	Green or sustainable approaches
1	Large window opening to allow for natural lighting and ventilation. Some of the openings are designed to be full height.
2	Wide roof overhang for shading and protection from heavy rain and direct sunlight.
3	Built on stilt with concrete or stone base to allow for working space, livestock, as well as to protect from wet ground, flood, and dangerous animals/insects.
4	Steep roof to allow an uninterrupted flow of rainwater and to clean the roof. Also, it is common to have a two-tier roof to allow natural ventilation through gaps between the two roofs.
5	The area with open space such as verandah for activities including entertaining visitors.
6	Opening at roof attic to cool the space under the roof.
7	Roof made of leaves or clay (Singgora) for better heat protection from sunlight.
8	Opening on walls (Examples: wooden craft and timber louver) to allow for natural lighting and ventilation, as well as for security reasons.
9	Natural construction materials such as timber and bamboo that are readily available from the site.
10	Open floor plan for uninterrupted airflow and daylight to penetrate into the house.
11	House plan with a small width for greater efficiency of natural ventilation and lighting.
12	A modular house plan that allows easy construction, future extension, and relocation.
13	Rain harvesting is an alternative water source to water well for domestic water usage.
14	Different floor levels to differentiate between various purposes of spaces, to suit timber joint construction, and to allow ventilation between the floor gap.
15	Seating area below the house for resting, meeting visitors, storage, or large event or gathering.
16	Construction using a wood joint for stronger construction as well as easy assemble and disassemble.

Table 1 (Continued)

No	Green or sustainable approaches
17	More than one entrance to allow for access or exit for various purposes, such as in the event of a large gathering, normally, a woman will use the entrance at the kitchen.
18	Connecting space between two indoor space for natural lighting and ventilation, as well as various activities such as drying clothes and other family activities.
19	Flooring with gaps to allow for natural ventilation as well for cleaning & draining.
20	High indoor spaces that allow better natural ventilation airflow.
21	Attic space with ladder for various purposes such as protection (hiding), room for women (daughter), and storage
22	Orientation towards Qibla due to religious factors and subsequently lead to the housing to face west sunlight (the house protected from the direct sunlight by building elements such as long overhang and verandah), and some respond to site context such as road.
23	Reuse of materials especially hardwood such as Cengal wood that can last for hundreds of years and the reuse purpose is supported by the jointing system (without nail) that allow the house to be easily disassembled and relocated to other locations.

The approaches listed in Table 1 are common design approaches in traditional Malay house design and are influenced by many social, economic, and environmental factors. This study observed that the cultural differences among many states contributed to the varying architectural aesthetic of traditional Malay houses according to the states in Malaysia. Despite the differences, the basic design approaches are generally similar.

Thermal Comfort

Thermal comfort occurs in a narrow range of body temperatures and the skin moisture is kept low with minimum physiological effort (ASHRAE, 1992). Occupants in naturally ventilated buildings have a wide range of comfort temperatures which may exceed the limitations of ASHRAE Standard 55, and also a response to local outdoor climate changes. On the contrary, occupants who stay in air-conditioned buildings have a very narrow range of comfort temperature as they develop a high expectation for cool temperature, and as a result, they become critical to any changes (Brager & de Dear, 2000).

The objective of this study is to investigate the effect of the application of selected green passive cooling strategies of typical traditional Malay houses on a double-storey modern detached home on the aspects of indoor temperature and thermal environment.

METHODOLOGY

In this study, several design strategies that traditional Malay house used were selected and

applied to the tested house (Case 1) to investigate the effect of these strategies on internal thermal environment. The investigation was conducted and compared under two conditions which were fully ventilated and totally without ventilation. The selected approaches are as follows:

- a. Clay roof tiles: to investigate the roof materials with different thermal transmittance (U-value).
- b. Ventilated attic space: to allow ventilation within the attic for heat removal.
- c. No ceiling underneath the roof: to allow higher clear vertical height but without protection from the heat originated from the roof tile.
- d. Extended overhang for all the four facades: to further protect façade from direct sunlight.

The tested house was located at the Faculty of Engineering and Built Environment, Universiti Kebangsaan Malaysia (UKM), Bangi, 27 km from Kuala Lumpur, the capital city of Malaysia. Figure 3 shows the double-storey house that was built using common modern construction building materials (sand brick and cement for the walls and concrete roof tiles for the roof). This house was built purposely for research activities. Therefore, easy access of the house enabled the researcher to enter the house easily for site investigation.

The house includes a living room, kitchen, and a bathroom on the ground floor, and three bedrooms and two toilets on the first floor. In this study, Room 1 refers to the living room, Room 2 refers to the second bedroom, and Room 3 refers to the master bedroom. The windows of the house faced the four directions. Figure 4 shows the ground and first floors' plans of the tested house.

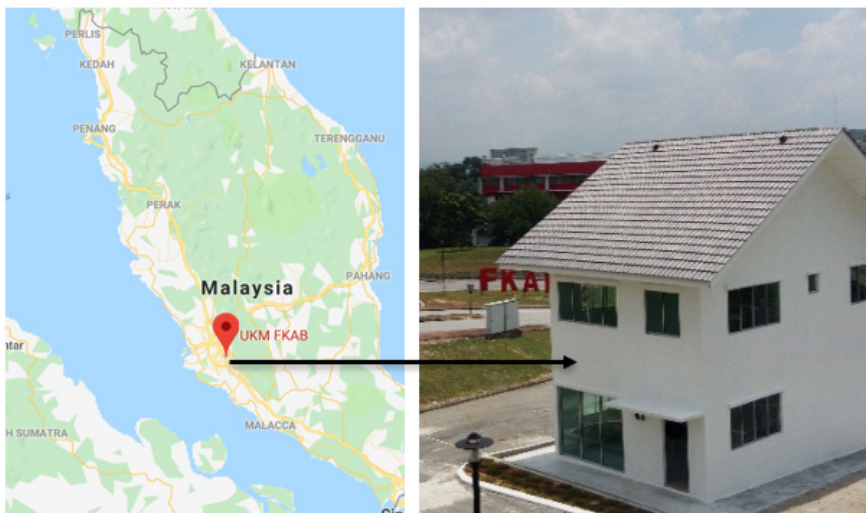


Figure 3. The tested house

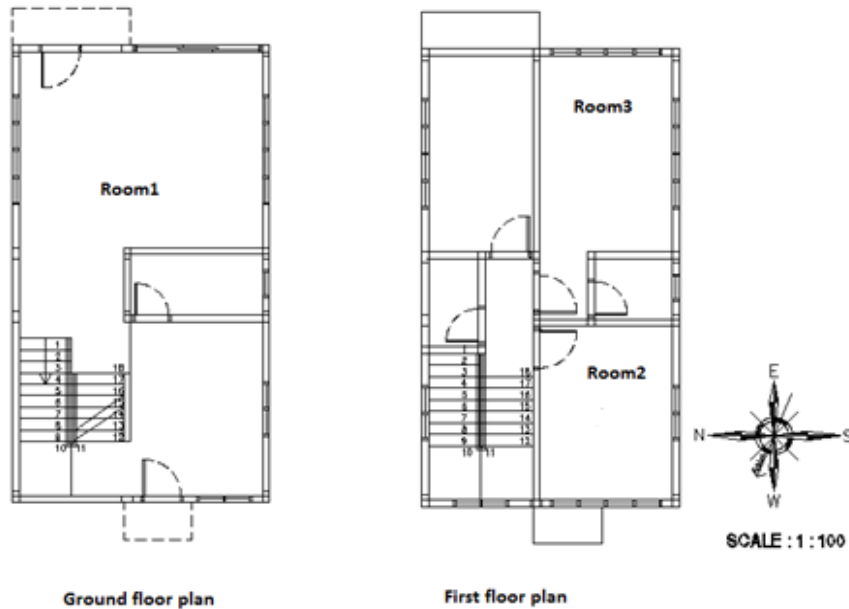


Figure 4. The plans of the tested house

The geometry of the computational model for the tested house was created using the IES-VE (2019) software. The building's geometry, dimensions, and areas of solid and transparent surfaces were defined using [ModelIT]. The location of the house was defined by [APlocate], which was set as Kuala Lumpur suburbs as it was the nearest to the site. Windows and doors openings profile and windows' types were set using [MacroFlo]. Then, fabric materials and their thermal properties and simulation time were defined using [Apache]. Table 2 shows the details of construction materials used in the simulation of the tested house.

The simulation was done in April which falls in the hottest period in the year in Malaysia as reported by the MMD (2017) and Sulaiman (2017). The weather data of this study were provided by IES-VE which can be determined according to the nearest location that was defined earlier.

The main simulation output obtained from IES-VE was the indoor and outdoor air temperature (T_a) and the mean radiant temperature (T_{mrt}) for each case. The results were converted into Excel sheets to facilitate data analysis. IES-VE was selected as it is a software that has a simple input interface that creates the prototype and modifies its layout. The output can be easily interpreted and represented in tabular data or graphs. Finally, IES-VE is user-friendly, which makes this software useful for this study. A three-dimensional view of the building modelled in IES-VE is illustrated in Figure 5.

Table 2

Details of construction materials for the tested house. The values are the standard values in the software.

Element	Construction	U-Value (W/m ² k)
External walls	Cement plaster (19 mm), sand brick with cement (114 mm), cement plaster (19 mm)	3.4
Internal walls	Cement plaster (19 mm), sand brick with cement (114 mm), cement plaster (19 mm)	3.4
Roof (concrete tiles)	Concrete tiles (5 mm), aluminium sisalation (1 mm), UCO Superflex ceiling board (3.2 mm)	6.9
Roof (clay tiles)	Clay tiles (5 mm), aluminium sisalation (1 mm), UCO Superflex ceiling board (3.2 mm)	6.85
Floor	Cast concrete (220 mm)	3.6
Doors	Plywood	2.2
Windows	Glass (one-layer clear float 5 mm)	5.2

IES-VE permits the user to change the types of construction materials, dimensions, thicknesses for all house components (roof, ground, and walls), and windows opening profile.

Firstly, the concrete roof tiles that were used in the tested house were changed to clay roof tiles (Case 2), while other original conditions of Case 1 remained. In Case 3, two small openings were added to the attic, while other original conditions of Case 1 remained. In Case 4, the ceiling between the upper floor and the attic was removed to determine any difference, while other original conditions of Case 1 remained. In Case 5, the overhang of the house was extended by 50% of its length for all four facades (from 1.2 to 1.8 meter) to increase its shading on the walls and windows, while other original conditions of Case 1 remained. All the previous strategies were done with full-day natural ventilation (all windows and internal doors were open for 24 hours) and no ventilation (all windows and

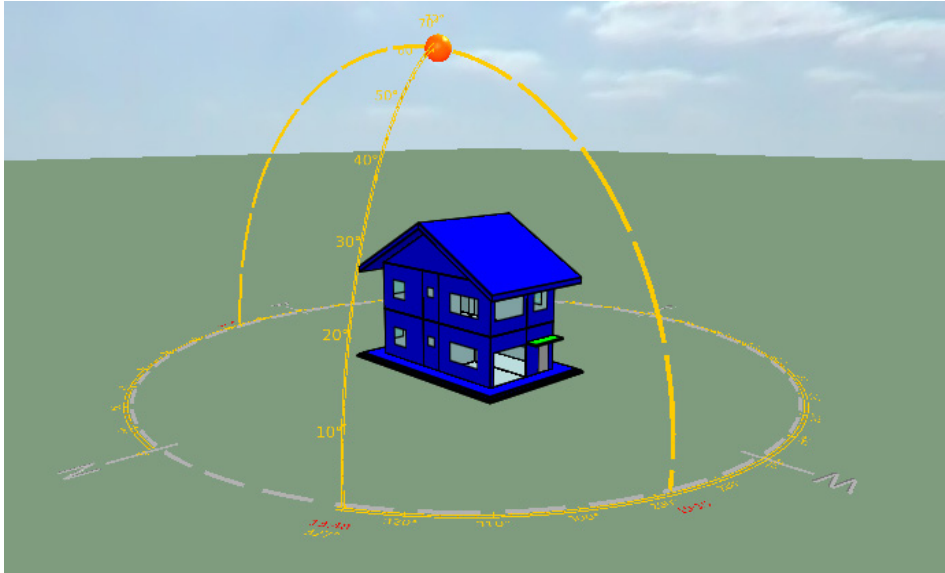


Figure 5. The tested house in IES-VE software

doors were closed for 24 hours) separately. All five cases were simulated on the same date of the year using the validated model of the tested house at IES-VE software. The validation process of IES-VE software was previously performed by Oleiwi et al. (2019).

The main simulation output obtained from IES-VE was the indoor air temperature (T_a) and the mean radiant temperature (T_{mrt}) for the three rooms. The results were converted into Excel sheets to facilitate data analysis.

Operative temperature (T_o) was calculated based on the Equation 1 that was proposed by ASHRAE Standards (ASHRAE, 2010):

$$T_o = \frac{(T_a + T_{mrt})}{2} \quad (1) \quad (3.1)$$

Where T_o is the operative temperature in °C, T_a is the air temperature in °C, and T_{mrt} is the mean radiant temperature in °C.

RESULTS AND DISCUSSION

A comparison of the operative temperature of the five cases was conducted by applying full-day ventilation (all windows and internal doors were opened for 24 hours) and no ventilation (all windows and doors were closed for 24 hours) to identify the best strategy to achieve thermal comfort for the tested house. Thereafter, strategies that could potentially achieve the best thermal comfort were combined to test if the combination can increase thermal comfort inside the tested house.

The effect of different strategies on indoor thermal environment using the validated model (with ventilation)

Table 3 shows the average operative temperature of April of the simulated five cases with full natural ventilation. Figure 6 shows the average hourly operative temperature of April to demonstrate the patterns of the five cases with full-day ventilation.

Table 3

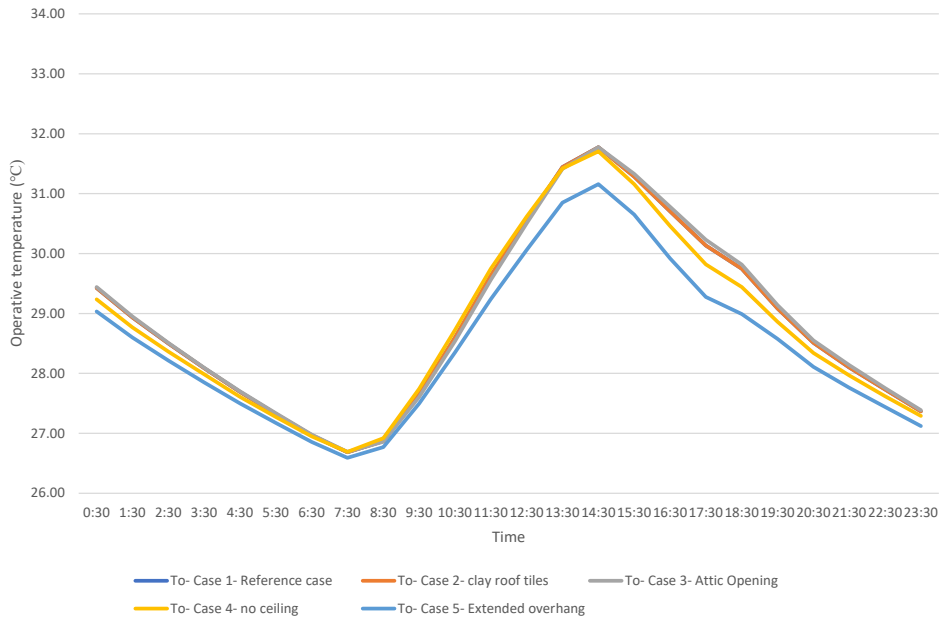
Operative temperature of the simulated five cases with full natural ventilation.

Thermal variables	Statistics	Case 1	Case 2	Case 3	Case 4	Case 5
Room 1						
Operative temperature (°C)	Mean	28.87	28.87	28.88	28.78	28.48
	Maximum	31.78	31.78	31.78	31.71	31.16
	Minimum	26.69	26.69	26.69	26.69	26.59
Room 2						
Operative temperature (°C)	Mean	28.84	28.84	28.86	28.64	28.22
	Maximum	31.81	31.81	31.81	31.51	30.72
	Minimum	26.92	26.92	26.93	27.01	26.86
Room 3						
Operative temperature (°C)	Mean	28.88	28.88	28.89	28.69	28.24
	Maximum	31.76	31.76	31.75	31.53	30.73
	Minimum	27.00	27.00	27.01	27.05	26.87

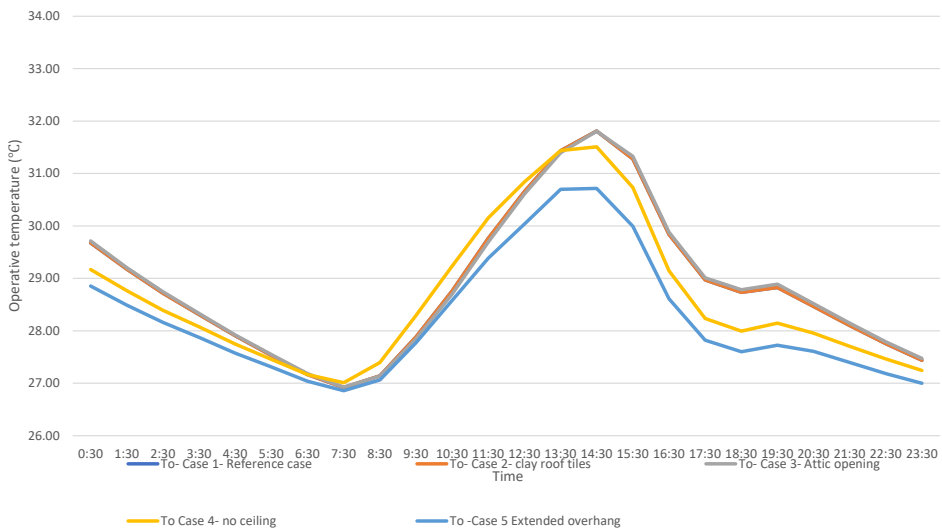
Table 2 and Figure 6 depict that no changes happened in the mean, maximum, and minimum operative temperature when the concrete roof tiles were changed to clay roof tiles (Case 2). This could be attributed to the small difference between the u-value of clay roof tiles for 5mm thickness (6.85 W/m²K) and the u-value of concrete roof tiles (6.85 W/m²K) with the same thickness.

It was observed that the addition of two small windows to the attic in Case 3 was an ineffective strategy when full-day ventilation was applied for all the three rooms. The highest maximum operative temperatures for Room 1, Room 2 and Room 3 were 28.88, 28.86, and 28.89°C during the afternoon hours, respectively.

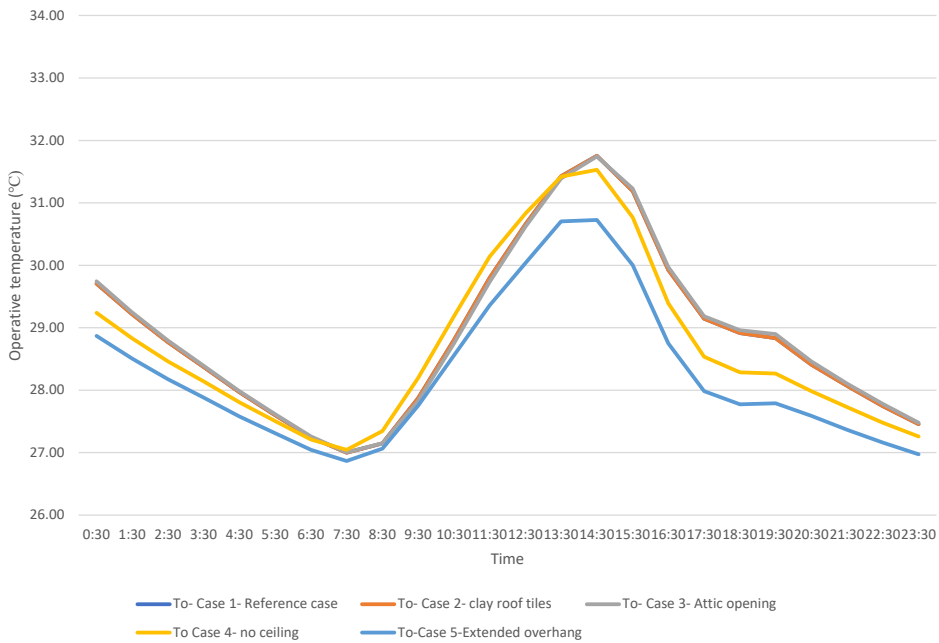
Effects of Adaptive Malay Traditional House's Passive Cooling Strategy



(a) Room 1



(b) Room 2



(c) Room 3

Figure 6. The hourly operative temperature patterns of the five cases with full-day ventilation strategy

In Case 4, the removal of the ceiling between the rooms in the upper floor and the attic contributed to a slight decrement from the reference case (Case 1) in the mean, maximum, and minimum operative temperature of Room 1, and the mean and maximum operative temperature of Room 2 and Room 3.

The best decrement was obtained when the overhang was extended by 50% of its length for all the four facades (Case 5) in the mean, maximum, and minimum operative temperature of Room 1, Room 2, and Room 3.

The previous cases were tested with 24 hours of natural ventilation when outdoor air can enter the rooms and affect the indoor environment. Subsequently, no ventilation strategy was applied by closing all the windows and doors for the same five cases to remove the effect of outdoor air.

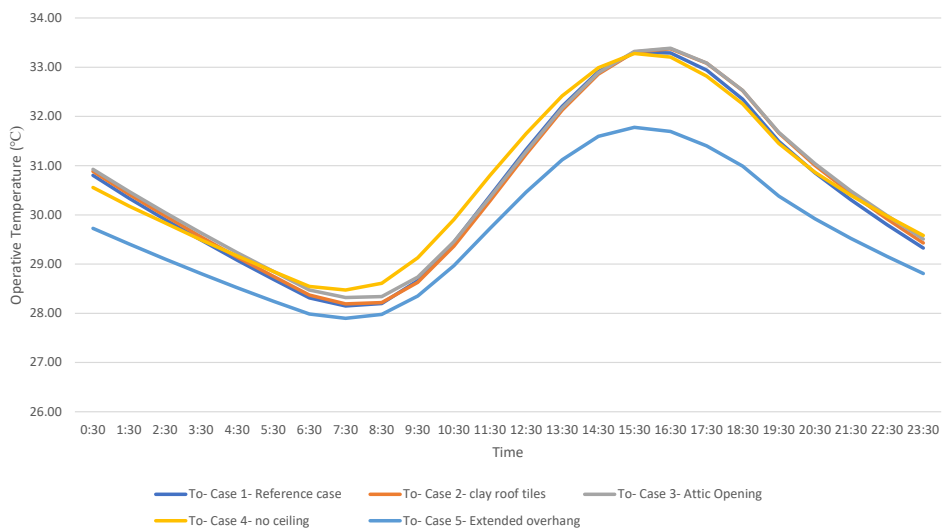
The effect of different strategies on the indoor thermal environment using the validated model (without ventilation)

Table 4 shows the average operative temperature of April of the simulated five cases with no ventilation. Figure 7 depicts the average hourly operative temperature of April and demonstrates the patterns of the five cases with no ventilation.

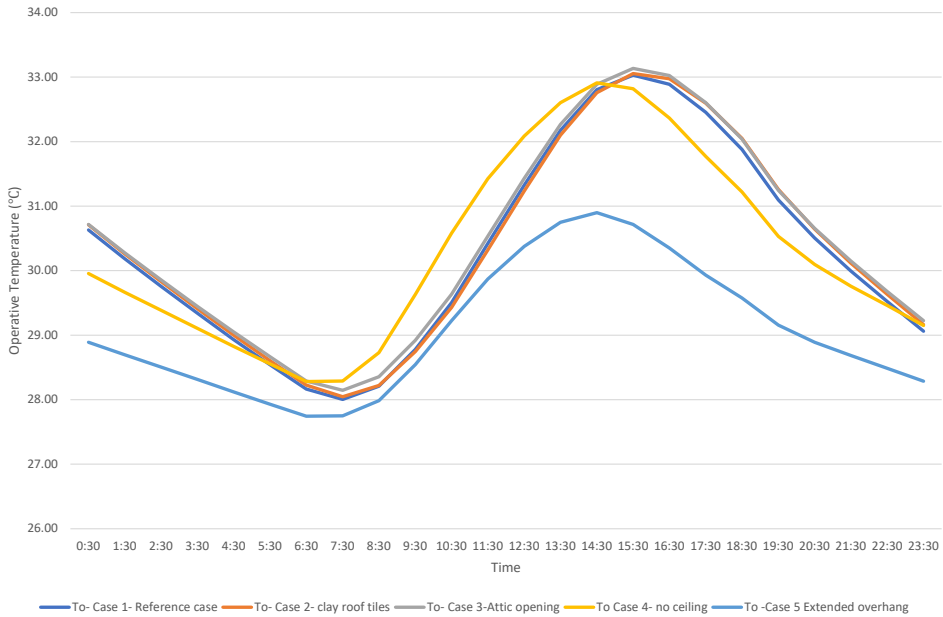
Table 4

Operative temperature of the simulated five cases without ventilation.

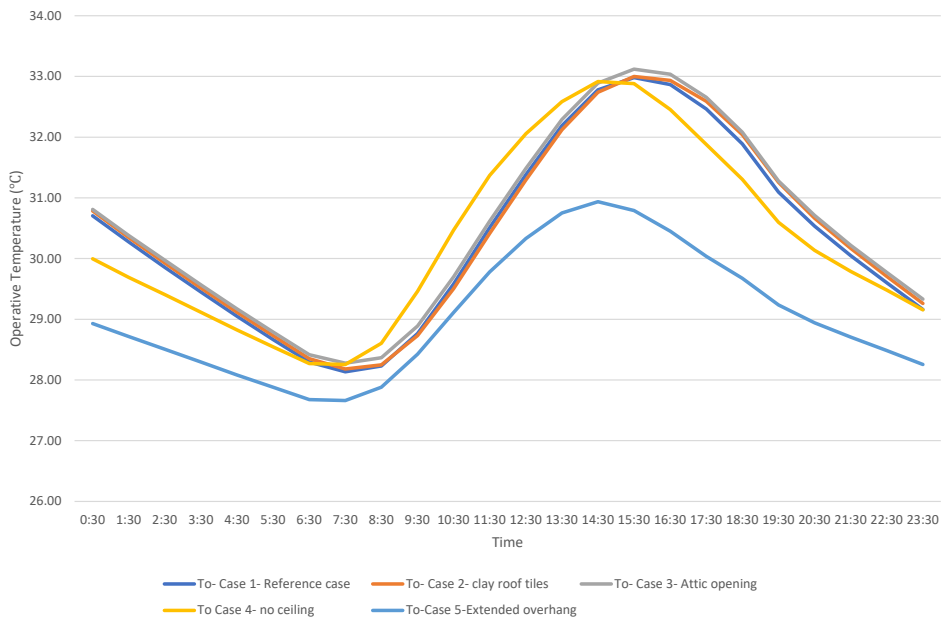
Thermal variables	Statistics	Case 1	Case 2	Case 3	Case 4	Case 5
Room 1						
Operative temperature (°C)	Mean	30.48	30.53	30.59	30.60	29.65
	Maximum	33.30	33.38	33.39	33.28	31.78
	Minimum	28.15	28.19	28.32	28.48	27.90
Room 2						
Operative temperature (°C)	Mean	30.30	30.35	30.43	30.32	29.07
	Maximum	33.03	33.06	33.14	32.91	30.90
	Minimum	28.01	28.05	28.15	28.28	27.75
Room 3						
Operative temperature (°C)	Mean	30.35	30.40	30.49	30.30	29.06
	Maximum	32.98	33.00	33.12	32.92	30.94
	Minimum	28.14	28.18	28.28	28.26	27.66



(a) Room 1



(b) Room 2



(c) Room 3

Figure 7. The hourly operative temperature patterns of the five cases without ventilation.

Table 3 and Figure 7 highlighted that there was a slight difference in the operative temperature of the three rooms for all the other cases when compared with the reference case (Case 1).

The mean, maximum, and minimum operative temperature of all the three rooms was slightly higher when the concrete roof tiles were changed to clay roof tiles (Case 2). Moreover, the maximum operative temperature of Room 1, the mean and maximum operative temperature of Room 2, and the mean, maximum and minimum operative temperature of Room 3 registered the highest value when two small windows were added to the attic (Case 3).

In addition, the mean and minimum operative temperature of Room 1 and the minimum operative temperature of Room 2 registered the highest operative temperature when the ceiling between the upper floor and the attic was removed (Case 4). However, the maximum operative temperature of Room 1 and Room 2 and the mean and maximum operative temperature of Room 3 was slightly less than the reference case (Case 1).

The best thermal comfort was achieved when the overhang of the house was extended by 50% of its length for all the four facades of the house (Case 5). The mean, maximum, and minimum operative temperature of Room 1, Room 2, and Room 3 registered the lowest operative temperature values.

The previous results concluded that there is no thermal comfort improvement when clay roof tiles were used (Case 2) or two small windows were added to the attic (Case 3) for both full-day ventilation and no ventilation. However, the removal of the ceiling between the rooms of the upper floor and the attic contributed to a slight decrement in the operative temperature inside the house for the two ventilation strategies. Additionally, extending the overhang to increase shading and protect the wall and windows from direct solar radiation for all the four facades of the house was identified as the best-tested strategy in this study to decrease the operative temperature inside the tested house in both full-day ventilation and no ventilation. Therefore, Case 4 and Case 5 were combined to examine their effect on thermal comfort inside the tested house for both full-day ventilation and no ventilation.

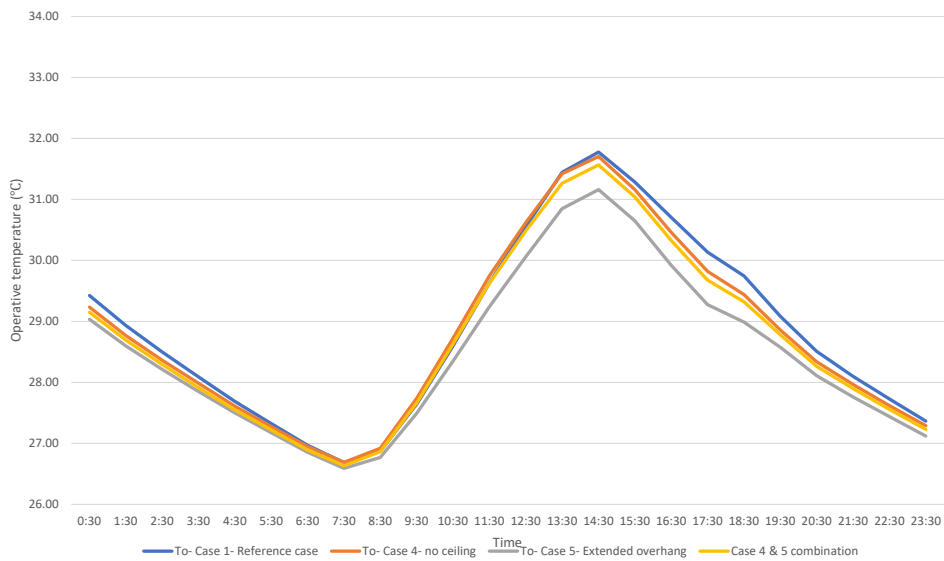
The effect of combined strategies on the indoor thermal environment (with ventilation)

Case 4 and Case 5 were reported to have achieved the best thermal comfort for all three rooms for ventilation and no ventilation strategies. Therefore, a simulation for a combination of the two cases was performed (Case 6) using the same validated model. Table 5 compares the average operative temperature of April of Case 6 and other related cases. Figure 8 shows the average hourly operative temperature of April and demonstrates the patterns of the selected cases with ventilation.

Table 5

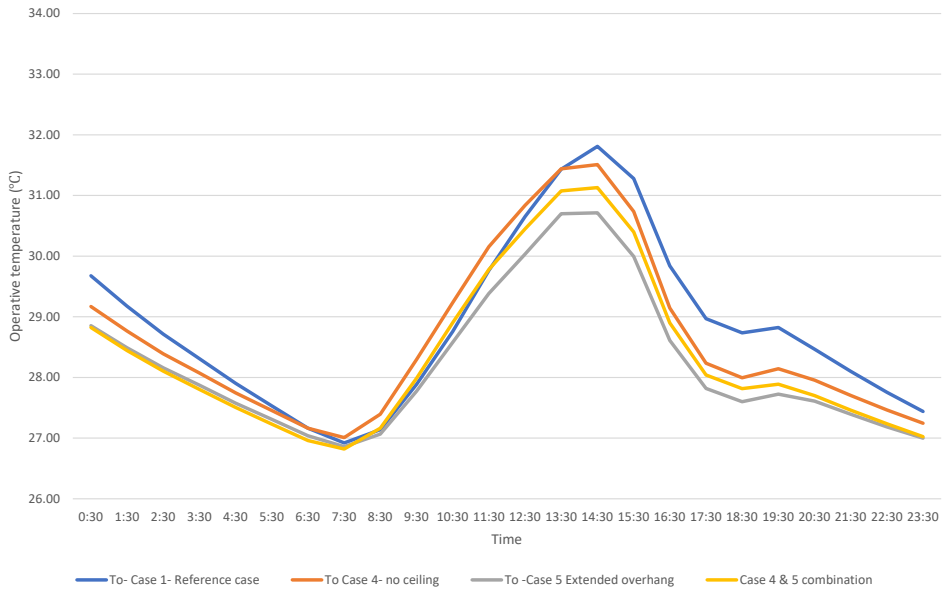
Operative temperature of the simulated selected strategies with ventilation.

Thermal variables	Statistics	Case 1	Case 4	Case 5	Case 6 (combination of case 4 and case 5)
Room 1					
Operative temperature (°C)	Mean	30.48	28.78	28.48	28.69
	Maximum	33.30	31.71	31.16	31.57
	Minimum	28.15	26.69	26.59	26.64
Room 2					
Operative temperature (°C)	Mean	30.30	28.64	28.22	28.36
	Maximum	33.03	31.51	30.72	31.13
	Minimum	28.01	27.01	26.86	26.82
Room 3					
Operative temperature (°C)	Mean	30.35	28.69	28.24	28.36
	Maximum	32.98	31.53	30.73	31.13
	Minimum	28.14	27.05	26.87	26.82

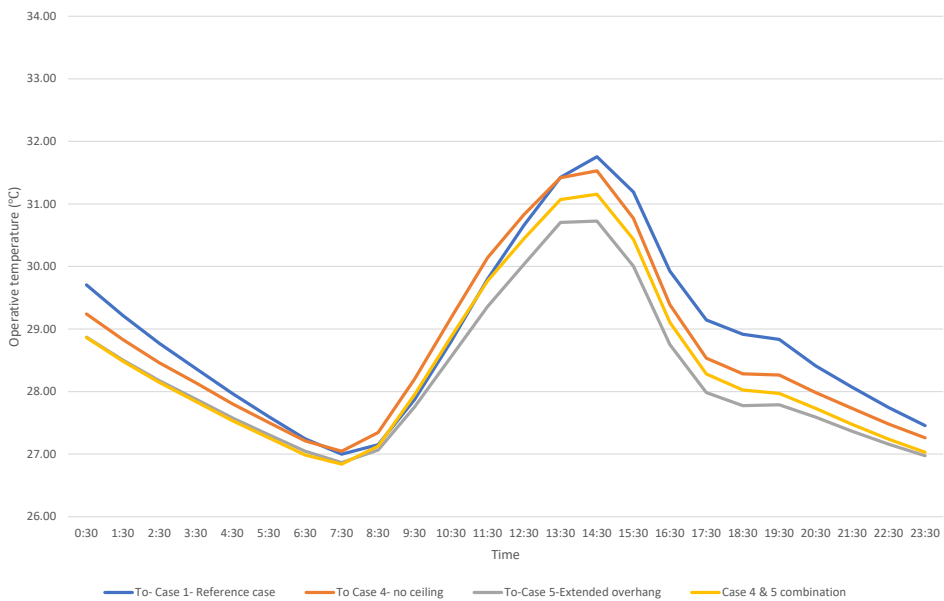


(a) Room 1

Effects of Adaptive Malay Traditional House's Passive Cooling Strategy



(b) Room 2



(c) Room 3

Figure 8. The hourly operative temperature patterns of the selected cases with ventilation.

Table 4 and Figure 8 established that the mean, maximum, and minimum operative temperature of Room 1 and the mean and maximum operative temperature of Room 2 and Room 3 when the overhang was extended (Case 5) was less than Case 6 (the combination of removing the internal ceiling and extending the overhang). The best thermal comfort was not achieved in Case 6 and this is attributable to the effect of the roof that receives a high amount of solar radiation during the daytime and there is no barrier or insulation to protect the rooms. Moreover, when full-day ventilation was applied, the windows and internal doors were open, which may increase the operative temperature by spreading the hot air inside the house. For the same reason, the minimum operative temperature of Room 2 and Room 3 in Case 6 was less as the cool outdoor air could enter the house easily and spread inside the rooms.

The Effect of Combined Strategies on the Indoor Thermal Environment (Without Ventilation)

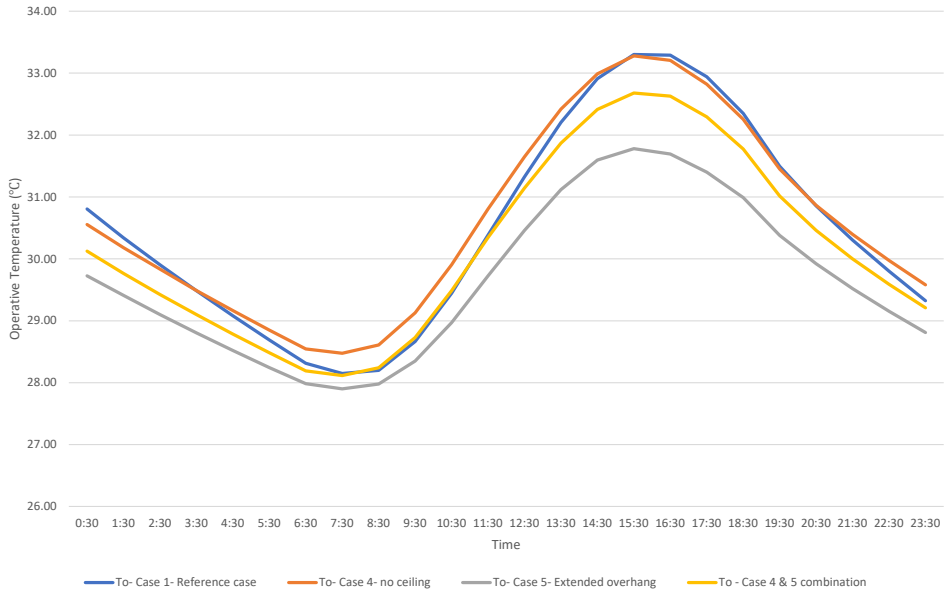
Table 6 shows the average operative temperature of April for Case 6 and other related cases. Figure 9 shows the average hourly operative temperature of April and also the patterns of the selected cases without ventilation.

Table 6

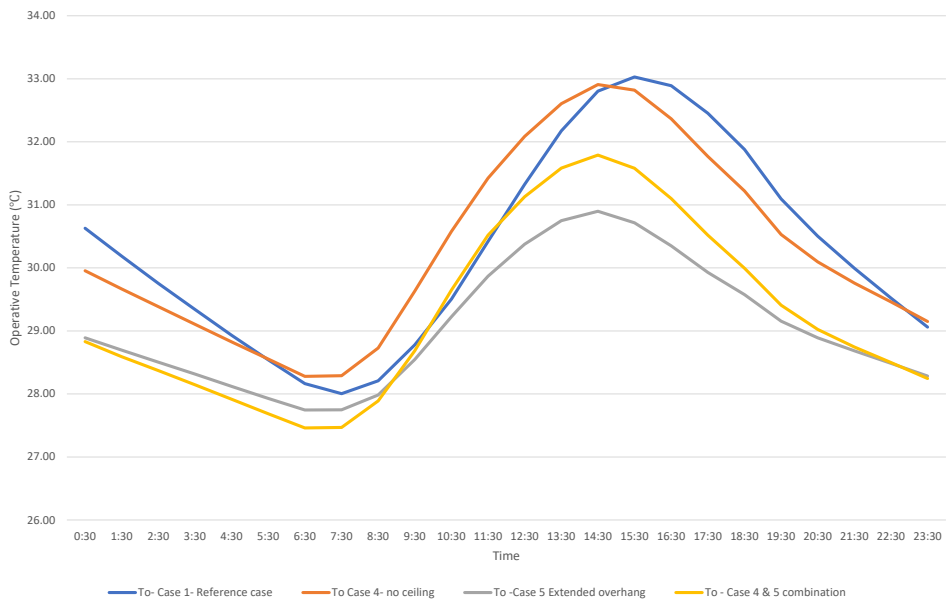
Operative temperature of the simulated selected strategies without ventilation.

Thermal variables	Statistics	Case 1	Case 4	Case 5	Case 6
Room 1					
Operative temperature (°C)	Mean	30.48	30.60	29.65	30.16
	Maximum	33.30	33.28	31.78	32.68
	Minimum	28.15	28.48	27.90	28.12
Room 2					
Operative temperature (°C)	Mean	30.30	30.32	29.07	29.28
	Maximum	33.03	32.91	30.90	31.79
	Minimum	28.01	28.28	27.75	27.46
Room 3					
Operative temperature (°C)	Mean	30.35	30.30	29.06	29.27
	Maximum	32.98	32.92	30.94	31.80
	Minimum	28.14	28.26	27.66	27.43

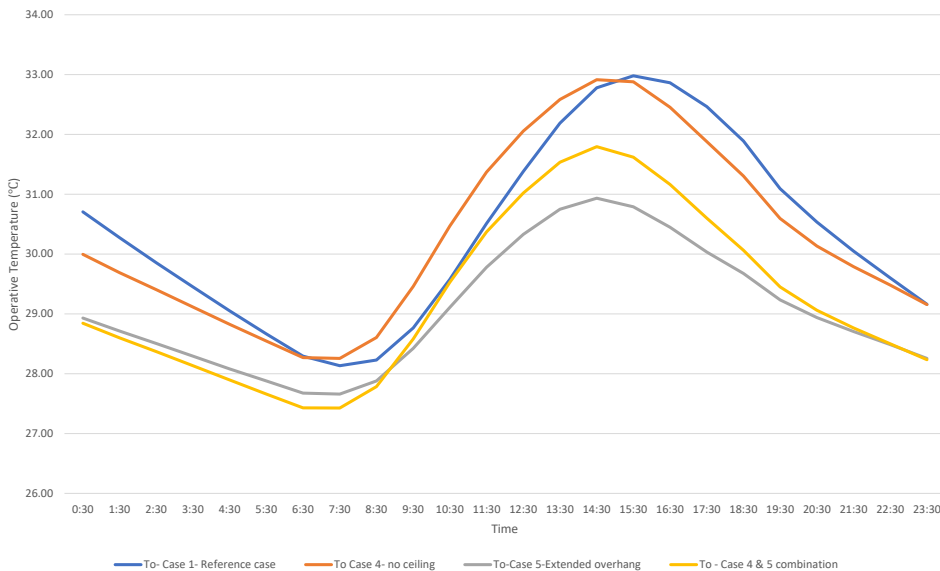
Effects of Adaptive Malay Traditional House's Passive Cooling Strategy



(a) Room 1



(b) Room 2



(c) Room 3

Figure 9. The hourly operative temperature patterns of the combination of case 4 and case 5 without ventilation.

Table 5 and Figure 9 reported that the mean, maximum, and minimum operative temperature of Room 1 and the mean and maximum operative temperature of Room 2 and Room 3 for Case 5 (when the overhang was extended) was less than Case 6 (the combination of removing the internal ceiling and extending the overhang). Meanwhile, the minimum operative temperature of Room 2 and Room 3 in Case 6 was less.

A possible reason for this situation is the effect of the high amount of solar radiation that the roof received during daytime. The removal of the ceiling creates no barrier to protect the rooms. However, the minimum operative temperature of Room 2 and Room 3 in Case 6 was less as the effect of solar radiation had disappeared during night-time and the effect of natural ventilation appeared.

Despite the best minimum operative temperature was received in Room 2 and Room 3 in Case 6 with and without ventilation, the improvement is minimal when compared with Case 5 for the same rooms. Therefore, this study identified that extending the overhang of the tested house by 50% of its length for all the four facades of the house (case 5) can reduce the operative temperature of the three rooms and enhance indoor thermal comfort. It must also be highlighted that low wind speed could be the main reason behind low-temperature changes when natural ventilation was applied due to low heat flushing through the openings.

CONCLUSION

The outcomes of this study showed that many sustainable strategies can be acquired from traditional Malay houses. These include strategies that were examined in this study. Strategies that were tested were clay roof tiles instead of concrete tiles, naturally ventilating the roof by adding several openings to the attic, changing the design of the roof by removing the ceiling, and extending the roof overhang to provide shade to the walls and windows and decrease the effect of solar radiation. These selected strategies were investigated and compared in this paper with full-day ventilation and without ventilation. The IES- VE simulation software was used to simulate the tested house and apply all the mentioned strategies.

The results proved that the use of clay roof tiles instead of concrete roof tiles or adding small openings to the attic would not enhance thermal comfort inside the tested house when using full-day ventilation and no ventilation. In addition, the removal of the ceiling between the attic and the upper floor contributed to a small amount of decrement in the operative temperature inside the house when full-day ventilation was applied. However, this strategy brought a negative effect when there was no ventilation. An interesting finding in this study is that the indoor operative temperature of test house reduced by extending the overhang with 50% of its length for all the four facades of the house which is ideal to protect the building's envelope from direct solar radiation. This can subsequently improve indoor thermal environment and reduce energy consumption of the modern house.

It is recommended for future research to examine other cooling strategies such as adding shading devices to the windows and using different materials for the walls and the glass of the windows to determine other effective strategies that can enhance thermal comfort inside modern houses.

ACKNOWLEDGEMENT

Maryam Qays Oleiwi would like to acknowledge and thank the Ministry of Education, Malaysia, for their support through the Fundamental Research Grant Scheme (FRGS/1/2020/TK0/UKM/02/26 – Roof Design in Controlling Mosque's Indoor Thermal Comfort). The authors would also like to thank Universiti Kebangsaan Malaysia (UKM) for the financial support provided through Dana Cabaran Perdana (DCP-2017-008/5).

REFERENCES

- Al-Obaidi, M. A. A. H., & Woods, P. (2006). Investigations on effect of the orientation on thermal comfort in terraced housing in Malaysia. *International Journal of Low-carbon Technologies*, 1(2), 167-176. <https://doi.org/10.1093/ijlct/1.2.167>
- ASHRAE. (1992). *Thermal environmental conditions for human occupancy*. American Society of Heating, Refrigeration and Airconditioning Engineers.

- ASHRAE. (2010). *Standard 55-2010: Thermal environmental conditions for human occupancy. 2010*. American Society of Heating, Refrigerating and Airconditioning Engineers.
- Brager, G. S., & de Dear, R. (2000). *A standard for natural ventilation*. American Society of Heating, Refrigerating and Airconditioning Engineers.
- Ibrahim, S., Baharun, A., & Nawi, M. M. (2014). Analytical studies on levels of thermal comfort in typical low-income houses design. *Journal of Civil Engineering, Science and Technology*, 5(1), 28-33. <https://doi.org/10.33736/jcest.130.2014>
- Lim, J. Y. (1987). *The Malay house: Rediscovering Malaysia's indigenous shelter system*. Institut Masyarakat.
- MMD. (2017). *General climate of Malaysia*. Malaysia Meteorological Department. Retrieved August 20, 2017, from <http://www.met.gov.my/en/web/metmalaysia/home>
- Mohamed, M. F. (2018a, November 13). *21 Green Approaches in Traditional Malay House* [Paper presenttaion]. Universiti Kebangsaan Malaysia Integrated International Conference (UKM-IIC) 2018. Putrajaya, Malaysia.
- Mohamed, M. F. (2018b, January 17-18). *Learning Traditional Malay House Through Working Drawing Course* [Poster presentation]. K-Inovasi 2018. Selangor, Malaysia.
- Nasir, A. H. (1985). *Pengenalan rumah tradisional Melayu semenanjung Malaysia* [Introduction to traditional malay houses in Peninsular Malaysia]. Loyal Press.
- Nasir, A. H., & Teh, W. H. W. (2011). *The traditional Malay house*. Institut Terjemahan Negara Malaysia.
- Nasir, A. H., & Teh, W. H. W. (1997). *Warisan seni bina Melayu* [Malay architecture heritage]. Penerbit Universiti Kebangsaan Malaysia.
- Nugroho, A. M., Ahmad, M. H., & Ossen, D. R. (2007). A preliminary study of thermal comfort in Malaysia's single storey terraced houses. *Journal of Asian Architecture and Building Engineering*, 6(1), 175-182. <https://doi.org/10.3130/jaabe.6.175>
- Oleiwi, M. Q., Mohamed, M. F., Sulaiman, M. K. A. M., Che-Ani, A. I., & Raman, S. N. (2019). Thermal environment accuracy investigation of integrated environmental solutions-virtual environment (IES-VE) software for double-story house simulation in Malaysia. *Journal of Engineering and Applied Sciences*, 14(11), 3659-3665. <https://doi.org/10.36478/jeasci.2019.3659.3665>
- Sadafi, N., Salleh, E., Haw, L. C., & Jaafar, M. F. Z. (2012). Potential design parameters for enhancing thermal comfort in tropical Terrace House: A case study in Kuala Lumpur. *ALAM CIPTA Journal*, 3(1), 15-24.
- Sulaiman, M. K. A. M. (2017). *Cooling effect performance of indirect green facade on building in tropical climate of Malaysia* [Doctor of Philosophy thesis]. Universiti kebangsaan malaysia, Malaysia.
- Surat, M. (2016, December 24-25). *Peranan budaya bangsa dalam pembangunan senibina Malaysia: Garapan dari teras senibina warisan Melayu* [The role of nation culture in Malaysia Architecture Development: A study from the thrusts of Malay heritage architecture]. In *Philosophy of Science and Technology Seminar II*. Anwar Ibrahim Research Institute, Kajang, Malaysia.
- Surat, M. (2018, November 13-14). Cadangan taksonomi pengukuran taraf senibina identiti kebangsaan Malaysia [Proposed rating measurement taxonomy for Malaysian national architectural identity]. In

Proceedings of Conference on Regional Architecture and Built Environment 2018 (SeniBINA2018), 'Theme: National Architectural Identity'. Palm Garden Hotel Putrajaya, Malaysia.

- Surat, M., Baharum, M., Usman, I., Musa, A., & Tawil, N. (2012). Mengenalpasti tahap kesejahteraan seni bina warisan Melayu melalui konsep Islam [Identifying the level of wellness in Malay heritage architecture through the concept of Islam]. *Journal of Design+ Built*, 5(1), 1-11.
- Tinker, J., Ibrahim, S., & Ghisi, E. J. A. B. (2004). An evaluation of thermal comfort in typical modern low-income housing in Malaysia. *Buildings*, 9, 1-5.
- Yau, Y., & Hasbi, S. J. E. P. (2017). A comprehensive case study of climate change impacts on the cooling load in an air-conditioned office building in Malaysia. *Energy Procedia*, 143, 295-300. <https://doi.org/10.1016/j.egypro.2017.12.687>
- Yusoff, W. F. M., & Mohamed, M. F. (2017). Building energy efficiency in hot and humid climate. In *Encyclopedia of Sustainable Technologies* (pp. 159-168). Elsevier.



Assessment by Simulation of Different Topological Integration of Solar Photovoltaic Plant in Medium Voltage Distribution Networks

Md. Milon Uddin*, Mushfiqur Rahman, Md. Tanzid Ridwan Hossain and Md. Habibur Rahman

Institute of Energy, University of Dhaka, Dhaka-1000, Bangladesh

ABSTRACT

Fossil fuels are diminishing day by day and are being utilized for various purposes like transportation, electricity generation, cooking and other uses which emits harmful gases. Global warming is increasing due to greenhouse gases. Sunlight is abundant compared to fossil fuels. Photovoltaic (PV) power plant could be one of the best techniques to lessen using of fossil fuels. Effect of different topological integration of solar photovoltaic plant in medium voltage distribution network had been analyzed in this research. Analysis had been done based on 1. loadability (capability of an electrical grid to run the connected load) 2. maximum power input. Fractioning (a way of splitting the total photovoltaic power generation into different size and number of plant) generation effect of PV generators in distribution systems had been found by a comparative study. Results show that change in loadability was due to placement and fraction of solar power plant in various buses. In addition, maximum input power of PV system varied with the placement of generation units among picked buses. Interaction among PV systems, and induction machine had been carried out and the result shows that interaction among PV and induction machine was different in terms of loadability because of fractioning. It had been found that change of loadability occurred due to fractioning and distribution of photovoltaic systems on different types of designed topology.

ARTICLE INFO

Article history:

Received: 05 September 2020

Accepted: 16 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.25>

E-mail addresses:

milon.brac456@gmail.com (Md. Milon Uddin)

mushfiqb91@gmail.com (Mushfiqur Rahman)

tanzid_ridwan@yahoo.com (Md. Tanzid Ridwan Hossain)

mhabib@du.ac.bd (Md. Habibur Rahman)

* Corresponding author

Keywords: Fossil fuel, greenhouse gases (GHG), loadability, medium voltage (MV), Photovoltaic (PV)

INTRODUCTION

Loadability is the capability of an electrical grid to run the connected load. Effect of integration of same size photovoltaic

generation plant into different buses of electrical medium voltage distribution network may vary in terms of loadability. A feasibility study using photovoltaic (PV), as an alternate source of electricity generation in Bangladesh, was done by analyzing the stability of electrical power systems with the penetration of photovoltaic based generation. System-loading margin was studied without and with PV based generator. The contribution of PV based generator on solving under voltage problem and improving bus voltage was studied. Solution of overloading problem of power transformers with solar PV generator was described (Khan et al., 2013). A study has been done about optimal placement and sizing method to improve the voltage stability margin in a distribution system using distributed generation (DG). In this study it had been observed that operating condition of Distributed generation (DG) affected the sizing and placing when the distributed generation units operated at the unity power factor. If it is necessary, they need to put in the most impressionable voltage buses to ameliorate the stability margin of voltage (Al Abri et al., 2013). A study was done for site selection of plug in hybrid vehicle (PHEV) charging station based on voltage sensitivity of commercial distribution system. The analysis was done in a 16-bus system in power factory simulation software. Voltage sensitivity factor was used to find an optimal place for the PHEV charging station (Rahman et al., 2013).

Voltage stability is an important part of power system stability. Voltage instability is occurred due to lack of reactive power supplied to the electrical grid. Definition of voltage stability is given by IEEE power system engineering committee in the following way (Begovic et al., 1995). Voltage stability is the ability of a system to maintain voltage so that when load admittance increased, load power will increase, and voltages are controllable.

The definition of power system stability given by IEEE\CIGER task force is “Power system stability is an ability of an electric power system, for a given initial operating condition, to regain a state of operating equilibrium after being subjected to a physical disturbance, with most system variables bounded so that practically the entire system remains intact” (Ingelsson et al., 1997).

The impact of Distributed Generation (DG) technology and penetration level on the dynamics of a test system has been presented in a research. From this research it has been found that effects of distributed generations on the dynamics of a power system depend on the technology of the Distributed Generators (Stootweg et al, 2002). A practical investigation of the impacts of Distributed Generation (DG) units on system stability has been studied on a Brazilian real network. It was found that the distribution generation enhanced the overall performance (Londero et al., 2009). A research showed an assessment of the impact of the Distribution Generation unit size and location under a change in the loading conditions due to a contingency on unbalanced distribution systems (Kotamarty et al., 2008). A study has analyzed the potential impacts of Distributed Generation (DG) on the stability on the electrical networks. The investigation was carried out at a constant load

demands but with different contributions from fuel cells and micro-turbines. As a result, the rated and supplied powers of the traditional synchronous generators were adjusted to gain the power balance in the network (Azmy & Erlich, 2005)

The static voltage stability impact of solar photovoltaic generation on power networks using PowerWorld simulator power-voltage (P-V) and voltage-reactive (V-Q) curves to investigate the renewable energy generator model performance suitability has been examined. From the simulation result, it has been recommended that in a steady-state analysis of the grid power system the effects of power factor (pf) and voltage drop control should be considered by power grid engineers (Muhammed & Rawa, 2020).

Basic phenomenon behind voltage stability:

- Voltage level decreases with loading
- No increase of voltage with addition of a new generation unit if the system reaches to its maximum transfer limit
- High reactive loading causes reduction of voltage in that area
- Load quality has an effect on voltage change
- Transfer capacity is reduced due to reactive loading

It is important to study the V-P (voltage vs power) curve for analyzing the voltage level stability of a power distribution system. Greater the output power from photovoltaic or solar thermal greater the load ability of a specific grid system. A sample P-V curve has been showed in the Figure 1.

Load quality has an effect on loadability. The part below the critical voltage in the curve cannot be found from practical system operation. This part is mathematical expression. In this condition collapse of voltage occurs in a grid. In case of higher load demand, control

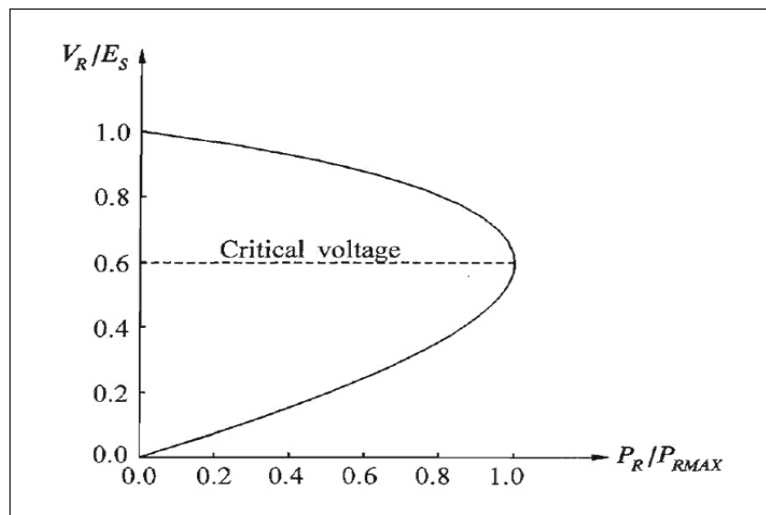


Figure 1. P-V curve (Al Abri et al., 2003)

of power by varying loads would be unstable. The system shows to be stabilized at voltage and power level minor than the intended value when there is a certain impedance static load characteristics. On the contrary, the system exhibits instability through collapse of the load voltage when there is a certain power load characteristics. With other characteristics, the voltage is determined by the composite characteristic of the transmission line and load. If the load is supplied by transformers with automatic under load tap-changing, the tap-changer action will try to raise the load voltage (Kundur et al., 1994).

Continuation power flow (CPF) is a method to find out some voltage stability conditions of a grid system. It is also used to find out the load margin. At several load quality and load level, the effect in voltage level is observed. Thus, an area is found from voltage vs. load curve where the voltage of a system is collapsed (Ajjarapu & Christy, 1992).

METHODS

The following procedure had been conducted for research purpose:

- Examining the impact of solar PV integration on IEEE 14-bus system.
- Analyzing the variation of loadability (λ) for placement and fraction of PV plant in IEEE 14 bus system condition
- Finding the maximum limit of PV integration for each bus for IEEE 14-bus system
- Examining the interaction between Photovoltaic and induction motor in an electrical grid.

For the analysis purpose, Power System Analysis Toolbox (PSAT) simulation software was used. From PSAT a test system was chosen named d_014.mdl bus system which was derived from IEEE14-bus system. Figure 2 is the 14 bus PSAT generalized test model.

Generalized IEEE 14-bus system consists of 14-buses including the slack bus. The rating of these buses varies from 69 KV to 13.8KV. Bus 1 to bus 5 is garish in color because this color signifies high voltage and power rating of 69 KV and 100MVA where rest of the buses (except bus-8) is blue in color signifying comparatively medium voltage of 13.8 KV. Bus number-8 is unique because it is the only bus of rating 18KV the color these buses is typically green. There are other components like transmission lines and step-up or step-down transformers connecting and establishing a network of fourteen buses. Apart from these, there are two static synchronous compensators, one is in bus-8 another is in bus-6.

Static synchronous compensator: Static synchronous compensator (STATCOM) also called Static synchronous condenser is a regulating device used on alternating current electricity transmission networks. It is based on a power electronics voltage-source converter and can act as either a source or sink of reactive AC power to an electricity network. If connected to a source of power, it can also provide active AC power. Usually, a STATCOM is installed to support electricity networks that have a poor power factor and often poor voltage regulation. There are, however, other uses, the most common use is for voltage stability.

Slack bus: In electrical power systems a slack bus (or swing bus), defined as a (Vδ) bus, is used to balance the active power |P| and reactive power |Q| in a system while performing load flow studies. The slack bus is used to provide for system losses by emitting or absorbing active and/or reactive power to and from the system.

$$\text{Power for slack bus} = \text{Total Power going into the system} - \text{Total Power going out of the System} + \text{Transmission line losses}$$

The slack bus is the only bus for which the system reference phase angle is defined. From this, the various angular differences can be calculated in the power flow equations. If a slack bus is not specified, then a generator bus with maximum real power |P| acts as the slack bus. A given scheme can involve more than one slack bus.

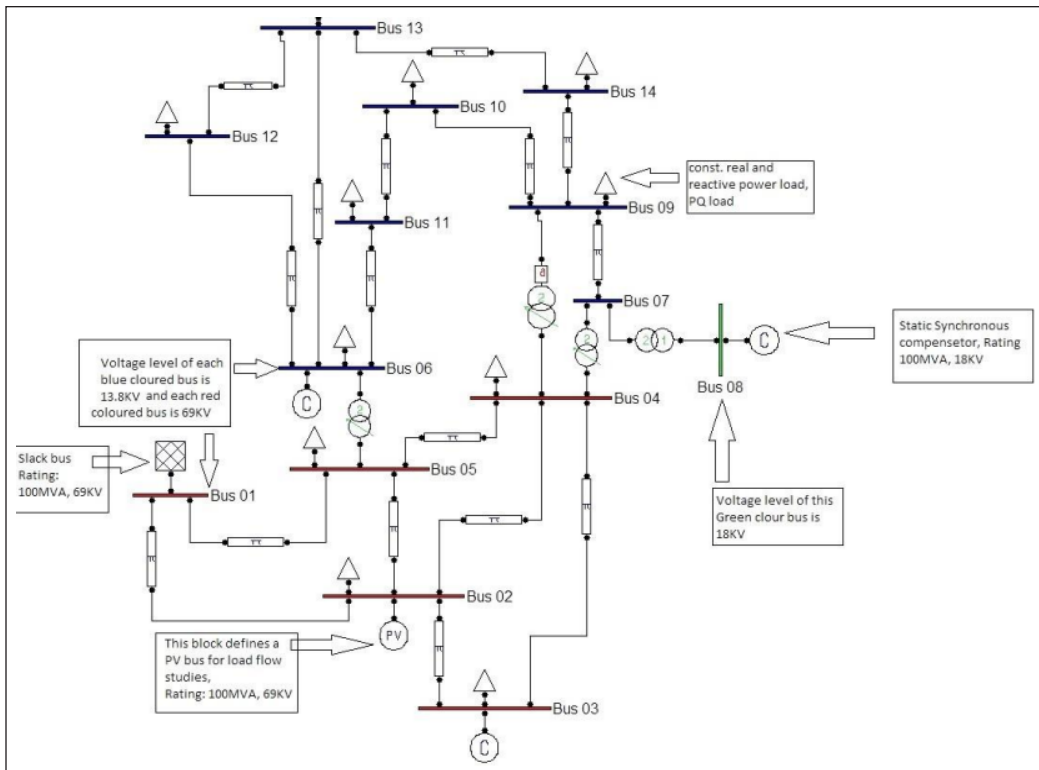


Figure 2. IEEE 14-bus PSAT model

RESULTS AND DISCUSSION

Load ability of IEEE 14-bus PSAT model varies with placement of PV plant at different buses. Load ability also varies with the size of a single generation plant and total number of generation plant after integrating different buses.

Value of loadability is represented by value of lamda (λ) in the simulation. The load ability of the default system which is taken from PSAT is $\lambda = 2.8286$.

Table 1 shows the results of loadability due to placement of solar PQ generator for every individual bus for different cases have been presented. In those cases, solar PQ generator of rating 30MW of active power and 0 MVAR of reactive power were placed at one bus for different time from bus number 2 - 14. Bus number-1 is slack bus, so no generator is added here.

In Figure 3, 14 bus default is the system where no extra generation unit was used. It was taken directly from PSAT without any modification. Table 1 shows that load ability increases (compared to 14 bus default) after integration of solar PQ generator, except at bus number 2, 6, 3, and 8 where load ability decreases. It is due to constant PV generator at bus 2 and synchronous condenser at bus number 3, 6 and 8. The increase rate at other buses is not same although the same size generator is placed.

For fractioning effect, size of single generation plant and number of total generation plant were changed maintaining same total generation capacity. In some cases, multiple

Table 1
Variation of loadability due to placement of the same sized generation unit for various buses

Generation unit installed bus number Name	Loadability (λ)
PQ generator at bus 14	2.9755
PQ generator at bus 10	2.9484
PQ generator at bus 9	2.9453
PQ generator at bus 13	2.9379
PQ generator at bus 11	2.9346
PQ generator at bus 12	2.9291
PQ generator at bus 7	2.9241
PQ generator at bus 4	2.8928
PQ generator at bus 5	2.8801
14 bus default	2.8286
PQ generator at bus 8	2.5175
PQ generator at bus 3	2.2129
PQ generator at bus 2	2.0344
PQ generator at bus 6	1.8792

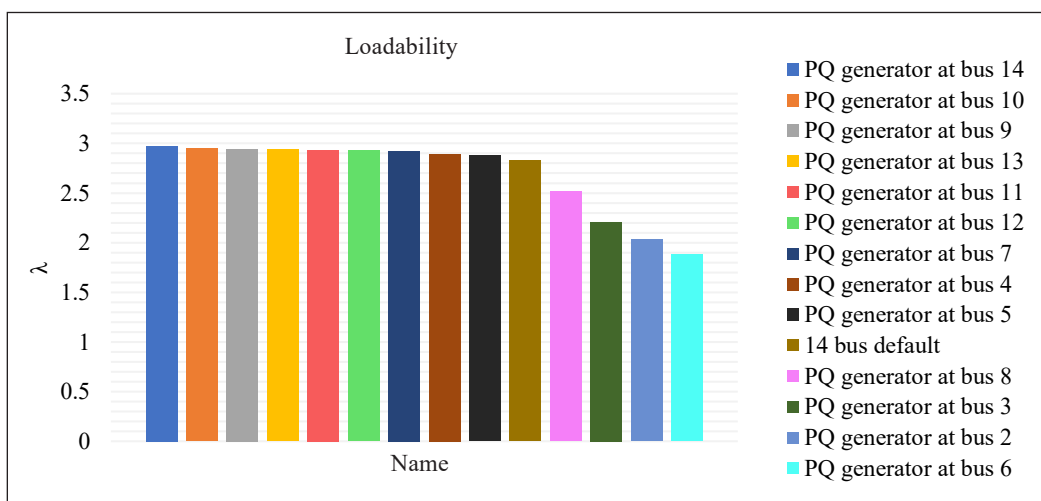


Figure 3. Variation of Load ability due to placement of PV generation unit for different buses

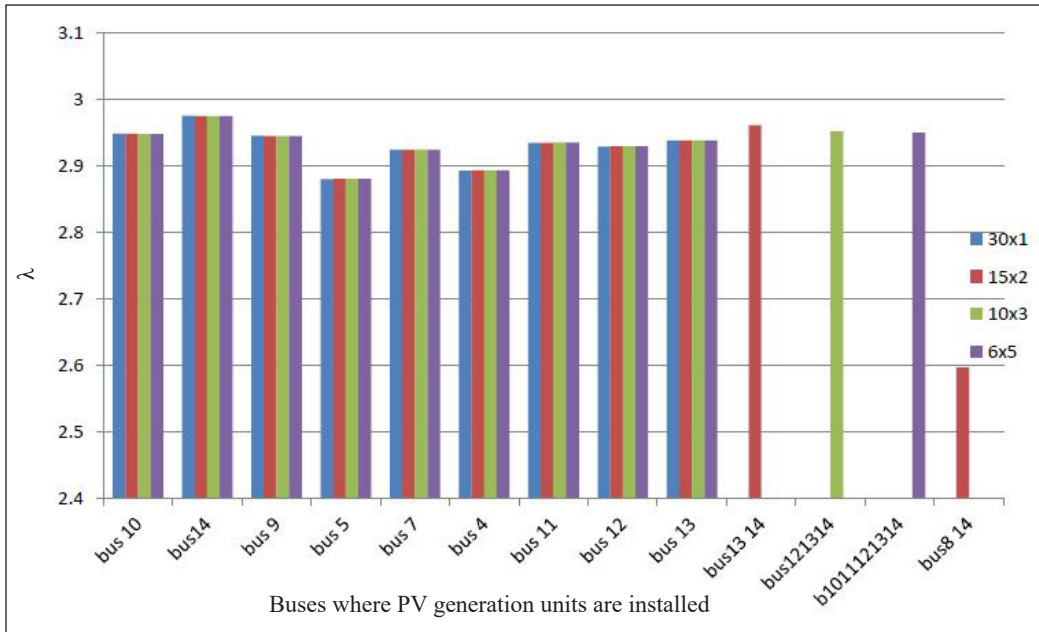


Figure 4. Variation due to fractioning and placement of generation unit at different buses

generation plants were integrated at same buses and in other cases plants were distributed at different buses. Effect of fractioning and placement of generation unit at different buses are shown in Figure 4.

As can be observed, the fractioning effect on load ability by chronological order was B14 (bus 14) = B10 = B09 > B12 = B11 > B13 = B07 > B04 = B05 (Figure 4). Generally, for integration of a new PV generation plant, the buses which exhibit increasing load are also sensitive to the fractioning generation unit. The three buses (bus 14, bus 10 and bus 9), where loadability was higher as compared to others, if we did fractioning in generation on those three buses, loadability decreased. Where on other buses, loadability increased if generation unit was fractioned. In case of distribution at different buses load ability decreased with increase of plant number.

The three buses (bus 14, bus 10 and bus 9), where loadability was higher as compared to others, if we did fractioning in generation on those three buses, loadability decreased. Where on other buses, loadability increased if generation unit was fractioned. In case of distribution at different buses, load ability decreased with increase of plant number.

In this case solar PQ generation unit was integrated at different buses of IEEE 14-bus system in PSAT. Ratings of the solar generation unit were increased until system became unstable. In each simulation power factor of PQ generation unit was considered 0.94.

In the Table 2, maximum input power was tested for each bus for integration PQ generation unit. For the bus number 5, active power limit was found maximum (i.e. 1198.5 MW)

Table 2
Maximum limit for input power of integrated PV for each bus

Bus number loadability	Max. Active Power (MW)	Max. Reactive Power (MVAR)
Bus 5	1198.5	76.5
Bus 4	1132.7	72.3
Bus 7	526.4	33.6
Bus 9	437.1	27.9
Bus 10	352.5	22.5
Bus 11	329	21
Bus 12	291.4	18.6
Bus 13	258.5	17.5
Bus 14	253.8	16.2

Figure 5 shows that Maximum limit is higher for bus which is closer to the slack bus. As the distance rose from slack bus the limit decreased. Significant change occurred in maximum limit from bus 4 to bus 7 due to change in bus voltage level from 69 KV to 13.8 KV. The voltage level of slack bus and bus number 2, 3, 4 and 5 was 69 KV where the voltage level of the rest of the buses was 13.8 KV except bus 8 where the voltage level was 18 KV.

To examine the interaction between photovoltaic plant and induction motor, solar PQ generation plant was integrated along with induction motor at different buses. The effect on load ability due to integration of any motor alongside PV generation plant was noted and is shown in the Table 3 and Figure 6. In this case induction motor rating was 3MVA, 11KV and 50Hz. Solar PV generation unit was also placed, and the rating was increased gradually, and the effect was monitored.

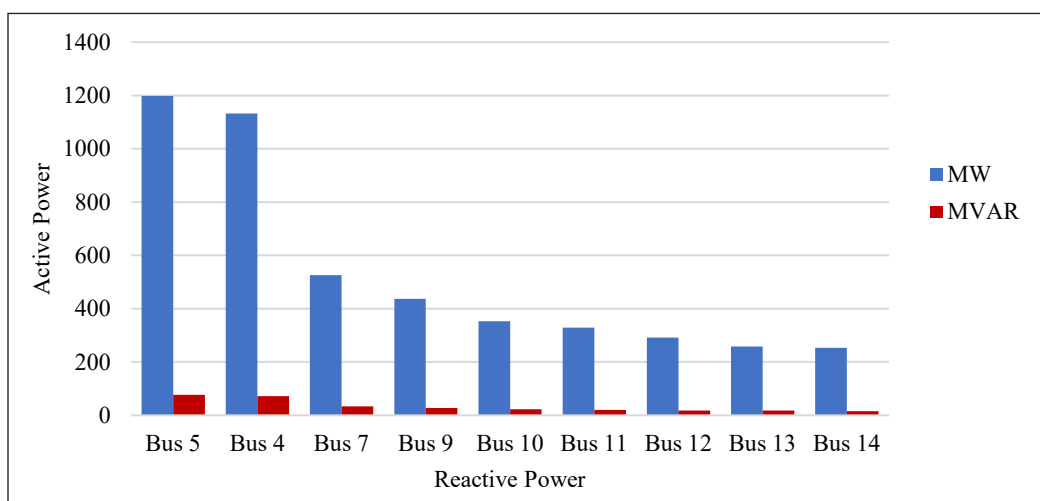


Figure 5. Maximum real power and reactive power limit

Table 3
Effect on loadability due to integration of induction motor and PV generation plant

Name	Loadability
14-bus default file	2.8286
14-bus default file	2.8286
Induction_Motor_Without_PQ Generation	1.8676
Induction_Motor_PQ1MWatB14	1.8695
Induction_Motor_PQ2MWatB14	1.8918
Induction_Motor_PQ3MWatB14	1.8938
Induction_Motor_PQ4MWatB14	1.875
Induction_Motor_PQ5MWatB14	1.8768
Induction_Motor_PQ6MWatB14	1.8996

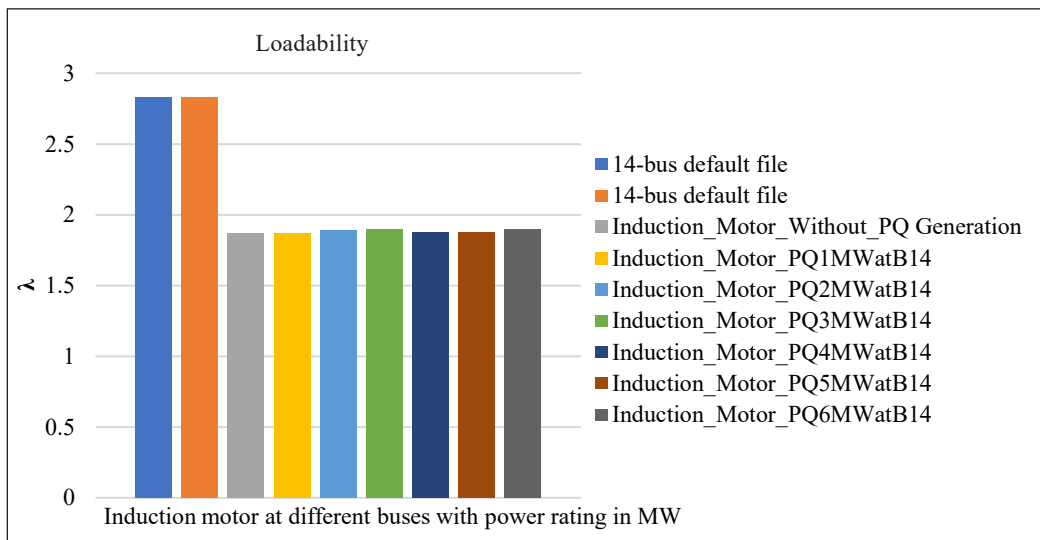


Figure 6. Effect on load ability due to integration of induction motor and PV generation plant

Effect of fractioning of induction motor and interaction with solar PQ generation plant: Table 4 and Figure 7 describe six induction motors of 0.5MVA were integrated with the bus. Rating of PQ generation unit was increased gradually from 1MW to 6MW.

Load ability was increased if the PQ generator was fractioned. Stability was better if the induction motor was fractioned. Loadability increased linearly with increase of rating of solar PQ generator in the case of fractioned motor. On the other side, load ability increased exponentially up to 3MW of solar PQ generator integration.

Modeling of solar photovoltaic integrated medium voltage distribution system was done using PSAT simulation tools. To find out the placement and fractioning effect of solar plants the placements of solar photovoltaic plant at different buses had been done.

Table 4
 Comparison between single induction motor and multiple induction motor connected to bus 14

Rating of PQ gen	Lodability for single induction motor	Lodability for multiple induction motor
without pq gen	1.8676	1.9254
PQ1MW	1.8695	1.9355
PQ2MW	1.8918	1.9396
PQ3MW	1.8938	1.944
PQ4MW	1.875	1.9465
PQ5MW	1.8768	1.9514
PQ6MW	1.8996	1.9562

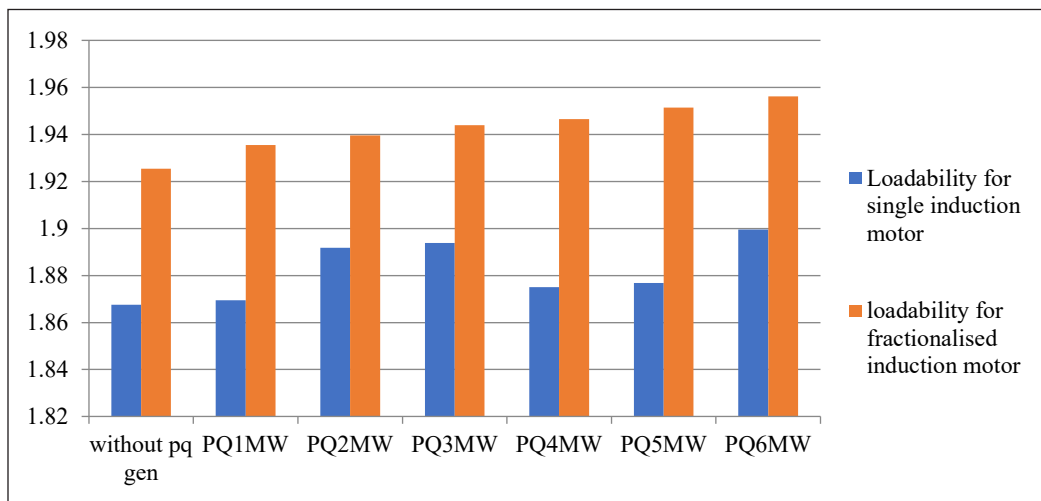


Figure 7. Comparison between single induction motor and multiple induction motor connected system.

CONCLUSION

It was found that loadability varied with fractioning and placement of solar plant at different buses. Maximum input rating of solar plant for an individual bus was not same for all buses. Loadability increased if the PQ generator was fractioned at bus 14 and stability was better if the induction motor was fractioned. Load ability increased linearly with increase of rating of solar PQ generator in the case of fractioned motor. On the other side, load ability increased exponentially up to 3MW of solar PQ generator integration. Due to research facility in laboratory the work had not been done exclusively.

The research work needs to be developed in future. Effects of placement & fraction of solar PV generation & induction motor in loadability can be analyzed using other simulation software for other test systems.

ACKNOWLEDGEMENT

This research work was submitted in partial fulfilment of the requirement for the degree of MS in Renewable Energy Technology at Institute of Energy, University of Dhaka. This paper has been put into the digital library (<https://repository.library.ac.bd>).

REFERENCES

- Ajjarapu, V., & Christy, C. (1992). The continuation power flow: A tool for steady state voltage stability analysis. *IEEE transactions on Power Systems*, 7(1), 416-423. <https://doi.org/10.1109/59.141737>
- Al Abri, R. S., El-Saadany, E. F., & Atwa, Y. M. (2013). Optimal placement and sizing method to improve the voltage stability margin in a distribution system using distributed generation. *IEEE transactions on Power Systems*, 28(1), 326-334. <https://doi.org/10.1109/TPWRS.2012.2200049>
- Azmy, A. M., & Erlich, I. (2005). Impact of distributed generation on the stability of electrical power system. In *IEEE Power Engineering Society General Meeting, 2005* (pp. 1056-1063). IEEE Conference Publication. <https://doi.org/10.1109/PES.2005.1489354>
- Begovic, M., Fulton, D., Gonzalez, M. R., Goossens, J., Guro, E. A., Haas, R. W., & Postforoosh, J. (1995). Summary of system protection and voltage stability. *IEEE Transactions on Power Delivery*, 10(2), 631-638. <https://doi.org/10.1109/61.400868>
- Ingelsson, B., Lindstrom, P. O., Karlsson, D., Runvik, G., & Sjodin, J. O. (1997). Wide-area protection against voltage collapse. *IEEE Computer Applications in Power*, 10(4), 30-35. <https://doi.org/10.1109/67.625371>
- Khan, M., Arifin, M., Haque, A., & Al-Masood, N. (2013). Stability analysis of power system with the penetration of photovoltaic based generation. *International Journal of Energy and Power Engineering*, 2(2), 84-89. <https://doi.org/10.11648/j.ijepe.20130202.18>
- Kotamarty, S., Khushalani, S., & Schulz, N. (2008). Impact of distributed generation on distribution contingency analysis. *Electric Power Systems Research*, 78(9), 1537-1545. <https://doi.org/10.1016/j.epr.2008.01.020>
- Kundur, P., Balu, N. J., & Lauby, M. G. (1994). *Power system stability and control*. McGraw-Hill Inc.
- Londero, R. R., Affonso, C. M., & Nunes, M. V. A. (2009). Impact of distributed generation in steady state, voltage and transient stability-Real case. In *2009 IEEE Bucharest PowerTech* (pp. 1-6). IEEE Conference Publication. <https://doi.org/10.1109/PTC.2009.5282016>
- Muhammed, A. O., & Rawa, M. (2020). A systematic PVQV-Curves approach for investigating the impact of solar photovoltaic-generator in power system using powerworld simulator. *Energies*, 13(10), Article 2662. <https://doi.org/10.3390/en13102662>
- Rahman, M. M., Barua, S., Zohora, S. T., Hasan, K., & Aziz, T. (2013). Voltage sensitivity based site selection for PHEV charging station in commercial distribution system. In *2013 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)* (pp. 1-6). IEEE Conference Publication. <https://doi.org/10.1109/APPEEC.2013.6837191>
- Slootweg, J. G., & Kling, W. L. (2002). Impacts of distributed generation on power system transient stability. *IEEE Power Engineering Society Summer Meeting*, 2, 862-867. <https://doi.org/10.1109/PESS.2002.1043465>



Volume Transport Variability in the Western Equatorial Pacific and its Relations to Halmahera Throughflow

Marlin Chrisye Wattimena^{1*}, Agus Saleh Atmadipoera¹, Mulia Purba¹,
I Wayan Nurjaya¹ and Fadli Syamsudin²

¹Department of Marine Science and Technology, Faculty of Fisheries and Marine Sciences, IPB University Bogor, 16680 Bogor, West Java, Indonesia

²The Agency for the Assessment and Application of Technology (BPPT), 10340 Jakarta, Indonesia

ABSTRACT

This study investigates the coherency of volume transport between Halmahera throughflow and current major system in the western equatorial Pacific Ocean (Mindanao Current – MC, New Guinea Coastal/Under Current – NGCC/NGCUC, and North Equatorial Counter Current – NECC). The validated daily ocean general circulation model datasets of INDES0 (2010-2014) were used in this study. The results showed that the estimated average transport volume was 25.6 Sv flowing southward through MC, 34.5 Sv flowing eastward through NECC, 18.3 Sv flowing northwestward through NGCC/NGCUC, and 2.5 Sv flowing southward through the Halmahera Sea. The variability of volume transport was dominated by intraseasonal, semiannual, and annual time-scales. The increased transport of NECC corresponded to the intensification of MC and NGCC/NGCUC transports. NGCC/NGCUC significantly controlled the South Pacific water inflow into the Halmahera Sea because of the positively high correlation between NGCC/NGCUC transport and Halmahera throughflow transport.

Keywords: Cross-PSD, Halmahera throughflow, INDES0 model, Western Equatorial Pacific

ARTICLE INFO

Article history:

Received: 12 November 2020

Accepted: 11 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.26>

E-mail addresses:

marlinchrisyewattimena@gmail.com (Marlin Chrisye Wattimena)

atmadipoera_itk@apps.ipb.ac.id (Agus Saleh Atmadipoera)

mpsiboro@yahoo.com (Mulia Purba)

wayan2011@gmail.com (I Wayan Nurjaya)

fadlihiro@yahoo.com (Fadli Syamsudin)

* Corresponding author

INTRODUCTION

The Pacific low-latitude western boundary current (LLWBC) plays a significant role in controlling Pacific circulation where the water mass from the northern and southern hemispheres accumulate (Lukas et al., 1996); hence, affecting the global climate due to mass and heat transport balances from sub-tropical area to tropical area and from

the Pacific Ocean to Indian Ocean (Hu et al., 2015). These oceanographic roles of LLWBC are of interest to regional oceanographic studies affected by LLWBC. The previous studies remarkably described LLWBC pathways such as Mindanao Current (MC), North Equatorial Counter Current (NECC), New Guinea Coastal Current (NGCC) and New Guinea Coastal Undercurrent (NGCUC), Mindanao Eddies (ME) and Halmahera Eddies (HE) (Kashino et al., 2011; Kashino et al., 2013).

Numerous studies have reported the importance of LLWBC in water transport in Indonesian waters via Indonesian throughflow (ITF). LLWBC has been found to accumulate in the western equatorial Pacific with complex water distribution (Figure 1). From the north, MC flows southward along the eastern Philippine coast (Lukas & Lindstrom, 1991; Gordon & Fine, 1996; Wijffels et al., 1995; Kashino et al., 2005) and contributes to Indonesian Throughflow (ITF) and NECC (Fine et al., 1994; Lukas et al., 1996; Kashino et al., 2001). NGCC/NGCUC flows northwest through the New Guinea coast from the southern hemisphere (Kuroda, 2000; Ueki et al., 2003; Kashino et al., 2007). These currents are jointed around eastern Halmahera islands, and some of these water masses entering Indonesian waters are known as ITF. At the same time, the rest flows eastward as NECC from the western Pacific to the east.

ITF is the only inter-ocean current present in low latitude. It flows from the Pacific Ocean to the Indian Ocean, being a part of the global thermohaline circulation cycle. The ITF water mass source is dominantly from North Pacific via MC and south Pacific via NGCC (Wyrтки, 1961; Gordon & Fine, 1996; Ilahude & Gordon, 1996; Gordon, 2005). Due to the location of the Halmahera Sea adjacent to the Western Pacific, the water mass of Halmahera Sea is likely to be affected by the LLWBC system despite the complexity of LLWBC current system close to the Halmahera Sea, which is challenging to study based on the model outputs, the relationship between the currents in the LLWBC system and inputs to the Halmahera Sea. Some previous studies have examined the structure and variability of MC, NGCC, NGCUC, NECC, HE, and ME, as part of LLWBC (Lukas & Lindstrom, 1991; Gouriou & Toole, 1993; Fine et al., 1994; Wijffels et al., 1995; Arruda & Nof, 2003; Atmadipoera et al., 2004; Kashino et al., 2007, 2013; Harsono et al., 2014, Wattimena et al., 2018). Many have been studied from the results of observation and modeling. Estimated volume transport from each LLWBC system branch has also been examined previously from the effects of observations and modeling. Estimates of Mindanao Current transport have ranged from 8 to 40 Sv (Lukas, 1988; Qu et al., 1998, Kashino et al., 2009). Most earlier studies focused only on one or two branches of the LLWBC system.

Despite studies in the Halmahera Sea related to LLWBC, the connection between LLWBC and the Halmahera Sea, such as the variability of Halmahera throughflow associated with LLWBC, is still unknown. The relationship of NGCC with input to the Halmahera Sea in the upper 200 m depth was previously reported by Wattimena et al. (2018).

Still, it did not account for the influence of other LLWBC current systems. The present study will estimate volume transport and describe the variability and connection between Halmahera throughflow and Pacific low latitude western boundary current (MC, NECC, and NGCC/NGCUC). The study area comprised of a transect in the eastern Philippine at 7°N (126°E – 130°E) representing MC, a transect at 133°E (2.0°N – 7.0°N) representing NECC, a transect in North New Guinea at 133°E (0.5°S – 2.0°N) representing NGCC/NGCUC, and a transect in the Halmahera Sea at 0.3°S (127.8°E – 130.4°E) (Figure 1).

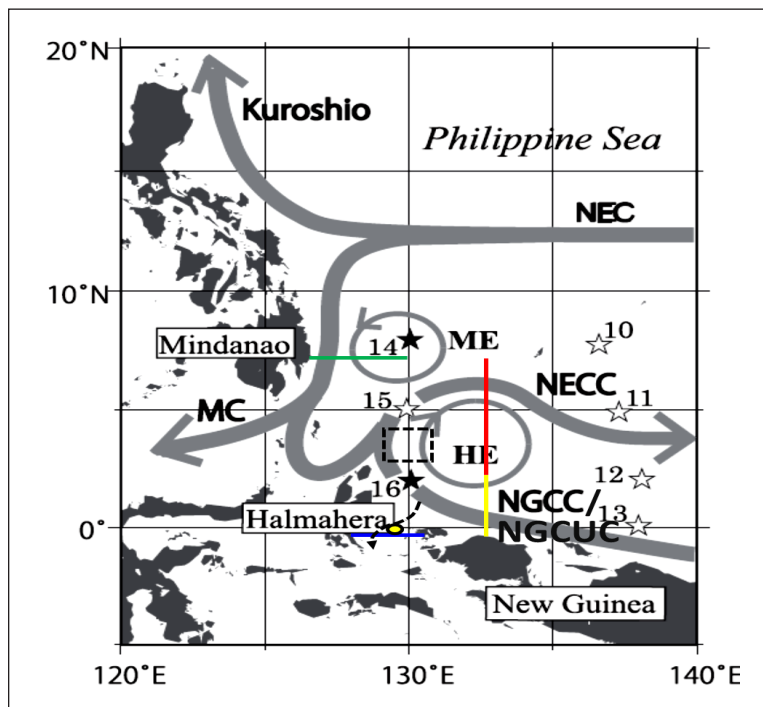


Figure 1. Main current circulation of LLWBC (adapted from Kashino et al., 2013). Blue line: representing transect of Halmahera throughflow. Yellow line: representing transect of NGCC/NGCUC. Red line: representing transect of NECC. Green line: representing transect of MC. Dashed line representing inflow to Halmahera Sea. Dash black box and yellow circle representing satellite and mooring INDOMIX location. Numbers and stars represent TRITON buoys array.

MATERIALS AND METHODS

INDES0 Model Output

The numerical outputs in this study used the simulations of Ocean General Circulation Model (OGCM) with Infrastructure Development for Space Oceanography (INDES0) configuration within the domain model in Indonesian Seas (90°E-145°E; 25°S-20°N), such as detailed bathymetry, direct tidal forcing and mixing parameterization. Mercator Ocean developed INDES0, a scientific program in cooperation with *Collecte Localisation*

Satellites (CLS) in Toulouse, France, to produce data for the Ministry of Marine Affairs and Fisheries. INDESO used Mercator Ocean System and Interface Relocatable Nesting (SIREN) tools to create a new and more significant configuration by combining bathymetry, initial condition, and boundary condition (Theetten et al., 2014). INDESO model used in-situ and satellite measurements as its simulation inputs. Data for the INDESO model were collected during the INDOMIX expedition. Baharel and Mercator (2006) detailing some data used to build, assimilate, and validate the model.

The domain model of this study covers the Western Equatorial Pacific and Halmahera Sea (125°E - 144°E; 4°S - 15°N), with 3-dimension parameters including temperature, salinity, and current (zonal and meridional components) and sea surface height from January 2010 to December 2014. Temporal resolution was daily, with a horizontal resolution of 1/12° (about 9.25 km). INDESO applied 50 vertical layers with the high-resolution discretization density near the surface layer. In the first 10 m, the layer thickness is less than 2 m, then raised to about 10 to 50 m depth. Modeled temperature and salinity are overlaid with the current in the same grids to describe the central circulation around the Low Latitude Western Boundary Pacific and the Halmahera Sea (study area). The estimated transport volume (S_v) is calculated using zonal and meridional velocity data according to the orientation of the current components that are dominant in each transect. The formula for transport estimation will be explained in the data analysis section.

Datasets Used for Validating INDESO Model

Observations used in the INDESO model are described as follows. For validating ocean currents, the study used meridional ocean currents observed from an ADCP INDOMIX mooring point in the Halmahera Sea (yellow circle in Figure 1) within 200 m - 700 m from July 2010 to December 2012 (Atmadipoera et al., 2013). For the sea surface height anomaly (SSHA) dataset, AVISO altimetry products (spatial resolution of 1/4°) were used with dataset range between January 2008 and December 2014. The GHRSSST-PP satellite observed measurements (spatial resolution of 1/20°) were used to validate the modeled SST.

Data Analysis

The characteristics and variability of ocean currents in the study area were assessed using time-series analysis through the determination of the average current, continuous wavelet transform (CWT), and cross-correlation as previously performed by Emery and Thomson (2014) and Bendat and Piersol (2010). The averaged ocean currents were determined to investigate the general spatial circulation pattern in the area. The climatological method was carried out to analyze the annual cycle from volume transports of MC, NECC, NGCC, and Halmahera throughflow. The CWT analysis was used to assess the temporal variability of MC, NECC, NGCC, and Halmahera throughflow transport.

Transport volume estimation of MC, NECC, NGCC/NGCUC, and Halmahera throughflow was calculated by establishing transects that passed through the eastern Philippines at 7 °N (126 °E – 128.5 °E) representing MC, at 133 °E (2.0 °N – 7.0 °N) representing NECC, in North New Guinea at 133 °E (0.5 °S – 2.0 °N), and in the Halmahera Sea at 0.3 °S (127.8 °E - 130.4 °E) (Equation 1 - 4). The estimation in each area was calculated by integrating zonal velocity (u) to determine the estimated volume transport of NECC and NGCC/NGCUC (Qu). The meridional rate (v) was incorporated to determine the volume transports of MC and Halmahera throughflow (Qv), considering the transect length and depth, as described by Emery and Thomson (2014):

$$Qv_{MC} = \int_{126E}^{128.5E} \int_{700m}^0 v dz \quad [1]$$

$$Qu_{NECC} = \int_{2N}^{7N} \int_{1142m}^0 u dz \quad [2]$$

$$Qu_{NGCC/NGCUC} = \int_{0.5S}^{2N} \int_{832m}^0 u dz \quad [3]$$

$$Qv_{Halmahera} = \int_{127.8E}^{130.5E} \int_{700m}^0 v dz \quad [4]$$

Where Qu_{NECC} (unit in Sv; 1 Sv = 10^6 m³/s) is the total volume transport of NECC that intersects the NECC transect line; $Qu_{NGCC/NGCUC}$ is the total volume of NGCC/NGCUC that intersects the NGCC/NGCUC transect line; Qv_{MC} is the total volume of MC that intersects the MC transect line; $Qv_{HALMAHERA}$ is the total volume of Halmahera throughflow that intersects transect line in eastern Halmahera. Vertically, the lower limits of the integration calculation of transport were 700 m, 1142 m, 832 m and 700 m, respectively for MC, NECC, NGCC/NGCUC and Halmahera at which the zero-velocity contour was located.

Cross-correlation analysis of the anomaly of transport in each area was performed with the following combinations: MC – NECC, MC – NGCC/NGCUC, MC – Halmahera, NECC–Halmahera, NGCC/NGCUC – NECC, and NGCC/NGCUC – Halmahera. The analysis resulted in energy co-spectrum, coherence, and phase difference. The energy co-spectrum indicates the magnitude of shared energy fluctuation at a specific frequency within two time-series. Meanwhile, coherence represents the correlation and phase difference, indicating the difference in time interval within both variables (Bendat & Piersol, 2010).

Model Validations

INDES0 model has significantly agreed with observations. For example, INDES0-TAO TRITON comparison, correlation coefficients and RMSE for zonal (meridional) velocities are 0.74 (0.7) and 0.25 m/s (0.21 m/s), respectively (Figure 2a). Similarly, the correlation coefficients and RMSE for SST (SSHA) are 0.85 (0.8) and 0.6 °C (0.03 m) between INDES0 model and satellite observations (Figure 2c and 2d). Besides, the depth-time distributions of meridional velocities of INDES0 are also qualitatively agreeable with a

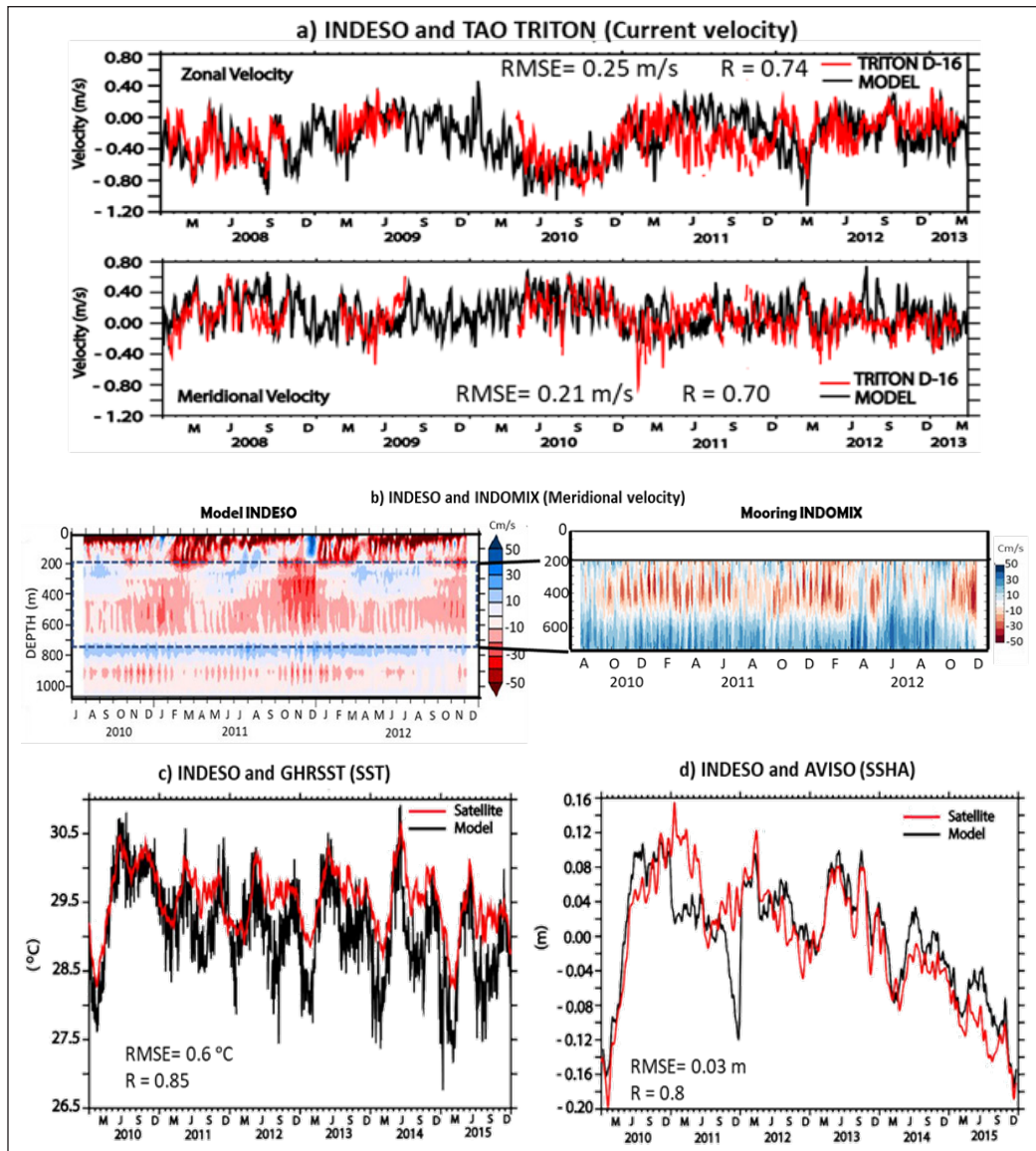


Figure 2. INDES0 model output compared with (a) TAO TRITON; (b) INDOMIX; (c) SST and (d) SSHA.

single ADCP INDOMIX (Figure 2b). It is noted that the single moored 75-kHz ADCP was placed unexpectedly much deeper on the seabed at about 760 m depth, thus the current data in the upper 200 m depth cannot be measured by the instrument. The validations of INDES0 model with ADCP show that there are southward flows within 200-400 m with the core flow at 400 m. This southward flow is due to the undercurrent of NGCUC, moving underneath NGCC along the back of Papuan coasts and ultimately reaching Halmahera. The validations of temporal changes between INDES0 model and ADCP data show that the model replicate southward flows within 200-400 m with the core flow at 400 m as shown by ADCP data with strong intraseasonal variability evident at sub-thermocline layer. The general model validation of the INDES0 model output has also been described in Tranchant et al. (2015).

RESULTS

General Circulation of LLWBC

Figure 3 shows the general circulations of LLWBC coupled with temperature and salinity at certain depths (25 m, 100 m, 200 m, and 300 m) and was long time averaged of all datasets. The LLWBCs area is a confluence region of current branches originating from the North Pacific and South Pacific. The overlays of averaged current vectors, salinity (left panel of Figure 3), and temperature (right panel of Figure 3) at depths of 25 m (a), 100 m (b), 200 m (c), and 300 m (d) from January 2010 to December 2014 showed two main streams of the western Equator Pacific current system: MC and NGCC / NGCUC. MC flowed southward along the east coast of Mindanao to the Pacific equator. MC's existence as the source input of ITF through the Sulawesi Sea and then forwarded to the Makassar Strait (and through the Maluku Sea) was visible in the flow vector model results from depths. MC transports a water mass characterized by low salinity from the subtropical region ($S = 34.6 - 34.9$) with temperature ranging from 20°C to 24°C . Below the thermocline (300 m) in MC, water mass was characterized by salinity and temperature of $34.3 - 34.5$ and $9^{\circ}\text{C} - 12^{\circ}\text{C}$, respectively. NGCC / NGCUC is the incoming inflow from the South Pacific Ocean. NGCC/NGCUC is part of the SEC that moves towards the Pacific region to the northwest along the back of Papua - New Guinea. NGCC carries high-salinity water ($35.0 - 35.5$) at temperatures of $18^{\circ}\text{C} - 26^{\circ}\text{C}$. In addition to the MC and NGCC/NGCUC that make up LLWBC, there are also NECC current branches that appear to result from the confluence region of two main current branches. NECC flows from the western edge of the Pacific to the eastern Pacific. Surface circulation (25 m) showed the presence of HE and ME. At a depth of 100 m, 200 m, and 300 m, in addition to HE and ME, New Guinea Eddies (NGE) was also seen (Kashino et al., 2007). However, at a depth of 300 m, the NGE began to weaken. The circulation pattern showed the South Pacific water entered through the Halmahera Sea and the Maluku Sea into the Seram Sea, then forwarded to the Banda Sea.

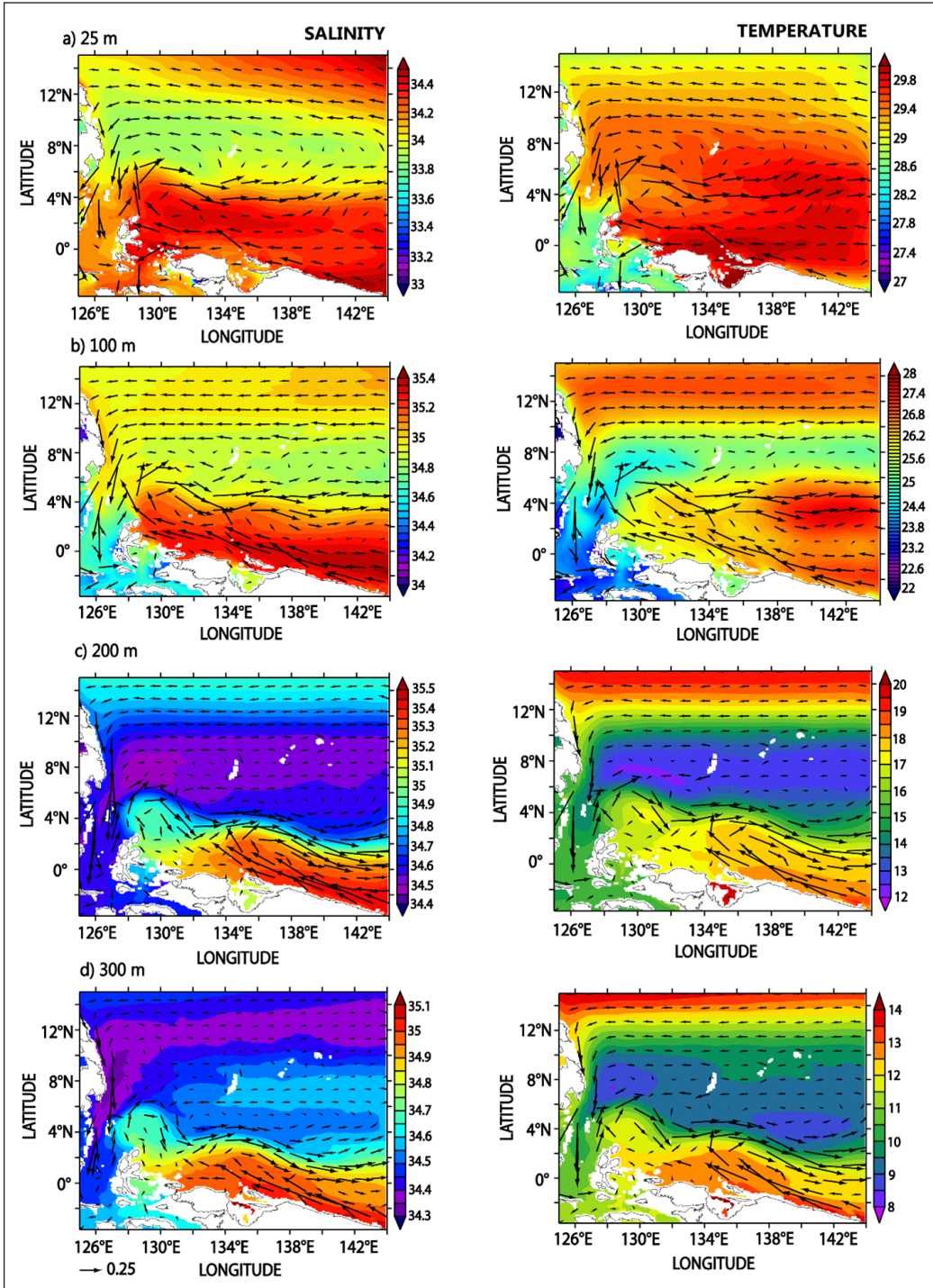


Figure 3. Mean ocean circulation pattern of Western Pacific, averaged from January 2010 to December 2014. The overlay result of average current vector with salinity (left) and temperature (right) at various depth levels: (a) 25 m; (b) 100 m; (c) 200 m; and (d) 300 m.

Seasonal Circulation of LLWBC

Similar to Figure 3, Figure 4a and 4b illustrate the mean seasonal circulations of LLWBC in winter (January) and summer (August) seasons, respectively. Strong Seasonal variation in LLWBC was evident from INDES0 model between winter and summer conditions. At the surface layer (25 m and 100 m), MC transports less salty, cooler water ($S=33.7 - 34.8$; $T= 28\text{ }^{\circ}\text{C} - 28.6\text{ }^{\circ}\text{C}$) in winter than in summer ($S=33.8 - 35$; $T= 28\text{ }^{\circ}\text{C} - 30\text{ }^{\circ}\text{C}$). MC is faster particularly at the flow along east coasts of Philippines in summer than in winter. The bifurcation of MC is located at 12°N during summer and winter. At the depths of 200 m and 300 m, the flow of MC weakens in the winter ($S=34.6-34.4$; $T= 14-16^{\circ}\text{C}$) compared to in the summer with salinity and temperature respectively $34.6 - 34.9$ and $10 - 20^{\circ}\text{C}$. NECC flow also shows a seasonal variability. In winter (Figure 4a), NECC flowing toward east Pacific at 4°N , 129°E . This flow is formed by the overturning of NGCC/NGUCC carrying salty water (34.5) that hence meets MC with low salinity. In contrast, in summer, the establishment of NECC is slightly northward (around 6°N , 128°E) with less salty, warmer water ($S = 34$; $T = 30\text{ }^{\circ}\text{C}$). NGCC/NGUCC varies seasonally. In summer (Figure 4b), at the top 100 m, NGCC flowing north-westward is intensified. In contrast, in winter, this flow weakens, causing this flow to turn to southeast. NGUCC (below 100 m) flowing to northwest is intensified in summer and its direction is also maintained in winter. NGE establishes at 4°N , 134°E with the larger flow magnitude found in summer than in winter. Halmahera throughflow is intensified in summer than in winter. The weak flow of this current system in winter causes some overflow in Pacific Ocean.

Vertical Structure of the Current

Characteristics of flows in each branch of LLWBC and Halmahera are shown from the mean and standard deviation presented in Figure 5. Dimensions of MC flow with an average meridional component at 7°N directing to the south appeared to widen from the eastern gate of Mindanao to 128°E , with a current thickness of 700 m. MC's core velocity reached $> 1\text{ m/s}$ from surface to 250 m depth (Figure 5a). The average zonal velocity along 133°E representing the NECC demonstrated a high current velocity to the east. The NECC flows between $2^{\circ}\text{N} - 6^{\circ}\text{N}$, reaching a depth of $> 1100\text{ m}$. The NECC core was found at 4°N , reaching a maximum velocity of $> 0.5\text{ m/s}$ a depth of 150 m (Figure 5b). The average zonal current component along 133°E showed that NGCC/NGCUC flow demonstrated a high speed, dominantly flowing to the west. Additionally, the core dimension of NGCC was observed to widen from North New Guinea to 2°N . It was found to be closed to surface up to 200 m depth. The maximum flow rate of NGCC was found at $> 0.5\text{ m/s}$, a depth of 100 m (Figure 5c). NGCUC flow was found below NGCC until up to 800 m depth. In terms of the current meridional component at 0.3°S , the flow east of the Halmahera Sea showed the current's high speed, dominantly entering into the south. This flow is an

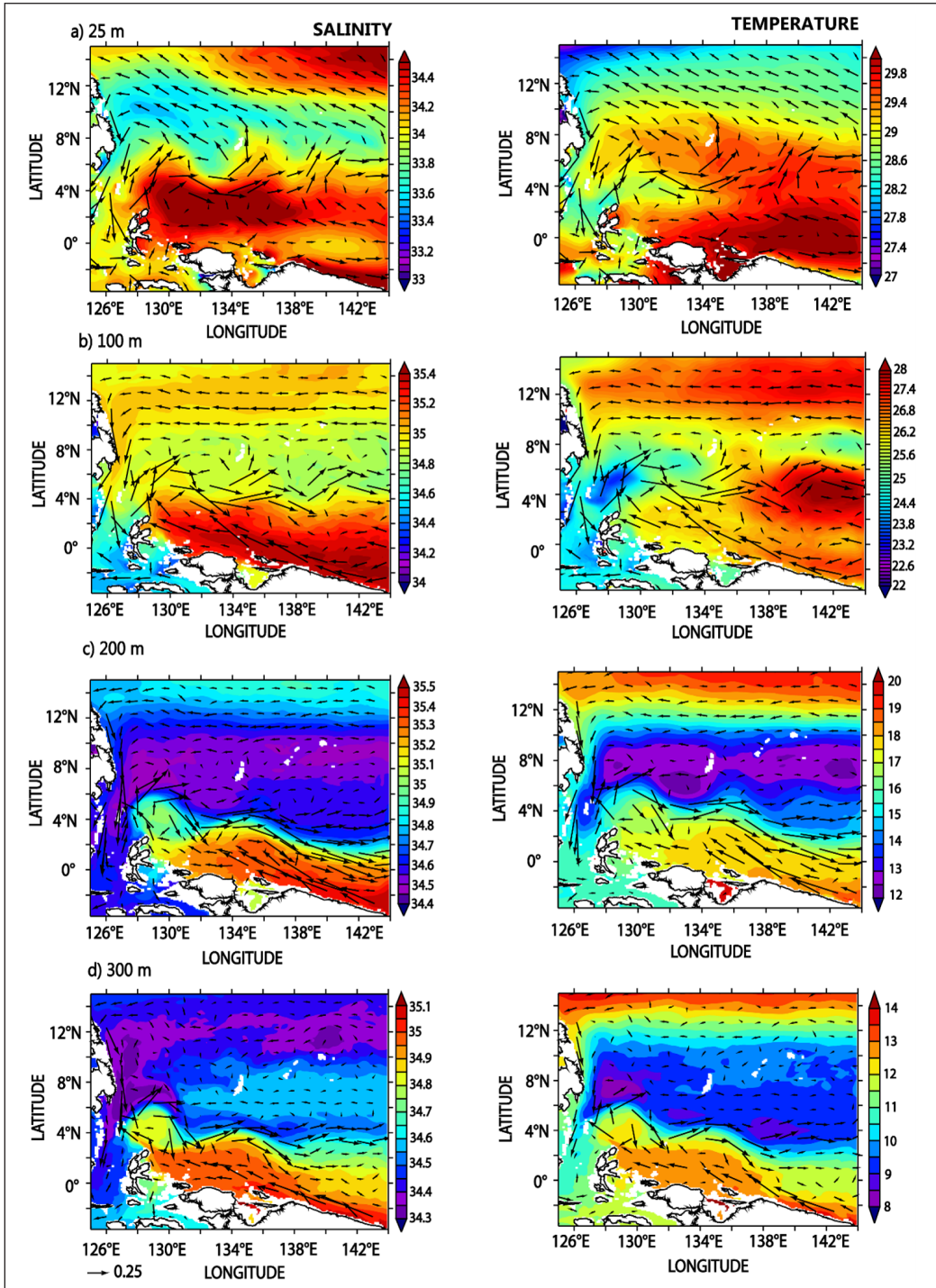


Figure 4a. Winter circulation pattern of Western Pacific. The overlay result of average current vector with salinity (left) and temperature (right) at various depth levels: (a) 25 m (a); (b) 100 m, (c) 200 m (c) and (d) 300 m. Note, the profiles were the average of December, January and February datasets.

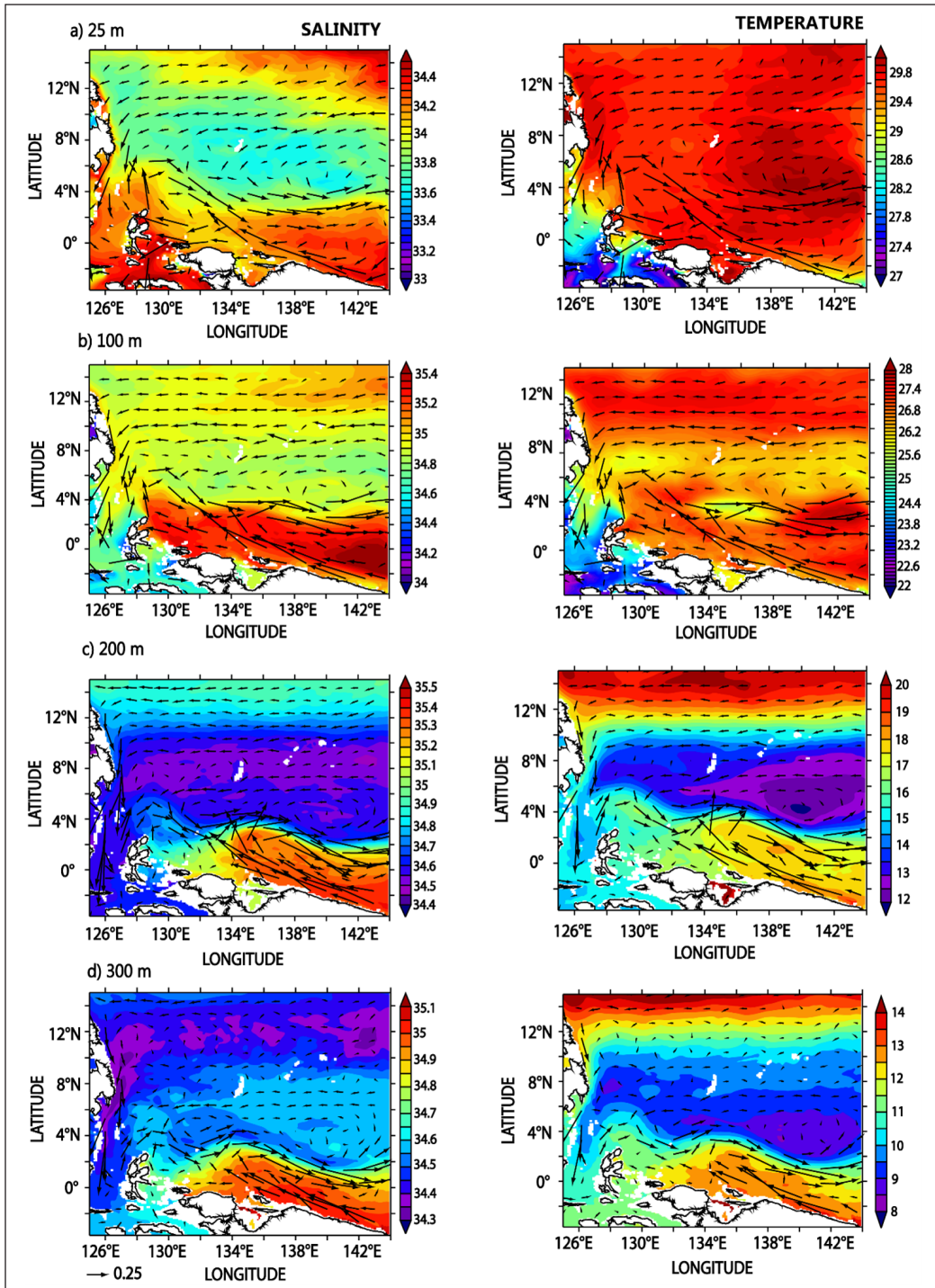


Figure 4b. Summer circulation pattern of Western Pacific. The overlay result of average current vector with salinity (left) and temperature (right) at various depth levels: (a) 25 m; (b) 100 m; (c) 200 m and (d) 300 m. Note, the profiles were the average of June, July and August datasets.

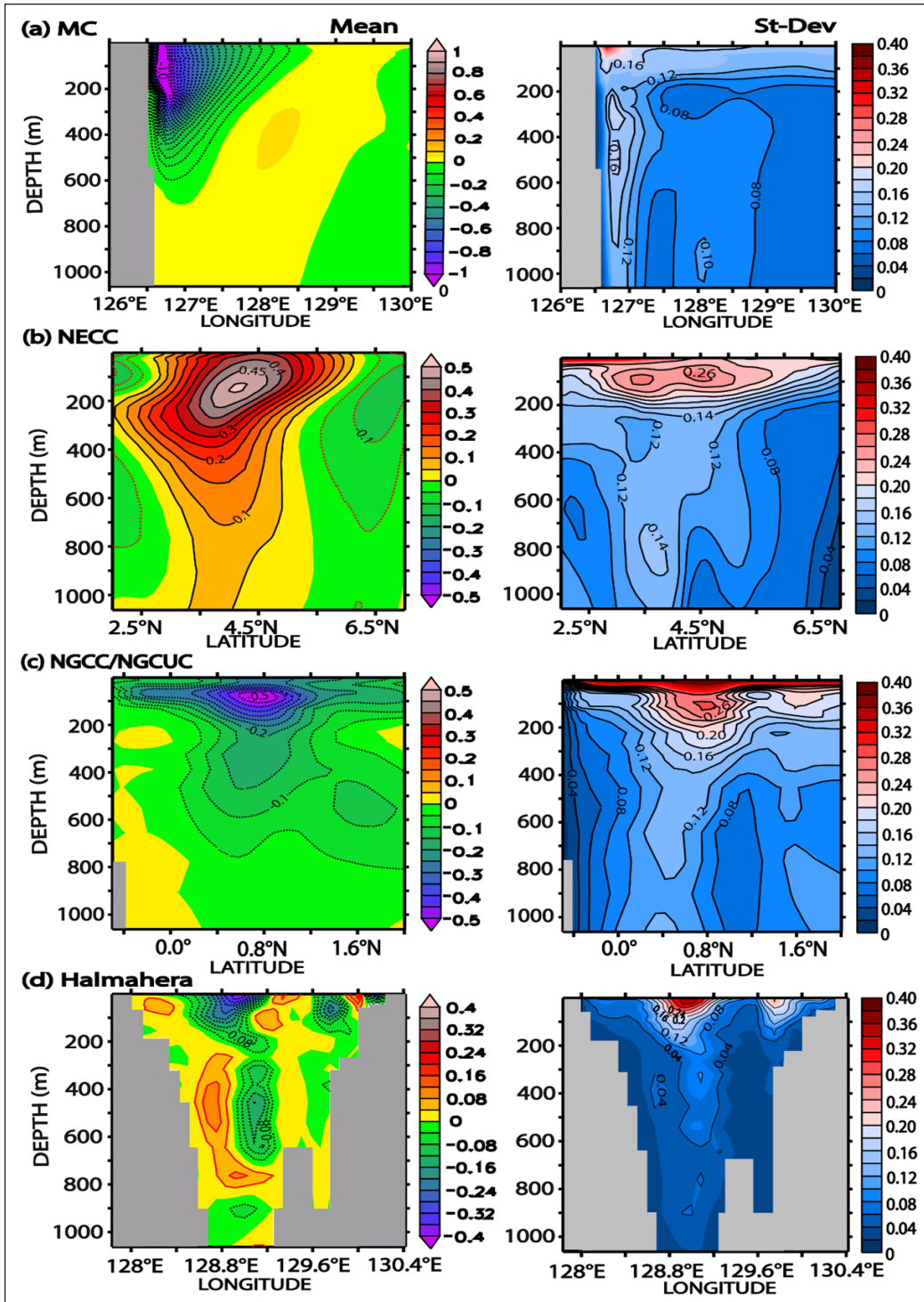


Figure 5. The cross section of mean (left) and standard deviation for (a) MC; (b) NECC, (c) NGCC/NGCUC, and (d) Halmahera throughflow.

ITF that passes through the Halmahera Sea, covering a width of 67.2 km (occurring up to a depth of 700 m), with the existence of two cores layer around 100 m and 400 m, the velocities at the sub-surface layer are more significant than a below-thermocline layer, with a maximum velocity of > 0.4 m/s from the surface to a depth of 100 m (Figure 5d). More detailed descriptions about the flow characteristics of MC, NECC, NGCC/NGCUC, and Halmahera throughflow will be further given in the discussion section.

Variability of Volume Transport

Based on time-series data of volume transport of MC, NECC, NGCC/NGCUC, and Halmahera throughflow from January 2010 to December 2014, the average volume transport of MC was at the range of 0-50 Sv (negative value indicates the direction to the south), with an average value of 25.6 ± 7.1 Sv. MC's annual cycle tended to be stronger in the south before the Southeast monsoon, weakening in the Northwest monsoon (Figure 6a). The average transport estimation of NECC to the east was 34.5 ± 17.5 Sv. Time-series data from 2010 to 2014 showed increased transport from mid-2013 until the end of 2014. The annual variation showed a maximum transport peak in November and a minimum transport peak in May and June (Figure 6b). Based on time-series data, the average transport of NGCC/NGCUC was -18.3 ± 8.3 Sv (negative value indicates westward flow). The annual variation achieved the maximum transport in November and January and reached the minimum transport in July (Figure 6c). The average transport of Halmahera was 2.5 ± 1.8 Sv to the south. The flow to the south could represent ITF that passed through the Halmahera Sea. Transport variability of Halmahera throughflow achieved the maximum level in October and the minimum level in January (Figure 6d). Temporal variability from time-series transport of MC, NECC, NGCC/NGCUC, and Halmahera throughflow was also found in CWT's analysis result (Figure 7). The CWT for MC, NECC, NGCC/NGCUC, and Halmahera throughflow transport during 2010-2014 showed transport fluctuation in intraseasonal, semiannual (except NECC) annual period. The detailed elaboration of temporal and spatial variability of MC, NECC, NGCC/NGCUC, and Halmahera throughflow transport is presented in the discussion.

Also, the study calculated the transport between Mindanao and Celebes island at the intersection of the Sangihe Ridge and between Sulawesi and the Halmahera in northern Maluku seas. The average volume transports (within the top 600 m) in transects across the Sangihe Ridge entrance into the Indonesian seas and across the north Maluku sea into Maluku Sea have values of $8.01 (\pm 2.5)$ Sv and $2.6 (\pm 2.1)$ Sv, respectively. The study will not discuss this aspect as this study's scope only focuses on Indonesian throughflow's eastern pathways. At the same time, Pacific inflow through Sangihe Ridge has contributed to ITF's western pathway downstream on Makassar Strait.

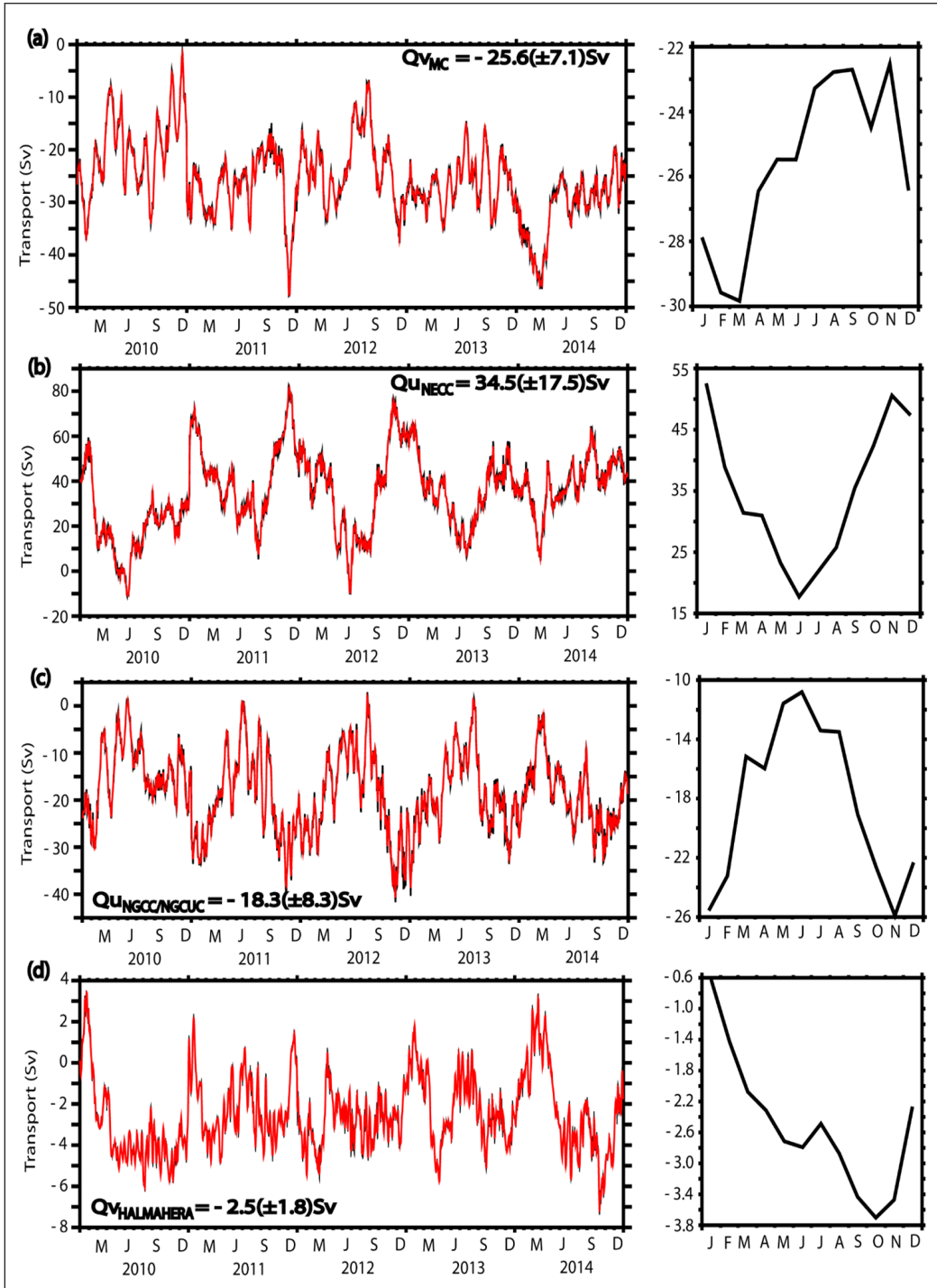


Figure 6. Time-series of volume transport (left) and annual variability (right) of (a) MC; (b) NECC; (c) NGCC/NGCUC; and (d) Halmahera throughflow. Black line indicates daily transport; red line indicates 20-days filtered transport.

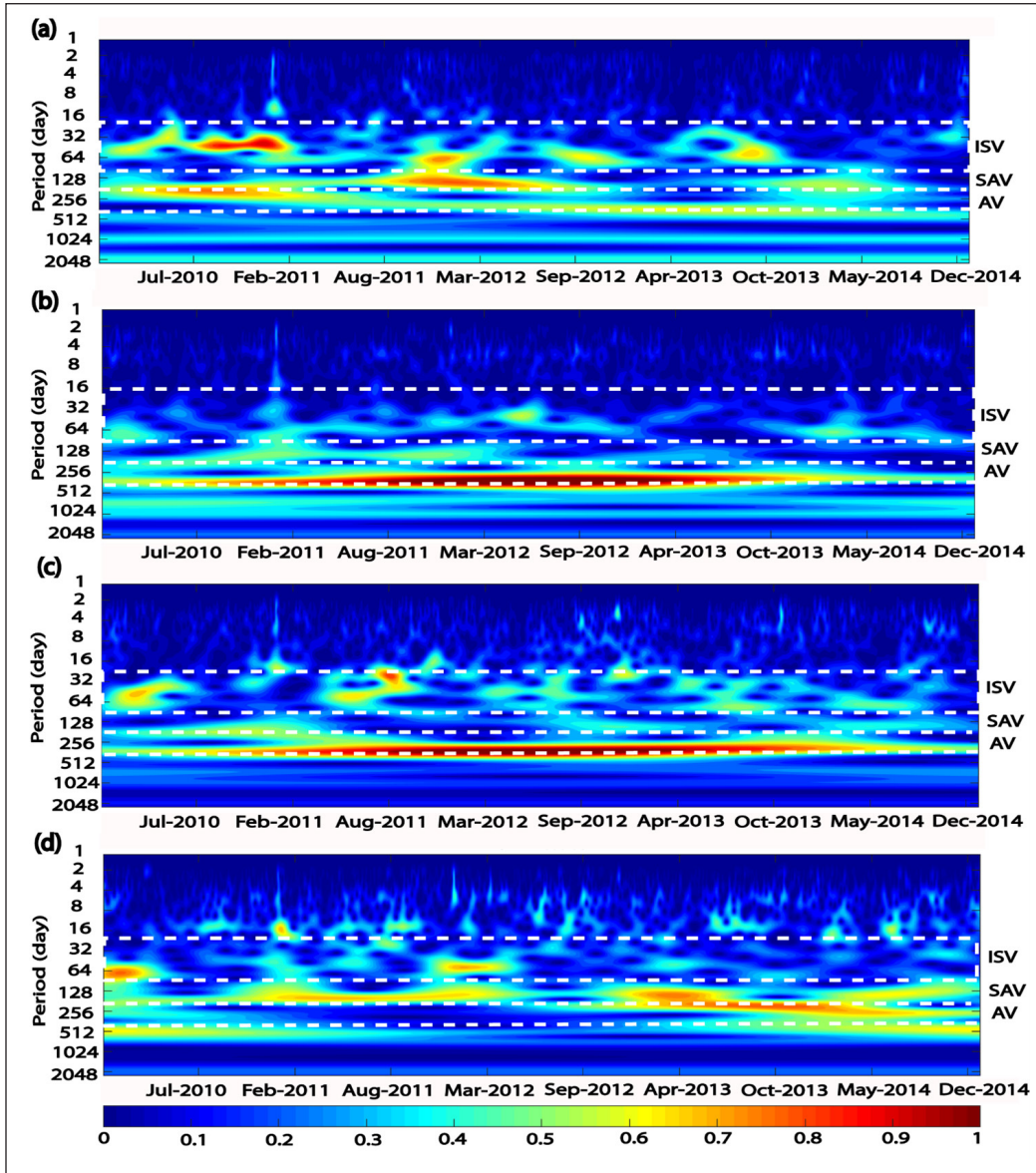


Figure 7. The result of CWT for transport of (a) MC; (b) NECC; (c) NGCC/NGCUC; and (d) Halmahera throughflow. ISV, SAV and AV stand for intraseasonal, semiannual and annual variability, respectively.

Coherence of the Connectivity System in the LLWBC Region and Halmahera Throughflow

The connection of volume transport between major currents of LLWBC (MC, NECC and NGCC/NGCUC) and Halmahera throughflow based on time series plots of transport anomaly (Figure 8) shows that when MC flows to the south, it is partially supplied to the Sulawesi Sea, while the rest changes direction to Pacific as NECC. NECC and NGCC/

NGCUC seem to have a noticeable relationship. The increase in NECC transport leads to a rise in NGCC/NGCUC, considering that the partial flow of NGCC/NGCUC eventually flows as NECC. This condition is also observed in the connection between NGCC/NGCUC transport and Halmahera transport. The anomaly fluctuation of both NGCC/NGCUC and Halmahera transport showed a substantial similarity. The rising transport of NGCC/NGCUC was responsible for the increasing transport of Halmahera. Detail of elaboration related to the transport relationship is presented in the discussion.

The relationship between MC, NECC, NGCC/NGCUC towards Halmahera throughflow transport was apparent in Figure 9 and Table 1. The cross-correlation analysis between time-series data of transport in each area demonstrated the presence of coherence, i.e., MC – NECC, MC – NGCC/NGCUC, MC – Halmahera, NECC – Halmahera, NGCC/NGCUC – NECC, and NGCC/NGCUC – Halmahera, significant at intraseasonal (20-90 days) and semiannual variabilities. The connection of transport fluctuation of MC – NECC in intraseasonal time-scales, i.e., 49 days and 60 days, contributed to the highest coherence value of 0.8 at 60 days, with a positive phase difference (6.42 days) and current signal in MC preceding NECC. The second highest fluctuation was found at 49-days with a coherence value of 0.79 and a positive phase difference of 3.33 days. The coherence value of MC – NGCC/NGCUC was found a semiannual period of 93 days and 128 days, i.e., 0.62 and 0.72, respectively. This output resulted in a positive phase difference of 9.78 days and 10.2 days, with a current signal in MC preceding NGCC/NGCUC. The relationship of transport fluctuation between NGCC/NGCUC and NECC was very close in 171 days and 341 days, as presented by a coherence value of 0.8 (171 days) and 0.79 (341 days). The phase difference was positive, i.e., 10.7 and 23.4 for 171 days and 341 days, respectively, with a current signal in NGCC/NGCUC preceding NECC. The fluctuation of MC – Halmahera transport was found at a semiannual period of 114 days and 256 days, with a coherence value of 0.69 and 0.57 and a positive phase difference of 13 days and 22.7 days, respectively. Meanwhile, the current signal was at MC preceding Halmahera throughflow. NECC's fluctuation – Halmahera transport was observed at a semiannual period of 171 days with a coherence of 0.69. In contrast, both areas' phase difference was at a positive value (12.9 days), with a current signal in NECC preceding Halmahera throughflow. The relationship between fluctuation in NGCC/NGCUC and Halmahera was observed at intraseasonal and semiannual time-scales of 34 days, 43 days, 128 days, and 171 days. The highest coherence (0.78) occurred at 128 days, with a positive phase difference of 14.8 days. Furthermore, the second-highest coherence (0.76) was found at periods of 43 days and 171 days, with a positive phase difference of 5.02 days and 20.3 days. This result suggests that the current NGCC/NGCUC precedes the signal in Halmahera throughflow between 5.02 days and 20.3 days.

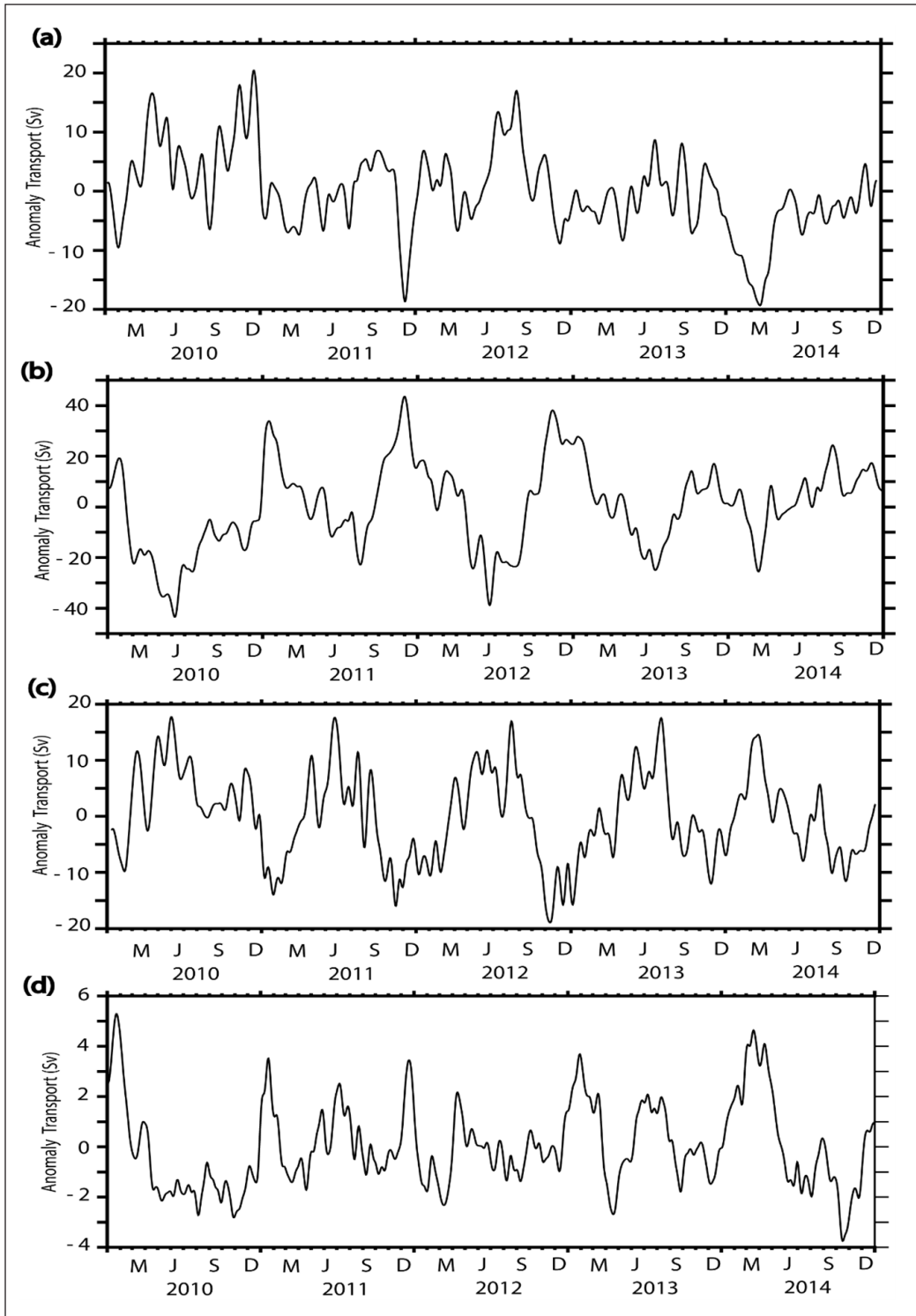


Figure 8. Transport anomaly: (a) MC; (b) NECC; (c) NGCC/NGCUC; and (d) Halmahera.

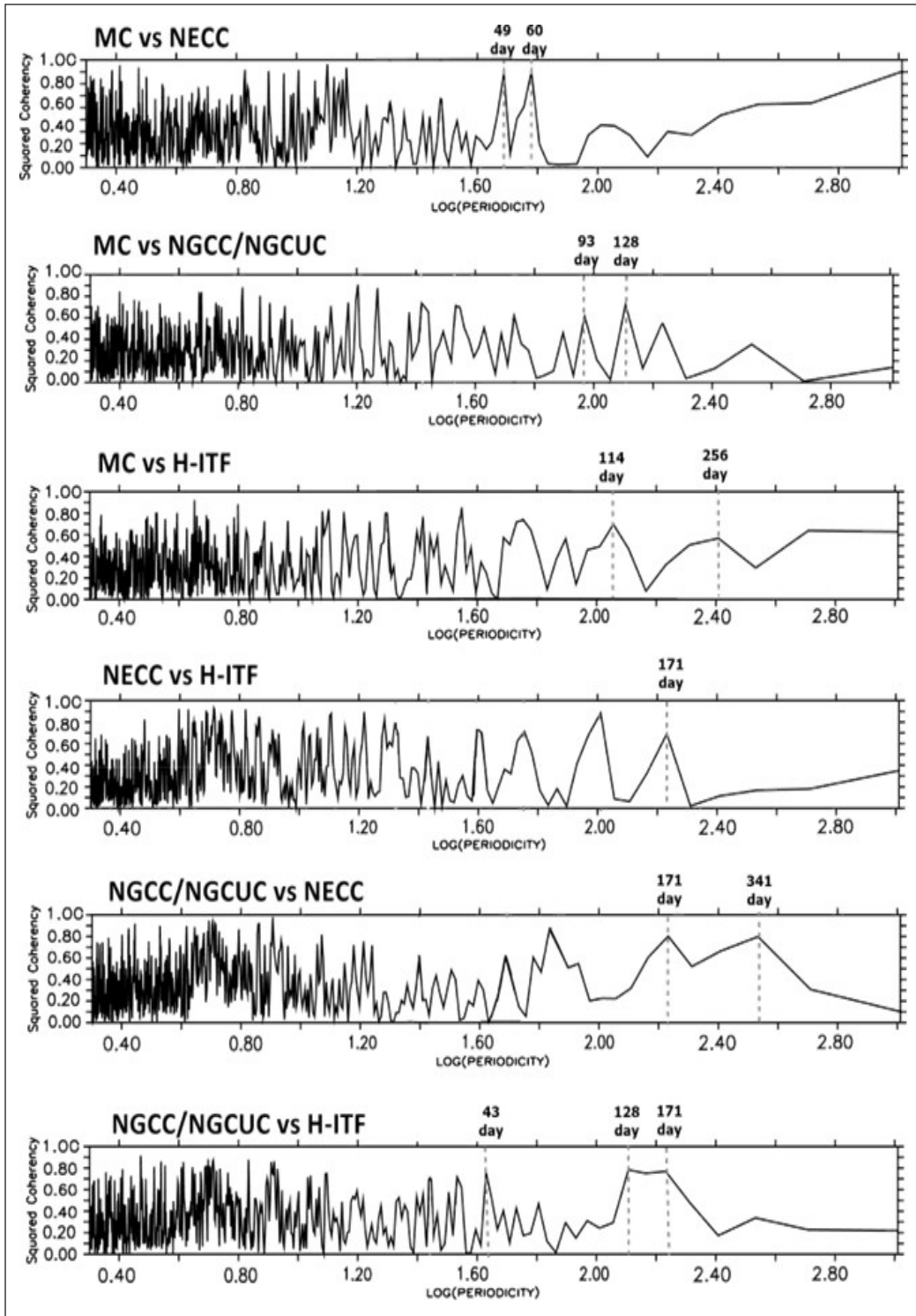


Figure 9. Coherence between MC, NECC, NGCC/NGCUC and Halmahera

Table 1
Coherence and phase lags observed in MC, NECC, NGCC/NGCUC and Halmahera

Cross PSD	Period (day)	Coherency	Phase (day)
MC – NECC	49	0.79	3.33
	60	0.80	6.42
MC – NGCC/NGCUC	93	0.62	9.78
	128	0.72	10.2
MC – Halmahera	114	0.69	13
	256	0.57	22.7
NECC – Halmahera	171	0.69	12.9
NGCC/NGCUC - NECC	171	0.80	10.7
	341	0.79	23.4
NGCC/NGCUC – Halmahera	43	0.76	5.02
	128	0.78	14.8
	171	0.76	20.3

DISCUSSION

The LLWBC is an intersection region of the water masses from the North Pacific (NP) and the South Pacific (SP) with distinct characteristics. The characteristics of the NP brought by MC are dissimilar to those from SP brought by NGCC/NGCUC. MC is a part of LLWBC in NP, formed when NEC impairs Philippine's eastern coast (Kashino et al., 1996; Kashino et al., 2013; Kashino et al., 2015). MC transports NP's water mass and mixes them when flowing to the south from the Philippine Sea, supplying the tropical circulation in the North Pacific and Indian Oceans (via ITF). The maximum core velocity of MC from INDESO model data (January 2010 to December 2014 along 7 °N) was above 1 m/s. Kashino et al. (2013) reported that according to SADPC assessment at 7 °N during a cruise from December 2006 to January 2007, the MC appeared to be very strong with a velocity of 2.0 m/s. The MC core velocity was high during boreal summers (Kashino et al., 2005). MC almost reached the maximum speed adjacent to 7°N with an average rate of 1.2 m/s nearby the surface (Qu et al., 2012). The cross-section demonstrates that MC existed on the surface until a depth of 700 m. This result follows a previous report by Qu et al. (2012), finding that MC disappeared below 600 m. In general, the current flows to the north closely relates to the southeastern anticyclonic eddies from Mindanao. The anticyclonic eddies consequence the slope towards the polar from HE in sub-thermocline (Qu et al., 1999). The fluctuation of MC transport was dominant in the annual period. Many studies on MC have focused on intraseasonal (Kashino et al., 2005), semiannual and yearly periods (Kim et al., 2004; Kashino et al., 2009, Kashino et al., 2011; Ren et al., 2018), as a local response to wind perturbation (Wyrтки, 1961; Qu et al., 2012) and remote forcing through propagating to the westward by Rossby waves (Kessler, 1990; Qiu & Lukas, 1996; Tozuka et al., 2002).

NECC is a flow originating from West Pacific that flows to the east. It was found to pass through the tropical Pacific between $2^{\circ}\text{N} - 10^{\circ}\text{N}$, mainly supplied by Mindanao current, which is recirculation of Mindanao Current from Sulawesi Sea entering the current of New Guinea (Wyrтки & Kendall, 1967; Lukas & Lindstrom, 1991; Donguy & Meyers, 1996; Hsin & Qiu, 2012; Zhao et al., 2016). The average zonal velocity from the INDESO model along 133°E was found to spread between $2^{\circ}\text{N} - 6^{\circ}\text{N}$. The core of NECC is found at 4°N . Donguy and Meyers (1996), obtaining data from expendable bathythermograph (XBT), reported that the NECC center was close to 5°N in the western Pacific and shifted to the north about 7°N in the center of the Pacific. NECC transport from the INDESO model was 34.5 Sv. Donguy and Meyers (1996) found the NECC transport of 21 Sv. It increased with increasing longitude from east to west. In terms of seasonality, the observed NECC transport is smaller in boreal winter–spring and larger in summer–fall (Johnson et al., 2002; Heron et al., 2006).

NGCC and NGCUC flow along the New Guinea coast, which is a part of South Equatorial Current. The core dimension of NGCC was spread from northern New Guinea up to 2°N , detected near-surface until 200 m. The maximum speed was > 0.5 m/s at a depth of 100 m. NGCUC occurs below NGCC, which flows to the northwest (Kuroda, 2000; Ueki et al., 2003). NGCC flows to the northwest and turns back to the eastern Halmahera Island, joining Mindanao current and then flowing to the east as NECC. The NGCC shows significant seasonal variations related to the monsoon. Ueki et al., (2003) reported that NGCC was stronger in Southeast monsoon than Northwest monsoon. Kuroda (2000) found that NGCC was a surface current caused by seasonal influences. Although the direction of NGCC varies on a seasonal basis (Wyrтки, 1961), it consistently flows towards the equatorial line. The subsurface NGCUC also changes seasonally, but the current direction does not change throughout the year (Kuroda, 2000; Ueki et al., 2003). During Southeast monsoon, NGCUC is strongly transported to the northwest (Kashino et al., 2007).

Halmahera throughflow reached 150 m in depth with a maximum speed of > 0.4 m/s from the surface to the depth of 100 m. The average transport of Halmahera current was 2.5 Sv, achieving the maximum value in September and October and minimum value in January. The fluctuation of Halmahera transport was dominated by semiannual (180 days) and intraseasonal (20 – 90 days) periods, followed by an annual period. The volume transport of ITF was higher in Southeast monsoon (May – September) than the Northeast Monsoon (October–March). This result was consistent with previous studies (Potemra, 2005, Sprintall et al., 2009). In the Monsoon break (October - November), Halmahera throughflow's velocity begins to diminish because changes in direction from the Southeast shifts to the Northwest monsoon winds, thus directing Ekman transport to the north (Wyrтки, 1961). The physical characteristics of water mass (temperature, salinity, and current) from observation data in Halmahera seas are relatively scarce. Cresswell & Luick (2001), after measurement using current mooring meter at a depth of 400 m, 700 m, and 900 m in the

Halmahera Sea, reported that the current in each depth level demonstrated distinct velocity direction. Nevertheless, the current change in the Halmahera Sea was dominantly influenced by NGCC. Wattimena et al. (2018) showed coherency in the transport volume between Halmahera throughflow and NGCC in the upper 200 m depth from daily INDES0 model output (2008-2014). They found the mean volume transport of Halmahera throughflow was $2.35 (\pm 1.94)$ Sv (southward), and volume transport of NGCC was $13.47 (\pm 6.6)$ Sv (westward). The model indicated that Halmahera throughflow is supplied by 17.45% of NGCC.

Transport of MC, NECC, NGCC/NGCUC, and Halmahera throughflow displays a clear connection but very complicated. The schematic of average transport LLWBC and partitions to Indonesian seas based on the INDES0 model are presented in Figure 10. Likewise, this pattern can also be seen in anomaly transport shown in Figure 8 in which the intensified MC increases NECC transport. A rising in the transport of NGCC leads to an increase in NECC transport. This aspect is explainable since the change in the direction of NGCC flow also contributed to NECC. This aspect also explains the connection between NGCC and Halmahera transport. With increasing NGCC transport, Halmahera transport was also rose. The comparisons of the estimations of volume transport of MC, NECC, NGCC/NGCUC and Halmahera throughflow between this current study and previous work are shown in Table 2. Burnett et al. (2003) reported that NECC transport flow increased with the decreasing MC transport entering the Sulawesi Sea. The role of NGCC to induce transport to Halmahera could be observed during Southeast monsoon, in which wind stress is almost in line with the NGCC direction. This aspect strengthens Ekman flow along the New Guinea coast, contributing to elevating transport to Halmahera. The roles of MC, NECC, NGCC/NGCUC, and Halmahera throughflow were in a strong relationship as they affected the water mass transport passing through Halmahera throughflow. MC transport to the south affects NECC transport.

Similarly, NECC transport is also affected by NGCC/NGCUC transport. The reversed direction of NGCC in the northeast of Halmahera Island eventually joins with reversal MC current, then flowing to the east as NECC. The relationship between transport of MC – NECC was found to have the highest coherence (0.8) at 60 days, with a positive phase difference (6.42 days) and current signal in MC preceding NECC up to 6.42 days. The shifting of NGCC/NGCUC into NECC in the northeast of Halmahera Island generated Halmahera Eddies (HE), also contributed to regulate the composition of ITF passing through the east gate (Aruda & Nof, 2003; Kashino et al., 2013; Harsono et al., 2014). The connection of transport fluctuation between NGCC/NGCUC – NECC was strongly correlated, as indicated by the highest coherence value (0.8) for 171 days with a positive phase difference of 10.7 days, meaning that signal in NGCC/NGCUC is preceding NECC up to 10.7 days. Transport of water mass to Halmahera through the east gate of ITF

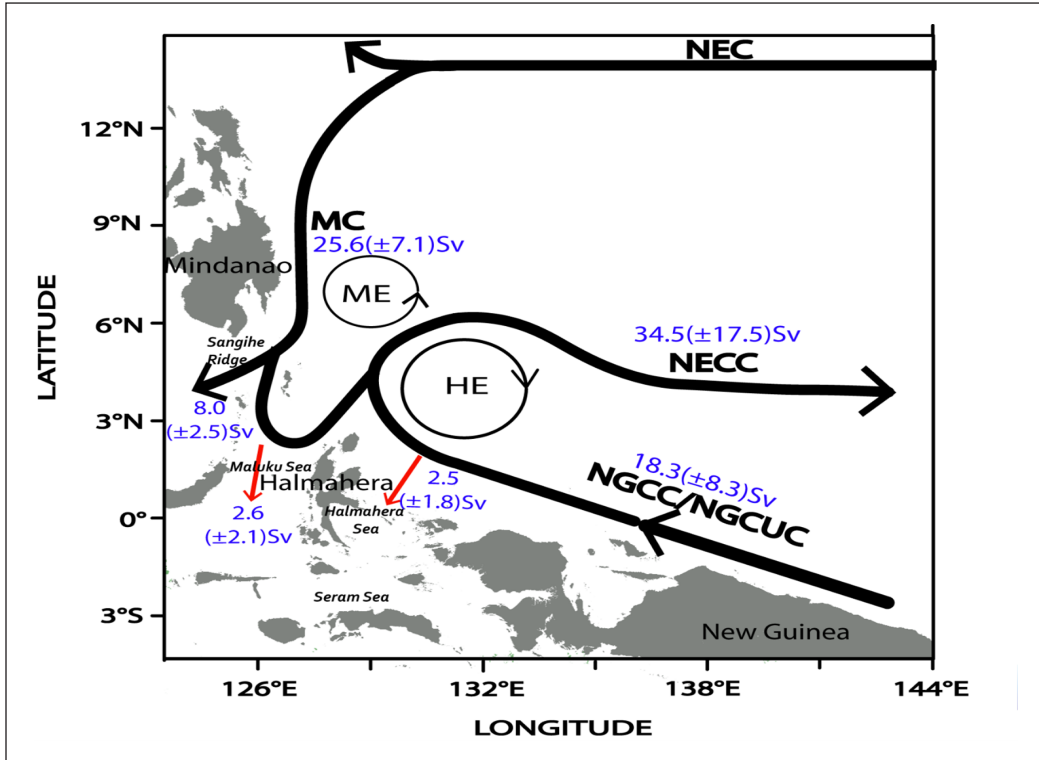


Figure 10. Schematic description of LLWBC. The blue value represents the transport of MC, NECC, NGCC/NGCUC and Halmahera throughflow, Sangihe Ridge and north Maluku Sea

Table 2
Comparison of LLWBCs system transport estimates between INDESO output model (this study) and past studies

Ocean Current	This Study	Past Studies
NECC	25.6 (±7.1) Sv	15-20 Sv (Webb, 2018) 10–30 Sv (Gouriou & Toole, 1993); Johnson et al., 2002); Wyrcki & Kendall, 1967) 21 Sv (Donguy & Meyers, 1996)
MC	34.5 (±17.5) Sv	8-40 Sv (Lukas, 1988) 31 Sv (Kashino et al., 2009) 27 Sv (Qu et al., 1998)
NGCC/ NGCUC	18.3 (±8.3) Sv	9-18 Sv (Lindstrom et al., 1990) 11-32 Sv (Kuroda, 2000) 13.47 Sv (Wattimena et al., 2018)
Halmahera	2.5(±1.8) Sv	3 Sv (Morey et al., 1999) 1.5 Sv (Cresswell & Luick, 2001) 2.2 Sv (Potemra et al., 2003) 2.35 Sv (Wattimena et al., 2018)

was dominantly affected by NGCC/NGCUC transport. The coherence analysis result demonstrates that the phase difference of signal fluctuation between NGCC/NGCUC and Halmahera throughflow was at a range of 0.76 and 0.78, with a positive phase difference of 5.02 days to 20.3 days. This result suggests that the signal fluctuation of NGCC/NGCUC flowing to Halmahera throughflow takes between 5.02 days and 20.3 days, indicating the presence of the Rossby wave's signal propagation in the Halmahera Sea.

CONCLUSION

The present study examined the vertical structures and their volume transport variabilities in the western equatorial Pacific current system (LLWBC). In particular, it looked at the coherence between the northwestward NGCC/NGCUC and Halmahera throughflow.

The conclusions of the present study can be summarized as follows:

1. MC is characterized by a core current occupying the surface layer to 250 m with a maximum current larger than 1 m/s. The core current of NGCC is located at the top 200 m with a maximum current larger than 0.5 m/s, while NGCUC exists below the core current of NGCC. The maximum currents of the cores of MC and NGCC/NGCUC are visible in the NECC current structure. NECC's core is present within the top 300 m with a maximum speed larger than 0.5 m/s.
2. Likewise, the maximum current of the core of NGCC/NGCUC is present in the ocean current's vertical structure in the Halmahera throughflow. Halmahera throughflow is characterized by two cores of current located within the top 100 m and around 400 m. The shallow core current has a stronger ocean current (> 0.4 m/s) compared to that in the deep core (up to 0.2 m/s).
3. There are strong relationships between NGCC/NGCUC and Halmahera throughflow based on the coherence value (coefficient coherence: ~ 0.8). This high coherence is positively related to Halmahera throughflow directly connected to NGCC/NGCUC. Based on this coherence analysis, the variability of NGCC/NGCUC will affect Halmahera throughflow after 5 to 20 days (i.e., the phase difference).

ACKNOWLEDGMENTS

The authors are very grateful to the INDESO project of the Indonesian Ministry of Marine Affairs and Fisheries for the data, the Group for High-Resolution Sea Surface Temperature (SST) (GHRSSST) (<https://www.ghrsst.org/ghrsst-data-services/products/>), *Segment Sol Multi missions d'Altimétrie, d'Orbitographie et de localisation précise* (SSALTO) and the Data Unification and Altimetry Combination System (DUACS) for the sea surface height anomaly (SSHA) data (<http://aviso.oceanobs.com/duacs/>), and INDOMIX program for the near-bottom current velocity data from ADCP. They also thank the Anonymous Reviewers, whose comments helped in the improvement of the manuscript.

REFERENCES

- Arruda, W. Z., & Nof, D. (2003). The Mindanao and Halmahera eddies: Twin eddies induced by nonlinearities. *Journal of Physical Oceanography*, 33(12), 2815-2830. [https://doi.org/10.1175/1520-0485\(2003\)033<2815:TMAHEE>2.0.CO;2](https://doi.org/10.1175/1520-0485(2003)033<2815:TMAHEE>2.0.CO;2)
- Atmadipoera, A., Koch-Larrouy, A., Cuypers, Y., Molcard, R., Jaya, I., & Harsono, G. (2013, June 24-27). *New current measurement in the Halmahera Passage, Indonesia [Poster Presentation]. Asian Oceania Geosciences Society Symposium, Australia.*
- Atmadipoera, A., Kuroda, Y., Pariwono, J. I., & Purwandana, A. (2004). Water mass variation in the upper layer of the Halmahera eddy region observed from a TRITON buoy. *The Institute of Electrical and Electronics Engineers*, 3, 1496-1503. <https://doi.org/10.1109/OCEANS.2004.1406342>
- Bahrel, P., & Mercator, P. T. (2006). Mercator ocean global to regional ocean monitoring and forecasting. In E. P. Chassignet & J. Verron (Ed.), *Ocean weather forecasting* (pp. 381-395). Springer.
- Bendat, J. S., & Piersol, A. G. (2010). *Random data analysis and measurement procedure*. John Wiley and Sons Inc.
- Burnett, W. H., Kamenkovich, V. M., Gordon, A. L., & Mellor, G. L. (2003). The Pacific/Indian Ocean pressure difference and its influence on the Indonesian Seas circulation: Part I—The study with specified total transports. *Journal of Marine Research*, 61(5), 577-611. <https://doi.org/10.1357/002224003771815963>
- Cresswell, G. R., & Luick, J. L. (2001). Current measurement in the Halmahera Sea. *Journal of Geophysical Research*, 106(C7), 13945-13952. <https://doi.org/10.1029/2000JC000688>
- Donguy, J. R., & Meyers, G. (1996). Mean annual variation of transport of major currents in the tropical Pacific Ocean. *Deep Sea Research, Part I*, 43(7), 1105-1122. [https://doi.org/10.1016/0967-0637\(96\)00047-7](https://doi.org/10.1016/0967-0637(96)00047-7)
- Emery, W., & Thomson, R. E. (2014). *Data analysis methods in physical oceanography* (3rd Ed.). Elsevier Science.
- Fine, R. A., Lukas, R., Bingham, F., Warnar, W., & Gammon, R. (1994). The western equatorial Pacific: A water mass crossroads. *Journal of Geophysical Research*, 99, 25063-25080. <https://doi.org/10.1029/94JC02277>
- Gordon, A. L. (2005). Oceanography of the Indonesian seas and their throughflow. *Oceanography*, 18(4), 14-27.
- Gordon, A. L., & Fine, R. (1996). Pathways of water between the Pacific and Indian oceans in the Indonesian seas. *Nature*, 379(6561), 146-149. <https://doi.org/10.1038/379146a0>
- Gouriou, Y., & Toole, J. (1993). Mean circulation of the upper layers of the western equatorial Pacific Ocean. *Journal of Geophysical Research*, 98(C12), 22495-22520. <https://doi.org/10.1029/93JC02513>
- Harsono, G., Atmadipoera, A. S., Syamsudin, F., Manurung, D., & Mulyono, S. B. (2014). Halmahera eddy features observed from multisensor satellite oceanography. *Asian Journal of Scientific Research*, 7(4), 571-580. <https://doi.org/10.3923/ajsr.2014.571.580>
- Heron, S. F., Metzger, E. J., & Skirving, W. J. (2006). Seasonal variations of the ocean surface circulation in the vicinity of Palau. *Journal of Oceanography*, 62(4), 413-426. <https://doi.org/10.1007/s10872-006-0065-3>

- Hsin, Y. C., & Qiu, B. (2012). Seasonal fluctuations of the surface North Equatorial Counter Current (NECC) across the Pacific basin. *Journal of Geophysical Research*, *117*, Article C06001. <https://doi.org/10.1029/2011JC007794>
- Hu, D., Wu, L., Cai, W., Gupta, A. S., Ganachaud, A., Qiu, B., Gordon, A. L., Lin, X., Chen, Z., Hu, S., & Wang, G. (2015). Pacific western boundary currents and their roles in climate. *Nature*, *522*(7556), 299-308. <https://doi.org/10.1038/nature14504>
- Ilahude, A. G., & Gordon, A. L. (1996). Thermocline stratification within the Indonesian seas. *Journal of Geophysical Research*, *101*(12), 401-409. <https://doi.org/10.1029/95JC03798>
- Johnson, G. C., Sloyan, B. M., Kessler, W. S., & McTaggart, K. E. (2002). Direct measurements of upper ocean currents and water properties across the tropical Pacific during the 1990s. *Progress in Oceanography*, *52*(1), 31-61. [https://doi.org/10.1016/S0079-6611\(02\)00021-6](https://doi.org/10.1016/S0079-6611(02)00021-6)
- Kashino, Y., Aoyama, M., Kawano, T., Hendiarti, N., Anantasena, Y., Muneyama, K., & Watanabe, H. (1996). The water masses between Mindanao and New Guinea. *Journal of Geophysical Research*, *101*(12), 391-400. <https://doi.org/10.1029/95JC03797>
- Kashino, Y., Atmadipoera, A. S., Kuroda, Y., & Lukijanto. (2013). Observed feature of the Halmahera and Mindanao Eddies. *Journal of Geophysical Research Oceans*, *118*, 6543-6560. <https://doi.org/10.1002/2013JC009207>
- Kashino, Y., Espana, N., Syamsudin, F., Richards, K., Jensen, T., Dutrieux, P., & Ishida, A. (2009). Observations of the North Equatorial Current, Mindanao Current, and Kuroshio current system during the 2006/07 El Nino and 2007/08 La Nina. *Journal of Oceanography*, *65*, 325-333. <https://doi.org/10.1007/s10872-009-0030-z>
- Kashino, Y., Firing, E., Hacker, P., Sulaiman, A., & Lukiyanto. (2001). Currents in the Celebes and Maluku Seas, February 1999. *Geophysical Research Letters*, *28*, 1263-1266. <https://doi.org/10.1029/2000GL011630>
- Kashino, Y., Ishida, A., & Hosoda, S. (2011) Observed ocean variability in the Mindanao Domeregion. *Journal of Physical Oceanography*, *41*, 287-302. <https://doi.org/10.1175/2010JPO4329.1>
- Kashino, Y., Ishida, A., & Kuroda, Y. (2005). Variability of the Mindanao current: Mooring observation results. *Geophysical Research Letters*, *32*, Article L18611. <https://doi.org/10.1029/2005GL023880>
- Kashino, Y., Ueki, I., & Sasaki, H. (2015). Ocean variability east of Mindanao: Mooring observations at 7N. *Journal of Geophysical Research Oceans*, *120*, 2540-2554. <https://doi.org/10.1002/2015JC010703>
- Kashino, Y., Ueki, I., Kuroda, Y., & Purwandana, A. (2007). Ocean variability north of New Guinea derived from TRITON buoy data. *Journal of Oceanography*, *63*, 545-559. <https://doi.org/10.1007/s10872-007-0049-y>
- Kessler, W. S. (1990). Observations of long Rossby waves in the northern tropical Pacific. *Journal of Geophysical Research: Oceans*, *95*(C4), 5183-5217. <https://doi.org/10.1029/JC095iC04p05183>
- Kim, Y. Y., Qu, T., Jensen, T., Miyama, T., Mitsudera, H., Kang, H. W., & Ishida, A. (2004). Seasonal and interannual variations of the North Equatorial Current bifurcation in a high-resolution OGCM. *Journal of Geophysical Research: Oceans*, *109*(C3). <https://doi.org/10.1029/2003JC002013>

- Kuroda, Y. (2000) Variability of the currents off the north coast of New Guinea. *Journal of Oceanography*, 56, 103-106. <https://doi.org/10.1023/A:1011122810354>
- Lindstrom, E., Butt, J., Lukas, R., & Godfrey, S. (1990). The flow through Vitiaz Strait and St. George's Channel, Papua New Guinea. In L. Pratt (Ed.) *The physical oceanography of sea straits*, (pp. 171-189). Kluwer Academic Publisher.
- Lukas, R. (1988). Interannual fluctuations of the Mindanao Current inferred from sea level. *Journal of Geophysical Research*, 93, 6744-6748. <https://doi.org/10.1029/JC093iC06p06744>
- Lukas, R., & Lindstrom, E. (1991). The mixed layer of the western equatorial Pacific Ocean. *Journal of Geophysical Research*, 96(S01), 3343-3357. <https://doi.org/10.1029/90JC01951>
- Lukas, R., Yamagata, T., & McCreary, J. P. (1996). Pacific low latitude western boundary currents and the Indonesian throughflow. *Journal of Geophysical Research Oceans*, 101, 12209-12216. <https://doi.org/10.1029/96JC01204>
- Morey, S. L., Shriver, J. F., & O'Brien, J. J. (1999). The effects of Halmahera on the Indonesian Throughflow. *Journal of Geophysical Research*, 104, 23281-23296. <https://doi.org/10.1029/1999JC000195>
- Potemra, J. T. (2005). Indonesian throughflow transport variability estimated from satellite altimetry. *Oceanography*, 18(4), 98-107.
- Potemra, J. T., Hautala, S. L., & Sprintall, J. (2003). Vertical structure of Indonesian throughflow in a large-scale model. *Deep Sea Research II*, 50(12-13), 2143-2162. [https://doi.org/10.1016/S0967-0645\(03\)00050-X](https://doi.org/10.1016/S0967-0645(03)00050-X)
- Qiu, B., & Lukas, R. (1996). Seasonal and interannual variability of the North Equatorial Current, the Mindanao Current, and the Kuroshio along the Pacific western boundary. *Journal of Geophysical Research: Oceans*, 101(C5), 12315-12330. <https://doi.org/10.1029/95JC03204>
- Qu, T. H., Mitsudera, H., & Yamagata, T. (1999). A climatology of the circulation and water mass distribution near the Philippine coast. *Journal of Physical Oceanography*, 29(7), 488-1505. [https://doi.org/10.1175/1520-0485\(1999\)029<1488:ACOTCA>2.0.CO;2](https://doi.org/10.1175/1520-0485(1999)029<1488:ACOTCA>2.0.CO;2)
- Qu, T. H., Mitsudera, H., & Yamagata, T. (1998). On the western boundary currents in the Philippine Sea. *Journal of Geophysical Research*, 103(C4), 7537-7548. <https://doi.org/10.1029/98JC00263>
- Qu, T., Chiang, T., Wu, C., Dutrieux, P., & Hu, D. (2012). Mindanao current/undercurrent in an eddy-resolving GCM. *Journal of Geophysical Research*, 117(C6), Article C06026. <https://doi.org/10.1029/2011JC007838>
- Ren, Q., Li, Y., Wang, F., Song, L., Liu, C., & Zhai, F. (2018). Seasonality of the Mindanao current/undercurrent system. *Journal of Geophysical Research Oceans*, 123(2), 1105-1122. <https://doi.org/10.1002/2017JC013474>
- Sprintall, J., Wijffels, S. E., Molcard, R., & Jaya, I. (2009). Direct estimates of the Indonesian throughflow entering the Indian Ocean: 2004-2006. *Journal of Geophysical Research*, 114(c7), Article C07001. <https://doi.org/10.1029/2008JC005257>
- Theetten, S., Thiebaul, B., Dumas, F., & Paul, J. (2014). Bmgtools: A community tool to handle model grid and bathymetry. *Mercator Ocean-Quarterly Newsletter*, (49), 94-98.

- Tozuka, T., Kagimoto, T., Masumoto, Y., & Yamagata, T. (2002). Simulated multiscale variations in the western tropical Pacific: The Mindanao Dome revisited. *Journal of Physical Oceanography*, 32(5), 1338-1359. [https://doi.org/10.1175/1520-0485\(2002\)032<1338:SMVITW>2.0.CO;2](https://doi.org/10.1175/1520-0485(2002)032<1338:SMVITW>2.0.CO;2)
- Tranchant, B., Reffray, G., Greiner, E., Nugroho, D., Koch-Larrouy, A., & Gaspar, A. (2015). Evaluation of an operational ocean model configuration at 1/12° spatial resolution for the Indonesian seas (NEMO2.3/INDO12) – Part 1: Ocean physics. *Geoscientific Model Development*, 9(3), 1037-1064. <https://doi.org/10.5194/gmd-9-1037-2016>
- Ueki, I., Kashino, Y., & Kuroda, Y. (2003). Observation of current variations off the New Guinea coast including the 1997-1998 El Niño period and their relationship with Sverdrup transport. *Journal of Geophysical Research*, 108(C7), Article 3243. <https://doi.org/10.1029/2002JC001611>
- Wattimena, M. C., Atmadipoera, A. S., Purba, M., Nurjaya, I. W., & Syamsudin, F. (2018). Indonesian throughflow (ITF) variability in Halmahera Sea and its coherency with New Guinea Coastal current. *IOP Conference Series: Earth and Environmental Science*, 176(1), Article 012011. <https://doi.org/10.1088/1755-1315/176/1/012011>
- Webb, D. J. (2018). On the role of the North Equatorial counter current during a strong El Niño. *Ocean Science*, 14(4), 633-660.
- Wijffels, S., Firing, E., & Toole, J. (1995). The mean structure and variability of the Mindanao Current at 8°N. *Journal of Geophysical Research: Oceans*, 100(C9), 18421-18435. <https://doi.org/10.1029/95JC01347>
- Wyrtki, K. (1961). *Physical oceanography of the Southeast Asian Waters*. NAGA Report 2 Scripps Inst.
- Wyrtki, K., & Kendall, R. (1967). Transports of the Pacific equatorial countercurrent. *Journal of Geophysical Research*, 72(8), 2073-2076. <https://doi.org/10.1029/JZ072i008p02073>
- Zhao, J., Li, Y. L., & Wang, F. (2016). Seasonal fluctuations of the WPO surface NECC. *Chinese Journal of Oceanology and Limnology*, 34(6), 1332-1346.



Solution of Collinearity Problem in Two-Dimensional Electrical Resistivity Tomography using Wenner Array

Mustapha Adejo Mohammed^{1,2}, Nordiana Mohd Muztaza^{1*} and Rosli Saad¹

¹*School of Physics, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia*

²*Department of Physics, Federal University of Lafia, Nasarawa, Nigeria*

ABSTRACT

Two-dimensional electrical resistivity tomography (2-D ERT) is one of the most common geophysical tools employed to satisfy the ever-growing need for obtaining subsurface information. Most of the conventional electrode arrays used for 2-D ERT survey are built with the theoretical assumption that the survey lines are straight to guarantee four collinear electrodes at every point of measurement. However, due to surface constraint associated with most survey areas, it is rarely possible to conduct a two-dimensional resistivity survey on a straight line. Therefore, 2-D ERT survey conducted on a surface constraint field requires shifting one or more electrodes off the survey line, which contrasts with the underlying assumption. Consequently, the result might be prone to false anomalies. Thus, this study aimed to devise a new approach that could mitigate the false anomalies posed by non-collinearity of electrodes in 2-D ERT result. In view of this, ABEM Terrameter SAS4000 using Wenner array configuration was adopted for the survey. The data was acquired with all electrodes inline and one or more electrodes offline at stepwise distances, respectively. Based on the result obtained, the new approach mitigates the offline electrodes effect, as the inverse resistivity tomograms resolves the geometries of the true model reasonably well. More so, it has high R-value >90% which is an indication of proximity to the true model. Hence, it is concluded that the new approach is effective in mitigating offline electrode effect on a 2-D ERT result.

ARTICLE INFO

Article history:

Received: 02 November 2020

Accepted: 19 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.27>

E-mail addresses:

mustyadejo@gmail.com (Mustapha Adejo Mohammed),

mmnordiana@usm.my (Nordiana Mohd Muztaza)

rosli@usm.my (Rosli Saad)

* Corresponding author

Keywords: Collinear electrodes, mitigate, resistivity tomography, Wenner array

INTRODUCTION

The quest for subsurface information in environmental and engineering management had become paramount. This is due to

the vital role the subsurface information plays in the design of civil structures to avoid foundation failure, which could lead to loss of lives and properties. More so, it plays an essential role in the delineation of groundwater potential zones to ensure optimum development and management of groundwater resources. Conventionally, a subsurface investigation is based on borehole exploration method (Anastasopoulos, 2013; Pando et al., 2013). However, the technique is attributed with several drawbacks such as high cost of drilling, limited data coverage, site damage and time-consuming. Therefore, to address these drawback, multidisciplinary research across the engineering, physical and social sciences is required (Fragaszy et al., 2011). Thus, geophysical methods offer solution to some of the inherent challenges in borehole exploration method (Abidin et al., 2017) Among the advantages of geophysical techniques are; fast data acquisition, low cost and ability to cover a large area (Cosenza et al., 2006; Godio et al., 2006)

As reported by Griffiths and Barker (1993), the electrical resistivity method is one of the most often used geophysical methods for subsurface investigation. Over time, the technique had witnessed great improvement in terms of availability of forward modelling and inversion schemes, efficient data acquisition and processing techniques (Ingeman-Nielsen et al., 2016; Rucker et al., 2012). This led to the evolvement of two, three- and four-dimensional electrical resistivity tomography. However, the two-dimensional electrical resistivity tomography (2-D ERT) is the best choice when the goal is to keep survey cost down and obtain an accurate result. As such, the technique is widely employed in geophysical investigations involving subsurface characterization, hydrogeology, karst and cave finding, mineral exploration, geotechnical and archaeological prospecting (Abudeif et al., 2020; Loke et al., 2013; Uhlemann et al., 2018).

Furthermore, most of the electrode arrays such as Wenner, Wenner-Schlumberger, Dipole-dipole and Pole-dipole employed for 2-D ERT survey are designed with the assumption that the survey line is straight to ensure collinear electrode pairs at each point of measurement. But, due to surface constraint associated with most survey areas, it is rarely possible to conduct a two-dimensional resistivity survey on a straight line. To this end, Szalai et al. (2008) investigated the effect of electrode positioning error on the inverted pseudo-section of a 2-D ERT survey. They found the positioning error to be insignificant except for a very rocky field. And they suggested neglecting a few electrode positions with the greatest positioning error. In another work by Zhou and Dahlin (2003), they examined the properties and effect of measurement error (off-line and in-line spacing error) on 2-D resistivity tomography survey. In their study, the offline electrodes causing the offline spacing error were omitted. Hence, they concluded that the offline spacing error was negligible. However, ignoring or omitting some electrodes could generate false anomalies, thus leads to misinterpretation.

This research, therefore, seeks to make known the effect of offline electrodes and develop a time and cost-effective approach that could mitigate the effect without ignoring or omitting any electrode. To achieve this, Wenner array was employed for the survey. The survey was conducted with all electrodes inline and some electrodes offline respectively, with the view to reveal the offline electrode effect. More so, the new approach attempted to recalculate the geometric factor and thus the apparent resistivity of those data points associated with the offline electrodes in an automated manner. The obtained result of this research will no doubt help to prevent misinterpretation caused by the offline electrodes on a 2-D ERT survey.

MATERIALS AND METHODS

Description of Study Area and Geology

The study area is located in Universiti Sains Malaysia, Penang Island, Malaysia. The area has a tropical climate with temperature ranging from 23.5°C to 31.4°C, average annual rainfall from 2670 mm to 3250 mm, and relative humidity ranging from 60.9 % to 96.8 % (Ali et al., 2011). Geologically, the island is underlain by granitic rock (Figure. 1). The island is divided mainly into two groups: the North Penang Pluton (NPP) and the South Penang Pluton (SPP). The NPP consists of orthoclase to intermediate microcline granite, while the SPP is dominated by microcline granite (Ahmad et al., 2006; Tan, 1994). The study

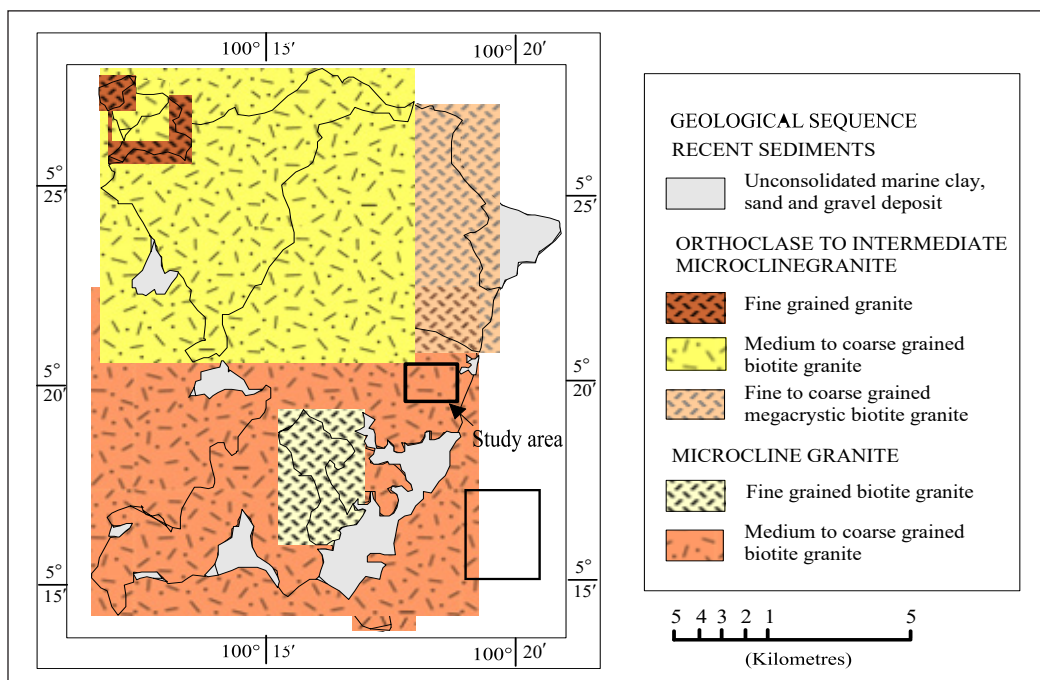


Figure 1. Geology map of Penang Island Malaysia showing the study area

area belongs to the South Penang Pluton, which is of late Carboniferous age. The SPP is composed of fined grained biotite and medium to coarse-grained biotite granite (Ali et al., 2011; Pradhan & Lee, 2010). Universiti Sains Malaysia (USM) consists mainly of residual soil originating from the parent rock; as such, it is relatively homogeneous in composition. This, therefore, reduces possible errors in our research as the geology is incomplex.

Data Acquisition

In general, a set of four electrodes were used for every measurement (Figure 2). The subsurface potential contrast (V) was measured between two electrodes (often called potential electrodes, indicated by M – N) upon injecting specified amount of electric current (I) into the ground between the other two electrodes (current electrodes, A – B). Since the principle of electrical resistivity method is embedded in Ohm’s law, the resistance (R) is then computed using Equation 1 (Muhammad & Saad, 2018). Subsequently, the apparent resistivity is obtained when array geometrical factor is known (Equation 2). The geometric factor (K) is a key component in determining the apparent resistivity. Its value varies, depending on the type of array (i.e., geometric location of current and potential electrodes) used for the measurement.

$$R = \frac{V}{I} \tag{1}$$

$$\rho_a = KR \tag{2}$$

And

$$K = \frac{2\pi}{\left(\frac{1}{r_1} - \frac{1}{r_2} - \frac{1}{r_3} + \frac{1}{r_4}\right)} \tag{3}$$

where r_1, r_2, r_3 and r_4 are the inter-electrode distances (Equation 3).

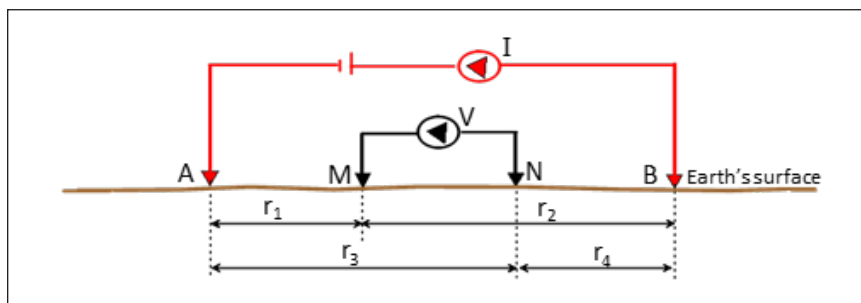


Figure 2. Four-point electrode configuration for Wenner array

Multi-electrode resistivity meter (ABEM Terrameter SAS4000) was used to acquire the 2-D ERT survey data. Wenner Array was employed for this study due to its high signal-to-noise ratio and sensitivity to vertical changes in resistivity (Dahlin & Zhou, 2004; Loke, 2018). The data acquisition was in two phases, these are; survey with all electrodes inline and survey with one or more electrodes offline. In both phases, the input current was set between 1 and 100 mA. The procedure involved for each phase is explained in the following subsection.

Wenner Array with All Electrodes Inline (WA-AEI). Measurement was made on a 40 m-long profile with 41 stainless steel electrodes planted at 1 m minimum electrode spacing (Figure 3). All the multi-electrodes were linearly planted to ensure the collinearity of electrodes pairs at each point of measurement, as stated theoretically (Loke et al., 2013). The measurement had a total of 260 data points. It is important to note that each electrode used for the measurement has a specified number of data points associated with it depending on its lateral position along the profile. More so, the WA-AEI survey was the true model.

Wenner Array with One Or More Electrodes Offline (WA-1MEO). In this phase, Wenner array was conducted with one, two and three electrodes offline (WA-1EO, WA-2EO, and WA-3EO) respectively. Figure 4 depicts survey site photos for WA-1EO, WA-2EO and WA-3EO with offline electrode(s) at 1 m off the survey line. The choice for the selected electrodes offline was discretionary, with the view to observe the possible effect on the inverse resistivity model. In WA-1EO survey, the electrode arrangement used for WA-AEI was maintained (40 m-long profile) except electrode number 21 that was perpendicularly shifted off the survey line in a stepwise distance of 0.1 m increment (0.1, 0.2, ... 1.2 m) for each acquisition. The shifted electrode (electrode number 21) was associated with 33 data points, and it implied that for all the measurements taken, 33 data points were affected for each increment. For WA-2EO

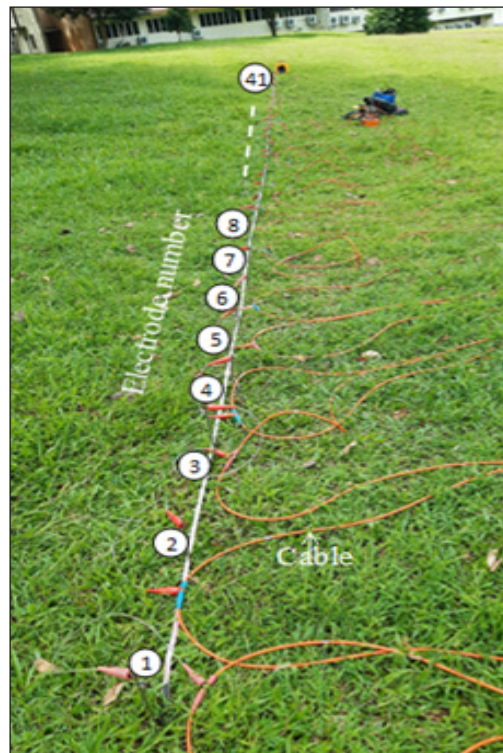


Figure 3. Site photos for a 2-D electrical resistivity tomography survey using Wenner array with all electrodes inline

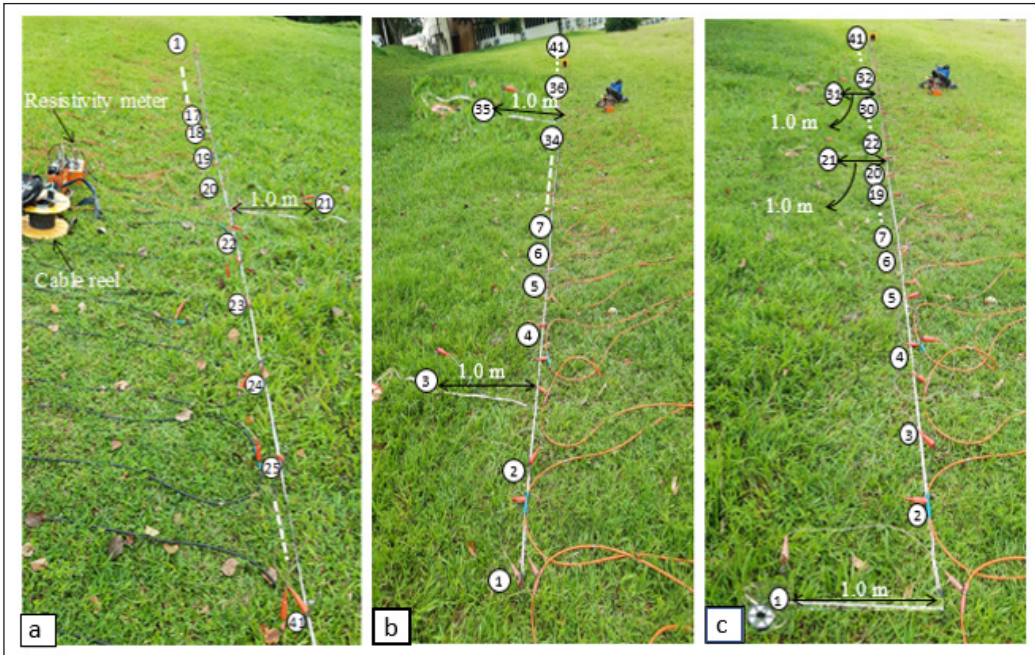


Figure 4. Site photos for a 2-D electrical resistivity tomography survey using Wenner array with (a) electrode number 21 offline (b) electrode number 3 and 35 offline and (c) electrode number 1, 21 and 31 offline at 1.0 m distance respectively

survey, all the electrodes used for WA-AEI measurement remained fixed (40 m-long profile) except electrodes number 3 and 35 that were orthogonally offline in a stepwise distance of 0.1 m increment (0.1, 0.2, ... 1.2 m) for each measurement. As mentioned earlier, each electrode had a range of data points associated with it. A total of 37 data points associated with the two shifted electrodes were found to be affected for each acquisition.

In a similar vein, WA-3EO measurements were made along 40 m-long profile as in WA-AEI, with 41 electrodes linearly planted except for electrodes (1, 21 & 31) that were shifted perpendicularly off the 2-D resistivity survey line in a stepwise distance of 0.1 m increment (0.1, 0.2, ... 1.2 m). Just like the previous phases, 68 data points were associated with these shifted electrodes. Therefore, 68 data points were affected in each of the measurements made for the respective increment.

Data Processing

The data was processed qualitatively and quantitatively. The qualitative assessment was done in two stages viz. using the conventional approach and a new approach. In the first stage, the conventional approach was built with the assumption of four collinear electrodes. Figure 5 depicts a typical scenario of four electrodes used to acquire the measurement at a time. Considering the four collinear electrodes (blue circles), the geometric factor

for Wenner array is thus defined as in Equation 4 and was used throughout for the transformation. The conventional approach (ABEM Terrameter SAS4000 utilities software) was used to transform all the acquired data to the desired output format such as ABEM multi-purpose format (*.AMP) and Res2Dinv format (*.DAT) for further processing and analysis. Thereafter, Res2Dinv Geotomo software was used to perform inversion of the transformed resistivity data and Surfer 13 Golden software was later used to enhance the visualization of the inverse resistivity models. Considering the relative homogeneous nature of the study area, smoothness-constrained least-square inversion method was used (Loke et al., 2003; Loke & Barker, 1996). The iterative inversion process was set at a maximum of five (5) iteration. For each iteration, the inversion tries to reduce the root mean squared (RMS) error in an effort to find a better model (Martínez-Moreno et al., 2013). However, a lower RMS error does not always equate to a better model, as more iteration tend to overfit the data.

$$K_{WA} = 2\pi \frac{1}{\left(\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AN} + \frac{1}{BN}\right)} \quad (4)$$

where AM , BM , and AN , BN represent the geometrical distance. The four inline electrodes are collinear, and the distance AM , MN and BN are all equal $AM = MN = BN = a$.

In the second stage, the new approach developed on Jupyter notebook environment using Python programming language was used to transform all the acquired data to Res2Dinv format (*.DAT) for further processing and analysis. Subsequent process is the same as in stage one above. The new approach was designed to account for the electrodes offline on a 2-D ERT survey. For instance, in Figure 5 above, consider electrode N is shifted perpendicular to a new position say X_n ($AM \neq MX_n \neq BX_n \neq a$), the theoretically assumed collinearity of electrodes in the conventional program is distorted, which in turn,

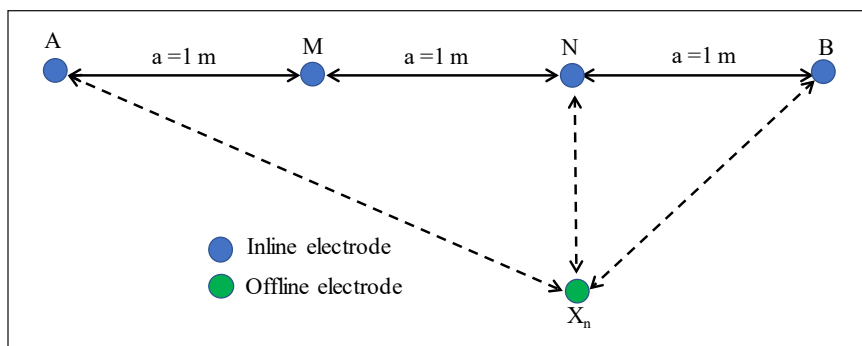


Figure 5. Typical scenario of four electrodes for measurement using Wenner array with electrodes inline (blue circle) and offline (green circle) arrangement

affects the K factor and, consequently, the apparent resistivity. Therefore, the geometric factor (Equation 4) defined above cannot hold. Hence, it must be modified as in Equation 5 and embedded in the template of the new approach. Basically, the new approach is a pre-processing template which enables the user to key in the offline electrodes and the offline distances thereby automatically recalculate the geometric factor and the apparent resistivity of the data points associated with the offline electrodes.

$$K_{WA_i} = 2\pi \frac{1}{\left(\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AX_n} + \frac{1}{BX_n}\right)} \quad (5)$$

$$AX_n = [(2a)^2 + (NX_n)^2]^{1/2}$$

$$BX_n = [(2a)^2 + (NX_n)^2]^{1/2}$$

Where $i = 0.1, 0.2, 0.3, \dots$ and $n = 1, 2, 3, \dots$

For the quantitative assessment, Pearson correlation coefficient (R) for the two phases was computed. The correlation coefficient (R) provides a simple measure of the degree of linear correlation between two images (Abdullah *et al.*, 2018). Its value ranges from -1 to +1. An identical or close comparison of the desired target is obtained for R-value approaching +1. R can be computed using Equation 6.

$$R = \frac{N(\sum \rho_t \rho_i) - (\sum \rho_t)(\sum \rho_i)}{\sqrt{\{N \sum (\rho_t)^2 - (\sum \rho_t)^2\} \{N \sum (\rho_i)^2 - (\sum \rho_i)^2\}}} \quad (6)$$

where; ρ_t is the resistivity of the true model and ρ_i is the resistivity of the estimated models, N is the total number of data points.

RESULTS AND DISCUSSION

The result was presented qualitatively and quantitatively. The qualitative assessments involved the visual evaluation of the inverse resistivity contour plots produced by the conventional program and new approach for 2-D ERT using Wenner array with all electrodes inline (true model), one offline electrode, two offline electrodes, and three offline electrodes (estimated models). A comparison was made between the resistivity contour plots produced using the conventional program and the new approach. The measured resistivity obtained for all the profiles varied from 0 to 4000 Ω -m. Six tomograms were considered in each case for the estimated model's (survey with one, two, and three electrode(s) offline). This became necessary in order to limit the number of displayed tomograms for better

visualisation and comparison. Therefore, tomograms with even number (0.2, 0.4, 0.6, 0.8, 1.0, 1.2 m) offline distances were considered. For the quantitative aspect, Pearson's correlation coefficient (R) was assessed.

Wenner Array with All Electrodes Inline (WA-AEI)

Figure 6 shows the result obtained for 2-D ERT Wenner array survey with all multi-electrodes inline using the conventional program. The inverse resistivity contour plot obtained after 5 iteration has RMS error of 8.7%. It indicated high resistivity zone ($> 900 \Omega\text{-m}$) within 0 to 4 m depth and low resistivity zone ($< 900 \Omega\text{-m}$) at depths greater than 4 m, which is in conformity with the findings of (Muhammad & Saad, 2018). The high resistivity zone defines the dry top residual soil which extends from the surface to a thickness of 4 m, whereas, the low resistivity zone corresponds to the residual soil saturated with water (Abdulrahman et al., 2016). It is important to note that only the conventional program was used to process the WA-AEI data because the theoretical assumption of collinear electrode pairs was achieved. Therefore, the inverse contour plot obtained for WA-AEI serves as the true model for which the other survey profiles with electrodes offline are compared to observe the effect of the electrodes offline.

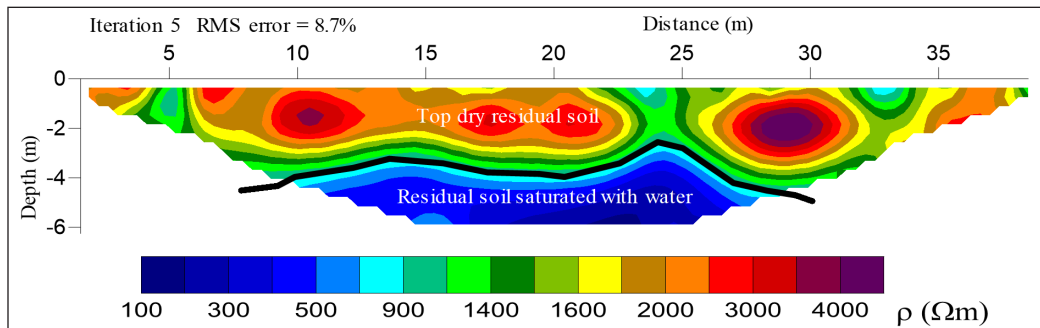


Figure 6. Two-dimensional resistivity contour plot for the true model

Wenner Array with One Electrode Offline (WA-1EO)

Prior to the inversion of the data, the percentage changes in the apparent resistivity due to the offline electrodes for the conventional and new approaches for WA-1EO, WA-2EO and WA-3EO is presented in Table 1. It is observed that the percentage change is directly proportional to the offline electrodes distance.

Figure 7 depicts juxtaposed two-dimensional tomograms for the true model and resistivity profiles with one electrode been offline at 0.2 m interval using the conventional approach and a newly developed approach. All the tomograms show fairly good result after five iteration of the inversion with RMS error ranging from 8.7% to 9.0%. They depict high resistivity zone ($> 900 \Omega\text{-m}$) at a depth of 0 to 4 m and low resistivity ($< 900 \Omega\text{-m}$) at a

Table 1

Summary of percentage change in apparent resistivity due to offline electrode for WA-1EO, WA-2EO and WA-3EO using conventional and new approaches

Offline electrode distance (m)	% change in apparent resistivity		
	WA-1ES	WA-2ES	WA-3ES
0.10	0.08	0.09	0.12
0.20	0.33	0.34	0.45
0.30	0.71	0.74	1.02
0.40	1.22	1.97	1.83
0.50	1.82	2.29	2.61
0.60	2.50	3.01	3.52
0.70	3.22	4.58	4.66
0.80	3.98	4.09	5.33
0.90	4.75	5.20	6.86
1.00	5.53	6.64	7.47
1.10	6.32	7.99	9.08
1.20	7.09	8.07	9.74

depth greater than 4 m. The tomograms produced using the conventional approach for the models with one offline electrode nearly match the true model except for a few instances (Figure 7e-g) that shows a very slight variation of anomalies. This close match of the tomograms is a consequence of the high signal-to-noise ratio of the array (Cubbage et al., 2017; Thabit & Khalid, 2016), whereas, the slight variation is likely due to the increase in offline electrode distance. However, the geometries of the true model are well resolved in all the tomograms obtained by using the new approach as shown in Figure 7b' - g'.

Wenner Array with Two Electrodes Offline (WA-2EO)

The juxtaposed tomographic images for the WA-AEI and WA-2EO at offline distances of 0.2, 0.4, 0.6, 0.8, 1.0, and 1.2 m using the conventional approach and newly developed approach are shown in Figure 8. All the tomograms obtained for the WA-2EO models using the conventional approach indicated high ($>900 \Omega\text{-m}$) and low ($<900 \Omega\text{-m}$) resistivity zones at depths below and above 4 m, respectively (Muhammad & Saad, 2018). The geometries of the true model are partially resolved in all the images of WA-2EO. It is observed that the tomograms (Figure 8b-c) give a close replica to the true model, while the remaining images (Figure 8d-g) produced by the survey (WA-2EO) with offline electrodes greater than $\frac{1}{2}$ of the minimum electrode spacing (a) indicates a slight variation of anomalies when compared to the true model (Figure 8a). This shows that the greater the offline electrode distance the more the effect. This outcome is contrary to that of Szalai et al. (2008), who found that offline electrodes effect on 2-D ERT result was insignificant. More so, the low resistivity zone or saturated zone along the distance (12 m - 27 m) at a depth of 4 m and

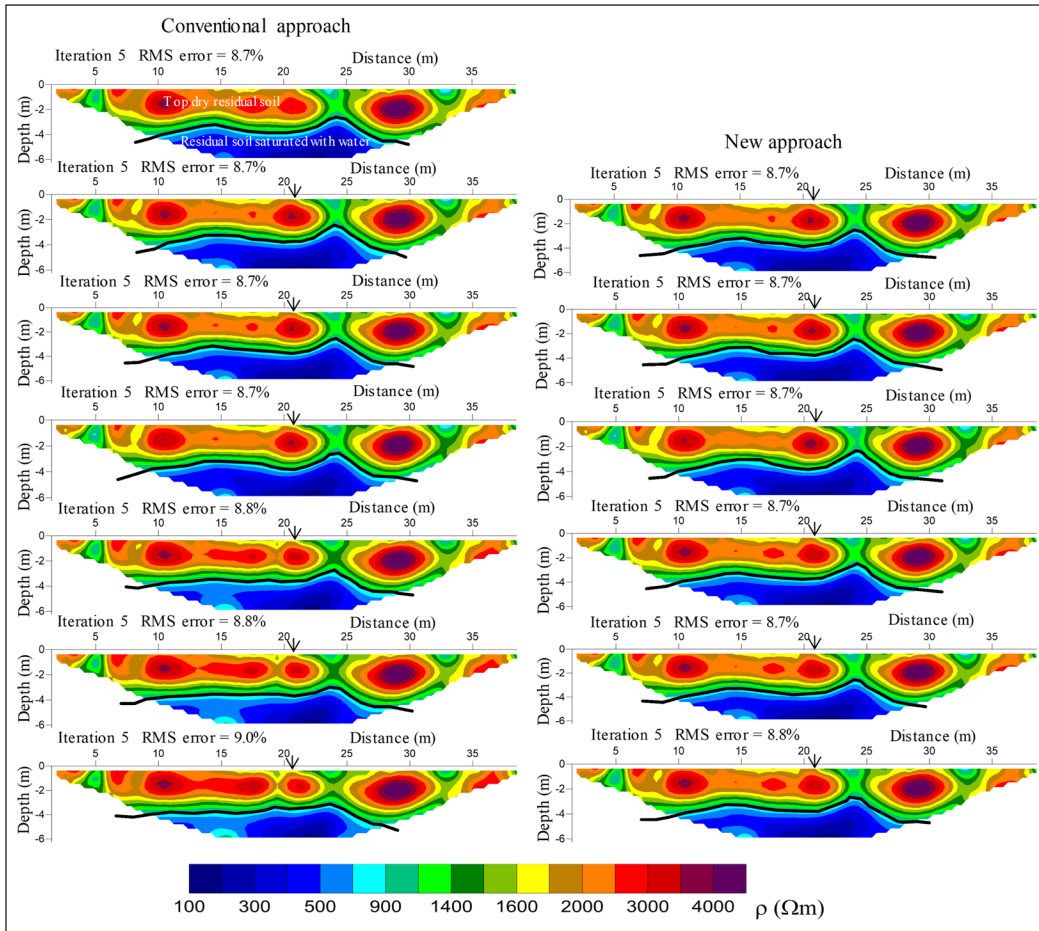


Figure 7. Two-dimensional resistivity contour plots comparison between (a) true model, and estimated models using conventional program and new approach for Wenner array with one electrode offline (WA-1EO) at (b/b') 0.2 m, (c/c') 0.4 m, (d/d') 0.6 m, (e/e') 0.8 m, (f/f') 1.0 m, (g/g') 1.2 m distances, respectively. The arrow sign indicates the position of the offline electrode

beyond (Figure 8a), decrease to 13 m - 24 m for WA- 2EO survey at 1.2 m offline electrodes distance (Figure 8g). Since the geology of the study area is relatively homogeneous, this variation of anomalies is attributed to the offline electrodes (3 and 35) and its corresponding distance of which the conventional approach cannot account for.

However, the limitation of the conventional approach which underestimated the anomaly for survey with electrodes offline has been demonstrated by the tomograms produced using the newly developed approach (Figure 8b'-g'). It is observed that the resistivity contour plots produced using the new approach fairly resolved the geometries of the true model than that of the conventional approach. This suggests that the newly developed approach can mitigate the false anomaly posed by offline electrode on a 2-D ERT survey.

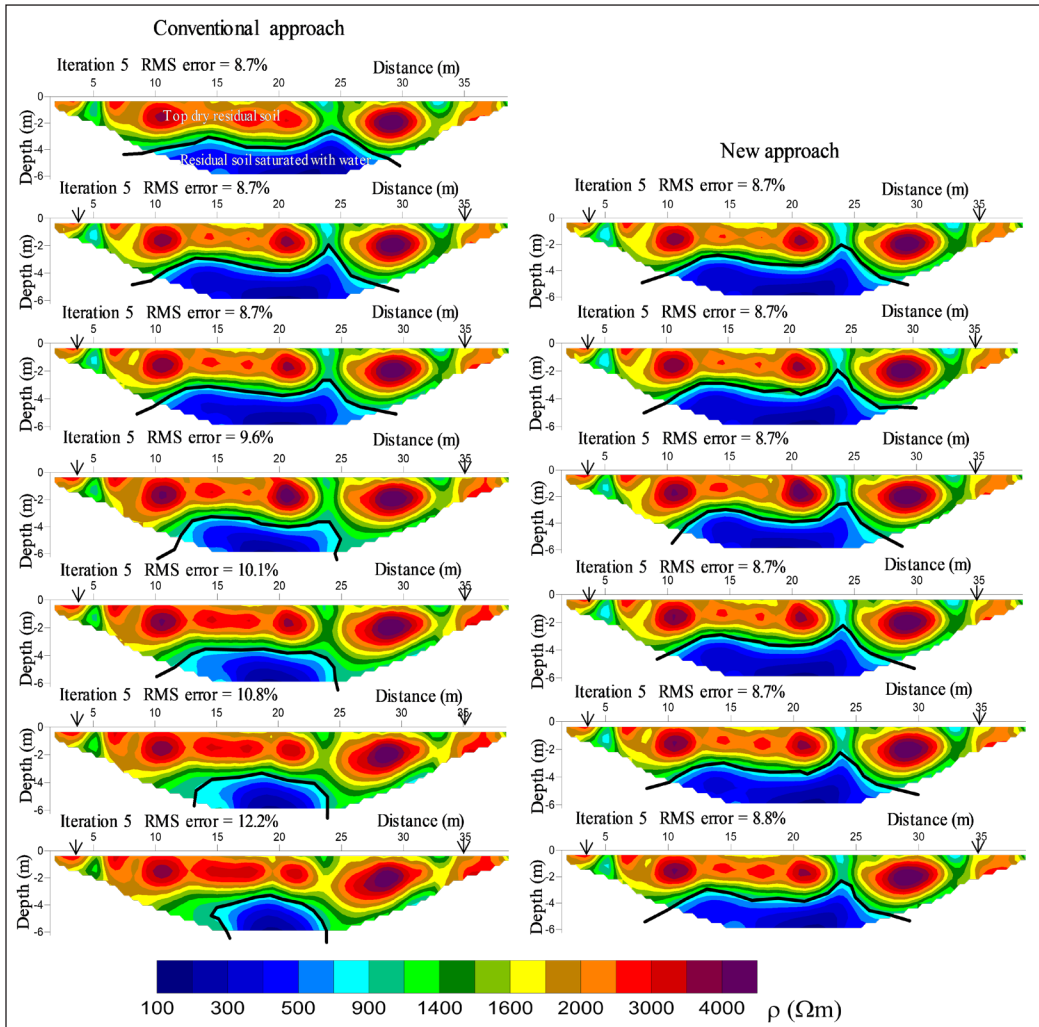


Figure 8. Two-dimensional resistivity contour plots comparison between (a) true model, and estimated models using conventional program and new approach for Wenner array with two electrodes offline (WA-2EO) at (b/b') 0.2 m, (c/c') 0.4 m, (d/d') 0.6 m, (e/e') 0.8 m, (f/f') 1.0 m, (g/g') 1.2 m distances, respectively. The arrow sign indicates the position of the offline electrodes

Wenner Array with Three Electrodes Offline (WA-3EO)

Figure 9 shows the tomographic images for the true model and resistivity profiles taken with three electrodes offline (WA-3EO) at distances of 0.2, 0.4, 0.6, 0.8, 1.0, and 1.2 m using the conventional approach and the newly developed approach. The range of the resistivity values obtained is similar to those of WA-1EO and WA-2EO tomograms. However, the tomograms obtained for WA-3EO after the predefine iteration, has high RMS error (10.1% to 18.2%) particularly those obtained using the conventional approach. These set of tomograms also show distinct variation of anomalies when compared to the true resistivity

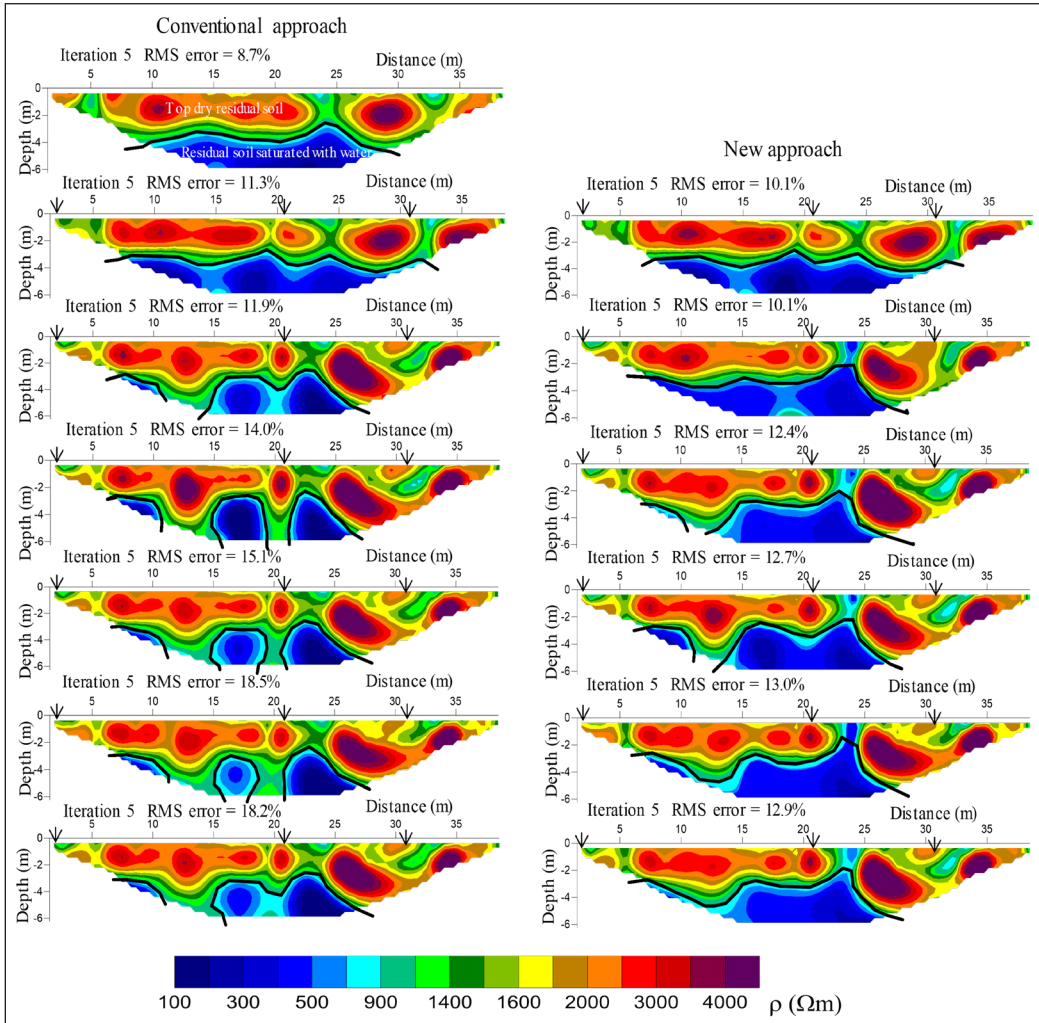


Figure 9. Two-dimensional resistivity contour plots comparison between (a) true model, and estimated models for Wenner array using conventional program and new approach with three electrodes offline (WA-3EO) at (b/b') 0.2 m, (c/c') 0.4 m, (d/d') 0.6 m, (e/e') 0.8 m, (f/f') 1.0 m, (g/g') 1.2 m distances respectively. The arrow sign indicates the position of the offline electrodes

model. For instance, the highly saturated zone ($< 900 \Omega\text{-m}$) for the true model (Figure 9a) which span from a distance of 12 m – 27 m at the bottom part of the tomogram (depth > 4 m) appears to split into two for almost all the tomograms (Figure 9b-g). This is probably due to the distance of the electrodes offline and a large number of data point (one-quarter of the total dataset) associated with the three electrodes offline, as the geology of the study area is relatively homogeneous. This outcome differs from the findings of Szalai et al. (2008) and Zhou and Dahlin (2003), who found that offline electrodes had insignificant effect on electrical resistivity tomography result.

Figure 9b'-g' depicts the tomograms obtained for Wenner array with three electrodes offline using the newly developed approach. The tomograms illustrate some level of match when compared to the true model (Figure 9a). However, it does not resolve the geometries of the true model well, as it present a slanted high resistivity section that extent from 24 m to 37 m contrary to the true model. This is likely due to the large number (26%) of contaminated data point involved. Comparing the tomograms produced by both approaches as juxtaposed in Figure 9, illustrated the effectiveness of the new approach in resolving the geometries of the true model. Hence, using the new approach in pre-processing a resistivity survey data with electrodes offline could significantly enhance the result, as it reduces the false anomalies introduced due to the electrodes offline.

In the quantitative assessment, the variation of the estimated models (WA-1EO, WA-2EO, and WA-3EO) from the true model (WA-AEI) was carried out. Pearson's correlation coefficient (R) of the estimated models using the conventional program and new approach was computed as presented in Tables 2, 3 and 4. The R-values obtained for processing with the new approach was appreciably high (0.957 – 0.963, 0.932 – 0.962 and 0.470 – 0.753 for WA-1EO, WA-2EO, and WA-3EO models respectively) as compared to that obtained for processing with conventional program (Table 2 - 4). This signifies that 95% – 96%, 93% - 96%, and 47% - 75% of correlation is achieved between the true and the estimated models (WA-1EO, WA-2EO, and WA-3EO), respectively for the new approach. Based on the established fact, R value approaching +1 gives an identical or close approximation to the desired true model (Abdullah et al., 2018). Hence, the R value obtained for WA-1EO, WA-2EO, and WA-3EO using the new approach suggest proximity to the true model than

Table 2
 Summary of Pearson's correlation coefficient between the conventional program (Conv) and new approach (NApp) for Wenner array with one electrode offline (WA-1EO)

Number of profiles	Offline electrode distance (m)	Pearson correlation coefficient	
		WA-1EO-Conv	WA-1EO-NApp
1	0.1	0.9627	0.9628
2	0.2	0.9613	0.9623
3	0.3	0.9593	0.9627
4	0.4	0.9579	0.9634
5	0.5	0.9527	0.9630
6	0.6	0.9346	0.9575
7	0.7	0.9257	0.9595
8	0.8	0.9108	0.9583
9	0.9	0.9023	0.9601
10	1.0	0.8953	0.9612
11	1.1	0.8772	0.9596
12	1.2	0.8693	0.9599

using the conventional approach. This is clearly demonstrated in the graph for each table as shown in Figure 10. It is observed that as the offline electrode(s) distance increases beyond half the minimum electrode spacing, the R-value decreases, particularly when using the conventional program and for WA-3EO (Figure 10a-c). Therefore, the quantitative assessment has further revealed the efficacy of the new approach.

Table 3

Summary of Pearson's correlation coefficient between the conventional program (Conv) and new approach (NApp) for Wenner array with two electrodes offline (WA-2EO)

Number of profiles	Offline electrode distance (m)	Pearson correlation coefficient	
		WA-2EO-Conv	WA-2EO-NApp
1	0.1	0.9598	0.9598
2	0.2	0.9618	0.9620
3	0.3	0.9600	0.9537
4	0.4	0.9615	0.9625
5	0.5	0.9428	0.9447
6	0.6	0.8932	0.9374
7	0.7	0.8802	0.9345
8	0.8	0.8824	0.9332
9	0.9	0.8689	0.9336
10	1.0	0.8511	0.9326
11	1.1	0.8441	0.9343
12	1.2	0.7912	0.9349

Table 4

Summary of Pearson's correlation coefficient between the conventional program (Conv) and new approach (NApp) for Wenner array with three electrodes offline

Number of profiles	Offline electrode distance (m)	Pearson correlation coefficient	
		WA-3EO-Conv	WA-3EO-NApp
1	0.1	0.7247	0.7535
2	0.2	0.7083	0.7449
3	0.3	0.7118	0.7469
4	0.4	0.7227	0.7467
5	0.5	0.6938	0.7076
6	0.6	0.6088	0.7008
7	0.7	0.5959	0.6756
8	0.8	0.5092	0.6341
9	0.9	0.4300	0.6039
10	1.0	0.3336	0.5872
11	1.1	0.2805	0.4901
12	1.2	0.2133	0.4706

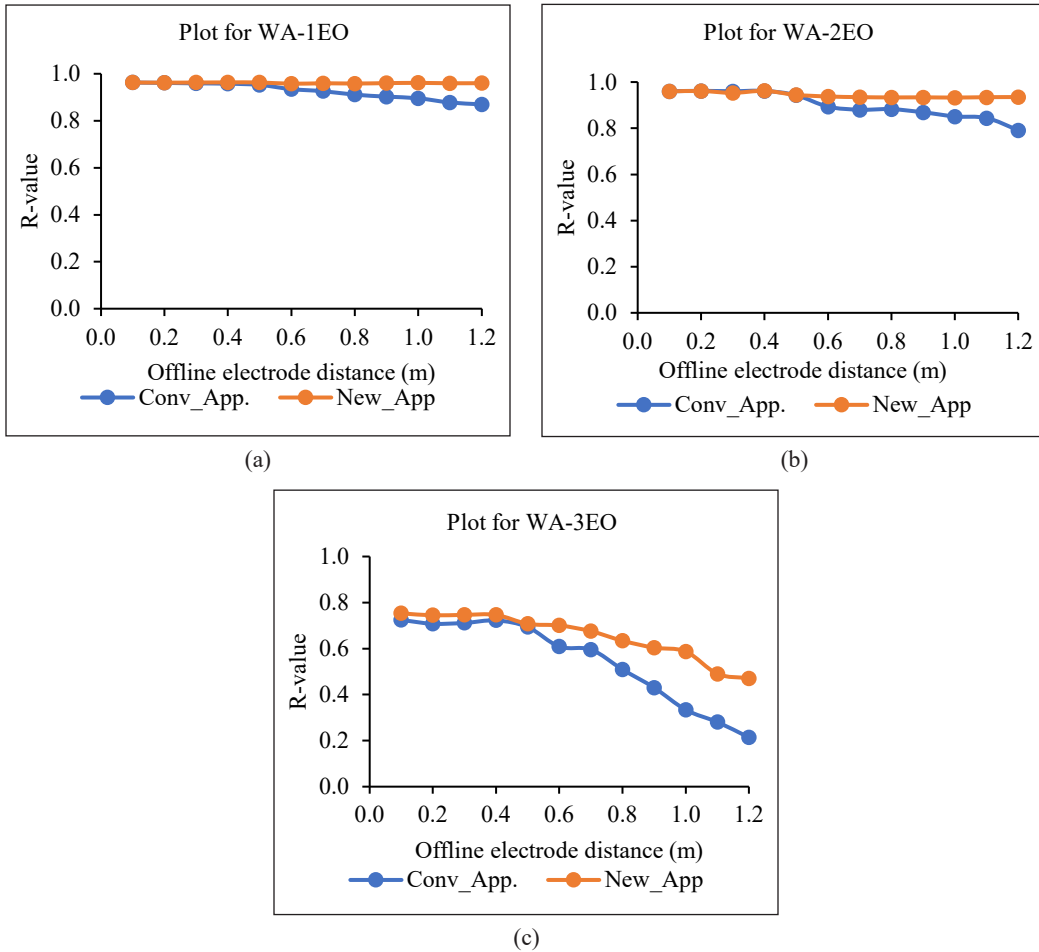


Figure 10. Graphs showing the variation of Pearson's correlation with offline electrode distance for (a) WA-1EO (b) WA-2EO and (c) WA-3EO

CONCLUSION

In this study, we proposed a new approach to enhance a 2-D ERT survey result by mitigating the effect of electrodes offline. The resistivity contour plots assessment performed between the true (WA-1EO) and the estimated models (WA-1EO, WA-2EO and WA-3EO) using the new approach, revealed insignificant variation of anomalies for all the estimated models. More so, most of the R values for the new approach is approaching +1 which demonstrated a close comparison to the true model. However, using the conventional program, the resistivity contour plots produced for WA-2EO and WA-3EO surveys showed slight to distinct variation of anomalies and the R value was relatively low (20%) as compared to that of the new approach >90%. This variation of anomalies obtained using the conventional program could misguide the environmental and engineering experts in their structural

design, thereby leading to collapse of structures, loss of lives and properties. Therefore, the study has proven the efficacy of the new approach in mitigating the effect posed by offline electrodes on a two-dimensional electrical resistivity tomography result.

ACKNOWLEDGEMENT

We would like to thank the technical staff and all our colleagues, Geophysics Section, School of Physics, Universiti Sains Malaysia (USM), for their help during the data acquisition. The authors wish to thank USM for providing the Fundamental Research Grant Scheme entitled Development of 2-D Linear Inversion Algorithm from Geophysical Approach for Soil or Rock Characteristics (203/PFIZIK/6711663). The first author expresses appreciation to the Federal University of Lafia and Tertiary Education Trust Fund (TetFund), Nigeria, for financial support.

REFERENCES

- Abdullah, F. M., Loke, M. H., Nawawi, M., & Abdullah, K. (2018). Assessing the reliability and performance of optimized and conventional resistivity arrays for shallow subsurface investigations. *Journal of Applied Geophysics*, 155, 237-245. <https://doi.org/10.1016/j.jappgeo.2018.06.018>
- Abdulrahman, A., Nawawi, M., Saad, R., Abu-Rizaiza, A. S., Yusoff, M. S., Khalil, A. E., & Ishola, K. S. (2016). Characterization of active and closed landfill sites using 2D resistivity/IP imaging: Case studies in Penang, Malaysia. *Environmental Earth Sciences*, 75(4), Article 347. <https://doi.org/10.1007/s12665-015-5003-5>
- Abidin, M. H. Z., Madun, A., Tajudin, S. A. A., & Ishak, M. F. (2017). Forensic assessment on near surface landslide using electrical resistivity imaging (ERI) at Kenyir Lake area in Terengganu, Malaysia. *Procedia Engineering*, 171, 434-444. <https://doi.org/10.1016/j.proeng.2017.01.354>
- Abudeif, A. M., Mohammed, M. A., Fat-Helbary, R. E., El-Khashab, H. M., & Masoud, M. M. (2020). Integration of 2D geoelectrical resistivity imaging and boreholes as rapid tools for geotechnical characterization of construction sites: A case study of New Akhmim city, Sohag, Egypt. *Journal of African Earth Sciences*, 163, Article 103734. <https://doi.org/10.1016/j.jafrearsci.2019.103734>
- Ahmad, F., Yahaya, A. S., Farooqi, M. A., & Tebal, N. (2006). Characterization and geotechnical properties of penang residual soils with emphasis on landslides. *American Journal of Environmental Sciences*, 2(4), 121-128.
- Ali, M. M., Ahmad, F., Yahaya, A. S., & Farooqi, M. A. (2011). Characterization and hazard study of two areas of Penang Island, Malaysia. *Human and Ecological Risk Assessment: An International Journal*, 17(4), 915-922. <https://doi.org/10.1080/10807039.2011.588156>
- Anastasopoulos, I. (2013). Building damage during nearby construction: Forensic analysis. *Engineering Failure Analysis*, 34, 252-267. <https://doi.org/10.1016/j.engfailanal.2013.08.003>
- Cosenza, P., Marmet, E., Rejiba, F., Cui, Y. J., Tabbagh, A., & Charlery, Y. (2006). Correlations between geotechnical and electrical data: A case study at Garchy in France. *Journal of Applied Geophysics*, 60(3-4), 165-178. <https://doi.org/10.1016/j.jappgeo.2006.02.003>

- Cubbage, B., Noonan, G. E., & Rucker, D. F. (2017). A modified Wenner array for efficient use of eight-channel resistivity meters. *Pure and Applied Geophysics*, 174(7), 2705-2718. <https://doi.org/10.1007/s00024-017-1535-9>
- Dahlin, T., & Zhou, B. (2004). A numerical comparison of 2D resistivity imaging with 10 electrode arrays. *Geophysical Prospecting*, 52(5), 379-398. <https://doi.org/10.1111/j.1365-2478.2004.00423.x>
- Fragaszy, R. J., Santamarina, J. C., Amekudzi, A., Assimaki, D., Bachus, R., Burns, S. E., Cha, M., Cho, G. C., Cortes, D. D., Dai, S., Espinoza, D. N., Garrow, L., Huang, H., Jang, J., Jung, J. W., Kim, S., Kurtis, K., Lee, C., Pasten, C., ... & Tsouris, C. (2011). Sustainable development and energy geotechnology - Potential roles for geotechnical engineering. *KSCE Journal of Civil Engineering*, 15(4), 611-621. <https://doi.org/10.1007/s12205-011-0102-7>
- Godio, A., Strobbia, C., & De Bacco, G. (2006). Geophysical characterisation of a rockslide in an alpine region. *Engineering Geology*, 83(1-3), 273-286. <https://doi.org/10.1016/j.enggeo.2005.06.034>
- Griffiths, D. H., & Barker, R. D. (1993). Two-dimensional resistivity imaging and modelling in areas of complex geology. *Journal of Applied Geophysics*, 29(3-4), 211-226. [https://doi.org/10.1016/0926-9851\(93\)90005-J](https://doi.org/10.1016/0926-9851(93)90005-J)
- Ingeman-Nielsen, T., Tomaškovičová, S., & Dahlin, T. (2016). Effect of electrode shape on grounding resistances - Part 1: The focus-one protocol. *Geophysics*, 81(1), WA159-WA167. <https://doi.org/10.1190/GEO2015-0484.1>
- Loke, M. H. (2018). *Tutorial: 2-D and 3-D electrical imaging surveys*. Geotomosoft.
- Loke, M. H., Acworth, I., & Dahlin, T. (2003). A comparison of smooth and blocky inversion methods in 2D electrical imaging surveys. *Exploration Geophysics*, 34(3), 182-187. <https://doi.org/10.1071/EG03182>
- Loke, M. H., & Barker, R. D. (1996). Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method1. *Geophysical Prospecting*, 44(1), 131-152. <https://doi.org/10.1111/j.1365-2478.1996.tb00142.x>
- Loke, M. H., Chambers, J. E., Rucker, D. F., Kuras, O., & Wilkinson, P. B. (2013). Recent developments in the direct-current geoelectrical imaging method. *Journal of Applied Geophysics*, 95, 135-156. <https://doi.org/10.1016/j.jappgeo.2013.02.017>
- Martínez-Moreno, F. J., Pedrera, A., Ruano, P., Galindo-Zaldívar, J., Martos-Rosillo, S., González-Castillo, L., Sánchez-Úbeda, J. P., & Marín-Lechado, C. (2013). Combined microgravity, electrical resistivity tomography and induced polarization to detect deeply buried caves: Algaidilla cave (Southern Spain). *Engineering Geology*, 162, 67-78. <https://doi.org/10.1016/j.enggeo.2013.05.008>
- Muhammad, S. B., & Saad, R. (2018). Linear regression models for estimating true subsurface resistivity from apparent resistivity data. *Journal of Earth System Science*, 127(5), Article 64. <https://doi.org/10.1007/s12040-018-0970-z>
- Pando, L., Pulgar, J. A., & Gutiérrez-Claverol, M. (2013). A case of man-induced ground subsidence and building settlement related to karstified gypsum (Oviedo, NW Spain). *Environmental Earth Sciences*, 68(2), 507-519. <https://doi.org/10.1007/s12665-012-1755-3>

- Pradhan, B., & Lee, S. (2010). Delineation of landslide hazard areas on Penang Island, Malaysia, by using frequency ratio, logistic regression, and artificial neural network models. *Environmental Earth Sciences*, 60(5), 1037-1054. <https://doi.org/10.1007/s12665-009-0245-8>
- Rucker, D. F., Crook, N., Glaser, D., & Loke, M. H. (2012). Pilot-scale field validation of the long electrode electrical resistivity tomography method. *Geophysical Prospecting*, 60(6), 1150-1166. <https://doi.org/10.1111/j.1365-2478.2011.01039.x>
- Szalai, S., Koppán, A., & Szarka, L. (2008). Effect of positional inaccuracies on multielectrode results. *Acta Geodaetica et Geophysica Hungarica*, 43(1), 33-42. <https://doi.org/10.1556/AGeod.43.2008.1.3>
- Tan, B. K. (1994). Engineering properties of granitic soils and rocks of Penang Island, Malaysia. *Bulletin of the Geological Society of Malaysia*, 35(July), 69-77. <https://doi.org/10.7186/bgsm35199408>
- Thabit, J. M., & Khalid, F. H. (2016). Resistivity imaging survey to delineate subsurface seepage of hydrocarbon contaminated water at Karbala Governorate, Iraq. *Environmental Earth Sciences*, 75(1), Article 87. <https://doi.org/10.1007/s12665-015-4880-y>
- Uhlemann, S., Chambers, J., Falck, W., Tirado Alonso, A., Fernández González, J., & Espín de Gea, A. (2018). Applying electrical resistivity tomography in ornamental stone mining: Challenges and solutions. *Minerals*, 8(11), Article 491. <https://doi.org/10.3390/min8110491>
- Zhou, B., & Dahlin, T. (2003). Properties and effects of measurement errors on 2D resistivity imaging surveying. *Near Surface Geophysics*, 1(3), 105-117. <https://doi.org/10.3997/1873-0604.2003001>



Potential Groundwater Exploration in Use of 2-D Electrical Resistivity Tomography (ERT) Techniques at the Department of Agriculture Kelantan Research and Developmental Platform Padang Raja Kelantan

Wan Fazilah Fazlil Ilahi^{1*}, Nur Hidayu Abu Hassan¹, Mohd Razi Ismail¹, Nik Norasma Che'Ya¹, Zulkarami Berahim¹, Mohamad Husni Omar¹, Nurul Idayu Zakaria¹ and Mohamed Azwan Mohamed Zawawi²

¹Department of Agriculture Technology, Faculty of Agriculture, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

²Department of Biological and Agricultural Engineering, Faculty of Engineering, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia

ABSTRACT

2-D electrical resistivity has been a proper investigation survey for determination of subsurface geophysical in describing the complex features geology profile. In this study, an electrical resistivity survey was conducted at paddy cultivation area located in Melor, Kelantan, Malaysia. Since the end plot of paddy field experiences water scarcity especially during dry season, there is a need to find other alternative water source. The study was conducted on 1st and 2nd February 2020 to identify zone area of groundwater

for Melor, Kelantan. Four resistivity lines using Induced Polarization (IP) and 2-D Electrical Resistivity Imaging Technique were conducted using a set of ABEM Terrameter SAS4000. Short resistivity survey was applied to gained detail of subsurface formation near the ground, while the longer resistivity survey was applied to obtain deeper subsurface delineation. Measured data obtained was analyzed using RES2DINV software and result of contrast resistivity values was used to determine the geological structures, while the chargeability values were analyzed

ARTICLE INFO

Article history:

Received: 5 January 2021

Accepted: 10 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.28>

E-mail addresses:

wanfazilah@upm.edu.my (Wan Fazilah Fazlil Ilahi)

hidayuhassan22@yahoo.com (Nur Hidayu Abu Hassan)

razi@upm.edu.my (Mohd Razi Ismail)

niknorasma@upm.edu.my (Nik Norasma Che'Ya)

zulkarami@upm.edu.my (Zulkarami Berahim)

mohdhusni@upm.edu.my (Mohamad Husni Omar)

idayu.nz@gmail.com (Nurul Idayu Zakaria)

mohdazwan@upm.edu.my (Mohamed Azwan Mohamed Zawawi)

*Corresponding author

accordingly to determine area of expected potential groundwater zone. Results from the resistivity profiles show a range values of 0 Ω m to 50,000 Ω m with total maximum acquired depth of 65.6 m below ground surface. The chargeability profiles show a range values of 0 msec to 500 msec, that shows potential of groundwater zone area lies at 0 to 4 msec. It was found that at a depth of 60 to 75 m, 30 m from center of Profile B was suitable for a production well which was expected to be a potential area for groundwater zone.

Keywords: 2-D electrical resistivity technique, chargeability, geological structures, groundwater, induced polarization, resistivity

INTRODUCTION

To date, worldwide is facing critical issues regarding surface water limitations especially in agriculture industry with approximately 70% of water usage used for agriculture purposes. As water is crucial to achieve a steady healthy crop, any consumption insufficiency at any crop growth stage may result inhibition of crop growth. Although Malaysia receives approximately 2,774 mm in average of annual rainfall that is considered high and more than the global average, however, the rainfall distributions patterns are uneven throughout the country causing some part of areas receiving low rainfall distributions compared to other areas (Hussin et al., 2020). Thus, may lead to changes in rainfall pattern and vegetation process of planted crop (Mohtar et al., 2014).

Rice as staple food for Malaysian has become one of the major crops planted all over the country. Conventional practice that uses continuous flooding system, has once been used to raise paddy crop, however when farmers encounter drought season, this practice can no longer be accessed as standing depth of almost 10 to 15 cm is hard to maintain because of water limitations. Therefore, rather than solely depending on surface water, alternative water supply especially with low cost yet sustainable is needed to be investigated to fulfill water demand for various purposes.

One of the major alternative water sources is groundwater. However, groundwater in Malaysia is yet to be exploited on a bigger scale (Azizan et al., 2018). Studies need to be properly done to investigate complex behavior of subsurface geophysical delineation to extract water from potential groundwater zone area. Hence, in line with technological development, resistivity survey has been frequently introduced in use of determination of subsurface delineation. The resistivity imaging profile specifically shows geological formation of survey line that indicates suitable layers that have the potential of groundwater (Hazreek et al., 2017). Therefore, this study was proposed to determine the zone area of potential groundwater using 2-D electrical resistivity technique and Induced Polarization (IP).

MATERIAL AND METHODS

Description of the Study Area

Melor, Kelantan which geographically lies at about 5° 57' N and 102° 17' E was chosen as the study area. Current main crop cultivation at the area is dominant by paddy crop. Since paddy is a high water-intensive crop, water scarcity is a major problem experienced during the dry season, especially the end plot of the paddy field which located far from the water sources. Hussin et al., (2020) reported Kelantan received less rainfall between April to October with the annual rainfall distribution of 2774 mm annually.

Generally, the area is covered by alluvium deposits of Quaternary age. The sediment quaternary can be divided into three major units: Gula Formation, Beruas Formation and Simpang Formation (Hutchison & Tan, 2009). The quaternary deposits that occupy the north of Kelantan river valley mainly consists of unconsolidated to semi consolidated gravel, sand, clay, and silts. The silty and clay deposit which lies at 13 m to 15 m is basically at recent age (Zakaria, 1972), while the thickness of alluvium may reach up to more than 200 m towards the seas (Suratman, 1997). Same pattern is shown by the unconsolidated deposit to form thick wedge towards the coastal that overlays granite and sediment (Stauffer, 1973).

Geographically, study area major potential zone for groundwater is at northern Kelantan as supported in Geological Map of Kelantan at Figure 1. Primarily, the area is underlain by the Quaternary alluvium. The alluvium region covers an area of 1500 km² that is 10% of the state area and mainly underlain by granitic and sedimentary/ metasedimentary bedrock, and the latter layer can be phyllite, slate, sandstone, or shale. The sedimentary/metasedimentary rocks which consist of phyllite, slate, shale and sandstone occur generally towards the western parts along the Kelantan River, while the granitic rocks exhibit existence at the eastern part (Pour & Hashim, 2017). Figure 1 presents the location of the study area at paddy cultivation, Melor Kelantan.

Data Acquisition

A 2-D electrical resistivity survey was conducted at the study area by measuring the potential of electrical between two electrodes that were injected with direct current by another pair of electrodes (Drahor et al., 2006; Osazuwa & Chii, 2010). The resistivity data and induced polarization were measured and automatically recorded by resistivity meter. RES2DINV software was used to processed measured data obtained by converting and distinguish the apparent resistivity data and true resistivity data using inversion method approach (Hazreek et al., 2017). The 2-D model separates the subsurface into several number of blocks and then produces the apparent resistivity which agrees with an actual measurement (Osazuwa & Chii, 2010). The generated contrast resistivity values are used to determine the subsurface geological structures, while the chargeability values are analyzed accordingly to determine area of expected potential groundwater zone (Rahmawati et al.,

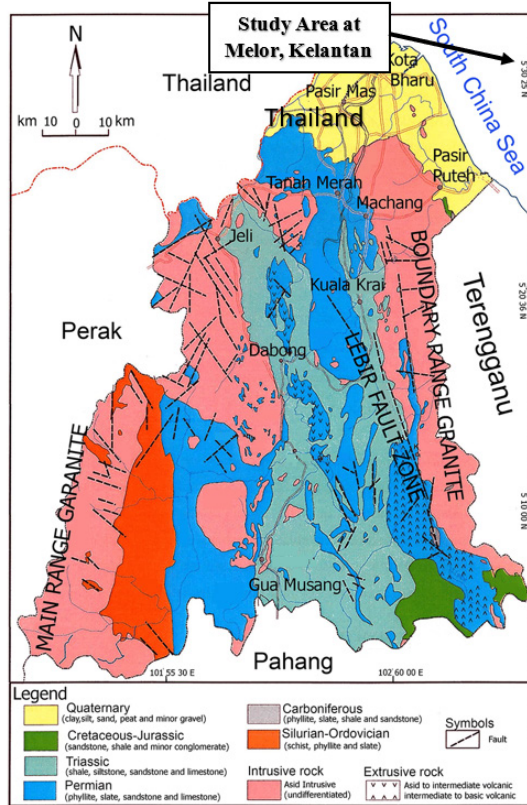
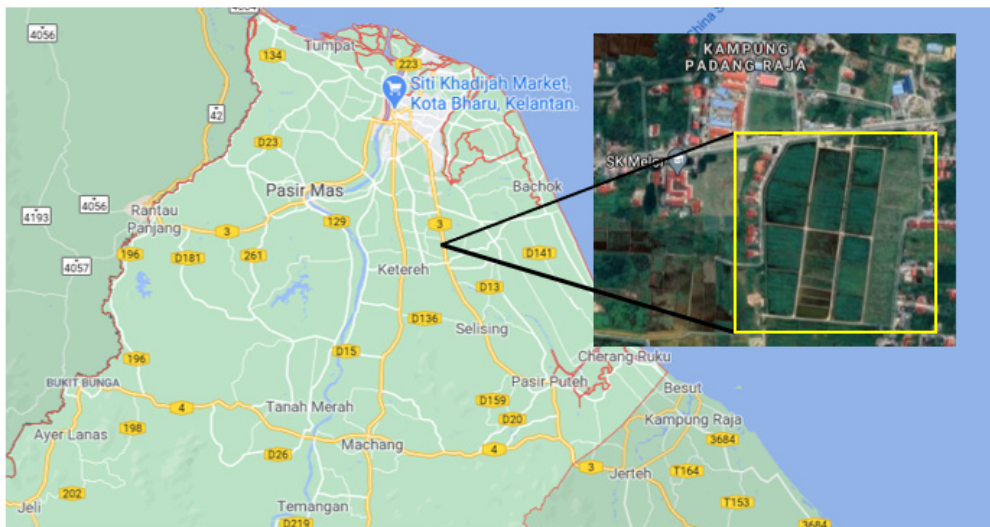


Figure 1. Geological map of Kelantan (Pour & Hashim, 2017)



Source: Google Map, 2021

Figure 2. Study area located at paddy cultivation field in Melor, Kelantan.

2018). Wenner-Schlumberger protocol was selected which basically an integration of Wenner and Schlumberger arrays for electrical resistivity technique (ERT). Resistivity surveys were carried out at four different location in the study area. Length of each lines varied from 200 – 400 m with electrode spacing of 2.5 m to 5 m. The exact position of each resistivity lines is shown in Figure 2. Shortest distance of resistivity survey was applied to obtain details of subsurface formation near the ground. Meanwhile longer distance of resistivity survey was conducted to gain deeper subsurface delineation.

In this study, a total of four resistivity lines using Induced Polarization (IP) and 2-D Electrical Resistivity Imaging Technique (ERT) were proposed with varied length range from 200 m to 400 m at random location to give better description of the study area. To control the induction of current and the electrodes potential readings, a set of ABEM Terrameter SAS4000 and a switcher unit was used and connected by multicore-cable along the survey line.

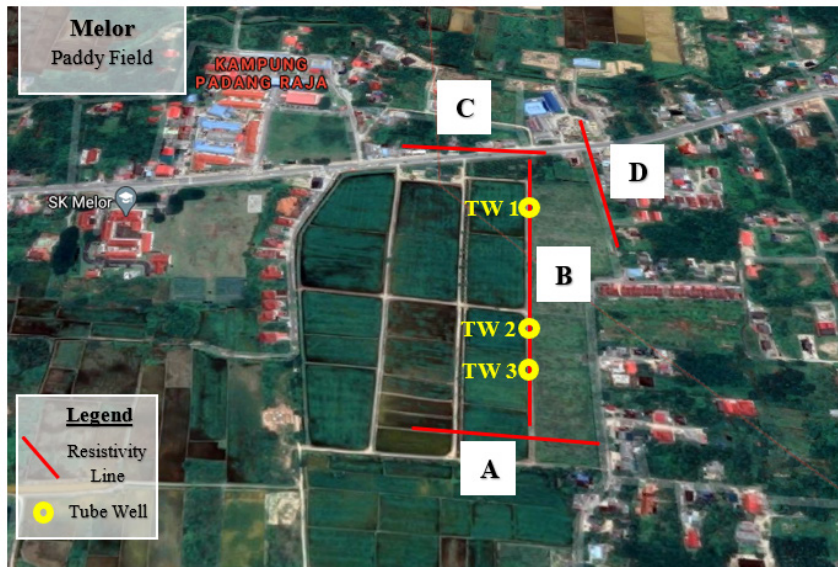


Figure 3. The location of the survey lines (Line A - D)

RESULTS AND DISCUSSIONS

In Figure 2, several tube wells can be seen (TW1, TW2 and TW3) owned by *Jabatan Pertanian Kelantan*, however, these tube wells were built using traditional methods thus no detail report was available.

Figure 4 to 7 represent the respective electrical resistivity and chargeability tomography at four different profiles (Profile A to D). The resistivity values obtained ranged from 0 Ωm to 50,000 Ωm while the chargeability values ranged from 0 msec to 500 msec. The maximum depth that had been acquired was 65.6 meters below ground surface for line B.

Profile A shows several zones with low resistivity with high chargeability that indicated presence of clay. While at high resistivity with low chargeability, there can be assumed to have a potential water zone, however since it is just around area of water table, it cannot be assumed as groundwater zone. Profile B indicates the presence of expected groundwater zone area at fractured zone appearing in high resistivity with low chargeability zone. 'JP' in this Profile B shows several tubes wells owned by *Jabatan Pertanian Negeri Kelantan* which had been built before. However, no detailed report was available from boring procedure as it was built by traditional method. 'New' indicates the new proposed tube well recommended by analysis from Profile B. At low resistivity with low chargeability in Profile C, clay is dominating the zone at that specific area. While at low resistivity and high chargeability, weathered rock where high impurities area is considered to cover up the zone area. At low resistivity with high chargeability in Profile D, it may indicate the present of clay. However, at high resistivity with high chargeability, it can be presumed that the area is a highly impurities area.

Results show the study area mainly consists of unconsolidated materials with resistivity value ranging from 10 to 80 Ω m. The resistivity reading and the colour indicate soil availability for the layer consists of variation of clay, silt, sand, and gravel. This layer was classified as alluvium that characterized the deposit of clay, silt, sand, and gravel. The resistivity value for alluvium is commonly ranging from 10 to 800 ohm.m. The lowest resistivity reading is ranging from 1 to 10 ohm.m and is represented by the dark blue colour. The colours specify the type of soil at this depth was soft clay or clay. High resistivity value ranging from 2000 Ω m to 50,000 Ω m portrays expected availability of granitic rock (Pour & Hashim, 2017). The granitic rock that underlay the layers of alluvial was at 40 m deep.

The measured chargeability is the discharged measure of polarization in the subsurface. Basically, it is related to the water content in the subsurface media and the minerals conductivity (Riwayat et al., 2018). Hence, the chargeability profile could be used to determine the characteristic of water either it has impurities or not. The profile, however, still needs to be supported by further water quality analysis to determine the groundwater quality.

It can be seen from the profile, that the chargeability distribution is ranging from very low chargeability reading; 0 msec to very high chargeability reading; 500 msec. Chargeability tomogram can indicate potential groundwater zone. It has been proven, that the groundwater occurrence is indicated with chargeability value ranging from 0 to 4 msec (Juanah et al., 2013). Thus, it was considered chargeability value of 0 to 4 msec specify the potential groundwater zone.

From the chargeability profile, high chargeability reading ranging from 4 msec to 100 msec represented by colour scheme between yellow to red shows highly contaminated area. Among the sources of groundwater contamination, may be agricultural pollutants

such as fertilizers and pesticides from the use in paddy field (Jeyaruba & Thushyanthy, 2009; Vandermaesen et al., 2016)

Commonly, fractures in subsurface indicate the present of groundwater. The lower the resistivity values may exhibit the greater the fractures are (Kim et al., 2011). From Profile B, it can be assumed that the fractured zone is a zone area with high potential groundwater source. The production tube well is recommended at 30 m from the centre line (Profile B) with depth of 60 – 75 m below ground surface.

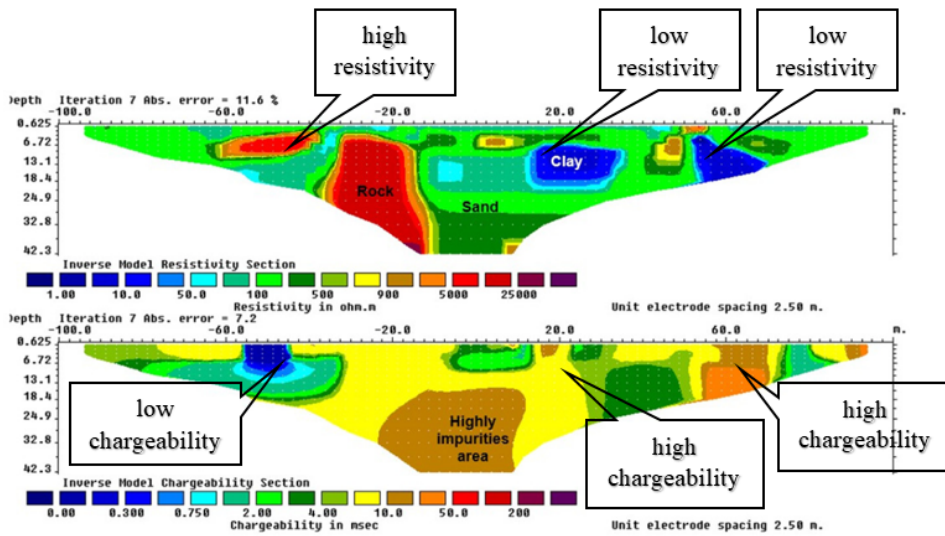
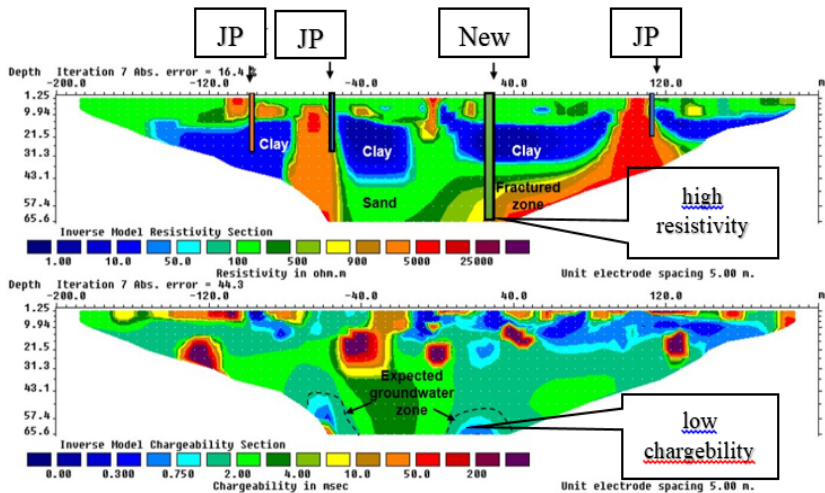


Figure 4. Electrical resistivity and chargeability tomography at Profile A (resistivity survey line of 200m)



**JP = *Jabatan Pertanian* Existing Tube Well; New = New Propose Tube Well

Figure 5. Electrical resistivity and chargeability tomography at Profile B (resistivity survey line of 400m)

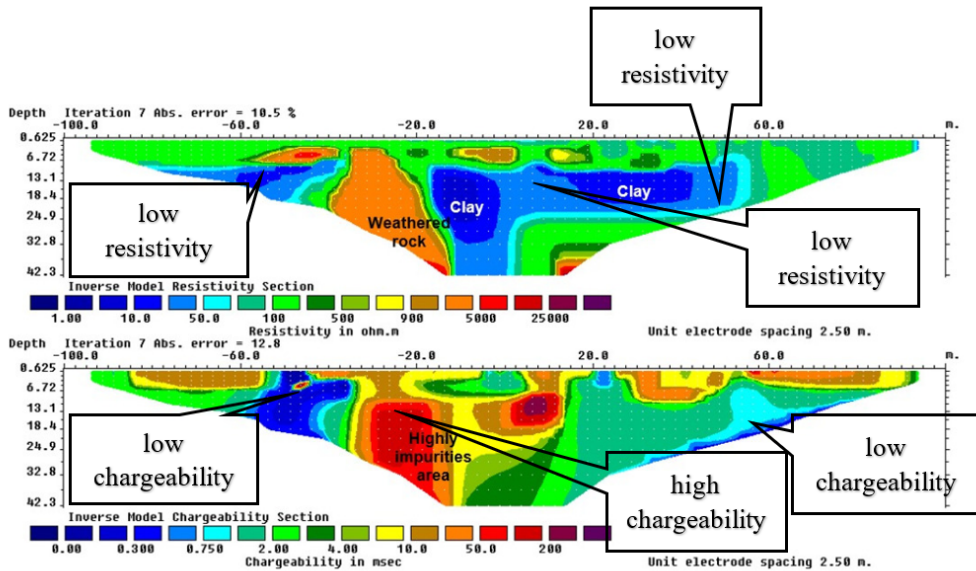


Figure 6. Electrical resistivity and chargeability tomography at Profile C (resistivity survey line of 200m)

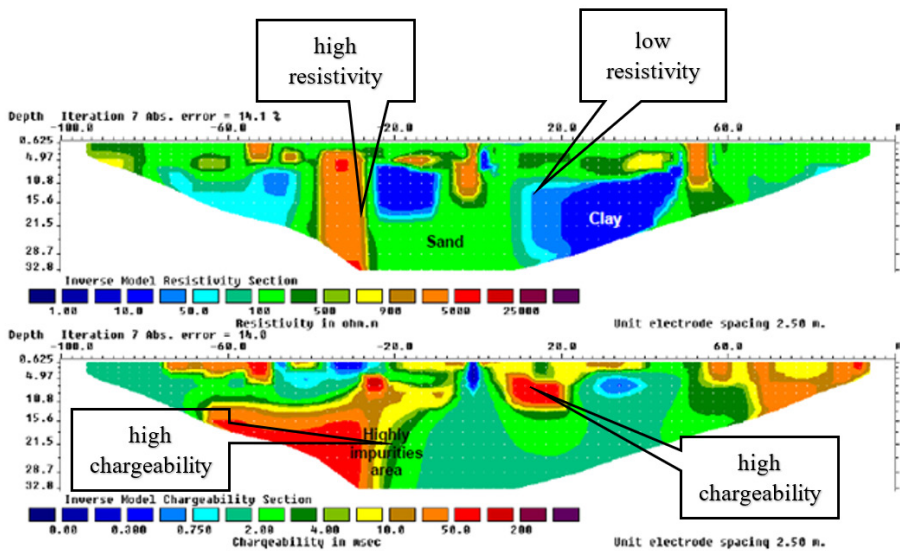


Figure 7. Electrical resistivity and chargeability tomography at Profile D (resistivity survey line of 200m).

CONCLUSIONS

The aquifer within study area has been identified as a fractured aquifer with the resistivity value ranging from 10 to 80 Ω m and chargeability value ranging from 0 to 4 msec. The

groundwater type in Melor, Kelantan is considered as unconfined aquifer. Thus, the suggested location for production well is located at 30 m from the centre line (Profile B) and at depth of 60 – 75 m beneath ground surface. The suggested area is considered as the possible location of potential groundwater zone. However, further study needs to be done to confirm the subsurface delineation, lithology borehole log and further test in terms of quality of potential groundwater zone.

ACKNOWLEDGEMENT

The authors would like to thank Ministry of Higher Education Malaysia for providing research fund from Long term Research Grant Food Security Niche Area LRGS/1/2019/UPM/01/2 Development of climate ready rice for sustaining food security in Malaysia and also to Department of Agriculture, Kelantan for allowing us to use their platform by LRGS Climate Ready Rice.

REFERENCES

- Azizan, F. A., Aznan, A. A., Ruslan, R., Nazari, M., & Jaafar, M. N. (2018). Groundwater assessment using geophysical survey at Insat, Perlis, Malaysia. *IOP Conference Series: Materials Science and Engineering* (Vol. 429, No. 1). Institute of Physics Publishing. <https://doi.org/10.1088/1757-899x/429/1/012026>
- Drahor, M. G., Göktürkler, G., Berge, M. A., & Kurtulmuş, T. Ö. (2006). Application of electrical resistivity tomography technique for investigation of landslides: A case from Turkey. *Environmental Geology*, 50(2), 147-155. <https://doi.org/10.1007/s00254-006-0194-4>
- Google Maps. (2021). *Google Maps directions in Melor, Kelantan*. Retrieved February 10, 2021, from <https://www.google.com.my/maps/place/Kelantan>
- Hazreek, Z. A. M., Raqib, A. G. A., Aziman, M., Azhar, A. T. S., Khaidir, A. T. M., Fairus, Y. M., Rosli, S., Fakhurrizi, I. M., & Izzaty, R. A. (2017). Preliminary Groundwater Assessment using Electrical Method at Quaternary Deposits Area. *IOP Conference Series: Materials Science and Engineering* (Vol. 226, No. 1). IOP Publishing. <https://doi.org/10.1088/1757-899x/226/1/012042>
- Hussin, N. H., Yusoff, I., & Raksmei, M. (2020). Comparison of applications to evaluate groundwater recharge at lower Kelantan River Basin, Malaysia. *Geosciences*, 10(8), Article 289. <https://doi.org/10.3390/geosciences10080289>
- Hutchison, C. S., & Tan, D. N. K. (Eds.). (2009). *Geology of peninsular Malaysia*. Published jointly by the University of Malaya and the Geological Society of Malaysia.
- Jeyaruba, T., & Thushyanthy, M. (2009). The effect of agriculture on quality of groundwater: A case study. *Middle-East Journal of Scientific Research*, 4(2), 110-114.
- Juanah, M. S., Ibrahim, S., Sulaiman, W. N. A., & Latif, P. A. (2013). Groundwater resources assessment using integrated geophysical techniques in the southwestern region of Peninsular Malaysia. *Arabian Journal of Geosciences*, 6(11), 4129-4144. <https://doi.org/10.1007/s12517-012-0700-9>

- Kim, M. I., Kim, J. S., Kim, N. W., & Jeong, G. C. (2011). Surface geophysical investigations of landslide at the Wiri area in Southeastern Korea. *Environmental Earth Sciences*, 63(5), 999-1009. <https://doi.org/10.1007/s12665-010-0776-z>
- Mohtar, Z. A., Yahaya, A. S., Ahmad, F., Suri, S., & Halim, M. H. (2014). Trends for daily rainfall in northern and southern region of peninsular Malaysia. *Journal of Civil Engineering Research*, 4(3A), 222-227.
- Osazuwa, I. B., & Chii, E. C. (2010). Two-dimensional electrical resistivity survey around the periphery of an artificial lake in the Precambrian basement complex of northern Nigeria. *International Journal of Physical Sciences*, 5(3), 238-245.
- Pour, A. B., & Hashim, M. (2017). Application of Landsat-8 and ALOS-2 data for structural and landslide hazard mapping in Kelantan, Malaysia. *Natural Hazards and Earth System Sciences*, 17(7), 1285-1303. <https://doi.org/10.5194/nhess-17-1285-2017>
- Rahmawati, D. R., Supriyadi, A. N., & Naufal, M. A. (2018). Groundwater potential prediction by using geoelectricity method a case study in Simpang Lima and around it. In *Journal of Physics: Conference Series* (Vol. 983). IOP Publishing. <https://doi.org/10.1088/1742-6596/983/1/012003>
- Riwayat, A. I., Nazri, M. A. A., & Abidin, M. H. Z. (2018). Application of electrical resistivity method (ERM) in groundwater exploration. In *Journal of Physics: Conference Series* (Vol. 995, No. 1, p. 012094). IOP Publishing. <https://doi.org/10.1088/1742-6596/995/1/012094>
- Stauffer, P. H. (1973). Kenny hill formation. In D. J. Gobbett, & C. S. Hutchison (Eds.), *Geology of the Malay Peninsula* (pp. 87-91). Wiley.
- Suratman, S. (1997). Groundwater protection in Kelantan. *Warta Geologi*, 23(6), 351.
- Vandermaesen, J., Horemans, B., Bers, K., Vandermeeren, P., Herrmann, S., Sekhar, A., Seuntjens, P., & Springael, D. (2016). Application of biodegradation in mitigating and remediating pesticide contamination of freshwater resources: state of the art and challenges for optimization. *Applied Microbiology and Biotechnology*, 100(17), 7361-7376. <https://doi.org/10.1007/s00253-016-7709-z>
- Zakaria, A. S. (1972). Morphometry of part of Kelantan River catchmen. *Sains Malaysiana*, 1(1), 59-76.

Efficacy of Intelligent Mosquito System (I.M.O.S) with Xmos Mini Aerosol Against *Aedes* in 17th College, Universiti Putra Malaysia

Latifah Saiful Yazan¹, Banulata Gopalsamy^{1*}, Siti Najiha Abu Bakar¹, Khairul Aiman Manan¹, Noranis Shahida Shahidan² and Lee Yean Wang²

¹Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia

²One Team Networks Sdn. Bhd. 1-1B, Block 6, Jalan Pahat 15/H, Section 15, 40200 Shah Alam, Selangor, Malaysia

ABSTRACT

This study reports the efficacy of I.M.O.S (Intelligent Mosquito System) by using Xmos mini aerosol in reducing *Aedes* mosquito population in 17th College, Universiti Putra Malaysia (UPM). Prior to the experiment, the *Aedes* mosquito population was determined in all blocks of the 17th College, UPM. The I.M.O.S. was installed above the entrance door of the hall and two rooms and was set to automatically spray at 6.30 am and 4.00 pm every day. No intervention was used in the control house. Adult efficacy study was conducted by placing 20 *Aedes* mosquitoes in each cage and was hung at a distance of 10 feet from the

I.M.O.S. The mortality caused by I.M.O.S throughout the exposed two hours and after 24 hours of exposure was recorded. The number mosquito eggs population were calculated throughout the installation of I.M.O.S. Data were analysed using two-way ANOVA and paired T-test, respectively. The mean number of *Aedes* mosquito eggs and ovitraps index showed no significant difference ($p > 0.05$) between control and treatment blocks. There were significant differences ($p < 0.05$) in the percentage of knockdown of adult *Aedes* mosquitoes (10, 20, 30, 60 and 120 minutes after exposure) and the mortality of adult *Aedes* mosquitoes

ARTICLE INFO

Article history:

Received: 21 October 2020

Accepted: 4 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.29>

E-mail addresses:

latifahsy@upm.edu.my (Latifah Saiful Yazan)

banulatagopalsamy@gmail.com, banulatagopalsamy@gmail.com

(Banulata Gopalsamy)

sitinajihaabubakar97@gmail.com (Siti Najiha Abu Bakar)

khai.manan@yahoo.com (Khairul Aiman Manan)

anisshahida11@gmail.com (Noranis Shahida Shahidan)

leeyeanwang@gmail.com (Lee Yean Wang)

*Corresponding author

after 24 hours exposure. Nano and slow-release technology of the I.M.O.S with Xmos mini aerosol proved that this type of intervention can kill adult *Aedes* mosquitoes. Thus, it is a potential intervention for vector control and management.

Keywords: *Aedes* mosquito; aerosol; dengue fever; dengue virus; intelligent mosquito system; Ovitrap

INTRODUCTION

Dengue fever (DF) is an endemic disease that critically affects the subtropical and tropical regions of the world. Female *Aedes* mosquitoes are the only known vector that can transmit dengue virus (DENV) and cause DF. DENV is transmitted and spread specifically by *Aedes aegypti* and *Aedes albopictus* (Medeiros et al., 2018). DENV is a single stranded RNA virus of Flaviviridae family. The flavivirus circulates in the blood of infected person for two to three days and develops symptoms such as sudden onset fever, severe flu, severe headache, nausea, vomiting, swollen glands, retro orbital pain, muscle, and joint pains (Ratini, 2019). People with a second or subsequent dengue infection as well as those with weakened immune systems are at higher risk for developing severe form of viral illness which are dengue hemorrhagic fever (DHF) and dengue shock syndrome (DSS) (Ratini, 2019).

According to the research that was done before 1970, there were only total of nine countries that was affected with dengue epidemics. But the disease today is widespread in over 100 countries including Western Pacific, South East Asia, Eastern Mediterranean, America and Africa. However, the most affected countries are in South East Asia and Western Pacific. Dengue cases in Malaysia has risen since the first major epidemic in 1973 (Wallace et al., 1980). In Malaysia, iDenggi (2019) reported 57,920 dengue cases from 29 December 2019 until 9 July 2020. To add, the cumulative number of death due to dengue infection from January until July were 94 deaths (iDengue, 2019). It clearly showed that demographic and social developments such as population development, urbanization and modern transportation significantly affected the increase in number of dengue cases. Higher infectious incidence due to different virus serotypes raised the rate of genetic modification in viruses. As a result, it will increase the chances of DENV genotypes with higher severity of the DF (Gubler, 2002).

Due to the failure of vector management, the increasing intensity of dengue has given more opportunity for scientist to develop the dengue vaccines, hence making it effective tetravalent dengue vaccine for global public health. For several decades, the dengue vaccine development has been in progress but the complex pathology of the illness, the importance to control four virus serotypes and low in investment by vaccine developers have inhibited the process (Guzman et al., 2010). While waiting for a treatment or cure for this disease

in terms of new tools of vaccines, antiviral drugs, and improved diagnostic to be found, it is of utmost importance that the spread of dengue is well-controlled and well-managed.

Generally, the most common vector management is categorised as physical, biological, chemical, and integrated vector management (combination of more than two interventions) (Bouzid et al., 2016). Physical controls consist of regular cleaning of mosquito breeding sites which are artificial containers, cover of the containers and ovitraps (Lima et al., 2015). Biological control involves the use natural predators such as Copepods (Lazora et al., 2015), single or multiple species larvivorous fish (Han et al., 2015) in mosquitoes' breeding sites. Furthermore, bacterium such as *Bacillus thuringiensis israelensis* (Bti) that develops toxic proteins resulting in high larval mortality after ingestion (Boyce et al., 2013). However, the practical side of this control remains highly questionable.

Insecticide spraying, insecticide treated curtains, nets, and screens, fogging and larvicide application are forms of chemical controls (Bouzid et al., 2016). Wilson and colleagues reported that materials treated with insecticide could minimize the spread of disease but recorded low *Ae. aegypti* mortality rates demonstrated substantial resistance to insecticides, which significantly reduces the efficacy of this form of control measure (Wilson et al., 2014).

Mosquito populations around the world are widespread and abundant despite decades of vector control programs (Marcondes & Ximenes, 2016). However, different types of interventions, resources, period of study will affect the effectiveness of any vector control program which may explain in part of varying degree of success between studies. On the other hand, due to poor reporting of the study design, observational methodologies, heterogeneity and indirect results, the quality of data falls within the range of low and very low (Bouzid et al., 2016).

As practiced in Malaysia, dengue vector control activities are strongly dependent on human resources. In fact, a strong and large human workforce is needed to carry out the numerous district-level activities of dengue vector management, surveillance, and prevention (Packierisamy et al., 2015). In addition, trained healthcare that performs inspections, fogging and larvaciding activities is one of the pillars of the national control activities in preventing the rise of dengue cases. Frontliners including medical doctors and entomologist provide technical support in this issue. As a result, human resources represent the most important community to combat against *Aedes* mosquitoes (Packierisamy et al., 2015). Fitzpatrick et al. (2017) reported six countries with middle-income economies and high number of dengue cases accounts for an estimated 15% of the global dengue burden.

Hence, there is a need for implementation of a new intervention which is less laborious to reduce the *Aedes* mosquito population. Intelligent mosquito system (I.M.O.S) with Xmos mini aerosol is a product developed by One Team Network Sdn. Bhd. as an alternative to combat this issue. I.M.O.S is an automatic system that releases the required amount of the

aerosol content (Personal communication Mr. Lim Chee Hwa 30 June 2019) at the time where mosquitoes' activities are high during the early hours in the morning and in the evening before dusk (WHO, 2017). Xmos mini aerosol contains 0.76% of metofluthrin as active ingredient and is highly effective to kill adult mosquitoes. Metofluthrin is ideal to be used in numerous current source products such as coil of mosquito, as well as in modern products such as ventilator vaporizers and paper strips (Ujihara et al., 2004). Metofluthrin has higher vapor pressure and deliver macro-particles into air lasting up to eight hours (Personal communication Mr. Lim Chee Hwa 30 June 2019). Matsuo et al., (2005) explained that metofluthrin shows high effectiveness of knockdown especially towards mosquitoes as well as other insects. Moreover, this ingredient is highly volatile and has low toxicity towards mammals (Matsuo et al., 2005).

Therefore, this study was designed to investigate the effect of I.M.O.S with Xmos mini aerosol in reducing *Aedes* mosquito population in 17th College, Universiti Putra Malaysia (UPM) as well as the knockdown efficacy of adult *Aedes* mosquitoes.

MATERIALS AND METHOD

Study Area

The research was carried out at the 17th College of Universiti Putra Malaysia located in Serdang, Selangor as this study site records high number of dengue cases (Abdul et al., 2016). It consists of four residential blocks namely Block A, B, C and D. The accommodation design is 'Apartment Style' with four rooms per house and two students per room. Block A, B and C are accommodated by female students while Block D is for male students. There are four wings with five levels in each block. A Lake can be found about 50 meters from the entrance of the college. All blocks were used for the pre-treatment assessment. After ensuring the mosquito population in all blocks was similar, Block B and C were randomly selected as treatment block and control block, respectively.

Ovitrap

Ovitrap were provided by One Team Networks Sdn Bhd. A total of 160 of ovitraps with QR code were installed along the front and back staircase of each level of Block A, B, C and D for the pre-treatment assessment. Non-woven tissue (17.5 cm x 7.3 cm) was placed in each ovitrap for six days to allow the mosquitoes to oviposit. All the tissues were collected, and the eggs deposited on the tissues were counted to determine the mosquito population for every block. This procedure was carried out every month for two months (referred hereafter as Intervention 1 and 2) at the required blocks.

Intelligent Mosquito System (I.M.O.S) with Xmos Mini Aerosol

I.M.O.S with Xmos mini aerosol was kindly sponsored by One Team Networks Sdn Bhd (Figure 1). A total of fifteen I.M.O.S and 45 cans of Xmos mini aerosol were used in five selected houses during the intervention. Each house contains three I.M.O.S installed at three different areas which are hall, room 1 and room 2. Before the trial, all the Xmos mini aerosol were pre-weighed and the weight loss was recorded after assessment and a new Xmos mini aerosol was replaced after 45 days.



Figure 1: I.M.O.S with Xmos mini aerosol

Pre-treatment Assessment

During the pre-treatment assessment, 160 ovitraps were installed at the front and back staircase of each level. The ovitraps were attached to the stairs by using cable tie (8 inches). The mosquito population was determined by counting the eggs of mosquito on the substrate inside the ovitrap after six days. All the substrates from the ovitrap were air-dried at room temperature before placing them under a magnifying glass to manually calculate the number of mosquito eggs. Two blocks with the highest population were chosen for the treatment. The mean number of eggs and ovitrap index (OI) were calculated by using the following formulas.

$$\text{Mean number of eggs} = \frac{\text{Total number of eggs in ovitraps}}{\text{Total number ovitraps examined}} \times 100$$

$$\text{Ovitrap index, OI} = \frac{\text{No.of positive ovitrap (with eggs)}}{\text{No.of ovitrap deployed}} \times 100\%$$

Intervention Study

The size of the hall was 20.4 m² while room 1 and room 2 were 13.7 m² as shown in Figure 2. For the intervention study, Block B was chosen as treatment block while Block C was the control block. A total of fifteen I.M.O.S with Xmos mini aerosol were placed inside the five selected houses in Block B. In each of the treated house, three I.M.O.S with Xmos mini aerosol were installed separately in the hall, room 1 and room 2. Xmos mini aerosol was set to spray at two time points, 6.30 am and 4.00 pm. At each spraying time, six sprays were released in room 1 and 2 while nine sprays were released in the hall. There was no instalment of I.M.O.S in the control houses.

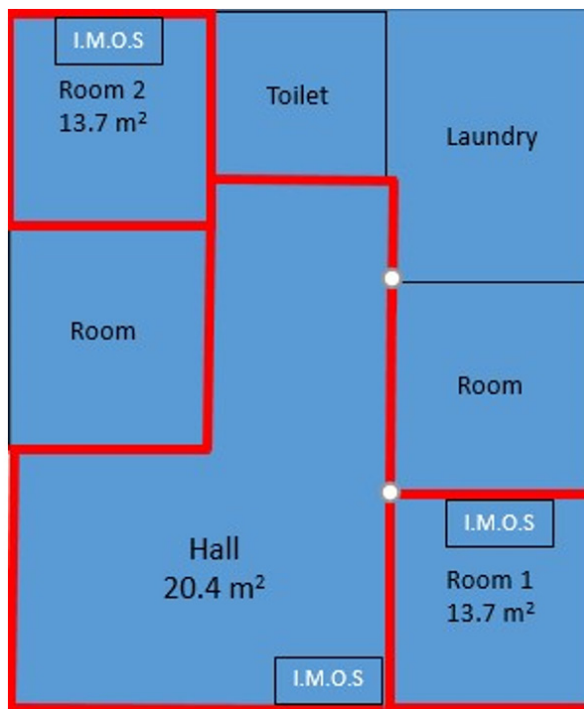


Figure 2: The size of hall, room 1 and room 2 at 17th College, UPM

Adult Efficacy Study

A total of 980 females non-gravid and non-blood fed, laboratory susceptible strain of *Ae. aegypti* were supplied in three batches of 360 mosquitoes each batch by Sumitomo Chemical Enviro- Agro Asia Pacific Sdn. Bhd. located in Senawang, Negeri Sembilan. Twenty *Ae. aegypti* adult female mosquitoes were placed in one net (10 cm x 17 cm x 32 cm). The nets were hung 10 feet (approximately 3.08 m) away from I.M.O.S in the five

treatment houses and control houses, with a total of three cages in each house (one in the hall, room 1 and room 2). Starting from 4:30 pm to 6:30 pm, the knockdown efficacy was calculated by measuring the number of knockdown mosquitoes in the net. After that, all nets were moved to the laboratory. The percentage of mortality after 24 hours of exposure was then calculated. This adult efficacy study was conducted thrice at an interval of 30 days during the intervention phase. The percentage of knockdown efficacy was calculated using the following formula:

$$\text{Percentage of knockdown efficacy} = \frac{\text{Number of mosquito knockdown}}{\text{Total number of mosquito in net}} \times 100$$

Temperature and Humidity

Three sets of humidity and temperature data logger from Temperature Technology of Australia (Model number: DS1921G) were used. The device was placed in the living room, room 1 and room 2 in the selected houses for hourly temperature and humidity recordings.

Data Analysis

Data were presented as mean \pm SE. Data were analyzed by using Statistical Package for Social Science (SPSS) Version 15. The number of mosquito eggs and ovitraps index data were compared among Block B and C using the Paired Sample T-test. Two-way ANOVA was used to compare the means on knockdown of adult *Aedes* mosquitoes between control and treatment blocks at different time points. The percentage of adult knockdown between room and hall and the percentage of adult *Aedes* mosquitoes after 24 hours were analysed by using unpaired T- test. A value was considered significant at $p \leq 0.05$.

RESULTS

The population of *Aedes* mosquito eggs (Figure 3) collected in 17th College during pre-treatment, intervention 1 and intervention 2 between control and treatment blocks were not significantly different ($p = 0.166$) with each other during the pre-treatment. During day 30 and 60 indicated as intervention 1 and 2, respectively, the mosquito population was not significantly ($p = 0.098$) different between the control and treatment blocks. The OI shows similar pattern with the mean number of eggs indicating that there was no significant difference ($p = 0.213$) between control and treatment blocks during pre-treatment, intervention 1 and intervention 2 as shown in Figure 4. The knockdown percentage of adult *Aedes* mosquitoes in the treatment block was significantly higher ($p < 0.05$) than the control block at all time points which were 10, 20, 30, 60 and 120 minutes of exposure (Figure 5). The mortality of adult *Aedes* mosquitoes after 24 hours of exposure was significantly

higher ($p=0.0383$) in the treatment group compared to the control block as represented in Figure 6. The percentage of knockdown of adult *Aedes* mosquitoes between room and hall at the treated areas in 17th College, UPM showed no difference ($p=0.705$) (Figure 7).

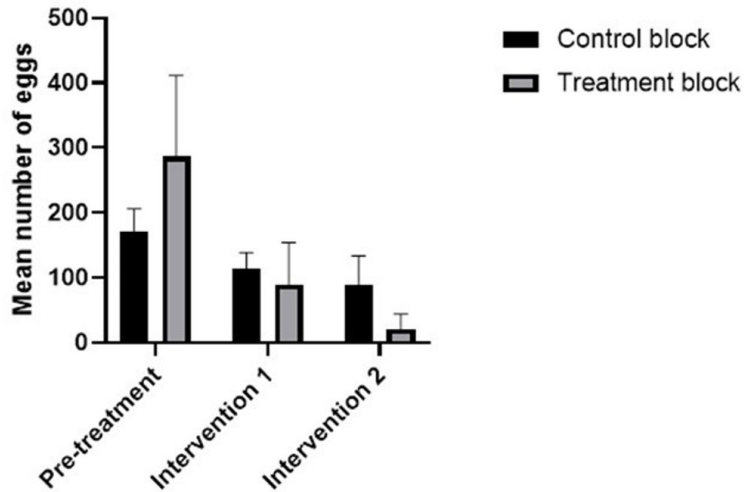


Figure 3: Number of eggs during pre-treatment, intervention 1 and intervention 2 between control and treatment blocks

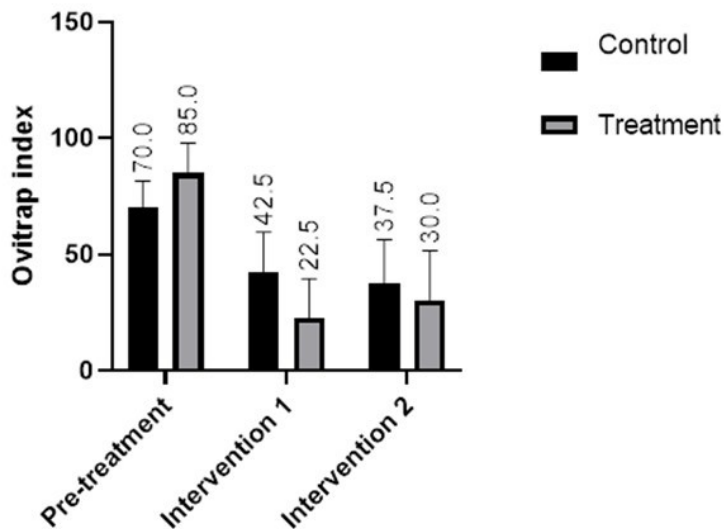


Figure 4: Ovitrap index (OI) during pre-treatment, intervention 1 and intervention 2 between control and treatment blocks

Xmos Mini Aerosol Effectively Knockdown *Aedes* Mosquitoes

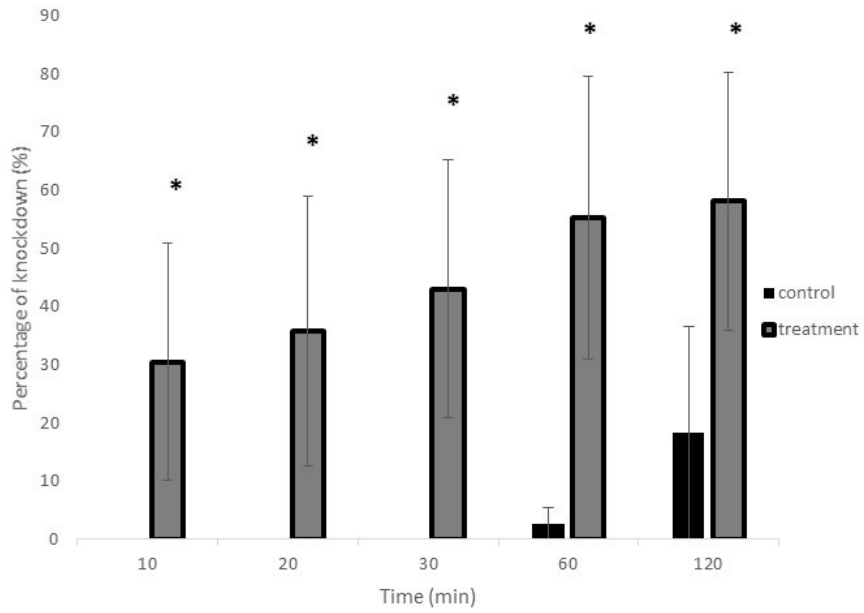


Figure 5: Percentage of knockdown at (10, 20, 30, 60 and 120 min) between control and treatment blocks. * significantly different at $p < 0.05$

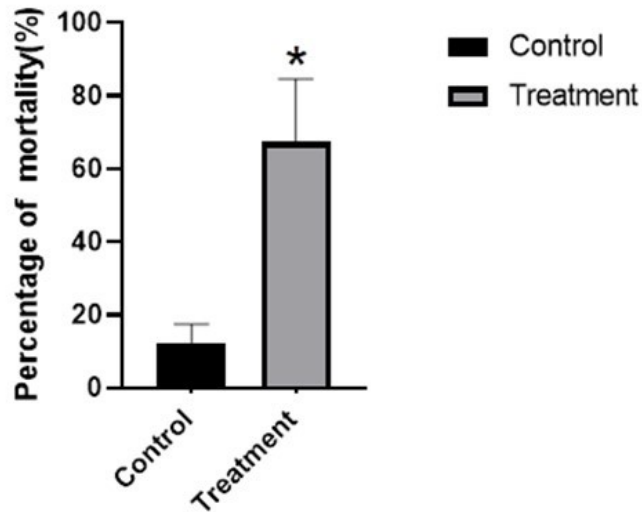


Figure 6: Percentage of mortality between control and treatment blocks after 24 hours of exposure. * significantly different at $p < 0.05$

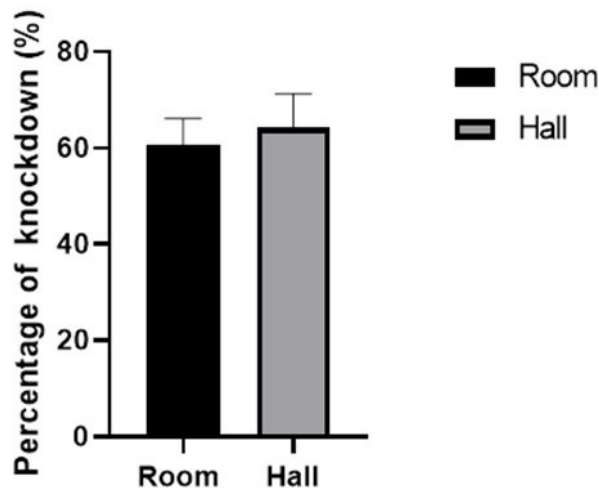


Figure 7: Percentage of knockdown of adult *Aedes* mosquitoes in the room hall area in the treatment block

DISCUSSION

Malaysia records frequent and consistent dengue cases throughout the years whereby the state of Selangor has the highest number of cases among all other states (iDenggi, 2019). Despite the intensive and extensive efforts by health and government-funded integrated vector control agencies, the outcome is still unsatisfactory as many new cases are still being reported. Therefore, this study focused on the use of insecticide, specifically I.M.O.S with Xmos mini aerosol as one possible method for vector control and management of *Aedes* mosquitoes.

The high population of *Aedes* mosquitoes and OI is consistent with our previous report by Latifah et al. (2020). This might be due to the presence of lakes in front of the 17th College and the presence of various containers, bottles, drains and students' pails in the study area that might provide ideal mosquito breeding sites. The use of I.M.O.S however did not effectively reduce the mosquito population, as the number and OI during the treatment months were still high. This is probably because the insecticide spraying was carried out inside the house repelling or killing mosquitoes that are present inside the house. Bibbs & Xue (2016) reported that the impact on mosquitoes was only at a very close proximity of approximately 0.3 m from the metofluthrin delivery system. The population of mosquitoes outside houses where ovitraps were installed were still high. We need to consider how long the efficacy of the treatment formulation lasts in the air when exposed against *Aedes* mosquitoes. *Ae. aegypti* shows an ability to escape from an open windows and doors (Moore et al., 2007). If this escape occurs, it will cause female *Aedes* mosquitoes to be able to oviposit as well as increases the eggs population.

I.M.O.S with Xmos mini aerosol is a chemical vector control that can kill adult mosquitoes where it contains 76% of metofluthrin as active ingredient (Matsuo et al., 2005). Previous study by Kawada et al. (2004) and Ujihara et al. (2004) reported multilayer paper strips impregnated with metofluthrin gave positive result as repellent against mosquitoes whether in field application or laboratory. Metofluthrin can reduce the biting activity of *Ae. aegypti* (Darbro et al., 2017). When the mosquitoes fly in the zone of metofluthrin, they will experience confusion leading to eventual knockdown of the mosquito (Buhagiar et al., 2017). Furthermore, metofluthrin mechanism of action focuses on causing disorientation affecting its biting activity, knockdown and killing the mosquitoes rather than exhibiting a repellent effect (Buhagiar et al., 2017).

Metofluthrin is highly lethal to *Ae. aegypti* in the confined space of an apartment at Queensland, Australia (Rapley et al., 2009). Shono (2004) reported that the use of metofluthrin coils effectively caused reduction in landing counts of *Ae. aegypti*. One Team Network Sdn Bhd claimed metofluthrin had high vapor pressure and deliver macro-particles into air and provided up to eight hours of protection. This statement was supported by Kawada et al., (2006) where metofluthrin had greater vapor pressure than d-allethrin and permethrin. Due to its characteristic, metofluthrin can vaporize at normal temperature while other pyrethroid are not able to vaporize due to their inability to evaporate effectively and require heating source for evaporation (Kawada et al., 2006).

Our outcome from the adult efficacy study shows that Xmos was able to cause knockdown at 10, 20, 30, 60 and 120 min and high mortality after 24 hours of exposure towards adult *Aedes* mosquitoes. We also analysed the mortality rate in two different size of spaces which were hall (20.4 m²) and room (13.7 m²) giving us the advantage to determine if the spray was effective at different space volumes. Our outcome was like Dabro et al. (2017) who reported that significant knockdown was reported as early as 10 min after exposure at a distance of 3 m. They also reported that the mosquito knockdown at 3 m was lower than the exposure at 1 m. The knockdown rates ranged from 20-90% when the mosquitoes were exposed at 1 m and 0-70% at 3 m indicating partial dependency of mosquito knockdown rates on the room volume. In our study when we placed the spray 3 m away from the mosquitoes resulted in a 67% of knockdown.

The advantage of this study is that it is carried out at the site of a hotspot area and therefore represents the actual scenario rather than testing the effectiveness of a product in an experimental setting with controlled environment. The ideal range of *Ae. aegypti* female's survival and ability to travel is between 15°C and 32°C with an optimal temperature of 27°C (Rowley & Graham, 1968). The temperature in this study fell within 27.3-30.3°C, which is ideal for the mosquitoes' optimum activity. In terms of relative humidity, it has little impact towards the activity of female *Ae. aegypti* and their flight activity if it still falls within 30 to 90% with an optimal relative humidity of 80% (Rowley & Graham, 1968).

The relative humidity of our study area was 67.3-80.2% also ideal for the survival of *Aedes* mosquitoes. This indicates that metofluthrin caused the mortality of the mosquitoes and the possibility of mosquitoes' knockdown due to unfavourable or harsh surrounding temperature is cancelled out.

CONCLUSION

I.M.O.S with Xmos mini aerosol was effective in killing adult *Aedes* mosquitoes following direct exposure and after 24 hours after withdrawal in 17th College, Universiti Putra Malaysia. I.M.O.S with Xmos mini aerosol was however not effective in reducing the number of mosquito eggs and ovitraps index.

ACKNOWLEDGEMENTS

The authors thank Mr. Lim Chee Hwa from One Team Network Sdn. Bhd. for his kind contribution in providing financial assistance, and other necessary materials for this project. We also thank Sumitomo Chemical Enviro- Agro Asia Pacific Sdn. Bhd. for providing us with adult *Aedes* mosquitoes for our adult efficacy study. Special thanks to the 17th College UPM, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia for providing the facilities.

REFERENCES

- Abdul, R. A., Hamizah, K., Olivia, L. J., Anita, A. R., & Rafee, B. B. M. (2016). Risk assessment on dengue among UPM students. *International Journal of Public Health and Clinical Sciences*, 3(3), 7289-7577.
- Bibbs, C. S., & Xue, R. D. (2016). OFF! Clip-on repellent device with metofluthrin tested on *Aedes aegypti* (Diptera: Culicidae) for mortality at different time intervals and distances. *Journal of Medical Entomology*, 53(2), 480-483. <https://doi.org/10.1093/jme/tjv200>
- Bouزيد, M., Brainard, J., Hooper, L., & Hunter, P. R. (2016). Public health interventions for *Aedes* control in the time of Zikavirus—A meta-review on effectiveness of vector control strategies. *PLoS Neglected Tropical Diseases*, 10(12), Article e0005176. <https://doi.org/10.1371/journal.pntd.0005176>
- Boyce, R., Lenhart, A., Kroeger, A., Velayudhan, R., Roberts, B., & Horstick, O. (2013). *Bacillus thuringiensis israelensis* (Bti) for the control of dengue vectors: Systematic literature review. *Tropical Medicine & International Health*, 18(5), 564-577. <https://doi.org/10.1111/tmi.12087>
- Buhagiar, T. S., Devine, G. J., & Ritchie, S. A. (2017). Effects of sublethal exposure to metofluthrin on the fitness of *Aedes aegypti* in a domestic setting in Cairns, Queensland. *Parasites & vectors*, 10(1), 1-7. <https://doi.org/10.1186/s13071-017-2220-7>
- Darbro, J. M., Muzari, M. O., Giblin, A., Adamczyk, R. M., Ritchie, S. A., & Devine, G. J. (2017). Reducing biting rates of *Aedes aegypti* with metofluthrin: investigations in time and space. *Parasites & vectors*, 10(1), 1-9. <https://doi.org/10.1186/s13071-017-2004-0>

- Fitzpatrick, C., Haines, A., Bangert, M., Farlow, A., Hemingway, J., & Velayudhan, R. (2017). An economic evaluation of vector control in the age of a dengue vaccine. *PLoS Neglected Tropical Diseases*, *11*(8), Article e0005785. <https://doi.org/10.1371/journal.pntd.0005785>
- Gubler, D. J. (2002). Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. *Trends in Microbiology*, *10*(2), 100-103. [https://doi.org/10.1016/S0966-842X\(01\)02288-0](https://doi.org/10.1016/S0966-842X(01)02288-0)
- Guzman, M. G., Halstead, S. B., Artsob, H., Buchy, P., Farrar, J., Gubler, D. J., Hunsperger, E., Kroeger, A., Margolis, H. S., Martínez, E., Nathan, M. B., Pelegrino, J. L., Simmons, C., Yoksan, S., & Peeling, R. W. (2010). Dengue: A continuing global threat. *Nature Reviews Microbiology*, *8*(12), S7-S16. <https://doi.org/10.1038/nrmicro2460>
- Han, W. W., Lazaro, A., McCall, P. J., George, L., Runge-Ranzinger, S., Toledo, J., Velayudhan, R., & Horstick, O. (2015). Efficacy and community effectiveness of larvivorous fish for dengue vector control. *Tropical Medicine & International Health*, *20*(9), 1239-1256. <https://doi.org/10.1111/tmi.12538>
- iDengue. (2019). *iDengue untuk komuniti* [iDengue for Community]. Retrieved July 10, 2020, from idengue.remotesensing.gov.my
- Kawada, H., Iwasaki, T., Loan, L. L., Tien, T. K., Mai, N. T. N., Shono, Y., Katayama, Y. & Takag, M. (2006). Field Evaluation of spatial repellency of metofluthrin-impregnated latticework plastic strips against *Aedes Aegypti* (L.) and analysis of environmental factors affecting its efficacy in My Tho City, Tien Giang, Vietnam. *The American Society of Tropical Medicine and Hygiene*, *75*(6), 1153-1157.
- Kawada, H., Maekawa, Y., & Takagi, M. (2004). Laboratory and field evaluation of spatial repellency with metofluthrin impregnated paper strip against mosquitoes in Lombok Island, Indonesia. *Journal of the American Mosquito Control Association*, *20*(3), 292-298.
- Latifah, S. Y., Paskaran, K., Gopalsamy, B., & Abd Majid, R. (2020). Aedestech mosquito home system prevents the hatch of aedes mosquito eggs and reduces its population. *Pertanika Journal of Science and Technology*, *28*(1), 263-278.
- Lazaro, A., Han, W. W., Manrique-Saide, P., George, L., Velayudhan, R., Toledo, J., Ranzinger, S. R., & Horstick, O. (2015). Community effectiveness of copepods for dengue vector control: systematic review. *Tropical Medicine & International Health*, *20*(6), 685-706. <https://doi.org/10.1111/tmi.12485>
- Lima, E. P., Goulart, M. O. F., & Neto, M. L. R. (2015). Meta-analysis of studies on chemical, physical and biological agents in the control of *Aedes aegypti*. *BMC public health*, *15*(1), Article 858. <https://doi.org/10.1186/s12889-015-2199-y>
- Marcondes, C. B., & Ximenes, M. D. F. F. D. M. (2016). Zika virus in Brazil and the danger of infestation by *Aedes* (Stegomyia) mosquitoes. *Revista da Sociedade Brasileira de Medicina Tropical*, *49*(1), 4-10. <https://doi.org/10.1590/0037-8682-0220-2015>
- Matsuo, N., Ujihara, K., Shono, Y., Iwasaki, T., Sugano, M., Yoshiyama, T., & Uwagawa, S. (2005). *Discovery and development of a novel pyrethroid insecticide 'Metofluthrin (SumiOne®/Eminence®)'*. Sumitomo Chemical Co., Ltd.
- Medeiros, A. S., Costa, D., Branco, M., Sousa, D., Monteiro, J. D., Galvão, S., Azevedo, P., Fernandes, J. V., Jeronimo, S., & Araújo, J. (2018). Dengue virus in *Aedes aegypti* and *Aedes albopictus* in urban areas in

- the state of Rio Grande do Norte, Brazil: Importance of virological and entomological surveillance. *PloS one*, 13(3), Article e0194108. <https://doi.org/10.1371/journal.pone.0194108>
- Moore, S. J., Davies, C. R., Hill, N., & Cameron, M. M. (2007). Are mosquitoes diverted from repellent-using individuals to non-users? Results of a field study in Bolivia. *Tropical Medicine & International Health*, 12(4), 532-539. <https://doi.org/10.1111/j.1365-3156.2006.01811.x>
- Packierisamy, P. R., Ng, C. W., Dahlui, M., Inbaraj, J., Balan, V. K., Halasa, Y. A., & Shepard, D. S. (2015). Cost of dengue vector control activities in Malaysia. *The American journal of tropical medicine and hygiene*, 93(5), 1020-1027. <https://doi.org/10.4269/ajtmh.14-0667>
- Rapley, L. P., Russell, R. C., Montgomery, B. L., & Ritchie, S. A. (2009). The effects of sustained release metofluthrin on the biting, movement, and mortality of *Aedes aegypti* in a domestic setting. *The American Journal of Tropical Medicine and Hygiene*, 81(1), 94-99. <https://doi.org/10.4269/ajtmh.2009.81.94>
- Ratini, M. (2019). *Dengue fever: Symptoms, causes, and treatments*. Retrieved June 28, 2019, from <http://www.denguevirusnet.com/dengue-virus.html>
- Rowley, W. A., & Graham, C. L. (1968). The effect of temperature and relative humidity on the flight performance of female *Aedes aegypti*. *Journal of Insect Physiology*, 14(9), 1251-1257. [https://doi.org/10.1016/0022-1910\(68\)90018-8](https://doi.org/10.1016/0022-1910(68)90018-8)
- Shono, Y. (2004). Field evaluation of papers strip and mosquito coil formulation impregnated metofluthrin for mosquito control in Malaysia. In *The Abstract Book, 70th Annual Meeting of the American Mosquito Control Association, 2004*. American Mosquito Control Association.
- Ujihara, K., Mori, T., Iwasaki, I., Sugano, M., Shono, X., & Matsuo, N. (2004). Metofluthrin: A potent new synthetic pyrethroid with high vapor activity against mosquitoes. *Bioscience, Biotechnology, and Biochemistry*, 68(1), 170-174. <https://doi.org/10.1271/bbb.68.170>
- Wallace, H. G., Lim, T. W., Rudnick, A., Knudsen, A. B., Cheong, W. H., & Chew, V. (1980). Dengue hemorrhagic fever in Malaysia: The 1973 epidemic. *The Southeast Asian Journal of Tropical Medicine and Public Health*, 11(1), 1-13.
- WHO. (2017). *Dengue/Severe dengue frequently asked questions*. Retrieved August 14, 2020, from <https://www.who.int/denguecontrol/faq/en/index5.htmlb>
- Wilson, A. L., Dhiman, R. C., Kitron, U., Scott, T. W., van den Berg, H., & Lindsay, S. W. (2014). Benefit of insecticide-treated nets, curtains and screening on vector borne diseases, excluding malaria: A systematic review and meta-analysis. *PLoS Neglected Tropical Disease*, 8(10), Article e3228. <https://doi.org/10.1371/journal.pntd.0003228>

Modelling *In-situ* Factors Affecting Bud's Growth of *Rafflesia kerrii* Meijer in Lojing Highlands, Kelantan, Peninsular Malaysia

Abdul Hamid Mar Iman¹, Nor Hizami Hassin^{2*}, Muhamad Azahar Abas² and Zulhazman Hamzah²

¹Faculty of Veterinary, Universiti Malaysia Kelantan, Pengkalan Chepa, 16100 Kota Bharu, Kelantan, Malaysia

²Faculty of Earth Science, Universiti Malaysia Kelantan, 17600 Jeli, Kelantan, Malaysia

ABSTRACT

Studies on the statistical approach to analyzing growth factors of bud's growth in the genus *Rafflesia* have been lacking. This study quantified the effects of eight selected ecological factors hypothesized to be influencing bud's growth (diameter and circumference) of *Rafflesia kerrii* Meijer. A non-experimental cross-sectional data collection was conducted between April and August 2018 by *in-situ* observation and measurements on eight ecological factors utilizing thirty-four sampled individual plants in Lojing Highlands, Kelantan, Peninsular Malaysia. The Ordinary Least Squares (OLS) and Heteroscedasticity-Consistent-Error (HCE) OLS regression models were employed to establish the statistical relationship between bud's growth and its influencing factors. Host plant's ecological ability, level of temperature, light shading, soil acidity, and interaction between plant survival condition and growth stage were found to be the significant and influential ecological factors to bud's growth of *Rafflesia kerrii*. The results also showed that, model wise, HCE OLS models outperformed the OLS models in explaining the cause-and-effect relationship under study. Due to some limitations in sampling and data collection, further studies were recommended to corroborate this study using a larger sample covering a larger geographic area – possibly across different localities.

ARTICLE INFO

Article history:

Received: 22 September 2020

Accepted: 30 December 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.30>

E-mail addresses:

hamid.m@umk.edu.my (Abdul Hamid Mar Iman)

hizami.h@umk.edu.my (Nor Hizami Hassin)

azahar.a@umk.edu.my (Muhamad Azahar Abas)

zulhazman@umk.edu.my (Zulhazman Hamzah)

* Corresponding author

Keywords: Bud growth, growth factors, *Rafflesia*, regression model

INTRODUCTION

The existence of rare and endangered gigantic plant *Rafflesia spp.* is geographically limited (Peters & Ting, 2016). The existing documentation indicates that it is found in

the lowland and highland of the tropical region of Southeast Asia in Thailand, Indonesia, Malaysia, and the Philippines (Beaman et al., 1998; Jamili, 2001a; Anon, 2003; Balete et al., 2010; Molina et al., 2014; Mahyuni et al., 2015; Mursidawati, 2015; Mursidawati, 2017; Peters & Ting, 2016; Nery et al., 2016). Even within a particular confine of a country, these species are a site-specific plant type whose distribution is spatially scarce (Nery et al., 2016). The actual number of varieties of this plant species is still not accurately known. To date, as many as twenty-four (Latiff, 2018), thirty-two (Jumaat et al., 2016); thirty-four (Kedri et al., 2018), forty (Nery et al., 2016), or even fifty-five (Faye, 2008) varieties of *Rafflesia* species were claimed to have been discovered in the Southeast Asian region with *Rafflesia kerrii* Meijer being one of them.

Studies on *Rafflesia* are numerous. Their attractive appearance with large petals, striking red colour, large-hollow perigone lobe, smell peculiarity, and parasitic in nature has attracted researchers, local and international alike, on their very nature. However, more aspects of *Rafflesia* are left to be investigated. Among others, bud growth is a critical aspect for survival of *Rafflesia* since the plant's abundance depends on bud density per unit area (Qayyum et al., 2012). Factors that influence the plant's distribution directly or indirectly influence the plant's bud growth itself. So far, published works examining the growth performance of *Rafflesia*'s bud are scanty, and thus, the mystery of the species' bud growth is more to be explored. No description of this variety is given in this paper since not much scientific characterization of it has been documented so far. Nevertheless, some good description of *R. kerrii* can be cited, for example, in Meijer (1984), Jamili (2001a) and Kanchanapoom et al. (2007). The results of this study will partly clarify the conjectural explanation about the factors that contribute to the success of this plant species to survive and develop (Nasihah et al., 2013; Kedri et al., 2018). This study differs from the previous ones since it provides some parametric explanation of the ecological factors hypothesized to be significantly influencing the bud's growth performance of these species, particularly on *R. kerrii*. Therefore, the objective of this study is to examine the *in-situ* factors significantly influencing bud's growth (growth in diameter and circumference) of *R. kerrii* by taking a site in Lojing Highlands, Kelantan, Peninsular Malaysia, as a case study.

Literature of Growth Pertinent Aspects of *Rafflesia*

In general, a biological organism grows in a particular territorial home called habitat which is ecologically governed by variations in factors such as latitude, altitude/elevation, soils, topography, canopy density, temperature, rainfall, biological interactions with other plants and animals (Martin, 1998). Within a habitat, there is a micro-habitat where temperature, humidity, and light intensity differ in influence from one organism to another (Inger & Lian, 1996). These are external factors that interact with the very architecture of an organism such as plant species. In this context, the growth aspects of *Rafflesia* can be studied by their

biological characterization (physiology, morphology, and phenology) and/or ecological characterization (the external factors) that determine their growth and growth. Each of these options has its specific purpose, methodology and assumptions. Biological characterization studies are reflected in some works such as Barkman et al. (2004), Nickrent et al. (2004), Ramamoorthy et al. (2013), Jumaat et al. (2016), Latifah et al. (2017), and Chu et al. (2018). In the current study, attention is given more on the ecological factors of *R. kerrii*.

Climatic factors (e.g. rainfall, humidity, wind, gases, temperature, light); edaphic/physiographic factors (e.g. relief, altitude/elevation, aspect, exposure, slope, niche); soil factors (e.g. structure, texture, nutrients, acidity); and biotic factors (e.g. interrelationship between organisms) are all the ecological factors that determine the growth and growth of plant species in general (Beaman et al., 1988; Bareja, 2011; Kumar, 2018). In particular, factors influencing the growth of *Rafflesia* species are very intricate and, so far, no studies have provided a satisfactory explanation on the factors that significantly determine the plant's growth success (Mursidawati, 2017). Out of such factors, there are some specific factors operating *in-situ* that have different degrees of influence and importance to *Rafflesia*. Further, there are some anomalies to these that are not easy to explain without a carefully designed quasi-experimental study. For example, *Rafflesia* species are not mainly influenced by soils since they can grow on various substrates (Jamili, 2001a). *Rafflesia* species are also a dioecious plant type (i.e., separate male and female flowers), with only a few colonies in a designated area, are mostly male while males and females rarely bloom at the same time; these factors contribute to the difficulty in their multiplication (Salleh, 2007; Balete et al., 2010; Susatya, 2011). It was also said that *Rafflesia* species have some specialized habitat where there is water abundance (river banks), close to bamboo clumps, dependent upon the distribution and characteristics (e.g. size and ability) of the host plants, endemic to different geographic locations, and grows at higher altitudes (Faye, 2008; Akhriadi et al., 2010; Galindon et al., 2016). *Rafflesia* species are known to be a site-specific plant type, and their propagation *ex-situ* is very difficult (Jamili, 2001a, Jamili, 2001b; Mursidawati & Risawati, 2009; Wicaksono & Da Silva, 2015). Scanty information on the *in-situ* and *ex-situ* factors believed to be influencing the plant's growth can be scanned through the existing body of knowledge.

However, many of these factors were studied in isolation, making the result conjectural in nature. An important characteristic of *Rafflesia* species is that they are holoparasites growing on vine's root tissues, such as *Tetrastigma leucostaphyllum* and *T. scariosum* (Veldkamp, 2009; Susatya, 2011). Therefore, edaphically, *Rafflesia* species can be found at sites where these host plants are distributed. Further, factors affecting *Tetrastigma* will directly affect *Rafflesia* although the relationships between these two partners are still poorly understood (Mursidawati, 2017). All these points to the need for some comprehensive studies on *Rafflesia* covering all of their growth factors *in-situ*.

Not much has been researched about *Rafflesia*'s bud growth from its dormancy, burst moment to flowering phases. *Rafflesia* undergoes nine-phase growth of which bud growth occurs in the third and fourth phases (Jamili, 2001b). During this period, pertinent *in-situ* factors operate to determine the growth of bud. Bud erupts from an infected host vine, swells slowly over months – usually four to nine – to the size of cabbages (Shaw, 2017; Latiff, 2018) and blooms into flower especially in the rainy season depending on the presence of large bud (Jumaat et al., 2016). Since *Rafflesia* has no leaves, stems, and chlorophyll, making them incapable of photosynthesis, the bud grows by living off the host plant's vine, draining water and nutrients from it for six to nine months before reaching the flowering phase. Since *Rafflesia* grows predominantly on vine's root, host plant's root density around a particular site is an important factor that encourages the production of bud, directly or indirectly. With *Rafflesia* species totally depend on their host plants for water and food, soil nutrients and soil acidity could be other factors directly influencing bud's growth. Besides, elevation/altitude which, in turn, indirectly determines other physical factors such as canopy shade, air temperature, humidity, and wind speed can also influence *Rafflesia*'s bud growth.

However, so far, the ecological factors influencing *Rafflesia*'s bud growth are yet to be statistically quantified. Such a study is vital to examine factors that might be statistically significant to the plant species' growth. More of such study is necessary to ascertain whether the relationship between plant growth and its determining factors is geographically and functionally stable. It is also important to ensure that while studies on *Rafflesia*'s biological characteristics are necessary, exploring their ecological characteristics is equally important.

MATERIALS AND METHODS

Study Area

Lojing is a highland area located in Gua Musang, Kelantan, Peninsular Malaysia, bordering Cameroon Highland, Pahang, on the eastern side (Figure 1a). It is popular for highland agriculture as well as one of the prime hotspots for *R. kerrii* in Malaysia. Ironically, *R. kerrii* had become an icon for conservation and a flagship of tourist destination to Lojing Highland.

Data Collection

A forest tract of about 3-ha. where *R. kerrii* thrived was selected for data collection. We adopted a non-experimental cross-sectional study with data collection conducted between April and August 2018 with the site being visited five times. A total of thirty-four individual plants were selected from a population of *Rafflesia spp.* that randomly distributed over the study area. The decision to have picked only thirty-four individual plants was made due to time and resource constraints. However, a sufficiently minimal sample size was managed

to be collected for the statistical analysis. Inspection proceeded in a linear traverse mode starting at the southwest corner of the study area where two baselines intersected (Figure 1b). One of the baselines was selected quite parallel to the main road, some 250 m away from it. Using a digital compass and Global Positioning System (GPS) device, each plant's site geo-coordinates were recorded for easy identification, systematic data records, and to avoid repeated data collection of the same spot. As we walked along each of the linear lines, we selected the *Rafflesia* plant, limiting it to 4-6 individuals, over a 25-m wide swath to the right of each traverse line. Data on the pertinent factors influencing plant growth were collected by *in-situ* observation and measurements where *R. kerrii* was found.

Site elevation/altitude was measured using a high-resolution real-time kinematic global positioning system device; for double-checking purposes, a topography map of the study area was also used. The temperature and humidity levels were measured using ThermoPro TP53 hygrometer humidity gauge. Light shading measurement was made by a modified "visible sky" approach (i.e. visible sky x 100 = % canopy openness) as per Montgomery (2004). In this approach, light shading (%) = (100 – canopy openness) %. It was found that canopy openness ranged between 13% to 29%. We followed Kedri et al. (2018) in classifying the growth stages of *R. kerrii* into stages I, II, or III. Each individual plant was roughly observed by assessing the morphological features of the flower bud (1 = if stages II or III; 0 = if stage I). The full description for assessing growth-stage morphology is contained in Jamili (2001b). The plant survival condition was roughly assessed by

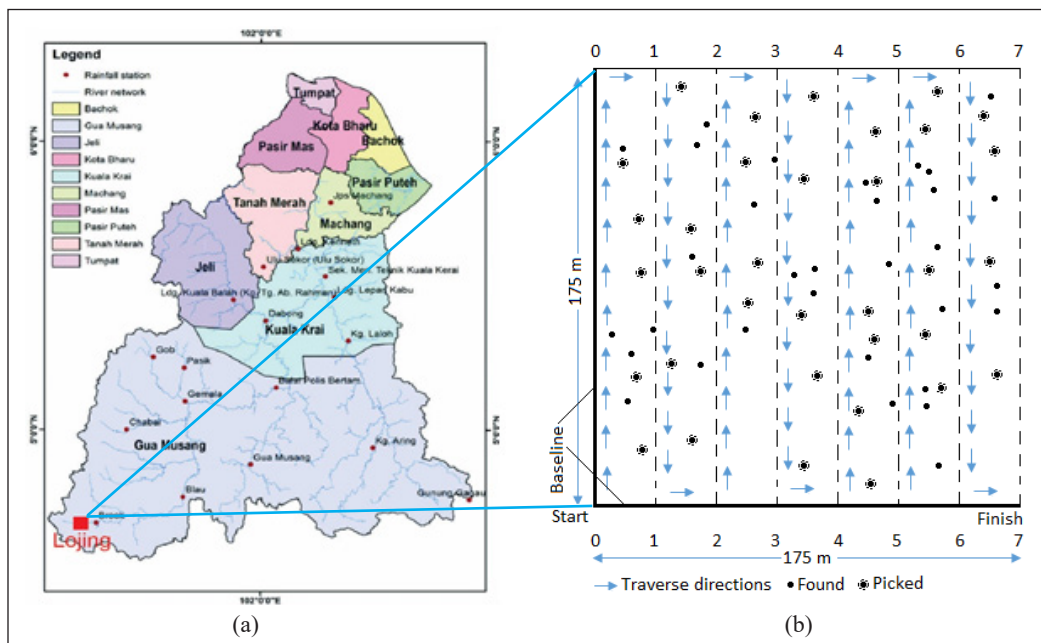


Figure 1. Location of the study area and sampling design

considering as to whether a particular individual plant was still alive or already dead/dying (1 = if alive; 0 = if dead/dying) during the fieldwork. A handy soil pH meter was used to detect the level of soil acidity measured at the base of the host plant. Although the data may not be as accurate as those obtained from a laboratory test, the purpose was to give an indicative state of *in-situ* soil acidity. Soil acidity was assumed to be directly related to soil's nutrient contents; a lower-level acidity (higher pH value) corresponds to a richer amount of nutrients. It was not accurately known as to the ideal level of soil acidity for *Rafflesia*, but the level of pH of 5.0-5.5 was perceived to be ideal for the plant's good growth performance. Bud growth was measured by its diameter (and circumference) as per Faust and Lewis (2005), using a measuring tape.

Data Analysis

Modelling Growth Factors. Factors that influence the growth of *Rafflesia* could not be explained satisfactorily without designing a sample for data collection, especially for parametric analysis. A sufficiently large geographic area across a country or a region may be necessary to ascertain the ecological characteristics of *Rafflesia* where the growth-influencing factors operate systematically such as the distribution and conditions of host plants *Tetrastigma* species., male-female bloom lag; canopy shading, growth stage and survival condition (Beaman et al., 1988; Jamili & Wilcock, 1998; Jamili et al., 2000; Jamili, 2001a, Jamili, 2001b; Akhriadi et al., 2010; Kedri et al., 2018). Unfortunately, within a small forest tract (say, a 3-ha. plot) and over a short period of study, only a limited number of factors could be studied ensuring, nonetheless, that there is a sufficient random variation of these factors in influencing the growth of the plant.

Given the array of growth factors, as mentioned earlier, statistical modelling is an objective way of simultaneous consideration of these factors for a parametric analysis under the assumptions of best linear unbiased estimates (BLUE). Since not all factors can be specified in a model, a statistical model will usually prescribe a limited number of pertinent factors. Given this premise, a parsimonious statistical relationship of *Rafflesia* against the pertinent *in-situ* growth factors must be established. Partial relationships between the growth of this plant species against its influencing factors such as host plant, soil, etc. have been described in several studies (Nasihah et al., 2013; Nasihah et al., 2016). However, due to such partial analyses, in most situations, the claimed relationship between *Rafflesia* and their growth-influencing factors was rather conjectural.

Despite the studies examining the habitat of *Rafflesia* (Faye, 2008; Hidayati & Walck, 2016; Kedri et al., 2018) there was no parametric analysis of factors that may significantly influence the plant's bud growth. In Faye (2008) study on *Rafflesia keithii*, canopy cover, temperature, and vegetation were studied. A variety of statistical models for evaluating

various issues on bud growth were employed in a number of studies where temperature and photoperiod were two main factors specified in the models (Osawa et al., 1983; Kramer, 1994; Hakkinen et al., 1998; Hakkinen, 1999; Mello & Tuan, 1999; Linksolo et al., 2000). However, none of these studies was investigating parasitic plants. Studies on parasite-host relationships were focused mainly on agricultural crops (e.g. Grenz et al., 2005; Moreau et al., 2016). Apart from *Rafflesia* works on parasitic plants can be cited in a limited number of studies (Medel, 2000; Hautier et al., 2010). Again, however, these studies were either non-statistical or based on controlled experiments and not focusing on bud growth. The current study is focusing on a non-experimental parasitic plant, i.e. *R. kerrii* Meijer whose bud growth is hypothesized to be influenced by its ecological factors and whose data are to be collected by cross-sectional *in-situ* observation and site measurements.

Basic Postulation of the Model. Plant growth is a fundamental ecological process with a variety of reasons for its growth and functional in nature (Paine et al., 2012). With a parsimonious statistical relationship, a group of ecological factors – physical and environmental – can be considered simultaneously by a regression model and, thus, the partial contribution of each factor on the plant's growth performance can be assessed more objectively. As found in the body of literature on statistical modelling, a parsimonious statistical relationship between a growth variable and its growth-influencing factors can be specified in general as Equation 1:

$$Y = X\beta + \varepsilon \quad (1)$$

where $Y = (n \times 1)$ matrix of dependent variable (e.g. bud diameter/circumference); $X = (n \times k)$ matrix of independent ecological variables (physical and environmental); $\beta = (k \times 1)$ matrix of regression parameters, including the intercept; and $\varepsilon = (n \times 1)$ matrix of error term. In a parsimonious statistical relationship, we are interested to quantify the regression parameter, β , which is derived by a matrix manipulation as Equation 2 (ignoring ε , for simplicity):

$$\beta = (X'X)^{-1}X'Y \quad (2)$$

Functionally, β is equivalent to the marginal effect of X on Y , i.e. $\beta = \delta Y / \delta X$ for a linear model. In other words, the marginal effect of a physical and environmental factor is quantified by the differentiation of the regression equation. In this context, our interest is to test the hypothesis (at $\alpha = 0.05$ and better) as follows: $H_0: \beta_1, \beta_2, \dots, \beta_k = 0$ against $H_1: \beta$'s $\neq 0$ at least for one factor. As with our case study, the general relationship between *Rafflesia*'s growth performance and the growth factors can be specified as Equation 3a and 3b:

$$\begin{aligned} \frac{DIAM^{\lambda_1} - 1}{\lambda_1} = & \alpha + \beta_1 \left(\frac{ELE^{\lambda_2} - 1}{\lambda_2} \right) + \beta_2 \left(\frac{TEMP^{\lambda_2} - 1}{\lambda_2} \right) + \beta_3 \left(\frac{SHADE^{\lambda_2} - 1}{\lambda_2} \right) \\ & + \beta_4 \left(\frac{HUMID^{\lambda_2} - 1}{\lambda_2} \right) + \beta_5 \left(\frac{PH^{\lambda_2} - 1}{\lambda_2} \right) + \beta_6 \left(\frac{ROOTS^{\lambda_2} - 1}{\lambda_2} \right) + \beta_7 \text{TERMIT} \\ & + \beta_8 \text{COND} + \beta_9 \text{STAGE} + \varepsilon \end{aligned} \tag{3a}$$

$$\begin{aligned} \frac{CIRCUM^{\lambda_1} - 1}{\lambda_1} = & \gamma + \varphi_1 \left(\frac{ELE^{\lambda_2} - 1}{\lambda_2} \right) + \varphi_2 \left(\frac{TEMP^{\lambda_2} - 1}{\lambda_2} \right) + \varphi_3 \left(\frac{SHADE^{\lambda_2} - 1}{\lambda_2} \right) \\ & + \varphi_4 \left(\frac{HUMID^{\lambda_2} - 1}{\lambda_2} \right) + \varphi_5 \left(\frac{PH^{\lambda_2} - 1}{\lambda_2} \right) + \varphi_6 \left(\frac{ROOTS^{\lambda_2} - 1}{\lambda_2} \right) + \varphi_7 \text{TERMIT} \\ & + \varphi_8 \text{COND} + \varphi_9 \text{STAGE} + \varepsilon \end{aligned} \tag{3b}$$

where DIAM = *Rafflesia* flower’s bud diameter (cm); CIRCUM = *Rafflesia* flower’s bud circumference (cm); ELEV = elevation above mean sea level (m); TEM = level of temperature (°C); SHADE = light shading/intensity (100-% of light penetration); HUMI = level of humidity (%); pH = soil acidity at the host plant’s base (0-14); ROOT = number of visible *Tetrastigma*’s roots/vines on ground surface within a 20-m radius; TERMIT = presence of termites on the plant (1 = yes; 0 = no); COND = state of survival (1 = alive; dead = 0); STAGE = growth stage at time of inspection (I, II, III); α , γ , β_s , φ_s are regression coefficients, ε is error term, $(.)\lambda_1$ and $(.)\lambda_2$ are the Box-Cox transformations of the variables over a range of values of λ_1 and λ_2 .

Equation 3a and 3b are specified because, in theory, plant growth performance is not likely to be influenced by its growth factors linearly (Koya & Ghosu, 2013). The body of literature on plant growth has indicated rather non-linear relationships between these two sets of factors. It follows that model’s functional forms need to be tested on a case by case basis. In this context, different values of λ can be chosen for the dependent and continuous explanatory variables. However, to avoid the estimation becoming too cumbersome, it is assumed that λ is equal for all variables (Greene, 1990). Maddala (1977) pointed out that the search procedure for the optimal regression model would not be efficient if there were more than two or three λ ’s. Considering all these points, therefore, the same value of λ was chosen for the above model, that is, $\lambda_1 = \lambda_2 = \lambda$. In general, the Box-Cox transformation indicates whether the model to which the data will best fit is a linear or non-linear model. The Box-Cox transformation, however, has one major caveat: the optimum equation (one with the smallest SSE, MSE, or RMSE) may not produce a model that can be easily used for estimating the effects of attributes (Milon et al., 1984). This occurs when the optimum

equation results in the Box-Cox parameter, λ_1 or λ_2 such that $0 > (\lambda_1, \lambda_2) > 1$, for the dependent variable. Furthermore, parameter estimates tend not to be stable, that is, they are susceptible to the inclusion of other variables in the regression equation. Therefore, the choice of model for estimating these effects is normally confined to the special cases of Box-Cox functions. In this context, if $\lambda_1 = \lambda_2 = 1$, Equation 5 is linear; if $\lambda_1 = \lambda_2 = 0$, it is double log; if $\lambda_1 = 1$ and $\lambda_2 = 0$, it is linear-log; and if $\lambda_1 = 0$ and $\lambda_2 = 1$, it is log-linear. In these special cases, the choice of best function is determined primarily by the standard statistical tests, two of which are the likelihood ratio test (Griffith et al., 1993; Maddala, 1992) and the Box-Cox test for model equivalence (Greene, 1990). However, the adjusted R^2 and F-value are consistent with these two measures so that in most cases, they can be used to compare between models.

Some modification in the modelling is proposed. First, variable manipulation necessitates creating a new variable $CONSTA = COND \times STAGE$. By assumption, growth stage and growth condition are interactive in terms of plant survival and performance. For instance, bud-stage growth will only be meaningful if the plant is alive while, in the same way, plant stage is important to give rise to plant survival condition. In particular, the chances of survival are considered higher if the plant reaches advanced stages of growth. Second, the effects of growth factors on the plant's growth performance can be expected to differ functionally.

Heteroscedasticity-Consistent-Error (HCE) OLS. Cross-sectional ecological data on plant growth may be subjected to heteroscedasticity due to changing variable interactions across a particular study area. For instance, there may be changing relationships of pH-nutrient, temperature-shade, temperature-humidity, etc. at different elevations. When modelled, such phenomena will give rise to the violation of OLS assumptions that $var(\varepsilon_i|X_i) = 0$ for all i and $E(\varepsilon_i\varepsilon_j) = 0$. When the assumptions of constant-variance and uncorrelated errors of a regression model are violated, the errors will be heteroscedastic, resulting in the non-minimum mean square error (MSE) of OLS estimator and inconsistent estimator for the variance of OLS estimates (Wikipedia, 2018). OLS estimates will also become inefficient while the estimated standard errors can be either too large or too small (Long & Ervin, 1998). Given that heteroscedasticity is a common problem in cross-sectional data analysis, corrective methods are important for prudent data analysis. This is solved by estimating a heteroscedasticity-consistent-error (HCE) model that will result in a statistically better model in terms of corrected regression-error estimates and higher degrees of variation in the dependent variable is explained.

From Equation 1, the i^{th} row of X can be written as Equation 4

$$y_i = x_i\beta + \varepsilon_i \quad (4)$$

With all BLUE assumptions being satisfied, the OLS estimator $\hat{\beta}=(X'X)^{-1}X'y$ has the covariance matrix as in Equation 5

$$\text{Var}(\hat{\beta})=(X'X)^{-1}X'\Phi X(X'X)^{-1} \tag{5}$$

where Φ is a diagonal matrix with $\phi_{ii} = \text{var}(\varepsilon_i)$. It can be proven that if the errors are homoscedastic, i.e. $\Phi =\sigma^2I$, Equation 5 becomes Equation 6

$$\begin{aligned} \text{Var}(\hat{\beta}) &= (X'X)^{-1}X'(\sigma^2I) X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1}X'X(X'X)^{-1} \\ &= \sigma^2(X'X)^{-1} \end{aligned} \tag{6}$$

where σ^2 is dependent variable's variance. Defining the residuals as $e_i = y_i - x_i\hat{\beta}$, the usual OLS covariance matrix, OLSCM, can be estimated as Equation 7

$$\text{OLSCM} = \frac{\sum e_i^2}{N-K}(X'X)^{-1} = \sigma^2(X'X)^{-1} \tag{7}$$

If the errors are heteroscedastic, OLSCM will be biased and the usual tests of statistical significance will be inconsistent. Since the form of heteroscedasticity is rarely known, we need a consistent estimator Φ in order to apply Equation 5. This is done by creating a heteroscedasticity consistent covariance matrix (HCCM) and a model having it is a heteroscedastic-corrected-error (HCE) OLS. Assuming that e_i^2 can be used to estimate ϕ_{ii} , let $\hat{\Phi} = \text{diag}[e_i^2]$, resulting in an estimator called HC0 as Equation 8

$$\begin{aligned} \text{HC0} &= (X'X)^{-1}X'\hat{\Phi}X(X'X)^{-1} \\ &= (X'X)^{-1}X'\text{diag}[e_i^2]X(X'X)^{-1} \end{aligned} \tag{8}$$

In Equation 8, $\hat{\Phi}$ is estimated based on OLS residuals, e , which may not likely be homoscedastic. If $h_{ii} = x_i(X'X)^{-1}x_i'$, then $\text{Var}(e_i) = \sigma^2(1-h_{ii}) \neq \sigma^2$. According to Belsley et al. (1980), since $0 \leq h_{ii} \leq 1$, $\text{Var}(e_i)$ tends to underestimate σ^2 . Further, since $\text{Var}(e_i)$ varies across observations, the OLS residuals will tend to be heteroscedastic too. Hinkley (1977) suggests inflating e_i by $\sqrt{N/(N-K)}$, where N = sample size and K = number of regression parameters, for 'correcting' Equation 8. But, since e_i^2 is still a biased estimator for σ^2 , weighing Equation 7 by $e_i^2/(1-h_{ii})$ or by $e_i^2/(1-h_{ii})^2$ will produce a less biased estimator. Without a further elaboration here, the literature suggests three variants of 'corrected' Equation 8 as Equation 9a, 9b and 9c (Hinkley, 1977; MacKinnon & White, 1985):

$$HC1 = \frac{N}{N-K}(X'X)^{-1}X\text{diag}[e_i^2]X(X'X)^{-1} \quad (9a)$$

$$HC2 = \frac{N}{N-K}(X'X)^{-1}X\text{diag}\left[\frac{e_i^2}{1-h_{ii}}\right]X(X'X)^{-1} \quad (9b)$$

$$HC3 = \frac{N}{N-K}(X'X)^{-1}X\text{diag}\left[\frac{e_i^2}{(1-h_{ii})^2}\right]X(X'X)^{-1} \quad (9c)$$

The default in *Gretl* software – the one used in this paper – is HC0 but it also includes the other options under a number of modeling choices, namely: (a) Model → Ordinary Least Squares → Robust standard errors (HC1); (b) Model → Instrumental variables → Two-Stage Least Squares → Robust standard errors (HC1); (c) Model → Other linear models → Weighted Least Squares → Robust standard errors (HC1); (d) Model → Other linear models → Heteroscedasticity corrected (presumably HC1); (e) by *gretl* scripting option; (f) by specifying either HC1, HC2, or HC3 in the main menu bar: Tools → Preferences → General → HCCME); there we can select from a set of possible robust estimation variants, and can also choose to make robust estimation the default (Cottrell & Luccetthi, 2005). In this paper, option (d) was selected.

RESULT AND DISCUSSIONS

Sample Profile

The profile of 34 individual sampled *R. kerrii* plants in the study area is presented in the form of a sample's descriptive statistics (Table 1) and variables' partial correlations (Table 2). Based on Table 1, the average altitude (ELEV) of the study plot was 1,044.03 m from the mean sea level with an average temperature (TEM) of 25.21°C. The study plot had a minimum-maximum difference of 28 m in elevation and of 0.55°C in temperature, which was rather large for a small-range vertical height. Still, the temperature difference could have also been interacting with *in-situ* air humidity (HUMI) and light shading (SHADE). This can be examined from the variables' partial correlations, ρ , namely TEM vs. HUMI, TEM vs. SHADE, and TEM vs. ELEV (Table 2). A low negative correlation ($\rho = -0.47$) occurred to *in-situ* TEM vs. ELEV and to TEM vs. HUMI ($\rho = -0.40$) while a high negative correlation occurred to TEM vs. SHADE ($\rho = -0.75$). In general, a lower temperature co-existed with a higher light shading, air humidity and elevation.

The average humidity was 85.18% while the moderate light shading under the canopy was 79.32%. However, the low-level correlations between TEM and HUMI; TEM and ELEV; and ELEV and SHADE were likely to be affected by other physical or environmental factors. For instance, forest clearing at higher altitudes or the global warming that increased the outer environment's temperature could have reduced the co-existing effects of these factors (Saxe et al., 2001). The soils in the study area have an average acidity of 4.84 on

Table 1
Sample's descriptive statistics

	ELEV	TEM	SHADE	HUMI	pH	TERMI	ROOT	CONST	Cond	Stage	Diam	Circum
Mean	1,044.03	25.44	79.32	85.18	4.84	0.24	6.56	0.32	0.85	0.41	11.46	36.00
Std. error	1.33	0.07	0.72	0.12	0.05	0.07	0.31	0.08	0.06	0.09	1.39	4.37
Median	1,045.00	25.53	78.00	85.50	4.93	0.00	6.00	0.00	1.00	0.00	9.15	28.75
Mode	1,051.00	25.30	78.00	85.50	4.98	0.00	5.00	0.00	1.00	0.00	9.75	30.63
Std. dev.	7.74	0.43	4.18	0.71	0.32	0.43	1.81	0.47	0.36	0.50	8.11	25.47
Variance	59.91	0.19	17.50	0.51	0.10	0.19	3.28	0.23	0.13	0.25	65.74	648.88
Kurtosis	0.35	2.63	-0.63	-1.40	-0.83	-0.32	0.88	-1.47	2.50	-1.98	14.64	14.64
Skewness	-0.94	-1.30	0.46	-0.50	-0.49	1.31	1.00	0.79	-2.09	0.38	3.35	3.35
Range	28.00	2.10	16.00	2.00	1.11	1.00	8.00	1.00	1.00	1.00	44.70	140.43
Minimum	1,026.00	24.20	71.00	84.00	4.19	0.00	4.00	0.00	0.00	0.00	4.80	15.08
Maximum	1,054.00	26.30	87.00	86.00	5.30	1.00	12.00	1.00	1.00	1.00	49.50	155.51
Sum	35,497.00	864.93	2,697.00	2,896.18	164.39	8.00	223.00	11.00	29.0	14.00	389.65	1,224.14
Count	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.00	34.0	34.0	34.00	34.00

Table 2
 Partial correlation coefficients, ρ , for linear (upper panel) and intrinsically linear (lower panel) variables

	ELEV	TEM	SHADE	HUMI	pH	TERMIT	ROOT	CONST	Cond	Stage
TEMP	-0.47									
SHADE	0.40	-0.75								
HUMID	0.93	-0.40	0.30							
pH	-0.27	0.24	-0.49	-0.25						
TERMIT	-0.05	0.05	-0.19	-0.09	0.21					
ROOTS	-0.11	0.23	-0.31	-0.08	0.12	-0.02				
CONSTA	0.07	0.03	-0.04	0.14	0.14	-0.24	0.77			
Cond	0.00	0.00	0.11	0.11	-0.07	-0.75	-0.01	0.29		
Stage	0.04	0.20	-0.23	0.06	0.13	0.10	0.84	0.83	-0.16	0.63
Diam	-0.08	0.12	-0.20	-0.09	0.13	-0.06	0.86	0.61	0.04	0.63
Circm	-0.08	0.12	-0.20	-0.09	0.13	-0.06	0.86	0.61	0.04	0.63
	lgELE	lgTEM	lgSHADE	lgHUMI	lgpH	TERMIT	lgROOT	CONST	COND	STAGE
lgTEMP	-0.47									
lgSHADE	0.40	-0.74								
lgHUMID	0.93	-0.40	0.30							
lgpH	0.35	-0.26	0.33	0.25						
TERMIT	-0.12	0.08	-0.06	-0.14	0.06					
lgROOTS	-0.09	0.22	-0.31	-0.06	-0.28	0.08				
CONSTA	0.26	-0.27	0.12	0.33	0.18	-0.24	0.05			
COND	0.12	-0.08	-0.05	0.17	-0.13	-0.75	-0.04	0.29		
STAGE	0.07	-0.24	0.22	0.09	0.33	0.10	0.00	0.83	-0.16	
lgDIA	-0.02	0.16	-0.25	-0.01	-0.23	0.08	0.93	0.05	0.04	-0.05
lgCIRCM	-0.02	0.16	-0.25	-0.01	-0.23	0.08	0.93	0.05	0.04	-0.05

the pH scale. This is considered good for the proliferation of *Rafflesia*. In spite of that, Table 2 shows that the effects of soil acidity have a low overall co-existence with the effects of other ecological factors in the environment. Other than that, as much as 24% of the sampled *R. kerrii*'s flower was infested by termites whereby they strongly co-existed with stages II and III (advanced stages) of plant growth.

This was quite logical because the organisms could have scavenged for foods and shelter on a more matured plant compared to a younger plant. On average, each host plant of *Tetrastigma* has about 6-7 visible roots/vines above the soil surface to enable infestation by the *R. kerrii*. Closely related to this was the 32% average proportion of individual plants growing at stages II and III that were alive (indicated by the variable CONST). Further, CONST was positively highly correlated with STAGE (Table 2) and was quite highly correlated with linear bud diameter (DIAM) and circumference (CIRCM). Bud diameter and circumference were highly positively correlated ($\rho = 0.84-0.93$) with the number of host plant's visible roots/vines above the ground surface within a 20-m radius around the host plant's base (ROOT). A quite strong positive correlation ($\rho = 0.77$) also occurred between the interacting growth status and growth stage (CONST) and ROOT in the linear data.

Comparing the upper and lower panels of Table 2, it is obvious that serious multicollinearity existed among variables in their linear form; the problem was much reduced in the intrinsically linear form. It signals to two important aspects. First, individual variables' partial relationships were not likely to be linear. Second, the appropriate statistical model was not likely to be linear in its functional form, rather, it needs to be intrinsically linear or non-linear. In this study, however, lower-power functional forms were tested on the basis of the theoretical background already discussed before which necessitated the Box-Cox variable transformation and intrinsically linear model specification.

Modelling Results

To economize space and discussion, only the results for bud diameter (DIAM) analysis are reported in this paper. The parametric effects of the eight specified ecological factors on the growth of both plant's bud diameter and circumference were largely symmetrical despite differences in figures and, therefore, the outputs for bud circumference are displayed only for comparison purposes. The smaller log-likelihood ratio, larger R^2 , larger adj. R^2 , larger F-value, and smaller standard error of regression indicates that the heteroscedasticity-corrected-error models (HCE OLS) were better than the ordinary least squares (OLS) models while on the basis of larger individual R^2 and larger adjusted R^2 , the log-log model slightly outperformed other models (Table 3 & 4). However, by considering all criteria, including the explanatory aspect (number of statistically significant factors), the lin-log model ($\lambda_1=1, \lambda_2=0$) was the best model to relate *R. kerrii*'s bud growth against its pertinent ecological factors. On the basis of this model, about 94% variation in bud growth (i.e.

Table 3
Outputs of Ordinary Least Squares (OLS) and Heteroscedasticity-Corrected Errors' (HCE) Regressions – Dep. Bud Diameter (DIAM)

	LOG-LOG ($\lambda_1=\lambda_2=0$)		LOG-LIN ($\lambda_1=0, \lambda_2=1$)		LIN-LOG ($\lambda_1=1, \lambda_2=0$)		LIN-LIN ($\lambda_1=\lambda_2=1$)	
	OLS	HCE	OLS	HCE	OLS	HCE	OLS	HCE
R ²	0.879	0.965	0.902	0.958	0.673	0.938	0.778	0.885
Adj. R ²	0.841	0.954	0.870	0.945	0.568	0.919	0.707	0.849
F-value	22.773	86.980	28.751	71.637	6.429	47.756	10.938	24.135
S.E. Regression	0.204	1.473	0.184	1.750	5.328	1.128	4.391	1.345
Log-likelihood	11.071	-56.175	14.601	-62.043	-99.896	-44.111	-93.325	-53.089
Sample size	34	34	34	34	34	34	34	34
Intercept	-56.5876	-75.7037	-10.715	-10.0122	-348.109	-763.244	-38.4796	-98.3742
	(-1.004)	(-1.704)*	(-1.435)	(-1.714)*	(-0.2362)	(-1.741)*	(-0.2156)	(-1.099)
Elevation (ELE), m	10.7611	11.0270	0.0116	0.0159	244.963	-11.436	0.1970	0.0243
	(0.7228)	(1.203)	(0.9016)	(1.887)*	(0.6293)	(-0.092)	(0.6377)	(0.1561)
Temperature (TEMP), °C	-0.0199	0.4154	0.0329	-0.0589	-56.7428	67.706	-1.4148	0.0783
	(-0.0058)	(0.1588)	(0.2702)	(-0.5995)	(-0.6382)	(2.248)**	(-0.4863)	(0.0691)
Light shading (SHADE), %	0.6590	0.6314	0.0152	0.0028	-6.0056	22.552	0.1745	0.1241
	(0.4944)	(0.8286)	(1.006)	(0.3199)	(-0.1723)	(1.755)*	(0.4817)	(0.8825)
Air humidity (HUMID), %	-5.3590	-1.9490	-0.0517	-0.0726	-268.081	108.150	-2.1030	0.4877
	(-0.4060)	(-0.3284)	(-0.3678)	(-0.9304)	(-0.7768)	(0.893)	(-0.6259)	(0.3366)
Soil acidity (pH)	0.9854	1.5315	0.2763	0.2743	5.6114	14.257	3.2086	2.2211
	(1.446)	(5.639)**	(2.108)**	(4.152)**	(0.3150)	(1.887)*	(1.024)	(1.644)
Number of host plant's visible roots (ROOTS)	1.9283	1.8977	0.2905	0.2807	27.3023	17.977	4.8819	3.0831
	(7.159)**	(14.420)**	(8.336)**	(16.82)**	(3.877)**	(7.606)**	(5.859)**	(7.224)**
Termites infestation (TERMIT), yes = 1, no = 0	-0.0412	-0.0155	-0.0159	0.0138	-2.1361	-0.761	-2.1717	-0.3386
	(-0.4418)	(-0.571)	(-0.1917)	(0.2737)	(-0.8765)	(-1.094)	(-1.095)	(-0.4252)
Survival condition x Growth stage (CONSTA)	-0.062	-0.0756	-0.0995	-0.0786	-1.3311	-2.235	-4.3904	-1.3996
	(-0.4088)	(-1.228)	(-0.7331)	(-1.162)	(-0.3363)	(-2.248)**	(-1.353)	(-1.072)

Notes: Figures in brackets are t-values as follows: ** significant at $\alpha = 0.05$ or better, * significant at $\alpha = 0.1$

Table 4
Outputs of Ordinary Least Squares (OLS) and Heteroscedasticity-Corrected Errors' (HCE) Regressions – Dep. Bud Circumference (circm)

	LOG-LOG ($\lambda_1=\lambda_2=0$)		LOG-LIN ($\lambda_1=0, \lambda_2=1$)		LIN-LOG ($\lambda_1=1, \lambda_2=0$)		LIN-LIN ($\lambda_1=\lambda_2=1$)	
	OLS	HCI	OLS	HCI	OLS	HCI	OLS	HCI
R ²	0.879	0.965	0.902	0.959	0.673	0.939	0.778	0.885
Adj. R ²	0.841	0.954	0.871	0.945	0.568	0.919	0.707	0.849
F-value	22.776	86.415	28.755	72.403	6.429	47.754	10.938	24.125
S.E. Regression	0.204	1.471	0.184	1.757	16.737	1.130	13.796	1.345
Log-likelihood	11.073	-56.140	14.603	-62.171	-138.817	-47.162	-132.246	-53.099
Sample size	34	34	34	34	34	34	34	34
Intercept	-55.4319	-74.5292	-9.5677	-8.8392	-1093.53	-2413.22	-120.855	-308.970
Elevation (ELE), m	(-0.9834)	(-1.6780)*	(-1.2820)	(-1.513)	(-0.2362)	(-1.755)*	(-0.2155)	(-1.098)
Temperature (TEMP), °C	10.7605	11.0363	0.0116	0.0160	769.594	-34.1989	0.6188	0.0759
Light shading (SHADE), %	(0.7228)	(1.2040)	(0.9016)	(1.899)*	(0.6293)	(-0.0885)	(0.6378)	(0.1553)
Air humidity (HUMID), %	-0.0225	0.4156	0.0328	-0.0603	-178.286	213.7030	-4.4457	0.2479
Soil acidity (pH)	(-0.0066)	(0.1590)	(0.2694)	(-0.6136)	(-0.6382)	(2.2590)**	(-0.4864)	(0.0697)
Number of host plant's visible roots (ROOTS)	0.6588	0.6305	0.0152	0.0027	-18.8691	71.1368	0.5481	0.3904
Termites infestation (TERMIT), yes = 1, no = 0	(0.4942)	(0.8628)	(1.0060)	(0.3099)	(-0.1723)	(1.7620)*	(0.4817)	(0.8835)
Survival condition x Growth stage (CONSTA)	-5.3585	-1.9683	-0.0517	-0.0738	-842.235	339.515	-6.6072	1.5345
	(-0.4060)	(-0.3314)	(-0.3678)	(-0.9640)	(-0.7768)	(0.8920)	(-0.6259)	(0.3372)
	0.9856	1.5288	0.2764	0.2744	17.6337	44.8569	10.0810	6.9844
	(1.4470)	(5.6150)**	(2.1080)**	(4.1640)**	(0.3150)	(1.8900)*	(1.0240)	(1.6450)
	1.9284	1.8970	0.2905	0.2808	85.7742	56.4202	15.3370	9.6900
	(7.160)**	(14.4000)**	(8.3370)**	(16.880)**	(3.8770)**	(7.5930)**	(5.8590)**	(7.225)**
	-0.0412	-0.0153	-0.0159	0.0139	-6.7108	-2.3748	-6.8225	-1.0694
	(-0.4418)	(-0.5610)	(-0.1917)	(0.2744)	(-0.8765)	(-1.0850)	(-1.095)	(-0.4252)
	-0.0619	-0.0755	-0.0996	-0.0781	-4.1822	-7.0254	-13.7930	-4.4087
	(-0.4091)	(-1.2240)	(-0.7334)	(-1.1154)	(-0.3363)	(-2.2480)**	(-1.353)	(-1.0750)

Notes: Figures in brackets are t-values as follows: ** significant at $\alpha = 0.05$ or better, * significant at $\alpha = 0.1$

growth in diameter and circumference) in *R. kerrii* was explained by the eight specified ecological factors leaving about 6% variation in the species' growth remained unexplained by the model.

Other factors that could have influenced bud's growth were soil moisture, soil temperature, soil nutrients, and anthropogenic activity, etc. but they could not be investigated due to resource constraints. Looking across all the models, the number of host plant's visible roots (ROOT) was a strong ($\alpha=0.05$ or better) ecological factor influencing *R. kerrii*'s bud growth. This variable is a proxy that measures the host plant's ecological ability (plant size, vigour, and capacity to provide base) for the proliferation of *Rafflesia*. This seems to be in support of the view that *Rafflesia* highly correlates with their host plants' conditions (Akhriadi et al., 2010). In a recent study, it was found that the number of *Rafflesia*'s buds increased with the class-size of the host plant (Kedri et al., 2018). The lin-log model also indicates that temperature (TEM) and plant's survival x growth condition (CONST) were strongly ($\alpha=0.05$ or better) influencing bud growth while light shading (SHADE) and soil acidity (pH) could have been important growth factors as well ($\alpha=0.10$).

Based on the lin-log HEC model, two related quantities of effects of ecological factors on *R. kerrii*'s bud growth can be computed. First, the marginal effect of factor, i.e. $\delta Y/\delta X = \beta \times 1/X$; and second, the growth sensitivity to ecological factor, i.e. $\zeta = (\delta Y/\delta X) \times (X/Y) = \beta/Y$, where β = regression coefficient; X = value of ecological factor; and Y = value of bud growth). Using the sample's average values (Table 1 & 4 for base figures), these effects were calculated and shown in Table 5.

From Table 5, a less acidic soil (higher pH value) could have contributed the most to *R. kerrii*'s marginal bud growth, followed by a better ecological ability of the host plant, and a higher temperature. Light shading exerted quite a small amount to bud's marginal growth while the interaction between the plant's survival condition and growth stage has tended to distract bud growth. However, this could have depended indirectly on some other ecological aspects. For example, apart from its ability to reach advanced growth stages, such as stages II and III, chances of plant's survival condition could have also been influenced by plant abundance on a particular site. In turn, the ability to reach such growth stages could have been determined by many other factors, such as the absence of human and/or animal disturbance throughout the growth period.

Bud growth was highly sensitive to temperature ($\zeta = 5.9$) that a 1% increase in temperature could have encouraged 2.66 cm and 8.40 cm in bud's growth. The latter was also sensitive to light shading ($\zeta = 1.97$) that a 1% increase in light shading could have increased 0.28 cm and 0.9 cm in bud's growth. This was followed by host plant's ecological ability ($\zeta = 1.57$) to support *R. kerrii* whereby a 1% increase in it could have induced 2.74 cm and 8.6 cm of bud's growth. Bud' growth was less sensitive to soil acidity ($\zeta = 1.24$), but a 1% increase in it could have induced 2.94 cm and 9.27 cm in bud's growth.

Table 5
Estimated effects of pertinent ecological factors on bud growth

Ecological factors	Bud diameter (cm)		Bud circumference (cm)	
	Marginal effect of ecological factor			
Host plant's ecological ability (ROOT)	17.977 x 1/6.56	= 2.74	56.4202 x 1/6.56	= 8.60
Temperature (TEM)	67.706 x 1/25.44	= 2.66	213.7030 x 1/25.44	= 8.40
Light shading (SHADE)	22.552 x 1/79.32	= 0.28	71.1368 x 1/79.32	= 0.90
Soil acidity (pH)	14.257 x 1/4.84	= 2.94	44.8569 x 1/4.84	= 9.27
Plant's survival condition x growth stage (CONST)	-2.235 x 1/0.32	= -6.98	-7.0254 x 1/0.32	= -21.95
Elevation		n.s.		n.s.
Humidity		n.s.		n.s.
Termite infestation		n.s.		n.s.
	Growth sensitivity* to ecological factor			
Host plant's ecological ability (ROOT)	17.977/11.46	= 1.57	56.4202/36.00	= 1.57
Temperature (TEM)	67.706/11.46	= 5.91	213.7030/36.00	= 5.94
Light shading (SHADE)	22.552/11.46	= 1.97	71.1368/36.00	= 1.98
Soil acidity (pH)	14.257/11.46	= 1.24	44.8569/36.00	= 1.25
Plant's survival condition x growth stage (CONST)	-2.235/11.46	= -0.20	-7.0254/36.00	= -0.20
Elevation		n.s.		n.s.
Humidity		n.s.		n.s.
Termite infestation		n.s.		n.s.

* Value > 1.0 was 'sensitive'; the larger the value the more sensitive was bud growth to a particular ecological factor. n.s. = not statistically significant

Rather sensitive to host plant's ecological ability ($\zeta = 1.57$), bud's growth can potentially increase by 2.74 cm and 8.60 cm respectively by a 1% increase in the number of host plant's roots/vines. Although not easy to explain, expanded host plant's ecological ability could have provided a better niche for this parasitic plant to grow. By contrast, bud's growth was insensitive to the interaction of *Rafflesia kerrii*'s survival condition and its growth stage. Still, the negative marginal effect of this ecological factor on bud's growth could have been substantial. The struggle of the plant to grow throughout its life stages tended to have suppressed the plant so much so causing the diameter and circumference of the bud to have reduced by 6.98 cm and 21.95 cm, respectively.

CONCLUSION

Based on the cross-sectional study, it was shown that the regression model can be applied to objectively estimate the effects of ecological factors on *Rafflesia kerrii*'s bud's growth. Host plant's ecological ability, level of temperature, light shading, soil acidity, and interaction between plant survival condition and growth stage were found to be significant

and influential ecological factors to *R. kerrii*'s bud growth. Future studies can further corroborate this finding by explicitly investigating the ideal levels of these factors that will maximize *R. kerrii*'s bud growth. Termite infestation was said to be the leading cause of failure in *Rafflesia*'s bud growth (Kedri et al., 2018). However, the small sample size has inhibited a positive result of this factor and, thus, future studies may focus on the role/effect of not only termites but also other organisms on *Rafflesia*'s bud growth. Further, with a small sample, there was insufficient evidence of statistically significant random variation in the influence of altitude and air humidity on *R. kerrii*'s bud growth. Further studies may re-assess the ecological impact of these factors using a larger sample covering a larger geographic area – possibly across different localities.

LIMITATION OF THE STUDY

A few factors deserve consideration for future study. First, the ecological factors other than those included in this study should be further investigated as to their role in influencing *Rafflesia*'s bud growth. There might be other pertinent *in-situ* micro factors determining bud growth of *R. kerrii*, but their cross-sectional measurements were difficult. Second, the length of time of experiment on soil moisture and light intensity, for example, need to be gauged over a long period of time (say a few hundred days) to cover periodic water-abundant and water-scarce periods (rainy and dry months) as well as possible diurnal variations. Light intensity/shading needs a specific experiment on how *Rafflesia* could have responded to it over a period of time. Perhaps, daily, weekly, or monthly micro temperature variations need also be recorded throughout the study period. The same goes for day-length photoperiod, which was difficult to measure due to resource constraints.

Third, the anthropogenic threats to bio-diversity which is understandable, but they are difficult to assess because the degree of trampling and disturbance by human or animals (e.g. intensity, frequency or severity, length of time, etc.) on individual *Rafflesia* plants were not easy to quantify. The negative marginal effect of this ecological factor on *Rafflesia*'s bud growth was quite substantial; it reduced bud's diameter and circumference by 6.98 cm and 21.95 cm, respectively, but this was considered an indirect factor. The possible direct factors will require some detailed anthropogenic studies to be carried out. Fifth, the forest soils are known to be deficient in many micro- and macro-nutrients, however, *in-situ* quantification of soil nutrients was quite troublesome due to resource constraints since it could have involved measurements and analysis of many nutrient elements. Sixth, the factors affecting the host plant, *Tetrastigma*, must be spatially correlated with the survival of *Rafflesia*'s over a large geographic area. The ratio of male and female flowers that may influence the rate of pollination needs to be modelled over a sufficiently large geographic region to provide a sufficient random variation to explain its influence on the plant species' multiplication. All these should provide an avenue for more statistical-based analyses in future.

ACKNOWLEDGEMENT

The authors would like to thank Universiti Malaysia Kelantan for the physical resources provided throughout this research. Special thanks to Majlis Daerah Gua Musang and Pejabat Tanah dan Jajahan Kecil Lojing for their helps and assistance to conduct this research at Tanah Tinggi Lojing, Gua Musang.

REFERENCES

- Akhriadi, P., Kiswanto, H. A., Taufiq, A., Alfajri, D., & Kardiman, R. (2010). Assessment of conservation status of *Rafflesia* in West Sumatra, Indonesia. *Final Report to Rufford Small Grant (For Nature Conservation)*. Rafflesia Monitoring Team (RMT) Padang, Indonesia.
- Anon. (2003). *Biological Diversity 2003 - The genus Rafflesia*. Biology 226 class project on the conservation of global biodiversity, Earlham College. Retrieved December 17, 2019, from <http://legacy.earlham.edu/~givenbe/Rafflesia/rafflesia/biodiv2.htm>.
- Balete, D. S., Pelsler, P. B., Nickrent, D. L., & Barcelona, J. F. (2010). *Rafflesia verrucosa* (Rafflesiaceae), a new species of small-flowered *Rafflesia* from eastern Mindanao, Philippines. *Phytotaxa*, 10(1), 49-57.
- Bareja, B. G. (2011). *Biotic factors and their interaction with plant*. Retrieved December 17, 2019, from <https://www.cropsreview.com/biotic-factors/html>.
- Barkman, T. J., Lim, S. H., Salleh, K. M., & Nais, J. (2004). Mitochondrial DNA sequences reveal the photosynthetic relatives of *Rafflesia*, the world's largest flower. *Proceedings of the National Academy of Sciences*, 101(3), 787-792. <https://doi.org/10.1073/pnas.0305562101>
- Beaman, R. S., Decker, P. J., & Beaman, J. H. (1988). Pollination of *Rafflesia* (Rafflesiaceae). *American Journal of Botany*, 75(8), 1148-1162. <https://doi.org/10.1002/j.1537-2197.1988.tb08828.x>
- Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression Diagnostics: Identifying influential data and sources of collinearity*. John Wiley. <http://dx.doi.org/10.1002/0471725153>.
- Chu, Z. F., Wen, J., Yang, Y. P., Nie, Z. L., & Meng, Y. (2018). Genome size variation and evolution in the grape family *Vitaceae*. *Journal of Systematics and Evolution*, 56(4), 273-82. <https://doi.org/10.1111/jse.12310>
- Cotrell, A. & Luccetthi, R. J. (2005). *Gretls user's guide*. Gnu Regression, Econometrics and Time Series.
- Faust, J. E., & Lewis, K. P. (2005). (66) Modeling Flower Bud Development of *Impatiens hawkeri* and *I. walleriana*. *HortScience*, 40(4), 1013A-1013. <https://doi.org/10.21273/HORTSCI.40.4.1013A>
- Faye, F. (2008). *Habitat comparison of Rafflesia keithii in village area* [Unpublished B.Sc. thesis]. Universiti Malaysia Sabah.
- Galindon, J. M. M., Ong, P. S. & Fernando, E. S. (2016). *Rafflesia consueloae* (Rafflesiaceae), the smallest among giants; a new species from Luzon Island, Philippines. *PhytoKeys*, (61), 37-46. <https://doi.org/10.3897/phytokeys.61.7295>.
- Greene, W. H. (1990). *Econometric analysis*. Mcmillan Publishing Company.

- Grenz, J., Manschadi, A. H., Uygur, N., & Sauerborn, J. (2005). Effects of environment and sowing date on the competition between faba bean (*Vicia faba*) and the parasitic weed *Orobanche crenata*. *Field Crop Research*, 93(2-3), 300-313. <https://doi.org/10.1016/j.fcr.2004.11.001>
- Griffiths, E. G., Hill, R. C., & Judge, G. G. (1993). *Learning and practicing econometrics*. John Wiley & Sons, Inc.
- Hakkinen, R. (1999). Statistical evaluation of bud growth theories: Application to bud burst of betula pendula leaves. *Tree Physiology*, 19(9), 613-618. <https://doi.org/10.1093/treephys/19.9.613>
- Häkkinen, R., Linkosalo, T., & Hari, P. (1998). Effects of dormancy and environmental factors on timing of bud burst in *Betula pendula*. *Tree Physiology*, 18(10), 707-712. <https://doi.org/10.1093/treephys/18.10.707>
- Hautier, Y., Hector, A., Vojtech, E., Purves, D., & Turnbull, L. A. (2010). Modelling the growth of parasitic plants. *Journal of Ecology*, 98(4), 857-866. <https://doi.org/10.1111/j.1365-2745.2010.01657.x>
- Hidayati, S. N., & Walck, J. L. (2016). A review of the biology of *Rafflesia*: What do we know and what's next? *Botanic Garden Bulletin*, 19(2), 67-78.
- Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics*, 19(3), 285-292.
- Inger, R. F., & Lian T. F. (1996). *The natural history of amphibians and reptiles in Sabah*. Natural History Publication (Borneo).
- Jamili, N., & Wilcock, C. C. (1998). The *Rafflesia* conservation incentive scheme in Sabah, Malaysian Borneo. *Sabah Parks Nature Journal*, 1, 9-17.
- Jamili, N., Baikan, B., & Sinun, W. (2000). Nature conservation: striking a balance between conservation and development for Sabah beyond 2000. *Sabah Beyond*, 181-202.
- Jamili, N. (2001a). *Rafflesia of the world*. Natural History Publications (Borneo).
- Jamili, N. (2001b). Life cycle and flower bud growth. In *Rafflesia of the world* (p. 13). Natural History Publications (Borneo).
- Jumaat, H. A., Afiq, A. J. M., Rahmah, M., Wahab, N. A. A., Syamsurina, A., Paiz, K. M., Raih, M. F. M., & Wan, K. L. (2016). *Rafflesia tuanku-halimii* (Rafflesiaceae), a new species from Peninsular Malaysia. *Sains Malaysiana*, 45(11), 1589-1595.
- Kanchanapoom, T., Kamel, M. S., Picheansoonthon, C., Luecha, P., Kasai, R., & Yamasaki, K. (2007). Hydrolyzable tannins and phenylpropanoid from *Rafflesia kerrii* Meijer (Rafflesiaceae). *Journal of Natural Medicines*, 61(4), 478-479. <https://doi.org/10.1007/s11418-007-0181-4>
- Kedri, F. K., Hamzah, Z., Sukri, N. S., Yaacob, S. H., Abd Majid, N. K. S., Mokhtar, N., & Amir, S. F. (2018). Distribution and ecology of *Rafflesia* in Royal belum state park, Perak, Malaysia. *International Journal of Engineering and Technology*, 7(2.29), 292-296.
- Koya, P. R., & Goshu, A. T. (2013). Generalized mathematical model for biological growths. *Open Journal of Modelling and Simulation*, 1, 42-53. <https://doi.org/10.4236/ojmsi.2013.14008>
- Kramer, K. (1994). Selecting a model to predict the onset of growth of *Fagus sylvatica*. *Journal of Applied Ecology*, 31(1), 172-181. <https://doi.org/10.2307/2404609>

- Kumar, S. (2018). *Ecological factors that affect the growth of plants*. Retrieved on December 17, 2019, from <http://www.biologydiscussion.com/growth-of-plants/ecological-factors-that-affect-the-growth-of-plants-with-diagrams/15288>.
- Latifah, D., Riswati, M. K., Handini, E., & Wawangningrum, H. (2017). Viability test on the seeds of *Rafflesia arnoldii* R. Br. & *R. patma* Blume. *Buletin Kebun Raya*, 20(1), 25-32.
- Latiff, A. (2018). Viability of having the gigantic *Rafflesia* flower in our park. *Open Access Journal of Science*, 2(2), 106-107.
- Linksolo, T., Carter, T. R., Hakkinen, R., & Hari, P. (2000). Predicting spring phenology and frost damage risk of *Betula spp.* under climatic warming: A comparison of two models. *Tree Physiology*, 20(17), 1175-1182. <https://doi.org/10.1093/treephys/20.17.1175>
- Long, J. S., & Ervin, L. H. (1998). Correcting for heteroscedasticity with heteroscedasticity consistent standard errors in the linear regression model: small sample considerations. *Indiana University, Bloomington, IN*, 47405, 1-33.
- MacKinnon, J. G., & White, H. (1985). Some heteroskedasticity-consistent covariance matrix estimators with improved finite sample properties. *Journal of Econometrics*, 29(3), 305-325. [https://doi.org/10.1016/0304-4076\(85\)90158-7](https://doi.org/10.1016/0304-4076(85)90158-7)
- Maddala, G. S. (1977). *Econometrics*. McGraw-Hill Book Company Publishing.
- Maddala, G. S. (1992). *Introduction to econometrics* (2nd Ed.). Maxwell Macmillan International Publishing.
- Mahyuni, R., Kusuma, Y. W. C., Wihermanto, W., & Veldkamp, J. (2015). Notes on *Rafflesia* (Rafflesiaceae) In Sumatra with A New Record *Rafflesia Gadutensis* Meijer. *Reinwardtia*, 14(2), 317-322. <https://doi.org/10.14203/reinwardtia.v14i2.1678>
- Martin, G. (1998). *Etnobotany*. Natural History Publications (Borneo).
- Medel, R. (2000). Assessment of parasite-mediated selection in a host-parasite system in plants. *Ecology*, 81(6), 1554-1564. [https://doi.org/10.1890/0012-9658\(2000\)081\[1554:AOPMSI\]2.0.CO;2](https://doi.org/10.1890/0012-9658(2000)081[1554:AOPMSI]2.0.CO;2)
- Meijer, W. (1984). New species of *Rafflesia* (Rafflesiaceae). *Blumea*, 30(1), 209-215.
- Mello, M. A., & Tuan, R. S. (1999). High density micromass cultures of embryonic limb bud mesenchymal cells: an in vitro model of endochondral skeletal growth. *In Vitro Cellular & Growth Biology-Animal*, 35(5), 262-269. <https://doi.org/10.1007/s11626-999-0070-0>
- Milon, J. W., Gressel, J., & Mulkey, D. (1984). Hedonic amenity valuation and functional form specification. *Land Economics*, 60(4), 378-87. <https://doi.org/10.2307/3145714>
- Molina, J., Hazzouri, K. M., Nickrent, D., Geisler, M., Meyer, R. S., Pentony, M. M., Flowers, J. M., Pelsler, P., Barcelona, J., Inovejas, S. A., Uy, I., Yuan, W., Wilkins, O., Michel, C. I., LockLear, S., Concepcion, G. P., & Purugganan, M. D. (2014). Possible loss of the chloroplast genome in the parasitic flowering plant *Rafflesia lagascae* (Rafflesiaceae). *Molecular Biology and Evolution*, 31(4), 793-803. <https://doi.org/10.1093/molbev/msu051>
- Montgomery, R. A. (2004). Effects of understory foliage on patterns of light attenuation near the forest floor. *Biotropica*, 36(1), 33-39. <https://doi.org/10.1111/j.1744-7429.2004.tb00293.x>

- Moreau, D., Gibot-Leclerc, S., Giardin, A., Pointurier, O, Reibel, C., Strbik, F., Fernandez-Aparicio, M., & Colbach, N. (2016). Trophic relationships between the parasitic plant species *Phelipanche ramosa* (L.) and different hosts depending on host phenological stage and host growth rate. *Frontiers in Plant Science*, 7, 1-12. <https://doi.org/10.3389/fpls.2016.01033>
- Mursidawati, S. (2015). Morphology of fruits and seeds of *Rafflesia patma* and *R. arnoldii*. *Buletin Kebun Raya*, 15 (1), 21-30.
- Mursidawati, S. (2017). *Ex-situ* conservation of *Rafflesia patma* (Rafflesiaceae) in Bogor Botanical Gardens (Indonesia). *Nature Conservation Research*, 2(2), 90-91.
- Mursidawati, S., & Risawati, M. K. (2009, July 14). *Conservation biology of holoparasit plant: in vitro seeds inoculation of Rafflesia patma* [Paper presentation]. Proceedings of a Seminary on Flora Coconservation in Addressing the Impact of Global Warming, Tabahan, Bali, Indonesia.
- Nasihah M., Zulhazman H., Siti-Munirah M. Y., Razak W., Fatimah, K., & Qayyum-Nadia, W. A. (2016). *Tetrastigma hookeri* (Laws.) Planch (*Vitaceae*), a host plant for *Rafflesia kerrii* Meijer in Peninsular Malaysia. *Malayan Nature Journal*, 68(Part 1 & 2), 33-39.
- Nasihah, M., Zulhazman, H., Siti-Munirah, M. Y., Razak, W., Fatimah, K., & Qayyum-Nadia W. A. (2013). *Soil nutrient dynamic at three defined elevations in relation to bud population of Rafflesia kerrii in Lojing Highland of Kelantan, Malaysia*. Universiti Malaysia Kelantan Press.
- Nery, S., Kamarudin, M. S., Khairil, M., Nor-Zuhailah, M. N., Hasein, M. R. A., & Burslem, D. F. R. P. (2016). *Rafflesia parvimaculata* (Rafflesiaceae), a new species of *Rafflesia* from Peninsular Malaysia. *Phytotaxa*, 253(3), 207-213. <https://doi.org/10.11646/phytotaxa.253.3.4>
- Nickrent, D. L., Blarer, A., Qiu, Y. L., Vidal-Russell, R., & Anderson, F. E. (2004). Phylogenetic inference in Rafflesiales: The influence of rate heterogeneity and horizontal gene transfer. *BMC Evolutionary Biology*, 4(40), 1-17. <https://doi.org/10.1186/1471-2148-4-40>
- Osawa, A., Shoemaker, C. A., & Stedinger, J. R. (1983). A stochastic model of balsam fir bud phenology utilizing maximum likelihood parameter estimation. *Forest Science*, 29(3), 478-490. <https://doi.org/10.1093/forests/29.3.478>
- Paine, C. T., Marthews, T. R., Vogt, D. R., Purves, D., Rees, M., Hector, A., & Turnbull, L. A. (2012). How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods in Ecology and Evolution*, 3(2), 245-256. <https://doi.org/10.1111/j.2041-210X.2011.00155.x>
- Peters, R. F., & Ting, Y. Y. (2016). Protection of *Rafflesia* through the appreciation of the Dusun's indigenous knowledge: A preliminary case study at Poring-Sabah. *Journal of Tropical Biology and Conservation*, 13, 27-42.
- Qayyum, W. A., Yunoh, S. M. M., Hamzah, Z., & Wahab, R. (2012, July 9-11). *A preliminary study on the mortality rates of Rafflesia kerrii in the Lojing Highlands, Gua Musang, Kelantan, Malaysia* [Paper presentation]. UMT 11th International Annual Symposium on Sustainability Science and Management, Terengganu, Malaysia.
- Ramamoorthy, R., Phya, E. E. K., Lim, S. H., Tan, H. T. W., & Kumar, P. P. (2013). Identification and characterization of RcMADS1, an AGL24 Ortholog from the Holoparasitic Plant *Rafflesia cantleyi* Solms-Laubach (Rafflesiaceae). *PLoS ONE*, 8(6), Article e67243. <https://doi.org/10.1371/journal.pone.0067243>

- Salleh, K. M. (2007). *Rafflesia*. Borneo Publishing Company.
- Saxe, H., Cannell, M. G., Johnsen, Ø., Ryan, M. G., & Vourlitis, G. (2001). Tree and forest functioning in response to global warming. *New Phytologist*, 149(3), 369-399. <https://doi.org/10.1046/j.1469-8137.2001.00057.x>
- Shaw, J. (2017, March-April). Colossal blossom - Pursuing the peculiar genetics of a parasitic plant. *Harvard Magazine*. Retrieved December 19, 2019, from <https://harvardmagazine.com/2017/03/colossal-blossom>.
- Susatya, A. (2011). *Rafflesia: The world's largest flower charm*. Directorate for conservation areas and protected forest growth. Department of Forestry, Jakarta, Indonesia.
- Veldkamp, J. F. (2009). Notes of the names of the *Tetrastigma* (*Vitaceae*) host of *Rafflesia* (*Rafflesiaceae*). *Reinwardtia*, 13(1), 75-78. <https://doi.org/10.14203/reinwardtia.v13i1.431>
- Wicaksono, A., & Da Silva, J. A. T. (2015). Short communication: Attempted callus induction of holoparasite *Rafflesia patma* Blume using primordial flower bud tissue. *Nusantara Bioscience*, 7(2), 96-101.
- Wikipedia. (2018). *Heteroscedasticity-consistent standard errors*. Retrieved December 24, 2019, from <https://en.m.wikipedia.org/>

Ecological Roles of Benthic Community in Seafloor Northwards of Pulau Indah, Klang, Malaysia

Mohd Sophian Mohd Kasihmuddin* and Zaidi Che Cob

Department of Earth Sciences and Environment, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 UKM, Bangi, Selangor

ABSTRACT

Trait assessment is an alternative but effective method of biomonitoring benthic community in ecosystem, but it was rarely done in Malaysia. An assessment was done to identify the dominant sub-traits, or modalities, within benthic community in seafloor northwards of Pulau Indah, Klang, Malaysia, and to establish possible correlation between modalities with environmental parameters measured in the area. Sampling was conducted in November 2018, during Northwest Monsoon season. Sediments were obtained by using a Ponar Grab and filtered through 500 µm sieve. Sediments were subjected into Loss of Ignition method to extract organic carbon, and filtration method to determine particle size distribution percentages. 37 taxa were recorded in this study, totalling to 848 individuals. Traits assessment indicated vermiform body shape, direct larval development, deposit feeder, non-motile, asexual reproduction, and second-order opportunistic group as most dominant modalities in the study area. Station 1 recorded highest macrobenthos density (21000 ind./m²), while station 5 recorded highest diversity ($H' = 2.19$). Functional diversity (FD) for reproduction method was highest in station 1 ($FD = 0.61$), ecological group in station 4 ($FD = 0.73$); body shape ($FD = 0.56$) and larval development ($FD = 0.63$) in station 5, and lastly, adult mobility ($FD = 0.70$) and feeding mode ($FD = 0.66$) in station 6. Principal component analysis showed Station 1 has most modalities attributed towards it. Pearson correlation analysis suggested particle size distribution as major connector for most modalities ($P = 0.05$). As BTA study is still new in Malaysian waters, it could serve as a good starting point for future assessment related to ecological functioning of benthic community in Malaysian seafloor.

ARTICLE INFO

Article history:

Received: 03 December 2020

Accepted: 05 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.31>

E-mail addresses:

sophianmkmaidin@gmail.com (Mohd Sophian Mohd Kasihmuddin)

zdcc@ukm.edu.my (Zaidi Che Cob)

* Corresponding author

As BTA study is still new in Malaysian waters, it could serve as a good starting point for future assessment related to ecological functioning of benthic community in Malaysian seafloor.

Keywords: Biological traits analysis, ecosystem monitoring, macrobenthos, modalities, polychaete, traits

INTRODUCTION

Benthic community distributes very differently across the seafloor of the world (Lomovasky et al., 2011). This variation occurs due to the nature of the benthos themselves. Characteristic of the benthic organisms can define what they are capable of, and how they respond to changes in the environment, as well as how do they engage on forage and procreation (Hou et al., 2020; Widdows et al., 2000). These, known as traits, are unique and only found in certain group of taxa (Verberk et al., 2013). Taxa from higher hierarchy taxonomy may possess similar traits, but lower taxonomic level taxon may display different traits.

Traits found in available taxa could serve as a useful tool to understand the nature of the biotope taxa is found in better perspective (Zhong et al., 2018). Unique traits displayed by the taxa can denote their role toward the ecosystem. Traits shown in taxa denote their feeding mode, territoriality, and reproductive strategy (Zhong et al., 2018). Benthic community depends heavily on the seafloor habitat and does not have effective displacement method to migrate to another seafloor region. Depending on localities and existing environmental factors, diversity of traits shown at the seafloor biotope may vary. Study by Paganelli et al. (2012) on the connection between existing traits in benthic community with their habitat's distance to the coastline of Emilia-Romagna, Italy, found majority of taxa in nearshore (~3 km) area was dominated by mobile vermiform organisms with tube-dwelling and deposit-feeding behavior. However, further off the coastline (< 10 km), the dominant taxa were either laterally compressed or globose body shape, and majority of them possess deposit-feeding and filter-feeding as dominant traits in terms of feeding behavior.

Trait assessment has become increasingly popular, but it has rarely been done in Malaysian waters (Hanamura et al., 2008; Hussin, 2014). This is because identifying the traits can be time-constraining, and details of targeted taxa can be very vague due to the lack of accurate reference about the benthic taxa (Korte & Ashley, 2013). Most of ecological assessments done in biodiversity in Malaysian waters were conducted in traditional approaches via identifying inhabiting taxa in-site but may or may not disregard important factor behind the taxa such as origin of taxa (nature vs. invasive), feeding behaviour, and body shape.

Resultant observation on traits from taxa can be further applied on discerning correlation between governing traits with surrounding area. Known as Biological Traits Analysis (BTA), this analysis aims to establish connection between traits, taxa, and the surrounding environment (Linden et al., 2012). BTA requires compilation of sub-traits within traits in form of matrices, followed by ordination analysis such as Fuzzy Correspondence Analysis (FCA) or Principal Component Analysis (PCA) (Dell et al., 2011). These sub-traits can further be interpreted by ecological mean as direct influence toward the related taxa, and ultimately, the taxon's influence toward or by corresponding environment (Xu et al., 2018).

Bolam et al. (2016) undertook BTA assessment towards benthic community in a licensed disposal dredged material site in northeast coast of England. Their assessment indicated significant differences between site affected by dredged material fluxes and non-affected site over four years period, where constant influx of dredged materials reduces free-living and predator traits over time. Both traits were associated with complex taxa such as crustaceans, errant polychaetes, ophiuroids, and sea star, where their larger body size and enhanced mobility may enable them to escape from burial by dredged materials. Conversely, influx of dredged materials showed increased taxa capable of performing bioturbation across sediment such as bivalves, gastropods, and sedentary polychaetes. BTA assessments showed large flexibility in terms of interpretation of data, thus this makes a very suitable assessment for various studies regarding ecological issues. Unfortunately, BTA assessment was rarely done in Malaysian waters. All ecological studies towards benthic community in Malaysian waters are all rather assessed via traditional means, through assessing the number of members on each taxon. BTA can be very useful ecological tool to understand the current state of benthic community in Malaysian waters effectively.

This study intends to focus on identifying dominant traits within taxa found in benthic community in various point of seafloor northwards of Pulau Indah, one of the busiest industrial areas in Klang Valley. The findings from these assessments can be applied further for studies related to ecology such as assessing invasive species in enclosed water (Gergs & Rothhaupt, 2015), impact of environmental stressors to environmentally sensitive benthic community over time (Bolam et al., 2017) and genetic drifts for migration of certain organisms over time (Padron & Guizien, 2016). Therefore, the objective of this study is to establish list of traits attributed by available taxa in benthic community inhabiting seafloor in waters northwards of Pulau Indah, Klang, and to assess Biological Trait Analysis of benthic community in the area.

MATERIALS AND METHODS

Procurement of Benthic Specimens and Sediments

The sampling took place in waters northwards of Pulau Indah, Klang. Pulau Indah is an island bustling with industrial plants and human settlements (Figure 1). It is located nearby mangrove island Pulau Ketam, and within waterways connecting important ports such as West Port, North Port and Southpoint. It is also connected with multiple point sources from inland river outflows such as Sungai Klang and Sungai Langat. During the sampling session, sediment samples were acquired by using a Ponar Grab with 0.023 m² mouth area. The salinity and water depth were measured by using a salinometer and a digital depth sounder respectively. Sediment samples were taken for taxa identification and another samples for sediment characteristic analysis. For taxa identification part, the sediment samples were sieved through a 500 µm sieve to obtain the macrobenthic specimen.

Specimens were sorted in laboratory and stored in containers with 70% ethanol for taxa identification. As for sediment analysis, samples were subjected into Total Organic Carbon (TOC) and Particle Size Distribution assessment via Loss of Ignition (LOI) method and filtration method, respectively.

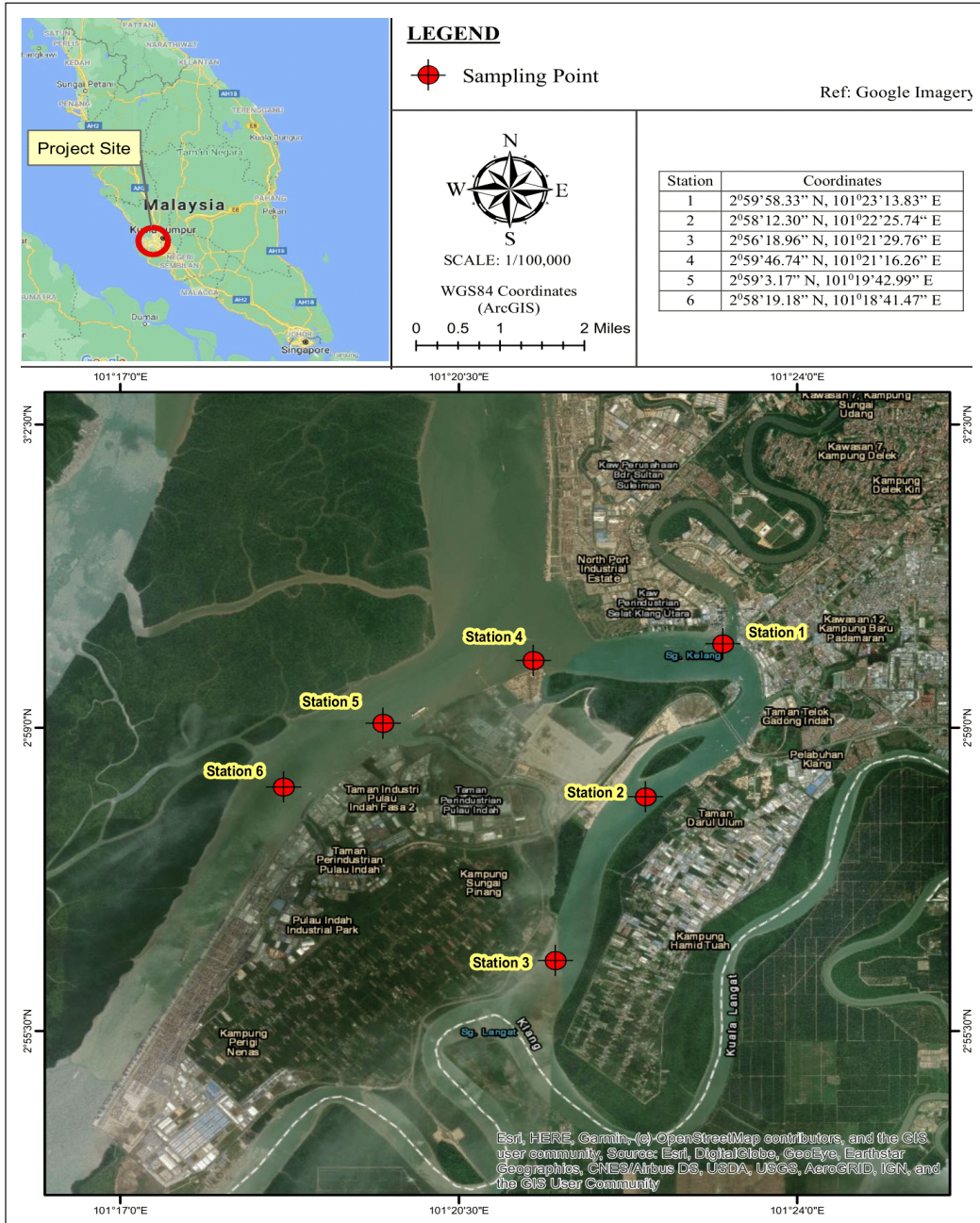


Figure 1. Location of sampling areas in this study

Assessment of Taxa and Their Respective Traits

Macrobenthos specimens were identified to the lowest taxonomic level as possible, preferably genera or species. As for traits assessment, each taxon was classified based on the following traits: body shape, feeding mode, adult mobility, reproductive technique, larval development, and ecological group (Table 1). These modalities or sub-traits were useful to define traits more accurately in terms of their biological and ecological characteristics and their possible role towards ecosystem they inhabit. In this stage, all taxa were compiled into a master list in which each taxon were matched to their corresponding sub-traits. Identification of each taxon and their respective traits were done based on checklist, database, and key descriptions from selected references such as Baharuddin et al. (2018), Fauchald (1977), Faulwetter et al. (2017), Fujita and Irimura (2015), Idris

Table 1
List of traits and modalities used in this study

Traits	Modalities	Label
Body Shape	Dorso-Ventral Compressed	DvC
	Laterally Compressed	LtC
	Globose	Gb
	Vermiform	Vf
Feeding Mode	Filter Feeder	FIFd
	Carnivore/Omnivore	CrOm
	Surface-Deposit Feeder	SDFr
	Sub Surface-Deposit Feeder	SSDFr
	Scavenger	ScV
Adult Mobility	Swimmer	Sw
	Burrower	Bw
	Crawler	Cw
	Walker	Wk
	Non-Motile/Semi Motile	nM
Reproductive Technique	Asexual	Asex
	Sexual, Gonocoric	Sex_G
	Sexual, Heramphrodite	Sex_H
Larval Development	Planktontrophic	Pl
	Lecitotrophic	Le
	Direct Development	DD
	Brooding	Br
Ecological Group (AMBI)	Sensitive	I
	Indifferent	II
	Tolerant	III
	Second Order Opportunistic	IV
	First Order Opportunistic	V

and Arshad (2013), Mikac (2015), Ng and Davie (2002) and Ong and Wong (2015) and online databases such as Polytrait, Animal Diversity (ADW) and World Register of Marine Species (WoRMS).

Trait assessment was done via fuzzy coding, in which each taxon was given a score based on its affinity to the sub-traits. Score 0 represents no affinity, score 1 represents low affinity, score 2 represents high affinity, and score 3 represents absolute affinity (exclusive) towards the sub-trait. The purpose of this coding is to exhibit degrees of variations and diversity within targeted benthic community in the region (Paganelli et al., 2012).

Determination of Ecological Indices and Functional Diversity

In this study, the following ecological indices were used: Shannon-Wiener Index (H'), Pielou's Index (J'), Simpson Index (D) and Margalef Richness Index (D_{mg}). H' was used to assess diversity of benthic taxa in station, J' for assessing evenness, D for taxon's dominance in the region, and D_{mg} for assessing taxa richness (Equation 1, 2, 3, & 4).

$$\text{Shannon-Wiener Index, } H' = -\sum \frac{n_i}{n} \ln \left(\frac{n_i}{n} \right) \quad [1]$$

$$\text{Pielou's Index, } J' = \frac{H'}{H_{\max}} \quad [2]$$

$$\text{Simpson Index, } D = \sum \left(\frac{n_i}{n} \right)^2 \quad [3]$$

$$\text{Margalef Richness Index, } D_{mg} = \frac{(S-1)}{\ln N} \quad [4]$$

n_i denoted individual number for a taxon i , H_{\max} denoted maximum diversity in the region and S denoted number of taxa in the region.

Rao's quadratic entropy was used to measure functional diversity of benthic communities based on traits assessed in this study. Rao's quadratic entropy calculates Functional Diversity (FD) based on taxa abundance and traits gathered in 'Taxa by Station' matrix. Excel's macro feature 'FunctDiv.xls' (<https://botanika.prf.jcu.cz/suspa/FunctDiv.php>) Lepš et al. (2006) was used to obtain the functional diversity for each station. Record of functional diversity (FD) for traits in each station was calculated via Excel macro file developed by Lepš et al. (2006) and shown in Table 4.

Statistical Analysis and Biological Traits Analysis (BTA)

Datasets from the physical parameters, ecological indices and functional diversity were subjected with one-way ANOVA test to identify differences between stations. Pearson

Correlation Analysis was performed on the 'Traits by Station' matrix to identify possible correlation between sub-traits of each trait toward the physical parameters.

To initiate BTA, two matrices were used i.e., 'Taxa by Station' matrix and 'Taxa by Trait' matrix. 'Taxa by Station' matrix comprised of taxa with individual numbers, while 'Taxa by Traits' matrix comprised of biological traits for each taxon found in the entirety of this study, derived from the fuzzy-coded database as mentioned previously. Both matrices were then combined to form 'Trait by Station' matrix, whereby denotes biological traits found in each station. The 'Trait by Station' matrix was used further for BTA analysis in form of ordination using Principal Component Analysis (PCA), to determine which sub-traits are associated with the stations. In this study, PAST 4.03 was used to assess one-way ANOVA test and PCA, whilst SPSS 1.0.0.1406 was used to assess Pearson Correlation Analysis.

RESULTS AND DISCUSSIONS

Assessment of Taxa and their Respective Traits

A total of 848 macrobenthos individuals were enumerated, with station 1 recorded the highest abundance ($n=483$) whilst station 4 recorded the lowest ($n=25$). In terms of phyla, the Annelida recorded the highest abundance ($n=789$), followed by Arthropoda ($n=15$), Echinodermata ($n=12$) and Mollusca ($n=32$). Thirty seven (37) taxa were found in this study (Table 2), with phylum Annelida accounted as the highest taxa count than other phyla ($n=23$). In terms of taxa count at each station, Station 1 recorded the highest taxa count ($n=11$), while station 4 recorded the lowest ($n=8$).

In terms of trait assessment, the modality vermiform (Vf) from body shape and direct development (DD) from larval development recorded the highest in all stations (Figure 2). Majority of taxa members in Phylum Annelida possess vermiform body shape in adult stage, as seen in all polychaete taxa found in this study (Fauchald, 1997; Mikac, 2015). Previous records for BTA analysis in Emilia-Romagna coast, Italy reported the dominance of vermiform trait amongst other body shape modality due to large numbers of polychaetes at coastline and intertidal zone (Pagenelli et al., 2012). Based on information gathered from Polytrait and ADW database, most Sedentaria polychaetes such as Capitellids, Cirratulids and Spionids undergo direct development growth instead of usual multiple growth stages like Errantia polychaetes (Faulwetter et al., 2017). Higher percentage of direct development modality in this study was highly attributed to dominance of these Sedentarian polychaetes throughout the stations, which are very common in sediment closer to estuaries (Guan et al. 2014; Mohamamad & Jalal, 2018).

In terms of feeding mode trait, sub-surface deposit feeder (SSDFr) modality was the highest in station 1 to 3, while surface deposit feeder (SDFr) modality was the highest in the next three stations. Like body shape and larval development trait, Sedentarian polychaetes

Table 2
Taxa recorded in seafloor of Pulau Indah, Klang

	St1	St2	St3	St4	St5	St6
<i>Alvania</i>			7			
<i>Ancistrosyllis</i>		7				
<i>Capitella</i>	4					
<i>Caulleriella</i>	83	98				4
<i>Cerithidea rhizophorarum</i>	5		5			
<i>Cerithidea rhizophorarum morchii</i> (var.)	8		2			
<i>Cerithium</i>		4				
<i>Cirriformia</i>	130	43				
<i>Cossura</i>						1
<i>Eunice</i>		2	3			
<i>Eurythoe</i>					1	
<i>Flabellina</i>	1					
<i>Glycera</i>				2	4	6
<i>Goniada</i>				4	5	
Maldanidae.			2	1	2	
<i>Mediomastus</i>	82		11			
<i>Micronephthys</i>	9			3		
<i>Micropanope spinipes</i>					2	
<i>Monoporeia</i>						3
<i>Nassarius</i>			1			
<i>Notomastus</i>			33			
<i>Ophiactis</i>				3	3	2
<i>Ophiothrix</i>		1				
Paraonopidae.				1		
<i>Phyllodoce</i>	1					
<i>Poecilochaetus</i>	3		2			1
<i>Prionospio</i>	157	3	3	10		4
<i>Pulliella</i>			11			
<i>Scina</i>					3	
<i>Scoloplella</i>		3	2			
<i>Sigambra</i>		5	34			
<i>Sthenelanella</i>		2			5	
<i>Syllidia</i>					1	
<i>Tanais</i>		1				3
Terebellidae.				1		
<i>Thyonidium</i>					3	
<i>Thysanoessa</i>						2

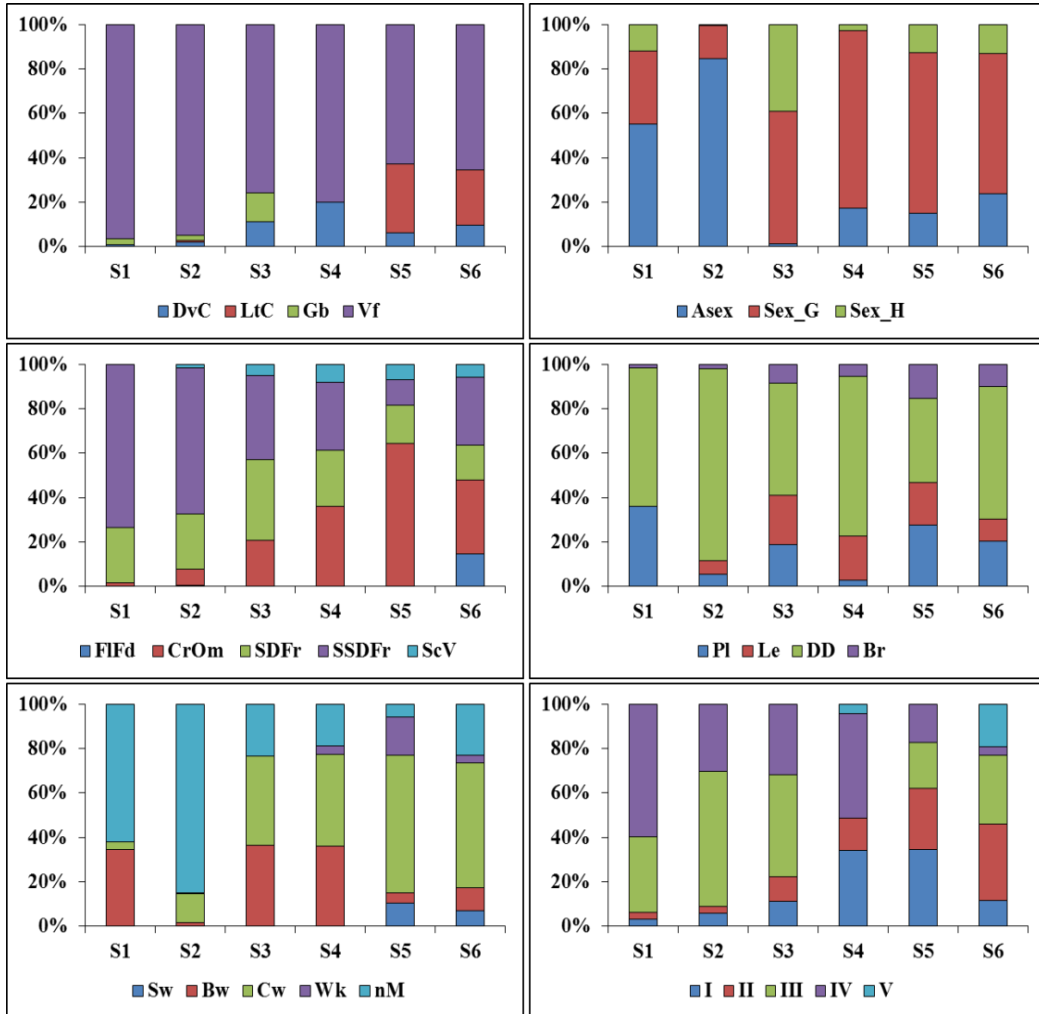


Figure 2. Distribution of traits for each station. Refer to Table 1 for labels

such as Capitellids, Cirratulids and Spionids are well-known to feed on detritus and absorb nutrients underneath sediment (Huang et al., 2012). They form tubes that transverse across sediment to absorb the nutrients, thus bearing the sub-surface deposit feeder (SSDFr) modality. These polychaetes are readily found in greater abundance in sediment with smaller size particle (Otegui et al., 2016). Guan et al. (2014) reported higher percentage of Sedentaria polychaetes in Sungai Pulai estuary which had less human intrusion. Mohamad and Jalal (2018) also reported similar percentage of Sedentaria polychaetes in Pahang estuary, but unlike Sungai Pulai, it was the centre of human activity and the Errantia polychaetes were only found further off towns in Sungai Pahang. Smaller sediment size allowed easier propagation and settlement of Sedentaria polychaetes. Coupled with direct development trait, the taxon’s larvae able to easily burrow into sediment and depend

on nutrients, ultimately enable them to grow indiscriminately without problem, regardless the condition of the sediment, unpolluted or otherwise (Fadhullah & Syakir, 2016).

Conversely, Surface deposit feeder (SDFr) modality encompassed all phyla in this study, but majority of these taxa associated with this trait were not entirely exclusive (fuzzy coding 3), but rather preferable (fuzzy coding 2) with only *Thyodinium* sea cucumber qualified as exclusive to this modality (González-Wangüemert et al., 2016). All polychaetes found in this study were associated with this modality (except Terebellid polychaete) but showed minor preference to this modality (fuzzy coding 1), including Sedentarian polychaetes as mentioned before. Brittle star, polychaete and gastropod taxa found in this study were commonly known for burrowing underneath sediment for wide range of purpose, from searching for food to protection from predators (Jumars et al., 2015; Roche et al. 2016). The idea of identifying macrobenthos' burrowing behaviour have long been investigated. Zorn et al. (2010) observed and reported movement of burrowing polychaetes in sediment. The polychaetes observed, were shown to be revolving throughout the sediment, forming tubes along their way, and occasionally returned to the same tubes to forage food particles that might come from surface above which dispersed into the sediment via bioturbation. Zorn et al. (2010) also noted these polychaetes' immediate burrowing response when detecting light intensity in the vicinity. Boos et al. (2010) focused similar behaviour, but towards brittle star, but unlike Zorn et al. (2010), variation of sizes in brittle star genera showed different burrowing strategies, in which smaller Ophiroids such as *Ophiactis* performed burrowing solely to avoid predation and deposit feeding, but larger genera done so to hunt for meiobenthos such as polychaetes and smaller arthropods.

Station 1 and 2 recorded the highest modality for non-motile (nM) in adult mobility trait and asexual (Asex) in reproduction technique trait. Similarly, Capitellid, Cirratulid and Spionid polychaetes contributed to higher taxa abundance in all stations especially in stations 1 and 2 (Faulwetter et al., 2017). These polychaetes are generally smaller-sized and prefer sediment with higher percentage of small-sized sediment for easier settlement and foraging (Figure 3) (Idris & Arshad, 2013). A study by Wouters et al. (2018) in Southeastern coast of Brazil showed an increasing trend of Cirratulids, Capitellids and Spionid polychaetes across the continental slope, and highest intensity at flat beaches with lesser wave exposure and smaller-sized sediment grain in which the polychaetes would find easier to settle.

Lastly, for ecological group trait, second order opportunistic (IV) modality showed the highest modality in all stations, particularly in station 1. Station 2 and 3 showed highest modality for tolerant (III) modality. Borja et al. (2009) introduced AZTI Marine Biotic Index (AMBI) as an ecological index suitable to define marine taxa based on their availability in certain ecological area under specific circumstances. The group number in AMBI index determines which type of ecological area the taxa prefer to, or more readily available on.

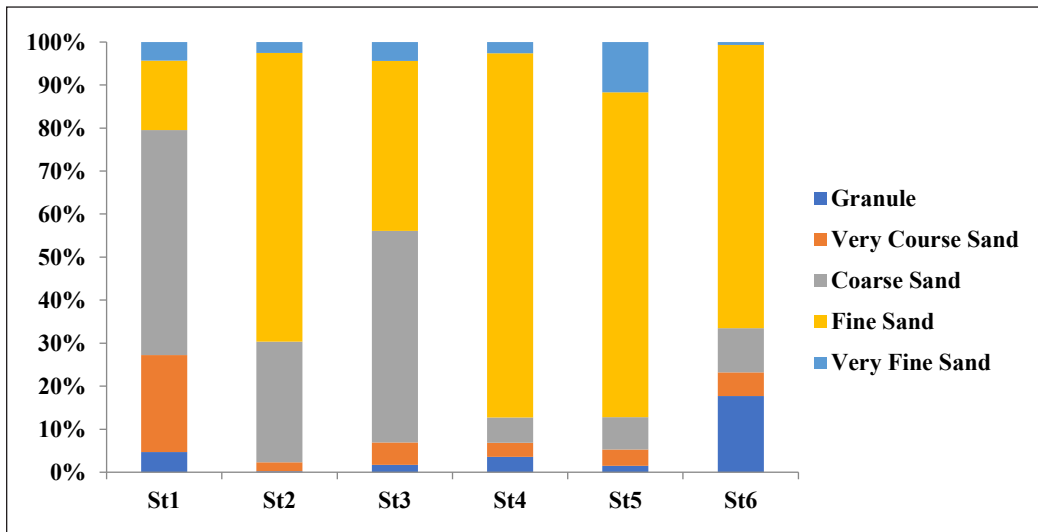


Figure 3. Particle size distribution for sediments in all stations in this study

For instance, *Caulleriella* polychaete in this study was classified as tolerant ecological group (III). This group indicates that the polychaete can be found in ecological zone where there are less ecological stresses (i.e., pollution problem). On the other hand, *Cirriformia*, another polychaete genus, was identified as second-order opportunistic group (IV), which means the genus was readily available in ecological zone that contained certain level of ecological stress (Galanidi et al., 2016). Notable study related to AMBI in Malaysia was conducted by Sany et al. (2015) toward benthic communities in Port Klang, focusing on assessing ecological quality of seafloor habitat. They measured the level of disturbance in the region by composing an assemblage of benthic taxa that had been ranked according to their respective ecological group (I to V), in which they concluded that the benthic communities in the area were overall still in tolerable condition due to lesser number of individuals in taxa ranked ecological I and II but higher number in taxa ranked III to V. Linares et al. (2018) assessed BTA toward benthic assemblages at a deactivated dam within Pandeiros River, Brazil. The study identified increase of non-native genera mollusks such as Asian Clam (*Corbicula fluminea*) and Red-Rimmed Melania Snail (*Melanoides tuberculata*), possibly originating from anthropogenic activities such as aquarium trade and ejection from removal of biofouling off boats into water. BTA assessment indicated that these taxa possessed scavenging traits, smaller bodies, and capable of tolerate changes of water quality in terms of pH and temperature (Ecological Group I).

Assessment for Ecological and Functional Diversity

Station 1 recorded the highest macrobenthos density (21000 ind./m²), while Station 4 showed the lowest (1086.96 ind./m²). While station 3 accounted highest taxa number in

all stations, station 5 accounted to highest in all ecological indices (Table 3). There was significant different in macrobenthos density between stations (one-way ANOVA, $p < 0.05$).

Unlike the diversity index, functional diversity's purpose is to assess the variation and range of traits or characteristic of taxa within the ecosystem (Tilman, 2001). Higher *FD* indicates higher variations of traits shown by existing taxa within the area, i.e., chances of finding two different taxa with two different modalities from a trait become higher. *FD* in body shape trait was highest in Station 5, followed by feeding mode trait and adult mobility in station 6, reproductive technique trait in station 1, larval development trait in station 5 and lastly ecological group trait in station 4 (Table 4).

The ecological group data was significantly differed between stations (one-way ANOVA, $P < 0.05$). All stations showed very large difference between all five modalities in ecological group trait. Large differences between AMBI ecological groups were also seen in other study such as Pilo et al. (2016), where there were distinct differences between benthic taxa sensitive, resistant, tolerant, and opportunistic in Tagus estuary, Portugal throughout the year. Pilo et al. (2016) noticed distinct changes of stress-tolerant ecological group and second-order opportunistic ecological group (IV) across the time.

Table 3
Ecological Indices for benthic communities in seafloor northwards of Port Klang

Station	Taxa No.	Taxa Density	<i>H'</i>	<i>J'</i>	<i>D</i>	<i>D_{mg}</i>
St1	11	21000.00	1.61	0.45	0.76	1.62
St2	10	7347.83	1.30	0.33	0.59	1.95
St3	13	5043.48	1.98	0.56	0.81	2.52
St4	8	1086.96	1.76	0.72	0.77	2.18
St5	10	1260.87	2.19	0.89	0.88	2.67
St6	9	1130.43	2.06	0.87	0.86	2.46

Table 4
Functional Diversity (FD) of traits in stations

Station	Body Shape	Feeding Mode	Adult Mobility	Reproductive Technique	Larval Development	Ecological Group
St1	0.05	0.18	0.60	0.61	0.20	0.56
St2	0.11	0.22	0.49	0.40	0.19	0.52
St3	0.31	0.23	0.44	0.38	0.59	0.69
St4	0.30	0.65	0.56	0.48	0.60	0.73
St5	0.56	0.57	0.55	0.50	0.63	0.72
St6	0.57	0.66	0.70	0.45	0.60	0.70

Ecological Functioning of Benthic Communities in Seafloor Northwards of Pulau Indah, Klang with Regards to Environmental Parameters

Principal component analysis and its eigenvalue showed ordination of modalities within all six stations in this study (Figure 4, Table 5). Station 1 was located at the mouth of Klang River and closest to Southpoint, one of the main ports in Port Klang. Station 1 was also attributed to multiple modalities, in which each of modalities were attributed to this station due to dominance of polychaetes and smaller percentage of gastropods (Table 6). Dominant polychaete genera in station 1 were *Mediomastus* (Family Capitellidae), *Caulleriella* and *Cirriiformia* (Cirratulidae), and *Prionospio* (Spionidae). Besides possessing vermiform body shape, these polychaete genera generally live-in burrow and non-motile throughout their lifetime, deposit feeder, reproduce asexually (except *Mediomastus*) and rather possess direct development from larvae to adult form (Fauchald, 1977; Faulwetter et al., 2017; Mikac, 2015). *Mediomastus* and *Caulleriella* were tolerant ecological group (III) while *Cirriiformia* and *Prionospio* were second-order opportunistic group (IV). Both *Cirriiformia* and *Prionospio* polychaetes were known to be very common in sediment closer to industrial zones and sewage area, where the zone's sediment content might restrict movement and settlement of benthic taxa sensitive and tolerant towards ecological stresses (Musale & Dessai, 2011). Rosli et al. (2016) showed similar report of higher abundance of Spionid, Capitellid and Cirratulid polychaetes at offshore Pekan, Pahang and Dungun, Terengganu coast and Kudat-Balambangan Island in Sabah, where smaller sediment size were more prevalent due to further proximity towards shoreline.

Station 2 was attributed towards globose modality (Gb), and two reproductive techniques i.e., sexual gonochoric (Sex_G) and sexual hermaphrodite gonochoric (Sex_H). Station 2 was located between the eastern end of Pulau Indah and western end of mainland Klang and known to be one of the main routes for ships originating from Southpoint (Sany et al., 2014). Globose modality in this station was contributed by abundance of *Cerithium* snail, which was common in sediment with higher percentage of smaller-sized particles (Reid & Claremont, 2014). Unlike in station 1, station 2 had both Errantian and Sedentarian polychaete groups, with most Errantian polychaetes such as *Eunice* (family Eunicidae), *Scolopella* (family Orbiniidae), *Ancistrosyllis* and *Sigambra* (family Pilargidae) and *Sthenelanelia* (family Sigalionidae) all utilized sexual reproduction via mating between two sexes instead of hermaphrodite (Blake & Arnosfsky, 1999).

Station 3 was attributed towards brooding modality (Br) and sensitive ecological group (I). Both modalities were closely attributed to the gastropod taxa, which were mostly abundant in this station than others. Station 3 was located south to station 2, and closer towards the mouth of Langat River. Next, station 4, located between waterways from three ports (Southpoint, West Port and North Point), was attributed towards dorsally-ventrally compressed (DvC), carnivore-omnivore (CrOm), scavenger (ScV), crawler (Cw),

lecithotrophic (Le) and tolerant ecological group (II). Benthic community in station 4 comprised some non-polychaete taxa such as *Ophiactis* and Braychuran crab which might contribute to dorsally ventral compressed modality (Fujita & Irimura, 2015). This body shape is exclusive in arthropods and echinoderms but also found in some polychaete taxa such as *Micronephthys* (family Nepthyidae) and *Syllidia* (family Eunicidae). All taxa found in this station were all exclusively predators and except *Ophiactis* which is scavenger (Faulwetter et al., 2017). Majority of polychaetes in this station was Errantian, and possess visible parapodia, and combined with other non-polychaete taxa (*Ophiactis* brittle star and Paraonopidae crab), contributed towards crawling modality in this study (Ng, 2017). Errantian polychaete are common in intertidal zone or in seafloor closer to estuaries, rocky shore, and coral reefs (Gholizadeh, 2012; Polgar et al., 2015), in which both studies reported higher Errantian polychaetes at rocky shores nearby beaches of Penang, and former coral reef site in Jeram Beach, Selangor.

Next, station 5 was attributed to dorsal-ventrally-compressed body (DvC), scavenger (ScV) and carnivore-omnivore (CrOm) feeding behaviour, crawler (Cw) and walker (Wk) adult mobility and lecithotrophic (Le) larval development modality. Station 5 was located between the west side of Pulau Indah, and the east side of Pulau Kelang. Majority of taxa found in this station included several Errantian polychaetes such as *Eurythoe*, *Glycera*,

Table 5
Eigenvalue scale for PCA analysis in Figure 3

PC	Eigenvalue	% variance
1	0.232	52.575
2	0.120	27.161
3	0.054	12.169
4	0.022	4.951
5	0.008	1.885
6	0.006	1.259

Table 6
List of modalities attributed to the closest sampling stations in this study. (Refer to Table 1 for labels)

Station	Trait	Modality
St1	Body Shape	Vf
	Feeding Mode	SDFr, SSDFr
	Adult Mobility	Bw, nM
	Reproductive Technique	Asex
	Larval Development	DD, PI
	Ecological Group	III, IV
St2	Body Shape	Gb
St3	Reproductive Technique	Sex_G, Sex_H
	Adult Mobility	Br
	Ecological Group	I
St4	Body Shape	DvC
	Feeding Mode	CrOm, ScV
	Adult Mobility	Cw
	Larval Development	Le
St5	Ecological Group	II
	Body Shape	DvC
	Feeding Mode	CrOm, ScV
	Adult Mobility	Cw, Wk
St6	Larval Development	Le
	Body Shape	LtC
	Feeding Mode	FIFd
St6	Adult Mobility	Sw, Wk
	Ecological Group	V

Table 7

*Pearson Correlation Table between trait modalities and environmental parameters assessed in this study. Refer to Table 1 for labels. * indicates significant correlation at 0.01 (two-tailed) and ** indicates significant correlation at 0.05 (two-tailed)*

	Salinity	Depth	TOC	VFS	FS	CS	VCS	G
DvC	-0.15	-0.33	-0.07	-0.46	-0.08	0.62	-0.32	0.10
LtC	0.50	0.93**	0.17	0.86*	-0.03	-0.62	0.38	-0.49
Gb	-0.58	-0.37	0.29	-0.47	-0.41	0.98**	-0.20	0.37
Vf	-0.77	-0.46	0.38	-0.41	-0.42	0.76	-0.03	0.60
FlFd	0.70	0.50	0.78	0.75	-0.71	-0.23	0.96**	-0.37
CrOm	-0.09	0.23	-0.32	-0.12	0.29	0.22	-0.51	-0.35
SDFr	-0.78	-0.48	0.37	-0.46	-0.43	.851*	-0.08	0.59
SSDFr	-0.75	-0.45	0.38	-0.39	-0.42	0.73	-0.01	0.61
ScV	0.12	-0.03	-0.30	-0.25	0.24	0.23	-0.41	-0.42
Sw	0.46	0.93**	0.14	0.84*	-0.01	-0.60	0.35	-0.40
Bw	-0.68	-0.37	0.42	-0.31	-0.51	0.76	0.05	0.76
Cw	-0.08	0.03	0.08	-0.20	-0.16	0.61	-0.24	-0.28
Wk	0.12	0.64	-0.34	0.46	0.44	-0.62	-0.15	-0.21
nM	-0.75	-0.48	0.34	-0.42	-0.36	0.69	-0.04	0.50
Asex	-0.74	-0.47	0.32	-0.41	-0.32	0.62	-0.04	0.46
Sex_G	-0.70	-0.36	0.41	-0.36	-0.51	0.87*	-0.03	0.69
Sex_H	-0.54	-0.26	0.39	-0.32	-0.51	0.92*	-0.06	0.53
Pl	-0.69	-0.32	0.44	-0.25	-0.50	0.69	0.08	0.73
Lc	-0.05	-0.12	-0.21	-0.35	0.13	0.44	-0.43	-0.31
DD	-0.77	-0.51	0.34	-0.46	-0.37	0.75	-0.07	0.51
Br	-0.56	-0.10	0.29	-0.28	-0.38	0.89*	-0.20	0.31
I	-0.94**	-0.49	-0.10	-0.66	0.01	0.81	-0.57	0.57
II	-0.31	0.17	0.66	0.11	-0.73	0.78	0.25	0.31
III	-0.77	-0.50	0.33	-0.49	-0.36	0.81*	-0.12	0.38
IV	-0.74	-0.42	0.39	-0.35	-0.44	0.72	0.01	0.70
V	0.88*	0.51	0.59	0.78	-0.54	-0.46	0.92*	-0.32

and *Tanais* were common at seafloor rather than the surface, and preferred sediment with smaller sediment size, in which they utilize the sediments' nature for foraging detritus via filter feeding (Navarro-Barranco & Hughes, 2015). The Tanaid shrimp (*Tanais*), Pontoporeiid amphipod (*Monoporeia*), and krill (*Thysanoessa*) were common in Malaysian water and were in higher abundance in nutrient-rich estuaries such as Pulau Kelang, Selangor (Norhayati et al., 2009) and Mengabang Telipot river, Terengganu (Ramli et al., 2018).

Coarse sand (0.5 mm) was correlated towards the following modality: globose (Gb), sexual gonocoric (Sex_G), sexual hermaphrodite (Sex_H), brooding (Br) and tolerant ecological group (III). Coarse sand (0.5 mm) is often considered as suitable substrate

for settlement of macro-sized benthic taxa such as gastropods and arthropods, and even Errantian polychaetes. This includes suitability for procreation of offspring for the taxa by general. Bannister et al. (2012) reported noticeable changes of benthic larval spawning across the Great Barrier Reef, Australia and this fluctuation was attributed to particle size distribution, with coarse sand (0.5 mm) as suitable substrate for macrobenthic taxa larva to settle and develop. Coarse sand (0.5 mm) also considered as suitable substrate for gastropods, as was shown in Genner and Michel (2003) where their study established link between particle size distributions and sea snail's assemblage in Cape Maclear, Malawi. They found that habitat with higher percentage of sediment size less than 0.5 mm contained more sea snails (Genner & Michel, 2003).

CONCLUSION

This study concludes that vermiform body shape, direct larval development, deposit feeder, non-motile, asexually reproductive and second-order opportunistic ecological group were the most dominant sub-traits in this study. While station 1 has the highest taxa density, station 5 is most diverse. Functional diversity showed that reproductive technique trait was more diverse in station 1, followed by ecological group in station 4, body shape and larval development in station 5, and lastly feeding mode and adult mobility in station 6. PCA analysis indicated stations 1 and 4 attributed majority of modalities assessed in this study, but none in station 5. Correlation analyses showed multiple modalities strongly correlated towards size particle distribution, more specifically very fine sediment (0.063 mm) and coarse sediment (0.5 mm), and depth.

Unlike other countries, BTA study is still a new field in Malaysia. Therefore, the data linked to BTA regarding benthic population in Malaysian water is very inadequate and lacking. The study could be a good starting point for the use of BTA, or at least could serve as a suitable and adequate comparative study for other areas in Malaysia such as in coral reefs, mangrove biotopes and even waters nearby industrial zones like this study. In similar fashion as identifying suitable occupation for human population in an area for a certain period of situations and time, if conducted more frequently and comprehensively, BTA can serve as time capsule for us to observe certain changes in benthic community in Malaysian's water over time.

ACKNOWLEDGEMENT

We sincerely thank Mr. Zuhaimi from the Marine Science Program, Universiti Kebangsaan Malaysia for his dedicated efforts on sample collection. This work was funded by Fundamental Research Grant, Ministry of Higher Education, Malaysia (project no. FRGS/1/2017/STG03/UKM/02/5).

REFERENCES

- Baharuddin, N., Basri, N. B., & Syawal, N. H. (2018). Marine gastropods (Gastropoda; Mollusca) diversity and distribution on intertidal rocky shores of Terengganu, Peninsular Malaysia. *Bioflux*, *11*(4), 1144-1154.
- Bannister, R. J., Battershill, C. N., & de Nys, R. (2012). Suspended sediment grain size and mineralogy across the continental shelf of the great barrier reef: Impacts on the physiology of a coral reef sponge. *Continental Shelf Research*, *32*, 86-95. <https://doi.org/10.1016/j.csr.2011.10.018>.
- Blake, J. A., & Arnofsky, P. L. (1999). Reproduction and larval development of the *Spioniform polychaeta* with application to systematics and phylogeny. *Hydrobiologia*, *402*, 57-106. <https://doi.org/10.1023/A:1003784324125>.
- Bolam, S. G., Garcia, C., Eggleton, J., Kenny, A. J., Buhl-Mortensen, L., Gonzalez-Mirelis, G., van Kooten, T., Dinesen, G., Hansen, J., Hiddink, J. G., Sciberras, M., Smith, C., Papadopoulou, N., Gumus, A., Van Hoey, G., Eigaard, O. R., Bastardie, F., & Rijnsdorp, A. D. (2017). Differences in biological traits composition of benthic assemblages between unimpacted habitats. *Marine Environmental Research*, *126*, 1-13. <https://doi.org/10.1016/j.marenvres.2017.01.004>.
- Bolam, S. G., McIlwaine, P. S. O., & Garcia, C. (2016). Application of biological traits to further our understanding of the impacts of dredged material disposal on benthic assemblages. *Marine Pollution Bulletin*, *105*(1), 180-192. <https://doi.org/10.1016/j.marpolbul.2016.02.031>.
- Boos, K., Gutow, L., Mundry, R., & Franke, H. (2010). Sediment preference and burrowing behaviour in the sympatric brittlestars *Ophiura albida* (Forbes, 1839) and *Ophiura ophiura* (Linnaeus, 1738) (Ophiuroidea, Echinodermata). *Journal of Experimental Marine Biology and Ecology*, *393*(2010), 176-181. <https://doi.org/10.1016/j.jembe.2010.07.021>.
- Borja, A., Muxika, I., & Rodríguez, J. G. (2009). Paradigmatic responses of marine benthic communities to different anthropogenic pressures, using M-AMBI, within the European water framework directive. *Marine Ecology*, *30*(2), 214-227. <https://doi.org/10.1111/j.1439-0485.2008.00272.x>.
- Dell, A. I., Pawar, S., & Savage, V. M. (2011). Systematic variation in the temperature dependence of physiological and ecological traits. *Proceedings of the National Academy of Sciences of the United States of America*, *26*, 10591-10596. <https://doi.org/10.1073/pnas.1015178108>.
- Fadhullah, W., & Syakir, M. I. (2016). Polychaetes as ecosystem engineers: Agents of sustainable technologies. In M. I. Ahmad, M. Ismail & S. Riffat (Eds.), *Renewable energy and sustainable technologies for building and environmental applications* (pp. 137-150). Springer. https://doi.org/10.1007/978-3-319-31840-0_8.
- Fauchald, K. (1977). *The polychaete worms: Definitions and keys to the orders, family and genera*. Natural History Museum of Los Angeles County, California.
- Faulwetter, S., Simbouna, N., Katsiaras, N., Chatzigeorgiou, G., & Arvanitidis, C. (2017). Polychaetes of Greece: An updated and annotated checklist. *Biodiversity Data Journal*, *2017*(5), 1-220. <https://doi.org/10.3897/BDJ.5.e20997>.
- Fujita, T., & Irimura, S. (2015). Preliminary list of ophiuroids (Echinodermata: Ophiuroidea) collected from the Johor Straits, Singapore. *Raffles Bulletin of Zoology*, *31*, 264-272.

- Galanidi, M., Kaboglu, G., & Bizsel, K. C. (2016). Predicting the composition of polychaete assemblages in the Aegean coast of Turkey. *Frontiers in Marine Science*, 3(8), 1-14. <https://doi.org/10.3389/fmars.2016.00154>.
- Gammal, J., Norkko, J., Pilditch, C. A., & Norkko, A. (2017). Coastal hypoxia and the importance of Benthic macrofauna communities for ecosystem functioning. *Estuaries and Coasts*, 40, 457-468. <https://doi.org/10.1007/s12237-016-0152-7>.
- Genner, M. J., & Michel, E. (2003). Fine-scale habitat associations of soft-sediment gastropods at Cape Maclear, Lake Malawi. *Journal of Molluscan Studies*, 69(4), 325-328. <https://doi.org/10.1093/mollus/69.4.325>.
- Gergs, R., & Rothhaupt, K. O. (2015). Invasive species as driving factors for the structure of benthic communities in Lake Constance, Germany. *Hydrobiologia*, 746(1), 245-254. <https://doi.org/10.1007/s10750-014-1931-4>.
- Gholizadeh, M. (2012). Effects of environmental factors on polychaete assemblage in Penang National Park, Malaysia. *World Academy of Science, Engineering and Technology*, 72, 669-672.
- González-Wangüemert, M., Valente, S., Henriques, F., Domínguez-Godino, J. A., & Serrão, E. A. (2016). Setting preliminary biometric baselines for new target sea cucumbers species of the NE Atlantic and Mediterranean fisheries. *Fisheries Research*, 179, 57-66. <https://doi.org/10.1016/j.fishres.2016.02.008>.
- Guan, W. S., Lee, D. M., Abd Ghaffar, M., Md Ali, M., & Che Cob, Z. (2014). Macrobenthos composition, distribution and abundance within Sungai Pulai estuary, Johor, Malaysia. In *AIP Conferences Proceedings* (pp 591-596). American Institute of Physics. <https://doi.org/10.1063/1.4895269>.
- Hanamura, Y., Siow, R., & Chee, P. E. (2008). Reproductive biology and seasonality of the Indo-Australasian mysid *Mesopodopsis orientalis* (Crustacea: Mysida) in a tropical mangrove estuary, Malaysia. *Estuarine, Coastal and Shelf Science*, 77(3), 467-474. <https://doi.org/10.1016/j.ecss.2007.10.015>.
- Hou, Y., Kong, F., Li, Y., Xi, M., & Yu, Z. (2020). Key factors of the studies on benthic macroinvertebrate in coastal wetlands: Methods and biodiversity. *Ecohydrology and Hydrobiology*, 20(3), 424-436. <https://doi.org/10.1016/j.ecohyd.2020.02.004>
- Huang, Y. C. A., Huang, S. C., Hsieh, H. J., Meng, P. J., & Chen, C. A. (2012). Changes in sedimentation, sediment characteristics, and benthic macrofaunal assemblages around marine cage culture under seasonal monsoon scales in a shallow-water bay in Taiwan. *Journal of Experimental Marine Biology and Ecology*, 422, 55-63. <https://doi.org/10.1016/j.jembe.2012.04.008>.
- Hussin, W. M. R. W. (2014). Marine fish farming in Bidong Island, Malaysia and its implications on benthic community structure and functional diversity. *Aquaculture, Aquarium, Conservation & Legislation*, 7(6), 431-440.
- Idris, I., & Arshad, A. (2013). Checklist of polychaetous annelids in Malaysia with redescription of two commercially exploited species. *Asian Journal of Animal and Veterinary Advances*, 8(3), 409-436. <https://doi.org/10.3923/ajava.2013.409.436>.
- Jumars, P. A., Dorgan, K. M., & Lindsay, S. M. (2015). Diet of worms emended: An update of polychaete feeding guilds. *Annual Review of Marine Science*, 7, 497-520. <https://doi.org/10.1146/annurev-marine-010814-020007>.

- Korte, A., & Ashley, F. (2013). The advantages and limitations of trait analysis with GWAS: A review self-fertilisation makes Arabidopsis particularly well suited to GWAS. *Plant Methods*, 9, 29. <https://doi.org/10.1186/1746-4811-9-29>.
- Lepš, J., de Bello, F., Lavorel, S., & Berman, S. (2006). Quantifying and interpreting functional diversity of natural communities: Practical considerations matter. *Preslia*, 78, 361-374.
- Linares, M. S., Callisto, M., & Marques, J. C. (2018). Thermodynamic based indicators illustrate how a run-of-river impoundment in neotropical savanna attracts invasive species and alters the benthic macroinvertebrate assemblages' complexity. *Ecological Indicators*, 88, 181-189. <https://doi.org/10.1016/j.ecolind.2018.01.040>.
- Linden, P., Patrício, J., Marchini, A., Cid, N., Neto, J. M., & Marques, J. C. (2012). A biological trait approach to assess the functional composition of subtidal benthic communities in an estuarine ecosystem. *Ecological Indicators*, 20, 121-133. <https://doi.org/10.1016/j.ecolind.2012.02.004>.
- Lomovasky, B. J., Firstater, F. N., Salazar, A. G., Mendo, J., & Iribarne, O. O. (2011). Macro benthic community assemblage before and after the 2007 tsunami and earthquake at Paracas Bay, Peru. *Journal of Sea Research*, 65(2), 205-212. <https://doi.org/10.1016/j.seares.2010.10.002>.
- Mikac, B. (2015). A sea of worms: polychaete checklist of the Adriatic Sea. *Zootaxa*, 3943(1), 1-172. <https://doi.org/10.11646/zootaxa.3943.1.1>.
- Mohamamad, A., & Jalal, K. (2018). Macrobenthic diversity and community composition in Pahang Estuary, Malaysia. *Journal of Coastal Research*, 82(4), 206-211.
- Musale, A. S., & Desai, D. V. (2011). Distribution and abundance of macrobenthic polychaetes along the South Indian coast. *Environmental Monitoring and Assessment*, 178, 423-436. <https://doi.org/10.1007/s10661-010-1701-3>.
- Navarro-Barranco, C., & Hughes, L. Z. (2015). Effects of light pollution on the emergent fauna of shallow marine ecosystems: Amphipods as a case study. *Marine Pollution Bulletin*, 94(1-2), 235-240.
- Ng, P. K. L. (2017). On the identities of the highland vampire crabs, *Geosesarma foxi* (Kemp, 1918) and *G. serenei* Ng, 1986, with description of a new phytotelmic species from Penang, Peninsular Malaysia (Crustacea: Decapoda: Brachyura: Sesarmidae). *Raffles Bulletin of Zoology*, 65, 226-242.
- Ng, P. K. L., & Davie, P. J. F. (2002). A checklist of the Brachyuran crabs of Phuket and Western Thailand. *Phuket Marine Biological Center Special Publication*, 23(2), 369-384.
- Norhayati, A., Shukor, M. N., Juliana, S., & Wan, J. W. A. (2009). Mangrove flora and fauna of Klang islands mangrove forest reserves, Selangor, Malaysia. *Malaysian Journal of Science*, 28(3), 275-288. <https://doi.org/10.22452/mjs.vol28no3.6>.
- Ong, J. Y., & Wong, H. P. (2015). Sea cucumbers (Echinodermata: Holothuroidea) from the Johor Straits, Singapore. *Raffles Bulletin of Zoology*, 31, 273-291.
- Otegui, M. B. P., Brauko, K. M., & Pagliosa, P. R. (2016). Matching ecological functioning with polychaete morphology: Consistency patterns along sedimentary habitats. *Journal of Sea Research*, 114, 13-21. <https://doi.org/10.1016/j.seares.2016.05.001>.

- Padron, M., & Guizien, K. (2016). Modelling the effect of demographic traits and connectivity on the genetic structuration of marine metapopulations of sedentary benthic invertebrates Mariana. *ICES Journal of Marine Science*, 73(7), 1935-1945.
- Paganelli, D., Marchini, A., & Occhipinti-Ambrogi, A. (2012). Functional structure of marine benthic assemblages using biological traits analysis (BTA): A study along the Emilia-Romagna coastline (Italy, North-West Adriatic Sea). *Estuarine, Coastal and Shelf Science*, 96(1), 245-256. <https://doi.org/10.1016/j.ecss.2011.11.014>.
- Piló, D., Ben-Hamadou, R., Pereira, F., Carriço, A., Pereira, P., Corzo, A., Gaspar, M. B., & Carvalho, S. (2016). How functional traits of estuarine macrobenthic assemblages respond to metal contamination? *Ecological Indicators*, 71, 645-659. <https://doi.org/10.1016/j.ecolind.2016.07.019>.
- Polgar, G., Nishi, E., Idris, I., & Glasby, C. J. (2015). Tropical polychaete community and reef dynamics: Insights from a Malayan Sabellaria (Annelida: Sabellariidae) reef. *Raffles Bulletin of Zoology*, 63, 401-417.
- Ramli, R., Kassim, Z., & Roslani, M. A. (2018). Meiofaunal responses to azoic sediment in a sandbar-regulated estuary in the East Coast of Peninsular Malaysia. In *Proceedings of the Second International Conference on the Future of ASEAN (ICoFA) 2017–Volume 2* (pp. 899-908). Springer. https://doi.org/10.1007/978-981-10-8471-3_89.
- Reid, D. G., & Claremont, M. (2014). The genus *Cerithideopsis* Thiele, 1929 (Gastropoda: Potamididae) in the Indo-West Pacific region. *Zootaxa*, 3779(1), 61-80. <https://doi.org/10.11646/zootaxa.3779.1.8>.
- Roche, K. R., Aubeneau, A. F., Min, W. X., Aquino, T. C., Bolster, D., & Packman, A. I. (2016). An integrated experimental and modelling approach to predict sediment mixing from benthic burrowing behaviour. *Environmental Science & Technology*, 50(18), 10047-10054. <https://doi.org/10.1021/acs.est.6b01704>.
- Rosli, N. S., Yahya, N., Arifin, I., & Bachok, Z. (2016). Diversity of Polychaeta (Annelida) in the continental shelf of Southern China Sea. *Middle-East Journal of Scientific Research*, 24(6), 2086-2092. <https://doi.org/10.5829/idosi.mejsr.2016.24.06.23650>.
- Sany, S. B. T., Hashim, R., Salleh, A., Rezayi, M., & Safari, O. (2015). Ecological quality assessment based on microbenthic assemblages indices along West Port, Malaysia coast. *Environmental Earth Science*, 74(2), 1331-1341. <https://doi.org/10.1007/s12665-015-4122-3>.
- Sany, S. B. T., Rezayi, M., Hashim, R., Salleh, A., & Safari, O. (2014). Diversity and distribution of benthic invertebrates in the West Port, Malaysia. *International Journal of Environmental, Ecological, Geological and Mining Engineering*, 8(7), 458-461.
- Tilman, D. (2001). Functional diversity. *Encyclopedia of Biodiversity*, 3(1), 109-120.
- Verberk, W. C. E. P., Van Noordwijk, C. G. E., & Hildrew, A. G. (2013). Delivering on a promise: Integrating species traits to transform descriptive community ecology into a predictive science. *Freshwater Science*, 32(2), 531-547. <https://doi.org/10.1899/12-092.1>.
- Widdows, J., Brown, S., Brinsley, M. D., Salkeld, P. N., & Elliott, M. (2000). Temporal changes in intertidal sediment erodability: Influence of biological and climatic factors. *Continental Shelf Research*, 20(10-11), 1275-1289. [https://doi.org/10.1016/S0278-4343\(00\)00023-6](https://doi.org/10.1016/S0278-4343(00)00023-6).

- Wouters, J. M., Gusmao, J. B., Mattos, G., & Lana, P. (2018). Polychaete functional diversity in shallow habitats: Shelter from the storm. *Journal of Sea Research*, 135(2018), 18-30. <https://doi.org/10.1016/j.seares.2018.02.005>.
- Xu, Y., Stoeck, T., Forster, D., Ma, Z., Zhang, L., & Fan, X. (2018). Environmental status assessment using biological traits analyses and functional diversity indices of benthic ciliate communities. *Marine Pollution Bulletin*, 131, 646-654. <https://doi.org/10.1016/j.marpolbul.2018.04.064>.
- Yap, T. K., Saleh, E., & Gallagher, J. B. (2020). Impacts of associated fauna on seagrass during the conditioning period in husbandry tanks: Gaya Island, Sabah, Malaysia case study. *Transactions on Science and Technology*, 7(2), 50-57.
- Zhong, X. X., Xu, G. J., Mamun, A. A., Warren, A., & Xu, H. L. (2018). Identifying bioindicators across trait-taxon space for assessing water quality in marine environments. *Marine Pollution Bulletin*, 131, 565-571. <https://doi.org/10.1016/j.marpolbul.2018.04.044>.
- Zorn, M. E., Gingras, M. K., & Pemberton, S. G. (2010). Variation in burrow-wall micromorphologies of select intertidal invertebrates along the Pacific Northwest Coast, USA: Behavioural and diagenetic implications. *Palaios*, 25, 59-72. <https://doi.org/10.2110/palo.2009.p09-026r>.

Radiographic Measurement of Cochlear in Sudanese Population using High Resolution Computed Tomography (HRCT)

Lubna Bushara¹, Mohamed Yousef², Ikhlas Abdelaziz¹, Mogahid Zidan^{3*}, Dalia Bilal¹ and Mohammed El Wathig⁴

¹Department of Medical Imaging and Radiation Sciences, College of Applied Medical Sciences, University of Jeddah, Jeddah, 21589 Saudi Arabia

²Radiologic Sciences Program, Batterjee Medical College, Jeddah, 23819-6700 Saudi Arabia

³Department of Radiology, Faculty of Radiology and Nuclear Medicine Sciences, National Ribat University, Khartoum, 11123 Sudan

⁴Siemens Healthcare limited, Jeddah, 21589 Saudi Arabia

ABSTRACT

This study aimed to determine the measurements of the cochlea among healthy subjects and hearing deafness subjects using a High Resolution Computed Tomography (HRCT). A total of 230 temporal bone HRCT cases were retrospectively investigated in the period spanning from 2011 to 2015. Three 64-slice units were used to examine patients with clinical complaints of hearing loss conditions at three Radiology departments in Khartoum, Sudan. For the control group (A) healthy subjects, the mean width of the right and left cochlear were 5.61 ± 0.40 mm and 5.56 ± 0.58 mm, the height were 3.56 ± 0.36 mm and 3.54 ± 0.36 mm, the basal turn width were 1.87 ± 0.19 mm and 1.88 ± 0.18 mm, the width of the cochlear nerve canal were 2.02 ± 1.23 and 1.93 ± 0.20 , cochlear nerve density was 279.41 ± 159.02 and 306.84 ± 336.9 HU respectively. However, for the experimental group (B), the mean width of the right and left cochlear width were 5.38 ± 0.46 mm and 5.34 ± 0.30 mm, the height

were 3.53 ± 0.25 mm and 3.49 ± 0.28 mm, the basal turn width were 1.76 ± 0.13 mm, and 1.79 ± 0.13 mm, the width of the cochlear nerve canal were 1.75 ± 0.18 mm and 1.73 ± 0.18 mm, and cochlear nerve density were 232.84 ± 316.82 and 196.58 ± 230.05 HU, respectively. The study found there was a significant difference in cochlea's measurement between the two groups with a p-value < 0.05 . This study had established baseline measurements for the cochlear for

ARTICLE INFO

Article history:

Received: 17 September 2020

Accepted: 25 January 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.32>

E-mail addresses:

arwa6067@gmail.com (Lubna Bushara)

mohnajwan@gmail.com (Mohamed Yousef)

ikhlasabdelaziz888@gmail.com (Ikhlas Abdelaziz)

phd.zidan@gmail.com (Mogahid Zidan)

abhamit@gc.edu.sa (Dalia Bilal)

nafi.m1978@gmail.com (Mohammed El Wathig)

* Corresponding author

the healthy Sudanese population. Furthermore, it found that HRCT of the temporal bone was the best for investigation of the cochlear and could provide a guide for the clinicians to manage congenital hearing loss.

Keywords: Cochlea, hearing loss, HRCT, measurement

INTRODUCTION

The hearing or auditory system consists of the inner ear, outer ear, middle ear, and central auditory nervous system. The function of the hearing system is to sense the acoustic environment thus allowing us to detect and perceive sound (Emanuel et al., 2009). The ears are paired organs, one on each side of the head with the sense organ itself, which is technically known as the cochlea, deeply buried within the temporal bones. It is a spiral-cavity shaped within the bony labyrinth, in humans making 2.5 turns around its axis. The primary function of the human cochlea is to transform the vibrations into a neural signal.

Innervations of the cochlea involve both efferent and afferent neurons. The efferent nerve supply includes a much smaller population of descending neurons that send signals from the nervous system to the cochlea. The afferent nerve supply is made up of ascending sensory neurons that send signals from the cochlea to the nervous system (Wageih, 2017). The human cochlea is challenging to explore due to its vulnerability and bordering capsule; it is composed of two and threefourth the turns, unusual anatomy with cochlear three turns has been described (Tian et al., 2006).

Congenital inner ear abnormality such as cochlear aplasia, cochlear hypoplasia, and incomplete cochlear partition is a major cause of sensorineural hearing loss (SNHL). Globally, hearing loss causes moderate to severe disability in 124 million people as of 2004 (108 million of whom are in low and middle-income countries) (WHO, 2008). The cochlear malformation is a major cause of hearing loss worldwide. The majority of hearing loss is sensor neural particularly in the elderly and sensor neural hearing loss (SNHL) is primarily a result of damage to the cochlea of the inner ear.

Radiology plays an essential role in the evaluation of congenital sensorineural hearing loss. In children who are candidates for cochlear implantation surgery, it provides vital preoperative information about the inner ear, the vestibulocochlear nerve, and the brain. Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) provides excellent delineation of the intricate anatomy of the inner ear: CT depicts the minute details of osseous structures, and MRI allows visualization of the fluid-filled spaces and the vestibulocochlear nerve (Joshi et al., 2012).

A retrospective study by Fernando et al. (2011) described the cochlear anatomy among Filipino through high resolution computed tomography (HRCT) imaging. Other retrospective study by Teissier et al. (2010) established normative measurements of the inner ear using temporal bone CT scan to aid in the diagnosis of inner ear malformations as

prospective measurements of the structures of the inner ear. Teissier et al. (2010) established CT measurements of the normal cochlea in children and determined radiological criteria correlated with SNHL. The study was a retrospective study of temporal bone CT performed in 159 children, age range from 3 days to 16 years between February 1999 and July 2004. A control group (n=88) comprised of children without SNHL; the SNHL group comprised of 71 children. The width of the second turn of the cochlea (CW), the cochlear height (CH), and the width of the bony canal for the cochlear nerve (WCN) were measured on a reference plane containing the modiolus, the posterior semicircular canal, the footplate, and the stapes arch (Teissier et al., 2010).

The standard treatment for patients who experience sensorineural hearing loss is cochlear implants (CIs). All recent studies have shown that hearing outcome correlated with intra-cochlear anatomy and electrode placement. Furthermore, there are no local studies has been reported for the establishment of a normal measurement for Sudanese cochlea comparing with cochlear measurements for deaf patients. So the significance of this study provides extra information about the cochlear dimensions for Sudanese to use, especially in cochlear implantation.

The purpose of this retrospective study was to characterise the anatomy of the cochlea on Sudanese through an HRCT scan to set reference values of the normal related to abnormal measurements.

MATERIALS AND METHODS

This cross-sectional retrospective study for normal measurements of the human cochlea, was performed in Al-Amal Hospital, Military Hospital and Royal Care International Hospital, from January 2011 until January 2015.

Study Population

A total of 230 HRCT of the temporal bone for Sudanese subjects were included, 460 right and left ears, in both genders. They were divided them into two groups, group A for normal hearing subjects, and group B for bilateral congenital hearing loss subjects.

Group A: Healthy Subjects Control Group

Two hundreds (200) normal hearing were examined by audiometric tests, 137 males and 63 females, who ranged between 1-84 years old.

Inclusion Criteria for Group A

The Sudanese male and female , aged between 1-90 year. Normal hearing grade exclusion criteria: all participants who had cochlear abnormality or sensorineural hearing loss an etiological investigation after or before cochlear implants were excluded.

Group B: Experimental Group

This group consisted of 30 volunteers with bilateral congenital deafness confirmed by audiometric tests (12 males and 18 females), in different ages between 10 to 30 years.

Data Collection

The data were collected by measurements of different variables which include, cochlea width (CW), cochlea height (CH), a width of cochlear nerve canal (CNCW), Basal turn width (BTW), density of the cochlear nerve, and cranium transverse diameter recorded from a reconstruction of brains in HRCT scanning. Ages and gender were recorded. A clinical data collection sheet was designed, and it contained all the variables of the study.

Data Analysis

Differences in cochlear measurements between Group A and B were analyzed using the ANOVA test. These analyses were performed using SPSS statistical software (SPSS for Windows, version 16; SPSS Inc, USA ,Chicago, IL,). A p-value < 0.05 was considered statistically significant.

Instrumentation

Both 64 slice spiral CT units (Toshiba (Aquilon) and Siemens (Somatom plus) Medical Systems) were used to examine patients for HRCT scans temporal bones.

Data Acquisition and Measurement Protocol

Volumetric acquisition parameters were kV settings 120 kV, 350 mAs, 0.5 mm slice thickness, 70 mm field of view, Effective pitch of 0.85, the effective thickness of about 0.75 mm, HRCT images were reconstructed with a thickness of sequential 2 mm, brain window levels of 35–45, brain window widths of 110–160, bone window levels of 450, and bone window widths of 1600.

Measurements

The protocol used in this study was designed to visualizing details of bone outline for accurate measurements. The dimensions of cochlear were calculated from the axial sections: the width of the bony canal for the cochlear nerve was measured at the entry of the cochlea (WCN), both height (CH) and width of the cochlea (CW) was defined as the second turn at the reference slice level also the second turn of the cochlea was measure parallel to the latter X (Figure 1). The cochlear canal was measured tangentially to two inferior extremities of the X-shaped modiolus (Figure 2). The length between orthogonal projection passing through the canal and the cochlear tip was defined as the cochlear height.

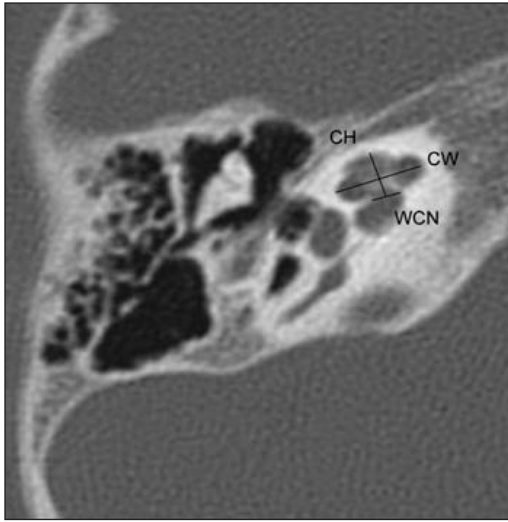


Figure 1. Bone window CT axial view of the right cochlea demonstrating the method of the measurement at the reference plane (X-shape)

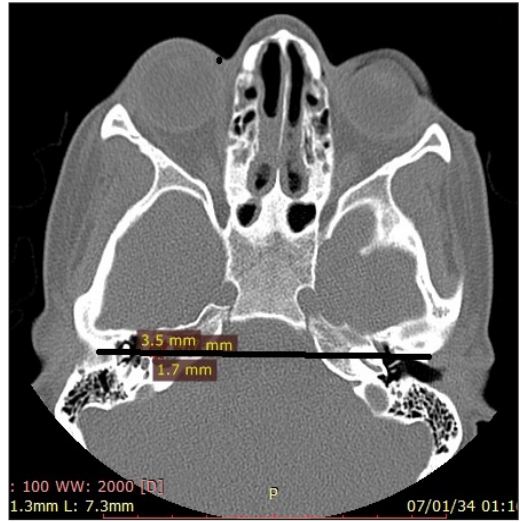


Figure 2. Bone window CT axial view temporal bone demonstrating transverse cranial width (TCW) measurement from the bony borders right to the light side

The basal turn width (BTW) measured from the sagittal section of the same slice obtained by MPR at the same slice level, bone window with filtering adjacent was used to provide an optimal image contrast to define the boundaries more precisely.

RESULTS

Two hundred and thirty ($n = 230$) study participants between 1 to 84 years of age met inclusion criteria for this study; 63 females (42.26 ± 17.85 years of age) and 137 males (37.26 ± 20.48 years of age). Table 1 shows measurements of cochlea according to gender. A statistical difference was not noted between females and males. Therefore, it was of interest to determine whether gender of the patient could not affect cochlea measurements.

Table 2 and 3 show measurements of cochlea in both sides for group both groups. However, there was a significant difference in CT numbers between the two groups where it was found that the CT Number in a normal people group was higher than in people who suffered from hearing loss .

Table 2 and 3 show measurements of cochlea in both sides for group both groups. However, there was a significant difference in CT numbers between the two groups where it was found that the CT Number in a normal people group was higher than in people who suffered from hearing loss .

Table 4 demonstrates that there are differences concerning the size of the cochlear, as it appears that both cochlear in people who have hearing loss seem smaller when compared to normal people .Thus the size of the cochlea in normal hearing subjects was significantly different from subjects with hearing loss.

Table 1
Measurements of the Cochlea in both gender

Variables	Gender	Mean	Std. Deviation
Age	Male	37.26	20.48
	Female	42.24	17.85
LT cochlea width	Male	5.58	0.66
	Female	5.52	0.38
LT cochlea height	Male	3.58	0.37
	Female	3.51	0.35
LT cochlea nerve canal width	Male	2.08	1.47
	Female	1.89	0.21
LT basal turn width	Male	1.87	0.19
	Female	1.89	0.17
LT means of CT number of the cochlea nerve	Male	277.78	153.11
	Female	282.95	172.41
RT cochlea width	Male	5.64	0.41
	Female	5.53	0.39
RT cochlea height	Male	3.57	0.36
	Female	3.47	0.34
RT cochlea nerve canal width	Male	1.95	0.21
	Female	1.89	0.19
RT basal turn width	Male	1.88	0.19
	Female	1.87	0.16
RT mean of CT number of the cochlea nerve	Male	322.26	394.61
	Female	273.30	145.83
Cranium/transverse diameter	Male	121.11	9.15
	Female	122.03	4.25

Table 2
Measurements of left and right cochlea in (Group A)

Variables	Left cochlea	Right cochlea
Cochlea width/mm	5.59±0.58	5.60± 0.40
Cochlea height/mm	3.56± 1.23	3.54± 0.36
Cochlea nerve canal width/mm	2.02±1.20	1.93± 0.20
Basal turn width/mm	1.87 ± 0.18	1.88 ± 0.18
CT number of Cochlea nerve/HU	279.41 ±159.02	306.84 ± 336.99

Table 3
Measurements of left and right cochlea in (Group B)

Variables	Right cochlea	Left cochlea
Cochlea width/mm	5.38± 0.46	5.34 ± 0.30
Cochlea height/mm	3.49± 0.28	3.53 ± 0.25
Cochlea nerve canal width/mm	1.73± 0.18	1.75 ± 0.18
Basal turn width/mm	1.79 ± 0.13	1.76 ± 0.13
CT number of Cochlea nerve/HU	196.58± 230.05	232.84± 316.82

Table 4
Measurements of Cochlea's in (Group B) and (Group A)

Variables	Group	Variables Mean	Std. Deviation
LT Cochlea width	B	5.34	0.30
	A	5.59	0.58
LT Cochlea height	B	3.53	0.25
	A	3.56	1.23
LT Cochlea nerve canal width	B	1.75	0.18
	A	2.02	1.20
LT Basal turn width	B	1.76	0.13
	A	1.87	0.18
RT Cochlea width	B	5.38	0.46
	A	5.60	0.40
RT Cochlea height	B	3.49	0.28
	A	3.54	0.36
RT Cochlea nerve canal width	B	1.73	0.18
	A	1.93	0.20
RT Basal turn width	B	1.79	0.13
	A	1.88	0.18

DISCUSSION

The familiarity of the anatomical measurement of the cochlea is insufficient compared to other sensory systems due to difficulties in examining inner ear structures. The measurement of cochlea dimensions using High resolutions computed tomography was implemented to evaluate the length of the cochlea using a CT scan. The length was described to be about 33.01 ± 2.31 mm (Nemzek et al., 1996). In this study cochlear dimension was measured for both normal subjects and subjects with congenital deafness in different ages, and for both genders using high resolution computed tomography.

The variables in group A demonstrated that the width of the right and left cochlear were found to be 5.60 ± 0.40 mm and 5.59 ± 0.58 mm, respectively, and the height of right and left cochlear were found to be 3.54 ± 0.36 mm and 3.56 ± 1.23 mm, the width of cochlear basal turn were a measure to be 1.87 ± 0.19 mm for the left cochlear and 1.88 ± 0.18 mm for the right cochlear. Also the widths of right and left Cochlear nerve canals were 2.02 ± 1.2 mm, 1.93 ± 0.2 mm, respectively.

In addition, CT attenuation for the cochlear nerve was measured, and it was found to be 279.41 ± 159.02 HU for the left cochlear nerve and 306.84 ± 336.9 HU for the right cochlear. These were the standard measures of normal Sudanese cochlea found by this study. In group B, the measurement was performed for Sudanese patients who have congenital hearing loss. For this group, the width of the right and left cochlear was found to be 5.38 ± 0.46 mm and 5.34 ± 0.30 mm, respectively. The height of right and left cochlear

was found to be 3.49 ± 0.28 mm and 3.53 ± 0.25 mm, respectively. The width of the cochlear basal turn was a measure to be 1.76 ± 0.13 mm for the left cochlear and 1.79 ± 0.13 mm for the right cochlear. Also the width of the right and left cochlear nerve canal was 1.73 ± 0.18 mm and 1.75 ± 0.18 mm, respectively. CT attenuation for the cochlear nerve in group B was 232.84 ± 316.82 HU for the left cochlear nerve and 196.58 ± 230.05 HU for the right cochlear nerve.

This study confirmed that there was no anatomical differences in the cochlear measurements on the left and right side in both groups (Table 1 & 2).

Furthermore, our results found that there was no significant differences between the measurement of the cochlea in both gender (Table 1). Mclay et al. (2002) explained that there was a possible variation between cochlear width, height with age. The height and width of cochlear in our study did not change with age; this was in agreement with the study done by Mori and Chang (2012) who concluded that there was no change in cochlear height from childhood to adulthood.

The importance of measuring cochlear height in this study is because all physicians believe in the use of cochlear height in the diagnosis of Sensorial hearing loss for the patients diagnosed with hypoplasia or hyperplasia. Concerning the cochlear height, we found a cochlear height difference in a normal population (Group A) and patients suffering congenital hearing loss (Group B). Using axial CT scans 3.54 ± 0.36 mm was found for the right cochlea and 3.56 ± 1.23 mm for the left cochlea in group A and 3.49 ± 0.28 mm for the right cochlea and 3.53 ± 0.25 mm for the left cochlea in group B. In the studies conducted by Purcell et al. (2006), they found the cochlear height for a normal population on coronal CT scans of the temporal bone was ranged from 4.4 mm to 6.2 mm, while cochlear height in control patient was ranged from 4.1 mm to 2.7 mm, $p > 0.001$). Our results on cochlear height in both group were in agreement with the results of Purcell et al. (2006).

In the current study, we noticed that there was a significant relation ($p < .05$) between age and cochlear nerve attenuation, a width of the right cochlear nerve, and a transverse cranial dimension. Our results differ from the results found by Teissier et al. (2010). This difference might be due to the fact that our study was preformed in the adult population, and Teissier et al. (2010) focused on the pediatric group.

Furthermore, Our results demonstrate that the size of cochlea increase with age and this is compatible with the study conducted by Marshall (1981), where they found that the cochlear and cochlear nerve size increased with increasing the age. Moreover, when comparing the size of the cochlea in both normal hearing subjects and subjects suffering from hearing loss, we reported that the size of the cochlea in normal hearing subjects was significantly different from subjects with hearing loss (Table 4).

Moreover, this occurred both when the cochlea was considered abnormal at visual inspection and when no major malformation was diagnosed.

Our results conformed to what was acknowledged by Teissier et al. (2010). Moreover, our results showed that the normal cochlear in both groups showed no noticeable difference radiographically. However, an embryological or genetic defect, from a clinical point of view, may influence the size of the ears, and this caused a congenital hearing loss. In the current study, the size of the cochleas in people who have hearing loss it appears seemed smaller when compared to normal people (Table 4). Our results which is similar to the study conducted by Fatterpekar et al. (2000) where they found the width and height of the cochlear neural canal in people with hearing loss was significantly smaller when compared to a group of normal people.

Furthermore, the results obtained in this study showed that there was a significant difference in CT numbers between the two groups where it was found that the CT Number in a normal people group was higher than in people who suffer from hearing loss (Table 2 & 3). To our knowledge, the auditory nerve CT numbers in our study decreased with increasing age. Our findings on Age-Related differences in CT numbers were compatible with the results of Stafford et al. (1988), where they found the CT density numbers decreased linearly with increasing age.

The heterogeneity of the population limits this study because of the randomized selection process. Other limitations of this study were: i) There was a relatively small cohort sample; ii) Lack of previous research studies on the topic iii) A high standard deviation was recorded due wider range of age variation. However, to the best of our knowledge, numerous authors have attempted to characterize normative measurements of inner ear structures when evaluating SNHL by CT.

However, this study reflects the necessity of coming across the normative values to differentiate between the normal from pathological conditions, additionally necessary for accurate interpretation of the inner ears radiographs, which thus signifies the importance of this study.

CONCLUSION

In conclusion, this study had established baseline measurements for the cochlear for the healthy Sudanese population. These baseline measurements should be included in the checklist when interpreting a CT scan in cases of SNHL. Furthermore, radiographically there is no significant difference in cochlear radiographic appearance regarding ages and gender. Besides, we found that HRCT of the temporal bone is the best for investigation of the cochlear and can provide a guide for the clinicians to manage congenital hearing loss.

ACKNOWLEDGMENTS

We thank all the hospitals from which we collected the data and for their patience during the preparation of this article.

REFERENCES

- Emanuel, D., Sumalai, M., & Letowski, T. (2009). Auditory function: Physiology and function of the hearing system. *Helmet-Mounted Displays: Sensory, Perceptual, and Cognitive Issues, Edition, 1*, 307-334.
- Fatterpekar, G. M., Mukherji, S. K., Alley, J., Lin, Y., & Castillo, M. (2000). Hypoplasia of the bony canal for the cochlear nerve in patients with congenital sensorineural hearing loss: Initial observations. *Radiology*, *215*(1), 243-246. <https://doi.org/10.1148/radiology.215.1.r00ap36243>
- Fernando, A., Jesus, B., Oplencia, A., Maglalang, G., & Chua, A. (2011). An anatomical study of the cochlea among filipinos using high-resolution computed tomography scans. *Philippine Journal of Otolaryngology Head and Neck Surgery*, *26*(1), 6-9. <https://doi.org/10.32412/pjohns.v26i1.591>
- Joshi, V. M., Navlekar, S. K., Kishore, G. R., Reddy, K. J., & Kumar, E. V. (2012). CT and MR imaging of the inner ear and brain in children with congenital sensorineural hearing loss. *Radiographics*, *32*(3), 683-698. <https://doi.org/10.1148/rg.323115073>
- Marshall, L. (1981). Auditory processing in aging listeners. *The Journal of Speech and Hearing Disorders*, *46*(3), 226-240. <https://doi.org/10.1044/jshd.4603.226>
- McClay, J. E., Tandy, R., Grundfast, K., Choi, S., Vezina, G., Zalzal, G., & Willner, A. (2002). Major and minor temporal bone abnormalities in children with and without congenital sensorineural hearing loss. *Archives of Otolaryngology-Head & Neck Surgery*, *128*(6), 664-671. <https://doi.org/10.1001/archotol.128.6.664>
- Mori, M. C., & Chang, K. W. (2012). CT analysis demonstrates that cochlear height does not change with age. *American Journal of Neuroradiology*, *33*(1), 119-123. <https://doi.org/10.3174/ajnr.A2713>
- Nemzek, W. R., Brodie, H. A., Chong, B. W., Babcock, C. J., Hecht, S. T., Salamat, S., Ellis, W. G., & Seibert, J. A. (1996). Imaging Findings of the Developing Temporal Bone in Fetal Specimens. *American Journal of Neuroradiology*, *17*(8), 1467-1477.
- Purcell, D. D., Fischbein, N. J., Patel, A., Johnson, J., & Lalwani, A. K. (2006). Two Temporal Bone Computed Tomography Measurements Increase Recognition of Malformations and Predict Sensorineural Hearing Loss. *The Laryngoscope*, *116*(8), 1439-1446. <https://doi.org/10.1097/01.mlg.0000229826.96593.13>
- Stafford, J. L., Albert, M. S., Naeser, M. A., Sandor, T., & Garvey, A. J. (1988). Age-related differences in computed tomographic scan measurements. *Archives of Neurology*, *45*(4), 409-415. <https://doi.org/10.1001/archneur.1988.00520280055016>
- Teissier, N., Van Den Abbeele, T., Sebag, G., & Elmaleh-Berges, M. (2010). Computed tomography measurements of the normal and the pathologic cochlea in children. *Pediatric Radiology*, *40*(3), 275-283. <https://doi.org/10.1007/s00247-009-1423-2>
- Tian, Q., Linthicum, F. H., & Fayad, J. N. (2006). Human cochleae with three turns: An unreported malformation. *The Laryngoscope*, *116*(5), 800-803. <https://doi.org/10.1097/01.mlg.0000209097.95444.59>
- Wageih, G. (2017). Ear Anatomy. *Global Journal of Otolaryngology*, *4*(1), Article 555630. <https://doi.org/10.19080/GJO.2017.04.555630>
- WHO. (2008). *The global burden of disease 2004*. World Health Organization.

Prevalence of Vitreous & Retinal Disorders among Sudanese Diabetic Patients: A B-Scan Ultrasonography Study

Mohamed Yousef^{1,2}, Safaa Bashir², Awadalla Wagealla³, Mogahid Zidan^{3*}, Mahmoud Salih Babiker⁴ and Mona Mohamed²

¹Radiologic Sciences Program, Batterjee Medical College, Jeddah, Makkah Province 23819-6700, Saudi Arabia
²College of Medical Radiological Sciences, Sudan University of Science and Technology, 79371 Khartoum, Sudan

³Al-Ghad International College of Applied Medical Science, Radiological Sciences Department, 62217 Abha, Saudi Arabia

⁴Department of Diagnostic Radiologic Technology, College of Applied Medical Sciences, Taibah University, Al-Madinah Al-Munawarah, 42221 Saudi Arabia

ABSTRACT

Retina and vitreous abnormalities represent the most common eye disorders in diabetic patients; they may be associated with severe complications. Therefore, this study aimed to study the prevalence of vitreous and retinal pathologies in diabetic patients using B-Scan ultrasound (U/S). A total of two hundred and three Sudanese diabetic patients with long diabetic disease duration (mean 16.28 ± 4.830) years were enrolled in a descriptive-analytical study. 55% (n = 112) were males and 45% (n = 91) were females. The mean age of the participants was 62.28 ± 8.041 (range between 30-79 years -old).

The study was conducted in a Sudanese ophthalmologic hospital in Khartoum, during the period from 2016–2019. A Nidek (Echoscan US–4000) - B-scan ultrasound unit with 10 MHZ transducer was used. A high-frequency direct contact technique was applied. The inclusion criteria included adult diabetic patients. The vitreous and retina disorders were more prevalent in diabetic hypertensive participants 55 % (n = 112). The high frequency of the disorders was observed in age groups: 60–69 and 50–59 years-old. The most common disorder was

ARTICLE INFO

Article history:

Received: 25 August 2020

Accepted: 19 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.33>

E-mail addresses:

mohnajwan@gmail.com (Mohamed Yousef)

jojo.zidan999@gmail.com (Safaa Bashir)

awagealla@gc.edu.sa (Awadalla Wagealla)

phd.zidan@gmail.com (Mogahid Zidan)

mozidan@gc.edu.sa (Mahmoud Salih Babiker)

jojo.7obi@yahoo.com (Mona Mohamed)

*Corresponding author

retinal detachment which was detected in 30.5% (n = 62) followed by vitreous changes in 16.3% (n = 33). Posterior vitreous was observed in 15.8% (n = 32), vitreous hemorrhage seen in 15.3% (n = 31), both retinal detachment with vitreous hemorrhage were detected in 11.3% (n = 23), retinal detachment with cataract were reported in 3.4% (n = 7), retinal detachment with Vitreous changes were seen in 3% (n = 6), and other changes were noted in 4.4% (n = 9) of the participants. There is no significant a statistical association between gender/diabetic duration and age with the disorders (P = 0.2, 0.43, and 0.5) respectively. Vitreous & Retinal disorders were more prevalent in diabetic hypertensive patients. The high frequency of the disorders was observed in the age group (50–70). The ultrasound is a useful method in diagnosing Vitreous & Retinal disorders among the diabetics.

Keywords: B-scan, diabetes, retina, ultrasonography, vitreous

INTRODUCTION

Retina and vitreous abnormalities represent the most common eye disorders in diabetic patients; they may be associated with severe complications. Diabetic eye disease and its complications are a leading cause of blindness and visual dysfunction in adults (Moss et al., 1988). The eye is a fluid-filled structure located in the anterior part of the orbit and embedded in the fat, and the tendon's capsule separates it from the orbital wall. The anterior segment forms 1/6th of the eyeball and the posterior segment forms 5/6th of the eyeball. The normal axial length of the eye is 22 mm (Srivastava, 2007). The vitreous body is bounded anteriorly by the lens zonular fibers and posterior lens capsule, anterolaterally by the non-pigmented epithelium of the ciliary body, and posterolaterally by the internal limiting membrane of the retina. The retroental space of Erggelet and the canal of Petit are potential spaces (Phillpotts, 2018). Diabetes mellitus (DM) is a chronic metabolic disorder characterized by increasing blood glucose over a long period.

There are three main diabetes types; Type 1 diabetes once known as juvenile diabetes or insulin-dependent diabetes, Type 1 diabetes occurs when the body fails to produce insulin, Type 2 diabetes is the most common type of diabetes, in which the body cells do not respond to the effect of insulin, this type of DM has a direct relation with obesity. The third is Gestational Diabetes Mellitus (GDM), which happens during the pregnancy when the body can become less sensitive to insulin. GDM generally does not continue after delivery and it does not occur for all women (Falck, 2020).

Ocular abnormalities represent approximately 2% to 3% of all visits to the emergency department (ED) (Tintinalli et al., 2020). These abnormalities can result in permanent vision loss or can be simple. The common eye disorders and diseases that are diagnosed in the ED are vitreous hemorrhage (VH), vitreous detachment (VD), and retinal detachment

(RD). RD is considered to be a true ophthalmologic emergency that requires immediate diagnosis and treatment (Haimann et al., 1982).

Patients with RD may have sudden, painless, monocular vision loss as well as flashes and floaters in the visual field. Similar to VH, and VD, symptoms. Distinguishing these conditions is clinically important for patients' management.

Patients with ophthalmologic abnormality undergo all of these examinations, including ophthalmoscopy, slit-lamp, tonometry, and visual acuity (Hollands et al., 2009). However, the criterion standard for ocular disease diagnosis establishment, such as RD is an ophthalmologic evaluation that includes several procedures such as ultrasound, optical coherence tomography, and dilated ophthalmoscopy (Corbett, 2003; Restori, 2008). These procedures are conducted to visualize the layers of the retina by scanning the posterior chamber of the eye.

Ultrasound has been used by ophthalmologists to evaluate ocular abnormality but has been used commonly by emergency medicine practitioners (Lizzi & Coleman, 2004). Many studies recommended that the physicians could use ocular point-of-care ultrasonography (POCUS) to identify RD in the ED (Baker et al., 2018; Blaivas et al., 2002; Shinar et al., 2011; Yoonessi et al., 2010). Ocular U/S is a modality that can be used to diagnose and identify ocular diseases (Jacobsen et al., 2016). Ultrasound is suitable because it is less expensive, simpler, and safer than other diagnostic imaging modalities. Utilization of ultrasound for ocular disorders evaluation is preferable not only due to eye location but also its fluid-filled structure. The emergency medicine practitioners confirm that ocular anomalies can be identified by using POCUS (Moore et al., 2004). In addition, permanent blindness can be avoided by the early diagnosis and treatment of ocular disorders processes, B-scan ultrasound can help in early detection of ocular diseases.

This study aimed to study the prevalence and types of the vitreous and retina pathologies detected by ultrasound examination in diabetic patients. Also, to compare the U/S findings with study gender and/or age groups.

MATERIALS AND METHODS

A descriptive-analytical study, the study took place in a Sudanese ophthalmologic hospital - in Khartoum, during the period from 2016 –2019.203 Sudanese diabetic patients were included. The inclusion criteria were included; all diabetic patients with a long duration of the disease (more than 10 years). The children and adults with the short duration of the disease were excluded. A Nidek (Echoscan US – 4000) B- Scan ultrasound machine with a frequency 10 MHZ transducer was used. A direct contact technique was applied. Initial examinations were done utilizing high gain ranged from 80 dB to 100 dB and low gain ranged from 60 dB to 70 dB.

Ultrasound Examination Protocol

Ultrasound evaluations of the eye and orbit were performed in the supine or sitting position. The probe was placed directly over the conjunctiva or cornea or placed over closed lids. The former has the advantage of reducing the sound attenuation caused by the lids; however, it requires sterilization of the probe between procedures. A coupling gel was used to provide standoff and avoid attenuation caused by air.

Ethics Consideration

This study was done after receiving the ethical approval from the local ethics committee of the medical radiological sciences college, Sudan University of Science and Technology, Khartoum, Sudan, and informed consent was also taken from the participants.

Statistical Analysis

The data collected during the study were stored in a compact disk in a personal computer. The statistical software SPSS Inc. Version 16.0. Chicago, USA, was utilized to interpret the data. Data were presented as percentages and frequency. A chi-square test was used to evaluate the association between age and gender with sonographic findings of the posterior segment of the eye. The statistical association considered significant when $P < 0.05$.

RESULTS

The study included 203 patients with Diabetic Mellitus (DM). A hundred and twelve of the participants (55%) were male and 45% ($n = 91$) were females. The mean age of the participants was 62.28 ± 8.041 (ranging between 30-79 years -old). Table 1 summarizes the common U/S findings; the most common disorders was retinal detachment which was detected in 30.5% ($n = 62$) followed by vitreous changes in 16.3% ($n = 33$), posterior vitreous detachment seen in 15.8% ($n = 32$), vitreous hemorrhage was seen in 15.3% ($n = 31$), retinal detachment with vitreous hemorrhage were detected in 11.3% ($n = 23$), retinal detachment with cataract were reported in 3.4% ($n = 7$), retinal detachment with Vitreous changes were seen in 3% ($n = 6$), and other changes were noted in 4.4% ($n = 9$) of the participants. The detail of other changes includes (other pathology including hyper mature cataract with vitreous change (5 patients), posterior vitreous detachment with vitreous change (1 patient), retinal cyst (1 patient), vitreous change with axial length defect (1 patient) and high myopia with vitreous change (1 patient).

A cross-tabulation represented in Table 2 presents the associated clinical history of the participants with the U/S findings revealed that the vitreous and retina disorders were more prevalent in diabetic hypertensive participants 55 % (out of 203). Table 3 shows a

cross-tabulation of the U/S findings with the gender of participants, the high frequencies of vitreous and retina disorders were noted in the male gender. No significant statistical association was observed between gender and disorders ($P = 0.2$). Table 4 summarizes a cross-tabulation between age groups and ultrasound findings, the high frequency of the disorders was observed in age groups 60–69, and 18.7 % of the retinal detachments were noted in this age group. Statistical analysis revealed, no significant statistical association between the participants' age and the disorders ($P = 0.5$) Table 5 and 6 demonstrate a cross-tabulation between duration of DM and the U/S findings, regarding Pearson Chi-Square test there is no significant association between sonographic findings and DM duration ($P = 0.4$).

The associated clinical history of the participants besides diabetes shown in Figure 1 which presents 55% of them were diabetic hypertensive. Figure 2 to 5 shows sample images of the B-scan U/S findings, a retinal detachment seen in Figure 2, a posterior vitreous detachment observed in Figure 3, a retinal detachment associated with vitreous hemorrhage noted in Figure 4, and vitreous changes presented in Figure 5.

Table 1

Frequency of ultrasound findings of the study

U/S findings	(n)	(%)
Retinal Detachment	62	30.5
Vitreous Changes	33	16.3
Vitreous Hemorrhage	31	15.3
Vitreous changes & Retinal detachment	6	3.0
Posterior Vitreous Detachment	32	15.8
Retinal Detachment & Vitreous Hemorrhage	23	11.3
Retinal Detachment & Cataract	7	3.4
Others	9	4.4
Total	203	100.0

Table 2
*Associated clinical history of the participants *ultrasound findings cross-tabulation*

Clinical Disorders	Ultrasound findings										Total
	Retinal Detachment	Vitreous Changes	Vitreous Hemorrhage	Retinal detachment & Vitreous changes	Posterior Vitreous Detachment	Vitreous Hemorrhage & Retinal Detachment	Retinal Detachment & Cataract	Other	Total		
No	8	3	1	3	3	1	0	1	20		
Hypertension (HTN)	46	19	5	1	21	8	7	5	112		
Trauma	3	8	20	1	4	6	0	1	43		
HTN+ Trauma	0	2	3	1	0	8	0	1	15		
Thyroid disease	2	0	0	0	1	0	0	0	3		
Heart Disease	2	0	0	0	1	0	0	0	3		
Renal Disease	0	1	0	0	2	0	0	1	4		
Others	1	0	2	0	0	0	0	0	3		
Total	62	33	31	6	32	23	7	9	203		

Table 3
Study gender * ultrasound findings cross-tabulation

Gender	Ultrasound Findings							Total
	Retinal Detachment	Vitreous Changes	Vitreous Hemorrhage	Retinal detachment & Vitreous changes	Posterior Vitreous Detachment	Retinal Detachment & Vitreous Hemorrhage	Retinal Detachment & Cataract	
Male	32	18	14	1	20	17	4	112
Female	30	15	17	5	12	6	3	91
Total	62	33	31	6	32	23	7	203

Table 4
Age groups * ultrasound findings cross-tabulation

Age Groups (years)	Ultrasound Findings							Total
	Retinal Detachment	Vitreous Changes	Vitreous Hemorrhage	Retinal detachment & Vitreous changes	Posterior Vitreous Detachment	Retinal Detachment & Vitreous Hemorrhage	Retinal Detachment & Cataract	
30-39	0	2	1	1	0	0	0	4
40-49	5	2	2	0	3	0	0	12
50-59	10	5	8	0	5	3	1	36
60-69	38	19	17	5	20	16	5	123
70-79	9	5	3	0	4	4	1	28
Total	62	33	31	6	32	23	7	203

Table 5
*Duration of DM * Ultrasound Findings Cross - tabulation*

Duration of the DM (years)	Ultrasound Findings							Total	
	Retinal Detachment	Vitreous Changes	Vitreous Hemorrhage	Retinal detachment + Vitreous changes	Posterior Vitreous Detachment	Retinal Detachment + Vitreous Hemorrhage	Retinal Detachment + Cataract		Others
10-14	27	11	12	6	17	7	3	4	87
15-19	12	10	10	0	6	6	0	1	45
20-24	20	12	7	0	7	7	4	4	61
25-29	1	0	2	0	2	2	0	0	7
30-34	2	0	0	0	0	1	0	0	3
Total	62	33	31	6	32	23	7	9	203

Table 6

Chi-Square Test for the association of the diabetic duration with the ultrasound Findings.

	Chi-Square Tests		
	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	34.762 ^a	34	.431
Likelihood Ratio	44.397	34	.109
N of Valid Cases	203		

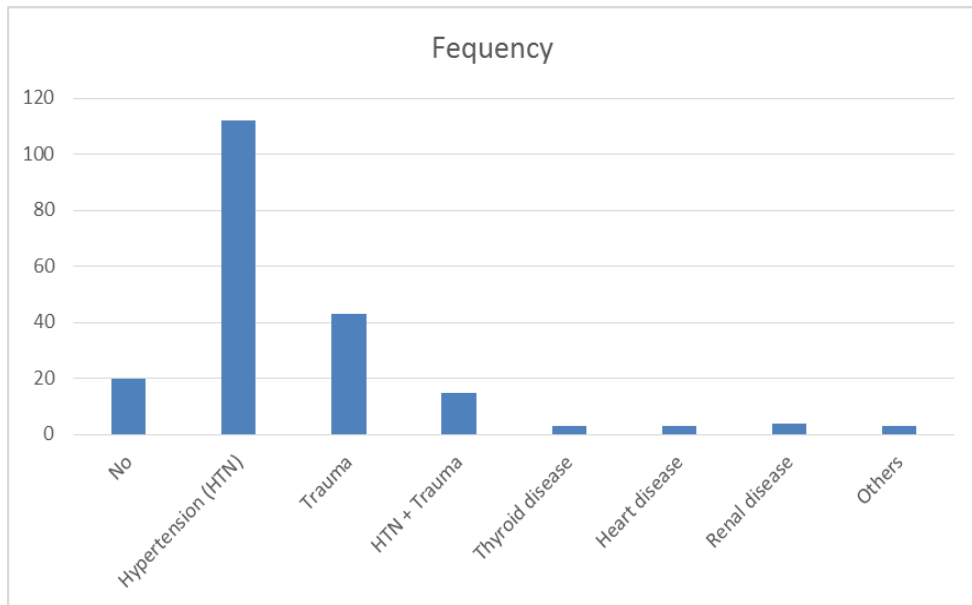


Figure 1. Clinical history of the participants

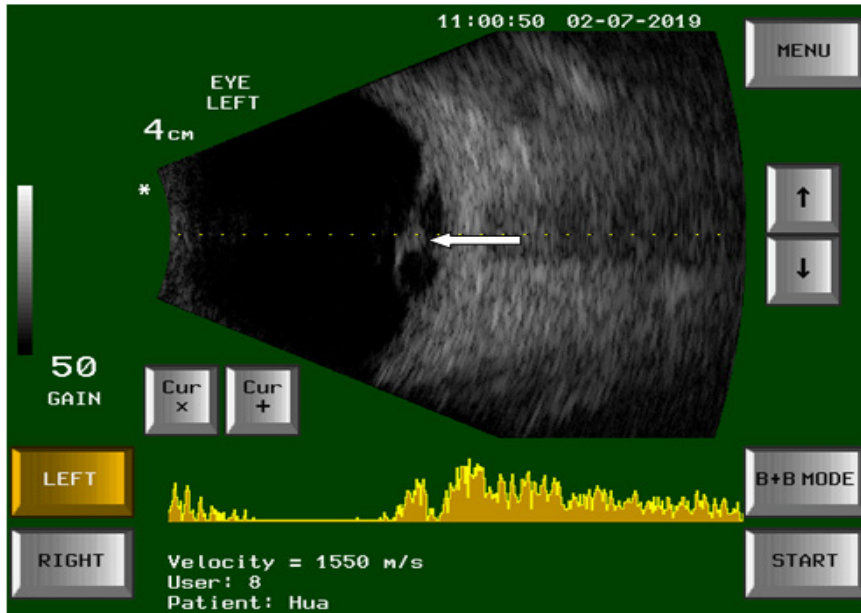


Figure 2. A transverse view of B-scan image showing a retinal detachment in 65 years- old female

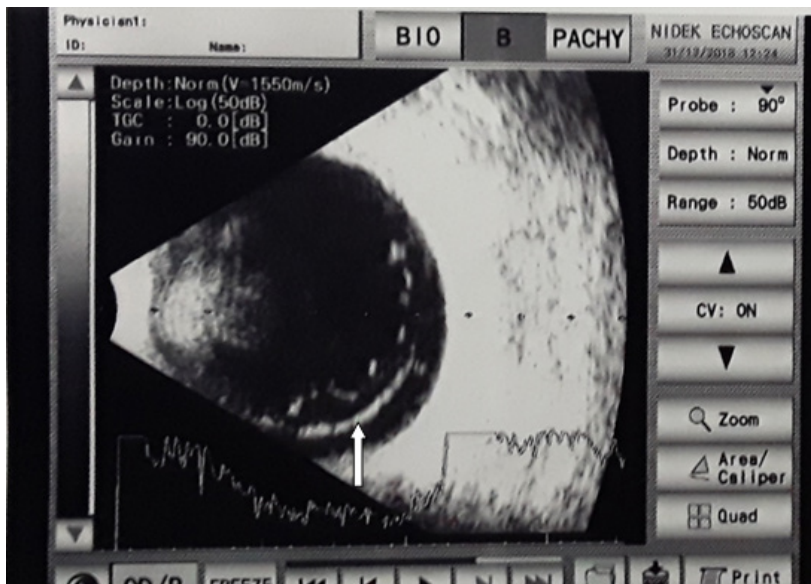


Figure 3. A transverse B-scan image for 61 years- old female showing a posterior vitreous detachment



Figure 4. A transverse B-Scan image for 59 years- old female showing a retinal detachment (white arrows) + vitreous hemorrhage (black arrows).

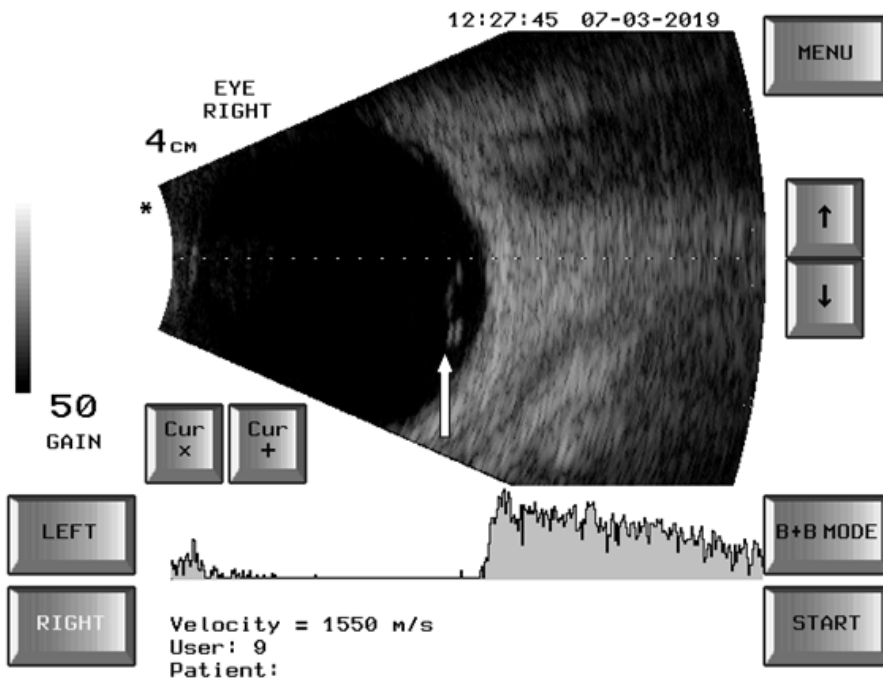


Figure 5. A transverse B-Scan image for 59 years- old female showing floaters (vitreous changes).

DISCUSSION

Ophthalmic ultrasound (U/S) has become an indispensable diagnostic modality that has increased the chance of detecting and differentiating many orbital and ocular disorders. According to our aim, this study was created to study the prevalence of vitreous and retina disorders among diabetic patients using B-Scan ultrasonography. Regarding the age, the study by Moss et al. (1988) revealed that the most affected age groups were 60–69 years (n = 123) and 50–59 years-old (n = 36). The result agrees with the study done by Nanda et al. (2017). In a study by Qureshi and Laghari (2010), maximum number of patients was also present in the age group of 60-69 years.

The study showed that the most common disorders of vitreous and retina were retinal detachment which was detected in 30.5% (n = 62) followed by vitreous changes in 16.3% (n = 33). Posterior vitreous detachment was observed in 15.8% (n = 32), vitreous hemorrhage was seen in 15.3% (n = 31) (Table 1). Dawood et al. (2008) have proposed that ultrasound examination has greatly advanced and this has enabled us to study the orbits even in the presence of opaque media. Rabinowitz et al. (2004) indicated that the most common causes of vitreous hemorrhage continue to be proliferative diabetic retinopathy and ocular trauma. These two causes account for 35% and 33% of all vitreous hemorrhage.

Retinal detachment (including patients of RD combined with vitreous disorders) was seen in 30 (15%) cases. Sharma (2005) concluded that the common finding was retinal detachment 21.31 % (out of 122). Javed et al. (2007) reported that out of 463 patients, the patients of retinal detachment were 68 (14.70%). According to the patient history, the study highlighted that the patient with a history of hypertension was more affected (55.2 %, n=112) (Table 2). These results agree with Jitendrakumar and Ram (2018) who had reported on the main risk factor for other retinal disorders such as occlusion of a retinal artery or vein and diabetic retinopathy. Besides, diabetes with hypertension combined extremely increases the risk of vision loss.

The current study found that the male was more affected than the female, the frequency in males was 112, and in females were 91 (Table 3). The findings agree with studies done by Nanda et al. (2017) and Jitendrakumar and Ram (2018).

Qureshi and Laghari (2010) showed that the high frequency of vitreous & retinal disorders was found in the age group of 60-69 years old. These results support the present study, which showed that 74.4% of the detected vitreous and retinal disorders were observed in the age group ≥ 60 years- old (Table 4). This agrees with the finding of the study done by Jitendrakumar and Ram (2018).

The main limitation is this study was conducted in a single center, another comparative study is needed. The study was performed by more than one sonographer. The study sample size is not large enough to generalize the results.

CONCLUSION

B-scan ultrasonography can be widely utilized in evaluating the prevalence of vitreoretinal diseases. It is a reliable, safe, cheap, rapid investigation, and a feasible option. Vitreous and retinal disorders were more prevalent in diabetic hypertensive participants 55%. The high frequencies of the disorders were observed in age groups 60–69 and 50–59 years-old. Further studies were recommended.

ACKNOWLEDGMENT

The authors would like to thank the staff of the Sudanese ophthalmologic hospital, Khartoum, Sudan, without whom this study would not have been possible.

REFERENCES

- Baker, N., Amini, R., Situ-LaCasse, E. H., Acuña, J., Nuño, T., Stolz, U., & Adhikari, S. (2018). Can emergency physicians accurately distinguish retinal detachment from posterior vitreous detachment with point-of-care ocular ultrasound? *YAJEM American Journal of Emergency Medicine*, *36*(5), 774-776. <https://doi.org/10.1016/j.ajem.2017.10.010>
- Blaivas, M., Theodoro, D., & Sierzenski, P. R. (2002). A study of bedside ocular ultrasonography in the emergency department. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, *9*(8), 791-799. <https://doi.org/10.1197/aemj.9.8.791>
- Corbett, J. J. (2003). The bedside and office neuro-ophthalmology examination. *Seminars in Neurology*, *23*(1), 063-076.
- Dawood, Z., Mirza, S. A., & Qadeer, A. (2008). Role of B-scan ultrasonography for posterior segment lesions. *Journal of the Liaquat University of Medical and Health Sciences*, *7*(1), 07-12.
- Javed, E. A., Ch, A. A., Ahmad, I., & Hussain, M. (2007). *Diagnostic Applications of? B-Scan? Pakistan Journal of Ophthalmology*, *23*(2), 80-83.
- Falck, S. (2020, June 17). *Diabetes: Symptoms, treatment, and early diagnosis*. Retrieved June, 17, 2020, from <https://www.medicalnewstoday.com/articles/323627#using-insulin>
- Haimann, M. H., Burton, T. C., & Brown, C. K. (1982). Epidemiology of retinal detachment. *Archives of Ophthalmology*, *100*(2), 289-292. <https://doi.org/10.1001/archophth.1982.01030030291012>
- Hollands, H., Johnson, D., Brox, A. C., Almeida, D., Simel, D. L., & Sharma, S. (2009). Acute-onset floaters and flashes: Is this patient at risk for retinal detachment? *Journal of American Medical Association*, *302*(20), 2243-2249. <https://doi.org/10.1001/jama.2009.1714>
- Jacobsen, B., Lahham, S., Lahham, S., Patel, A., Spann, S., & Fox, J. C. (2016). Retrospective review of ocular point-of-care ultrasound for detection of retinal detachment. *The Western Journal of Emergency Medicine*, *17*(2), 196-200. <https://doi.org/10.5811/westjem.2015.12.28711>
- Jitendrakumar, P. K. J., & Ram, A. (2018). A study of role of B-scan in evaluating posterior segment pathology of eye. *Journal of Dental and Medical Sciences*, *17*(10), 41-47.

- Lizzi, F. L., & Coleman, D. J. (2004). History of ophthalmic ultrasound. *Journal of Ultrasound in Medicine*, 23(10), 1255-1266. <https://doi.org/10.7863/jum.2004.23.10.1255>
- Moore, C. L., Gregg, S., & Lambert, M. (2004). Performance, training, quality assurance, and reimbursement of emergency physician-performed ultrasonography at academic medical centers. *Journal of Ultrasound in Medicine*, 23(4), 459-466. <https://doi.org/10.7863/jum.2004.23.4.459>
- Moss, S. E., Klein, R., & Klein, B. E. K. (1988). The incidence of vision loss in a diabetic population. *Ophthalmology*, 95(10), 1340-1348. [https://doi.org/10.1016/S0161-6420\(88\)32991-X](https://doi.org/10.1016/S0161-6420(88)32991-X)
- Nanda, D. R., Gupta, D., & Sodani, D. P. (2017). Role of B-scan ultrasonography in evaluating posterior segment of the eye in the event of non visualization of fundus. *Journal of Medical Science And Clinical Research*, 5(07), 25049-25055. <https://dx.doi.org/10.18535/jmscr/v5i7.124>
- Phillpotts, B. A. (2018, September 07). Vitreous hemorrhage. Retrieved September 17, 2018, from <https://emedicine.medscape.com/article/1230216-overview>
- Qureshi, M. A., & Laghari, K. (2010). Role of B-scan ultrasonography in pre-operative cataract patients. *International Journal of Health Sciences*, 4(1), 31-37.
- Rabinowitz, R., Yagev, R., Shoham, A., & Lifshitz, T. (2004). Comparison between clinical and ultrasound findings in patients with vitreous hemorrhage. *Eye*, 18, 253-256. <https://doi.org/10.1038/sj.eye.6700632>
- Restori, M. (2008). Imaging the vitreous: Optical coherence tomography and ultrasound imaging. *Eye*, 22(10), 1251-1256. <https://doi.org/10.1038/eye.2008.30>
- Sharma, O. (2005). Orbital sonography with it's clinico-surgical correlation. *Indian Journal of Radiology and Imaging*, 15(4), 537-554. <https://doi.org/10.4103/0971-3026.28792>
- Shinar, Z., Chan, L., & Orlinsky, M. (2011). Use of ocular ultrasound for the evaluation of retinal detachment. *The Journal of Emergency Medicine*, 40(1), 53-57. <https://doi.org/10.1016/j.jemermed.2009.06.001>
- Srivastava. (2007). Step by step ophthalmic ultrasound with photo CD-ROM (1st Ed.). Jaypee Brothers Medical Publishers (P) Ltd.
- Tintinalli, J. E., Ma, O. J., Yealy, D. M., Meckler, G. D., Stapczynski, J. S., Cline, D., & Thomas, S. H. (2020). *Tintinalli's emergency medicine: A comprehensive study guide*, 9e. McGraw-Hill Education.
- Yoonessi, R., Hussain, A., & Jang, T. B. (2010). Bedside ocular ultrasound for the detection of retinal detachment in the emergency department. *Academic Emergency Medicine*, 17(9), 913-917. <https://doi.org/10.1111/j.1553-2712.2010.00809.x>

Poly lactide and its Composites on Various Scales of Hardness

Abraham Kehinde Aworinde^{1*}, Eyere Emagbetere², Samson Oluropo Adeosun^{1,3}
and Esther Titilayo Akinlabi^{1,4}

¹Mechanical Engineering Department, College of Engineering, Covenant University, Ota, Nigeria

²Mechanical Engineering Department, Federal University of Petroleum Resources, Effurun, Nigeria

³Metallurgical and Materials Engineering Department, University of Lagos, Nigeria

⁴Pan African University for Life and Earth Sciences Institute (PAULESI), Ibadan, Nigeria

ABSTRACT

Poly lactide (PLA) has become a widely applied material. Its hardness property has, however, not been a subject of intense study. This study attempts to examine the hardness values of Poly lactide and its composites on ten hardness scales. Poly lactide composites were developed using three reinforcements (i.e., chitosan, chitin, and titanium powders). The compositing method was the melt-blending technique. Vickers microindentation test was carried out on all the developed samples. The experimental values obtained were related to nine (9) other scales of hardness via an online reference interface. Results showed that the Brinell and Rockwell hardness scales agreed, to a large extent, with the experimental values from several studies. Hence, this work can serve as a reference material on the Brinell and Rockwell hardness values of the unreinforced and reinforced composites considered in this study. The developed materials were also represented on the Mohs scale of hardness with unreinforced PLA having the least value of hardness which corresponds to the value of gypsum on the Mohs scale while the PLA reinforced with 8.33 weight (wt.) % of titanium powder has the highest value of hardness corresponding to the value of a material in-between calcite and fluorite. The hardness values obtained on Shore scleroscope could not agree with the experimental values from various studies. Succinctly, the three particulate fillers increased the hardness properties of PLA. The results of this study would go a long way in helping industrialists and researchers in the correct applications of PLA and its composites.

ARTICLE INFO

Article history:

Received: 19 October 2020

Accepted: 30 December 2020

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.34>

E-mail addresses:

abraham.aworinde@covenantuniversity.edu.ng

(Abraham Kehinde Aworinde)

emagbetere.eyere@fupre.edu.ng (Eyere Emagbetere)

sadeosun@unilag.edu.ng (Samson Oluropo Adeosun)

etakinlabi@uj.ac.za (Esther Titilayo Akinlabi)

* Corresponding author

Keywords: Annealed polylactic acid; biodegradable polymer; Knoop scale; poly lactide composites; Rockwell hardness scale

INTRODUCTION

Poly lactide (PLA) or polylactic acid (PLA) has been well researched in recent times because of some of its preferred properties compared to the petroleum-based polymers (Adeosun et al., 2016). Contingent upon its desirable characteristics, it has grown to become a widely applied biopolymer (Abreu et al., 2017; Deepthi et al., 2016). However, there is a need to improve on some of its drawbacks in order to give it a wider area of applications (Abreu et al., 2017). Researchers, in recent years, have worked to improve on its hydrophobicity (Aworinde et al., 2020a; Hendrick & Frey, 2014; Kurzina et al., 2020; Qi et al., 2019), modulus of elasticity (Adeosun et al., 2016; Aworinde et al., 2020b; Wang et al., 2016), brittleness (Kunwar et al., 2012; Liu & Zhang, 2011; Sennan & Pumchusak, 2014; Song et al., 2014), and so on. Notwithstanding the efforts that have been deployed to widen the areas of application of PLA through properties modification, there appears to be a need for more intense characterisation. The hardness properties of PLA, for instance, has not been seen as an advantage. Improving on it has, therefore, not been the focus of researchers. Yet studies have shown that improved hardness is an empirical result of reinforcing PLA with particulate fillers (Aworinde et al., 2020c; Aworinde et al., 2019a; Feng et al., 2016), and several particulate-reinforced PLA composites have been developed for both the soft (Adeosun et al., 2016; Deepthi et al., 2016; Gbenebor et al., 2018) and the hard (Aworinde et al., 2020c; Wang et al., 2016) tissues applications. Aside from the use of particulate reinforcements, developing PLA composites requires the use of certain polymer composites processing techniques (Aworinde et al., 2018; Aworinde et al., 2019b; Deepthi et al., 2016), which often impact the hardness properties of PLA significantly (Vian & Denton, 2018).

The microindentation properties of a material can be examined on several available scales of hardness. The understanding of the values of the hardness of a material such as PLA on various scales would enhance the right application of the material. The hardness test is one of the most performed mechanical tests because of its usefulness and versatility and Vickers hardness has been reported to be one of the oldest methods of determining the hardness properties of materials because of its wide hardness scale (Moore & Booth, 2015). From Vickers (HV) hardness test, for instance, a number of other helpful parameters can be measured to give a reasonable prediction of other mechanical properties such as fracture toughness (Bergner et al., 2007; Moradkhani et al., 2013) as described in our previous study (Aworinde, et al., 2020c). In addition to the Vickers hardness test, a few other available hardness tests include (though not limited to) Brinell hardness test (HB), Knoop hardness test (HK), Janka hardness test, Meyer hardness test, Rockwell hardness test (HR), Shore durometer hardness test and Barcol hardness test are a few of the available hardness tests.

Apart from the fact that a wide knowledge of material's hardness predisposes it to a wide area of applications, material for implants are expected to possess, among other critical mechanical properties, some degrees of wear resistance (Chandrasekaran, 2010).

In addition, recent studies are focusing on the abrasive/wear resistance of PLA for diverse applications (Eutionnat-Diffo et al., 2020; Mohd et al., 2019). Interestingly, research has shown that there exists a relationship between material hardness and its wear resistance (Faria et al., 2007; Luyckx & Love, 2004; Mezlini et al., 2009). The premise is that the harder the material, the greater the wear resistance. In the light of several possibilities ahead of PLA, this work assays to present its hardness values and those of its composites on various scales of hardness. This material of the future (i.e., PLA), as it is often called, has the propensity to replace many materials in the nearest future (Anderson & Shenkar, 2021; Cooper, 2013). It, therefore, becomes necessary to examine its hardness on various scales of hardness, because different materials possess widely varying hardnesses and require different scales for testing.

MATERIALS AND METHODS

The materials used in this study include PLA, chitosan, chitin, and titanium powder. Our earlier work (Aworinde et al., 2020a; Aworinde et al., 2020c) details the properties of the materials and how they were obtained. Melt-blending technique was used to produce three PLA composites with chitosan, chitin, and titanium (Ti-6Al-2Sn-2Mo-2Cr-0.25Si) powders as reinforcements. The three composites produced are designated as PLA/Ch (i.e., chitosan reinforced PLA), PLA/Ct (i.e., chitin reinforced PLA) and PLA/Ti (i.e., titanium-reinforced PLA). The maximum particle size of chitosan and chitin was 75 μm while the average particle size of titanium powder was 67.5 μm .

Chitosan and chitin are organic fillers, while titanium is an inorganic filler. The choice of these fillers thus helped to compare the effects of organic and inorganic reinforcements on the hardness value of PLA. Our previous studies explain, in detail, the composite fabrication method employed in the development of the samples (Aworinde et al., 2020a; Aworinde et al., 2020c), matrix-fillers formulation and the details of Vickers microhardness characterisation (Aworinde et al., 2020c). The microhardness indentation dwell times were chosen in accordance to ASTM standard (ASTM-E384, 2017).

In order to obtain the approximate equivalent hardness values of Polylactide and Polylactide composites on other hardness scales, the experimental values obtained from the Vickers hardness test was supplied to the interface of an online engineering reference (Efunda, 2020). Approximate values of hardness on Brinell hardness scale with 3000 kgf indentation load, Brinell hardness scale with 500 kgf indentation load, Knoop scale, Mohs scale, Rockwell B-scale, Rockwell Superficial 15T Scale, Rockwell Superficial 30T Scale, Rockwell Superficial 45T Scale and Shore Scleroscope were obtained.

RESULTS AND DISCUSSION

It was observed throughout the results obtained from this study that chitin improved the hardness of PLA better than chitosan. This is undoubtedly connected with the stronger structural property of Chitin over Chitosan, because of the difference in their degrees of de-acetylation. Also, an investigation between the hardness property of PLA/Ct and PLA/Ti showed superior hardness values of the later over PLA/Ct. This was in order since metallic powders are expected to produce polymer composites with higher mechanical properties than would a natural polymer filler (Premalal et al., 2002). However, further addition of titanium beyond 8.33 wt. % showed a drastic reduction in hardness, making PLA/Ti to be 27.66 % and 1.51% less hard than PLA/Ct and PLA/Ch, respectively. Traditionally, this is not expected since titanium is far structurally stronger than Chitin and Chitosan as observed in other composites with less concentration. These exceptions may be due to poor dispersion of the reinforcement in the matrix as Ti increased, thereby leading to poor miscibility: a condition peculiar to polymer-metal composites than polymer-polymer composites (Pozuelo et al., 2017; Souza et al., 2019).

Table 1 presents the Vickers hardness values of PLA and its composites obtained from the experiment. Table 2-10 give the approximate figures of hardness values on several hardness scales, namely: Brinell hardness scale with 3000 kgf indentation load (Table 2), Brinell hardness scale with 500 kgf indentation load (Table 3), Knoop scale (Table 4), Mohs scale (Table 5), Rockwell B-scale (Table 6), Rockwell Superficial 15T Scale (Table 7), Rockwell Superficial 30T Scale (Table 8), Rockwell Superficial 45T Scale (Table 9) and Shore Scleroscope (Table 10). The equivalent values on other hardness scales which could not be obtained from the online reference have been indicated by << (i.e., below the minimum obtainable value) and >> (i.e., above the maximum obtainable value).

The approximate value of PLA/Ti at 8.33 wt. % of titanium filler on the Brinell 10 mm Standard 3000 kgf Scale (i.e., 254 BRH) relates well with the value obtained from experiments on a Brinell scale of hardness (i.e., 265 BRH) when PLA was reinforced with 15 wt. % of carbon (Mohan et al., 2019). Although several studies did not state, in particular, the type of Rockwell scale of hardness used, the values of Rockwell B-scale, Rockwell Superficial 15T Scale and Rockwell Superficial 30T Scale obtained in this work form a close range with the experimental values obtained for annealed PLA (Farah et al., 2016). This shows that the values obtained in this study approximate, with minimal error, the hardness values of PLA and its composites on the various scale of hardness. The predicted values of hardness on the shore scleroscope scale seem not to be in agreement with the experimental values reported in various studies on Shore D scale (Byrne et al., 2009; Valerga et al., 2020). The values on Shore D scale from various studies are generally higher than the approximate equivalent values obtained from the online reference on Shore scleroscope scale. This may be attributed to the fact that there is no good correlation between the Vickers hardness scale and shore scleroscope scale.

The values of PLA and its composites on the Mohs scale imply that an unreinforced, heat-treated PLA was as hard as gypsum and can, therefore, be scratched with the fingernail (King, 2020). PLA composites, however, have higher hardness values. PLA/Ch, for instance, has a hardness value between gypsum (i.e. 2) and calcite (i.e. 3) while PLA/Ct with 8.33 wt. % has the same hardness value as calcite. On the other hand, PLA/Ti with 8.33 wt. % reinforcement has a value that is characteristic of materials with the hardness value between calcite and fluorite. Studies are very scanty on the hardness values of PLA and its composites on the Mohs scale of hardness as well as on the Knoop scale.

Table 1
Vickers hardness scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	72.8	72.8	72.8
1.04	76.1	150.6	151.8
2.08	90.2	130.2	161.8
4.17	107.2	148.1	180.5
8.33	148.4	203.9	268.1
16.67	167.7	166.6	165.2

Table 2
Brinell 10 mm standard 3000 kgf scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	69	69	69
1.04	72	144	145
2.08	86	124	154
4.17	102	141	172
8.33	141	194	254
16.67	160	159	157

Table 3
Brinell 10 mm standard 500 kgf scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	<<	131	132
2.08	81	114	140
4.17	95	129	154
8.33	129	170	>>
16.67	144	144	142

Table 4
Knoop scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	90	90	90
1.04	93	163	164
2.08	106	143	174
4.17	122	160	192
8.33	160	214	278
16.67	179	178	177

Table 5
Mohs scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	2.0	2.0	2.0
1.04	2.0	2.5	2.5
2.08	2.5	2.5	2.5
4.17	2.5	2.5	3.0
8.33	2.5	3.0	3.5
16.67	2.5	2.5	2.5

Table 6
Rockwell B-scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	32	79	79
2.08	47	72	82
4.17	60	78	87
8.33	78	92	95
16.67	84	84	83

Table 7
Rockwell superficial 15T scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	<<	87	87
2.08	76	84	88
4.17	80	86	90
8.33	86	91	94
16.67	88	88	88

Table 8
Rockwell superficial 30T scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	<<	70	70
2.08	48	65	72
4.17	56	69	76
8.33	69	79	>>
16.67	74	73	73

Table 9
Rockwell superficial 45T scale

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	<<	53	54
2.08	20	45	57
4.17	33	52	62
8.33	52	67	>>
16.67	58	58	58

Table 10
Shore scleroscope

Filler (wt. %)	PLA/Ch	PLA/Ct	PLA/Ti
0.00	<<	<<	<<
1.04	<<	22	22
2.08	<<	19	24
4.17	16	22	26
8.33	22	29	38
16.67	25	24	24

CONCLUSION

A holistic view of the consequence of particulate reinforcement of PLA, as it affects its hardness property, has been taken in this study. Unreinforced PLA and its three composites prepared by melt-blending technique were subjected to Vickers microhardness test. The results were examined on nine (9) other scales of hardness via an online reference interface and the outcome compared with the experimental results from other studies. The results form a database for comparing the hardness properties of PLA and its composites with the hardness values of various materials (e.g. petroleum-based plastics) for the possibility of getting a viable replacement. Also, it was observed that:

- i. in tandem with other research results, the particulate reinforcements used increased the hardness values of the matrix (i.e., PLA)
- ii. PLA composites as hard as calcite and fluorite can be produced using the simple melt-blending technique.
- iii. the Vickers hardness scale bears a good correlation with several other scales of hardness such as Brinell and Rockwell, but not with the shore scleroscope scale of hardness.
- iv. the availability of data as presented in this study would minimise the number of experimental runs and hence reduce material and time wastages in research and development.

ACKNOWLEDGEMENT

Authors are thankful for the financial support received from Covenant University, Ota, Nigeria.

REFERENCES

- Abreu, A. S. L. M., de Moura, I. G., de Sá, A. V., & Machado, A. V. A. (2017). Biodegradable polymernanocomposites for packaging applications. In A. M. Grumezescu (Ed.), *Food Packaging* (pp. 329-363). Academic Press. <https://doi.org/10.1016/B978-0-12-804302-8.00010-8>
- Adeosun, S. O., Aworinde, A. K., Diwe, I. V., & Olaleye, S. A. (2016). Mechanical and microstructural characteristics of rice husk reinforced polylactide nanocomposite. *The West Indian Journal of Engineering*, 39(2), 63-71.
- Anderson, G., & Shenkar, N. (2021). Potential effects of biodegradable single-use items in the sea: Polyactic acid (PLA) and solitary ascidians. *Environmental Pollution*, 268, Article 115364. <https://doi.org/10.1016/j.envpol.2020.115364>
- ASTM-E384. (2017). *Standard test method for microindentation hardness of materials*. ASTM International.
- Aworinde, A. K., Adeosun, S. O., Oyawale, F. A., Akinlabi, E. T., & Akinlabi, S. A. (2020a). Comparative effects of organic and inorganic bio-fillers on the hydrophobicity of polylactic acid. *Results in Engineering*, 5, 1-3. <https://doi.org/10.1016/j.rineng.2020.100098>
- Aworinde, A. K., Adeosun, S. O., & Oyawale, F. A. (2020b). Mechanical properties of poly (L-Lactide) -based composites for hard tissue repairs. *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, 9(5), 2152-2155.
- Aworinde, A. K., Adeosun, S. O., Oyawale, F. A., Emagbetere, E., Ishola, F. A., Olatunji, O., Akinlabi, S. A., Oyedepo, S. O., Ajayi, O. O., & Akinlabi, E. T. (2020c). Comprehensive data on the mechanical properties and biodegradation profile of polylactide composites developed for hard tissue repairs. *Data in Brief*, 32, Article 106107. <https://doi.org/10.1016/j.dib.2020.106107>
- Aworinde, A. K., Adeosun, S. O., Oyawale, F. A., Akinlabi, E. T., & Akinlabi, S. A. (2019a). The strength characteristics of chitosan- and titanium- poly (L-lactic) acid based composites. In *Journal of Physics: Conference Series* (Vol. 1378, No. 2, p. 022061). IOP Publishing. <https://doi.org/10.1088/1742-6596/1378/2/022061>
- Aworinde, A. K., Adeosun, S. O., Oyawale, F. A., Akinlabi, E. T., & Akinlabi, S. A. (2019b). Parametric effects of fused deposition modelling on the mechanical properties of polylactide composites: A review. *Journal of Physics: Conference Series*, 1378, Article 022060. <https://doi.org/10.1088/1742-6596/1378/2/022060>
- Aworinde, A. K., Adeosun, S. O., Oyawale, F. A., Akinlabi, E. T., & Emagbetere, E. (2018, October 29 - November 1). Mechanical strength and biocompatibility properties of materials for bone internal fixation: A brief overview. In *Proceedings of the International Conference on Industrial Engineering and Operations Management* (pp. 2115-2126). Johannesburg, South Africa.
- Bergner, F., Schaper, M., Hammer, R., Jurisch, M., Kleinwechter, A., & Chudoba, T. (2007). Indentation response of single-crystalline GaAs in the nano-, micro-, and macroregime. *International Journal of Materials Research (formerly Zeitschrift fuer Metallkunde)* 98(8), 735-741. <https://doi.org/10.3139/146.101531>

- Byrne, F., Ward, P. G., Kennedy, J., Imaz, N., Hughes, D., & Dowling, D. P. (2009). The effect of masterbatch addition on the mechanical, thermal, optical and surface properties of poly(lactic acid). *Journal of Polymers and the Environment*, 17(1), 28-33. <https://doi.org/10.1007/s10924-009-0119-x>
- Chandrasekaran, M. (2010). Forging of metals and alloys for biomedical applications. In M. Niinomi (Ed.), *Metals for Biomedical Devices* (pp. 235-250). Elsevier. <https://doi.org/10.1533/9781845699246.3.235>
- Cooper, T. A. (2013). Developments in bioplastic materials for packaging food, beverages and other fast-moving consumer goods. In N. Farmer (Ed.), *Trends in Packaging of Food, Beverages and Other Fast-Moving Consumer Goods (FMCG)* (pp. 108-152). Woodhead Publishing <https://doi.org/10.1533/9780857098979.108>
- Deepthi, S., Sundaram, M. N., Kadavan, J. D., & Jayakumar, R. (2016). Layered chitosan-collagen hydrogel/aligned PLLA nanofiber construct for flexor tendon regeneration. *Carbohydrate Polymers*, 153, 492-500. <https://doi.org/10.1016/j.carbpol.2016.07.124>
- Efunda. (2020). *Convert hardness: Vickers*. Retrieved September 21, 2020, from https://www.efunda.com/units/hardness/convert_hardness.cfm?HD=HV&Cat=Steel#ConvInto
- Eutionnat-Diffo, P. A., Chen, Y., Guan, J., Cayla, A., Campagne, C., & Nierstrasz, V. (2020). Study of the wear resistance of conductive poly lactic acid monofilament 3D printed onto polyethylene terephthalate woven materials. *Materials*, 13(10), Article 2334. <https://doi.org/10.3390/ma13102334>
- Farah, S., Anderson, D. G., & Langer, R. (2016). Physical and mechanical properties of PLA, and their functions in widespread applications - A comprehensive review. *Advanced Drug Delivery Reviews*, 107, 367-392. <https://doi.org/10.1016/j.addr.2016.06.012>
- Faria, A. C. L., Benassi, U. M., Rodrigues, R. C. S., Ribeiro, R. F., & de Mattos, M. D. G. C. D. (2007). Analysis of the relationship between the surface hardness and wear resistance of indirect composites used as veneer materials. *Brazilian Dental Journal*, 18(1), 60-64. <https://doi.org/10.1590/S0103-64402007000100013>
- Feng, P., Peng, S., Wu, P., Gao, C., Huang, W., Deng, Y., & Shuai, C. (2016). A space network structure constructed by tetra-needlelike ZnO whiskers supporting boron nitride nanosheets to enhance comprehensive properties of poly(L-lactid acid) scaffolds. *Scientific Reports*, 6(August), 1-15. <https://doi.org/10.1038/srep33385>
- Gbenebor, O. P., Atoba, R. A., Akpan, E. I., Aworinde, A. K., Adeosun, S. O., & Olaleye, S. A. (2018). Study on polylactide-coconut fibre for biomedical applications. In *TMS Annual Meeting & Exhibition* (pp. 263-273). Springer. https://doi.org/10.1007/978-3-319-72526-0_24
- Hendrick, E., & Frey, M. (2014). Increasing surface hydrophilicity in poly(lactic acid) electrospun fibers by addition of Pla-B-Peg co-polymers. *Journal of Engineered Fibers and Fabrics*, 9(2), 153-164. <https://doi.org/10.1177/155892501400900219>
- King, H. M. (2020). *Mohs hardness scale a rapid hardness test for field and classroom use*. Retrieved September 21, 2020, from <https://geology.com/minerals/mohs-hardness-scale.shtml>
- Kunwar, A., Gurung, R., Park, S. G., & Lim, J. K. (2012). Effect of hydrothermally prepared graft copolymer addition on a brittle matrix: A preliminary study on glass fiber reinforced PLA/LLDPE-g-MA composite. *Advanced Materials Research*, 530, 46-51. <https://doi.org/10.4028/www.scientific.net/AMR.530.46>

- Kurzina, I. A., Laput, O. A., Zuza, D. A., Vasenina, I. V., Salvadori, M. C., Savkin, K. P., Lytkina, D. N., Botvin, V. V., & Kalashnikov, M. P. (2020). Surface & coatings technology surface property modification of biocompatible material based on polylactic acid by ion implantation. *Surface & Coatings Technology*, 388, Article 125529. <https://doi.org/10.1016/j.surfcoat.2020.125529>
- Liu, H., & Zhang, J. (2011). Research progress in toughening modification of poly(lactic acid). *Journal of Polymer Science, Part B: Polymer Physics*, 49(15), 1051-1083. <https://doi.org/10.1002/polb.22283>
- Luyckx, S., & Love, A. (2004). The relationship between the abrasion resistance and the hardness of WC-Co alloys. *Journal of the South African Institute of Mining and Metallurgy*, 104(10), 579-582.
- Mezlini, S., Kapsa, P., Abry, J. C., Meille, G., Ribes, H., & Dif, R. (2009). Relationship between hardness and abrasive wear for some aluminium alloys. *Materials Science Forum*, 396, 1517-1524. <https://doi.org/10.4028/www.scientific.net/MSF.396-402.1517>
- Mohan, A. E., Habeeb, H. A., & Abood, A. H. (2019). Experimental and modeling stress concentration factor (SCF) of a tension poly lactic acid (PLA) plate with two circular holes. *Periodicals of Engineering and Natural Sciences*, 7(4), 1733-1742. <http://dx.doi.org/10.21533/pen.v7i4.916>
- Mohd, A. Z. F., Bavishi, V., Sharma, R., & Nagarajan, R. (2019). Barrier properties and abrasion resistance of biopolymer-based coatings on biodegradable poly(lactic acid) films. *Polymer Engineering and Science*, 59(9), 1874-1881. <https://doi.org/10.1002/pen.25187>
- Moore, P., & Booth, G. (2015). Mechanical testing of welds. In *The Welding Engineer's Guide to Fracture and Fatigue* (pp. 113-141). Woodhead Publishing Oxford.
- Moradkhani, A., Baharvandi, H., Tajdari, M., Latifi, H., & Martikainen, J. (2013). Determination of fracture toughness using the area of micro-crack tracks left in brittle materials by Vickers indentation test. *Journal of Advanced Ceramics*, 2(1), 87-102. <https://doi.org/10.1007/s40145-013-0047-z>
- Pozuelo, M., Hwang, I., Javadi, A., Yang, Y., Zhao, J., Lin, T. C., Cao, C., & Li, X. (2017). Stretching micro metal particles into uniformly dispersed and sized nanoparticles in polymer. *Scientific Reports*, 7(1), 3-7. <https://doi.org/10.1038/s41598-017-07788-3>
- Premalal, H. G. B., Ismail, H., & Baharin, A. (2002). Comparison of the mechanical properties of rice husk powder filled polypropylene composites with talc filled polypropylene composites. *Polymer Testing*, 21(7), 833-839. [https://doi.org/10.1016/S0142-9418\(02\)00018-1](https://doi.org/10.1016/S0142-9418(02)00018-1)
- Qi, Y., Ma, H. L., Du, Z. H., Yang, B., Wu, J., Wang, R., & Zhang, X. Q. (2019). Hydrophilic and Antibacterial Modification of Poly(lactic acid) Films by γ -ray Irradiation. *ACS Omega*, 4(25), 21439-21445. <https://doi.org/10.1021/acsomega.9b03132>
- Sennan, P., & Pumchusak, J. (2014). Improvement of mechanical properties of poly(lactic acid) by elastomer. *Malaysian Journal of Analytical Sciences*, 18(3), 669-675.
- Song, X., Chen, Y., Xu, Y., & Wang, C. (2014). Study on tough blends of polylactide and acrylic impact modifier. *BioResources*, 9(2), 1939-1952.
- Souza, P. J., Lira, S. H. A., & de Oliveira, I. N. (2019). Wetting dynamics of ferrofluids on substrates with different hydrophilicity behaviors. *Journal of Magnetism and Magnetic Materials*, 483, 129-135. <https://doi.org/10.1016/j.jmmm.2019.03.069>

- Valerga, A. P., Fernandez-Vidal, S. R., Girot, F., & Gamez, A. J. (2020). On the relationship between mechanical properties and crystallisation of chemically post-processed additive manufactured polylactic acid pieces. *Polymers*, 12(4), Article 941. <https://doi.org/10.3390/polym12040941>
- Vian, W. D., & Denton, N. L. (2018). Hardness comparison of polymer specimens produced with different processes. *ASEE Annual Conference and Exposition, Conference Proceedings*, 3, 1-14. <https://doi.org/10.5703/1288284316841>
- Wang, Wang, Y., Ito, Y., Zhang, P., & Chen, X. (2016). A comparative study on the in vivo degradation of poly(L-lactide) based composite implants for bone fracture fixation. *Scientific Report*, 6(1), 1-12. <https://doi.org/10.1038/srep20770>

SETS: A Seed-Dense-Expanding Model-Based Topological Structure for the Prediction of Overlapping Protein Complexes

Soheir Noori^{1,2*}, Nabeel Al-A'araji³ and Eman Al-Shamery¹

¹Department of Software, University of Babylon, Babylon, Hillah, Iraq

²Department of Computer Science, University of Kerbala, Karbala, Iraq

³Ministry of Higher Education, Baghdad, Iraq

ABSTRACT

Defining protein complexes by analysing the protein–protein interaction (PPI) networks is a crucial task in understanding the principles of a biological cell. In the last few decades, researchers have proposed numerous methods to explore the topological structure of a PPI network to detect dense protein complexes. In this paper, the overlapping protein complexes with different densities are predicted within an acceptable execution time using seed expanding model and topological structure of the PPI network (SETS). SETS depend on the relation between the seed and its neighbours. The algorithm was compared with six algorithms on six datasets: five for yeast and one for human. The results showed that SETS outperformed other algorithms in terms of F-measure, coverage rate and the number of complexes that have high similarity with real complexes.

Keywords: Common neighbours; density; protein complex; protein–protein interaction network; topological structure

ARTICLE INFO

Article history:

Received: 26 October 2020

Accepted: 18 February 2021

Published: 30 April 2021

DOI: <https://doi.org/10.47836/pjst.29.2.35>

E-mail addresses:

soheir.noori@uokerbala.edu.iq; soheirn.sw.hdr@student.

uobabylon.edu.iq (Soheir Noori)

nhkaghed@itnet.uobabylon.edu.iq (Nabeel Al-A'araji)

emanalshamery@itnet.uobabylon.edu.iq (Eman Al-Shamery)

* Corresponding author

INTRODUCTION

The key to exploring cell behaviour is understanding the mechanism of protein complexes. According to Pizzuti and Rombo (2014), protein complexes are a molecular aggregation of two or more proteins assembled by multiple PPIs. Protein-protein interaction (PPI) plays a central role in many biological functions and its network provides a global view of cellular functionality.

Advanced experimental techniques have generated a vast amount of data on PPI (Wang et al., 2017). A PPI network is represented as an undirected graph, where the nodes are proteins and the edges are the interactions between proteins. Many algorithms have been proposed for the analysis of PPI networks to discover the protein complex by defining a dense subgraph in the PPI network. Two of the earliest approaches that have been adopted include the Molecular Complex Detection (MCODE) (Bader & Hogue, 2003) and Markov (Van Dongen, 2000) algorithms. Many other algorithms search for cliques as researchers believe that a complete connected graph represents the protein complex such as Cliques Finder (CFinder) (Adamcsek et al., 2006), Clique Percolation- Distance Restriction (CP-DR) (Wang et al., 2010), Maximal Cliques (Liu et al., 2009) and Local Clique Merging Algorithm (LCMA) (Li et al., 2005). All these algorithms detect cliques and then merge them depending on different criteria to identify the protein complex. Most protein complex detection algorithms such as Graph Fragmentation Algorithm (GFA) (Feng et al., 2010), Dynamic Protein Complexes (DPC) (Li et al., 2014), Detect Module from Seed Protein (DMSP) (Maraziotis et al., 2007) use the topological properties of a graph or mix the PPI network with other information like the gene expression. A subgraph is a protein complex having high functional and structural consistency (Hartwell et al., 1999). Since proteins have multiple functions, they can belong to more than one dense subgraph (Palla et al., 2005; Rives & Galitski, 2003). Therefore, a protein complex can have an overlapping structure as is observed in the ClusterOne (Nepusz et al., 2012) and Near-Clique Mining (NCMine) (Tadaka & Kinoshita, 2016) algorithms.

Most of the existing algorithms can detect only highly dense regions as protein complexes and ignore low density complexes (Wang et al., 2018). Further, most of them cannot detect overlapping protein complexes (Zhao & Lei, 2019). In this study, the overlapping protein complexes with different densities are predicted through the seed expanding approach and the topological structure of the PPI networks (SETS) in an acceptable time period, namely less than five minutes for human and about one minute or less for yeast. A pre-processing step needs to be undertaken first to order the proteins according to their degrees. At the start, SETS choose the first node as a seed. Next, the direct neighbours of the seed will be added to the complex, depending on a common neighbour's technique. Notably, the proteins in a complex cannot be chosen as a next seed but can be added to another complex to generate the overlap between complexes. The predicted complex can be accepted if it is greater than the density threshold. Otherwise, all proteins of the complex will return to construct a new complex from a new seed. Finally, the preliminary complexes will be iteratively expanded according to the closeness score. This algorithm outperforms other algorithms.

METHODS

Preliminary Concepts

Generally, PPI is represented as an undirected and unweighted graph $G = (V, E)$, where V is the nodes representative of proteins and E is the edges which represent the interactions between the proteins. The algorithm uses several measurements.

For each vertex v , the degree of v is the summation of its connected edges (Equation 1).

$$d(v) = \sum_i e_i \quad [1]$$

The density of the set of vertices $S \subset V$ is the number of edges among them divided by the number of possible edges between the set nodes (i.e., how close the set to the clique is, ranging between 0 and 1) (Equation 2).

$$\text{density}(S) = \frac{2 \times |E|}{|V| \times (|V| - 1)} \quad [2]$$

The common neighbours (CN) between two proteins (P_i and P_j) are the number of proteins that indices to both divided by the square root of the product of the nodes' degrees (Equation 3).

$$CN = \frac{N_{p_i} \cap N_{p_j}}{\sqrt{d(p_i) * d(p_j)}} \quad [3]$$

The Algorithm

The results from some experiments (Goldberg & Roth, 2003; Peng et al., 2017) have shown that the methods that used the information of common neighbours are reliable. SETS is a technique that is employed to detect overlapping protein complexes based on common neighbours. Given an undirected and non-weighted graph, the goal of the algorithm is to identify overlapping protein complexes with different densities. The algorithm accomplishes this through the following steps (Appendix 1: Algorithm 1):

1. Those proteins with a degree higher than 1 are set in ascending order, put in a queue 'Q' and its visited label is set to 'False'.
2. The preliminary complex is built starting from the seed. The nodes that share a specific ratio of common neighbours are added iteratively and its visited label is set to true if it satisfies the predefined threshold of shared neighbours in order to avoid selecting it as a seed in the next iteration. The complex will be accepted as

a preliminary one if its density is greater than a predefined threshold. Otherwise, all the nodes visited labels will be set to 'false' and moved to the next node in the Q with a false visited label. After defining the preliminary complex, no nodes will be deleted from the Q, so that we can get overlapping complexes. The preliminary complex will be accepted as a candidate complex if it contains more than three proteins and was not previously defined from another seed to avoid redundancy (Appendix 1: Algorithm 1, steps 1-14).

3. The candidate complex (CC) will be iteratively expanded according to the closeness score (CS) as obtained through Equation 4 by adding its neighbours 'N_{CC}' that connect with half or more of the proteins of the candidate complex. The expansion will be done in rounds. In each round, the algorithm will search for proteins that satisfy the threshold of closeness score (T_{CS}) to add them to the complex. In the second round, the algorithm will search for proteins, such as those that relate to the updated complex, until no proteins can satisfy the T_{CS}. This step will assist in the identification of complexes with different densities (Appendix 1: Algorithm 1, steps 15-20).

$$CS(C, C u) = \frac{N_u \cap |V_{CC}|}{|V_{CC}|} \quad [4]$$

4. Redundant complexes will be removed by retaining only one of the exactly matched complexes.

Time Complexity

The execution time has been calculated for each dataset in order to analyse the time complexity of the SETS algorithm. As a pre-processing step, SETS receive a set of ordered nodes Q in increasing order that take O(n²). SETS process each node in Q that has a visible label set to false and adds its neighbours. This process takes O(n*m) and reduces it to O(N) since not all nodes will be processed. Where n and m represent the number of nodes and their neighbours respectively. The second part of SETS expands each candidate complex c. This takes O(c*m), where m stands for the neighbours of the proteins in a complex. The time complexity of SETS is O(n)+O(c*m). SETS is implemented in python on a 64-bit window system with a 2 GB memory and intel CPU i7 2.40 GHz. Table 1 reports SETS execution time in seconds by using the time package in python.

Table 1
SETS execution time in seconds

Dataset	Time in seconds
Collins	0.255
Gavin	0.104
Krogan	0.191
DIP	0.893
BioGRID	67.479
Human	261.536

Comparison of SETS with Other Algorithms

The performance of the algorithm has been compared to those of six others, namely MCODE (Bader & Hogue, 2003), ClusterONE (Nepusz et al., 2012), NCMine (Tadaka & Kinoshita, 2016), SPICi (Jiang & Singh, 2010), IPCA (Li et al., 2008), and PEWCC (Zaki et al., 2013).

MCODE is one of the seed-extension approaches, which identifies overlapping protein complexes in three steps. Step 1, based on the core clustering coefficient, assigns a weight to every node in the graph. Step 2, extending from seeds that have a high weight, finds a dense region in the weighted graph. Finally, subgraphs that are not dense are filtered.

ClusterOne is another algorithm that detects overlapping protein complexes by starting from the seed protein having the highest degree and then gradually adding and removing proteins to find a cohesive group of proteins that can be overlapped.

On the other hand, NCMine defines a near-complete subgraph as a functional module by using the centrality degree as the weight of the nodes, which then iteratively merges these like cliques to define overlapping modules.

SPICi is a fast heuristic clustering algorithm that selects the seed having the highest weight. The weight represents the degree of the node and then uses a support function to expand the way that the density of the cluster is saved.

IPCA is another algorithm that identifies a dense region in the PPI network as a protein complex. It starts from the seed that has the biggest weight, which is the summation of its weighted edge that represents the number of its common neighbours. IPCA then recursively adds the neighbours of the seed based on two criteria: the shortest path between the seed and the node as well as the probability of its interaction.

Another algorithm that evaluates the protein interactions reliability is PEWCC. It uses the weighted clustering coefficient to detect the protein complex.

All the aforementioned algorithms rely on the topological structure of PPI networks and most of them use the seed-extension approach to detect dense protein complexes.

RESULTS AND DISCUSSION

PPI and Benchmark Datasets

The algorithm has been analysed by concentrating on five PPI networks of *Saccharomyces cerevisiae* (yeast) and one network of *Homo sapiens* (human) (Ma et al., 2017). The latter is a combination of data from two databases: HPRD (Human Protein Reference Database) and BioGRID (version 3.2.109). The PPI datasets of the yeast are Collins and Gavin for ClusterONE (Nepusz et al., 2012), DIP (Xenarios et al., 2002), Krogan (Krogan et al., 2006) and BioGRID from SPICi. Table 2 explains the properties of these datasets. Each dataset contains a different number of proteins having a different number of interactions

that create a variety in the density of network to satisfy the diversity that is required in the PPI networks used with the algorithm. NewMIPS (Mewes et al., 2004) and CYC2008 (Pu et al., 2009) are used as benchmark complexes. All datasets are available online from authors and as Appendix 1 and 2.

Table 2
Number of proteins and intersections, and network density in PPI datasets

Datasets	No. of Proteins	No. of Intersections	Network density
Collins	1622	9074	0.007
Gavin	1855	7669	0.004
Krogan	2675	7084	0.002
DIP	4930	17201	0.001
BioGRID	5361	85866	0.006
Human	15459	144687	0.001

Evaluation Metrics

The quality of a predicted complex was evaluated using various metrics. The definitions of these metrics are introduced as follows:

Recall, Precision, and F-Measure. One of the metrics most commonly used to evaluate any algorithm is recall, precision and F-measure. The overlapping score (OS) is the matching score between the predicted complex (C_1) and benchmark complex (C_2), as expressed in Equation 5. C_1 and C_2 are considered matched if the OS between both is equal to or greater than 0.2 (Altaf-Ul-Amin et al., 2006; Bader & Hogue, 2003).

$$OS(C_1, C_2) = \frac{|C_1 \cap C_2|^2}{|C_1| \times |C_2|} \quad [5]$$

Recall and precision are defined as Equation 6 and 7:

$$Precision = \frac{N(C_1)}{|C_1|} \quad [6]$$

$$Recall = \frac{N(C_2)}{|C_2|} \quad [7]$$

$N(C_1)$ is the number of the predicted complex that satisfies the OS score with at least one complex in the benchmark. $N(C_2)$ is the number of the benchmark complex that satisfies the OS score with at least one predicted complex. The F-measure is a combination of recall and precision (Equation 8).

$$F - measure = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad [8]$$

Coverage Rate (CR). CR evaluates the number of proteins that have been covered by the predicted complexes (Brohée & van Helden, 2006; Friedel et al., 2008). CR is defined in Equation 9, where C_2 is the set of benchmark complexes, $maxcom_{ij}$ is the maximal common proteins between the i^{th} benchmark and j^{th} predicted complex divided by N_i protein numbers in i^{th} benchmark complex.

$$CR = \frac{\sum_{i=1}^{|C_2|} Max\{maxcom_{ij}\}}{\sum_{i=1}^{|C_2|} N_i} \quad [9]$$

Exact and High Matching with Real Complexes. The quality of predicted complexes was evaluated by reporting the number of real complexes that exactly match with the predicted complexes and that had an OS score greater than or equal to 0.8, excluding the exact match.

Selection of Parameters

The T_{CN} , DT and T_{CS} parameters had been used in SETS. Proteins that were not in the PPI network had been filtered from the benchmark complexes. Only those complexes with more than two proteins were retained and then filtered again to keep only the complexes that had all their proteins in the PPI network. CN was calculated between the proteins in the same filtered benchmark complex. We also calculated the number of complexes that at least two of their proteins satisfied the CN value (Appendix 2). According to the result of benchmark complexes analyses, the T_{CN} is set for each data. Liu et al. (2010) analysed the protein complexes of CYC2008 (Pu et al., 2009), MI PS (Mewes et al., 2004) and Aloy (Aloy et al., 2004). They found that almost 60% of the complexes had a density equal to or more than 0.5. Therefore, DT was set to 0.5 to define complexes that were dense enough to be the preliminary complexes. T_{CS} was set to at least 0.5 to let only the proteins that had a good closeness to the preliminary complex that was to be added. Table 3 explains the threshold of each dataset.

Quality of Predicted Complexes

The performance of SETS was compared with that of six other approaches using five datasets for yeast and one dataset for human. All datasets were unweighted except SPICi, which used weighted networks. Every parameter in all the algorithms was set to default. In addition, complexes with less than three proteins were ignored. All the algorithms were implemented in the Cytoscape software (Shannon et al., 2003) except SPICi, which was

Table 3
Threshold values for each dataset

Datasets	T _{CN}	DT	T _{CS}
Collins	0.3	0.5	0.6
Gavin	0.3	0.5	0.5
Krogan	0.2	0.5	0.5
DIP	0.1	0.5	0.6
BioGRID	0.2	0.5	0.7
Human data	0.1	0.5	0.7

implemented in its web site. The complex is considered matched if the OS with benchmark complex is greater than or equal to 0.2. SETS have the highest F-measure in all cases and competes with other algorithms in recall and precession (Tables 4 & 5). SETS obtain the highest CR in most cases except in Collins and BioGRID, where it obtained the second-highest CR. Besides a few exceptions where its prediction ranks behind that of the PEWCC, the exact and well-predicted complexes by SETS are the best in most cases (Figure 1). All the results are available in the Appendix 1.

The ProCope software tool (Schlicker et al., 2006) was used to evaluate the biological significance of predicted complexes and the data used in the evaluation process was set to 'default'. The evaluation was based on BP and CC. SETS detects more complexes that are

Table 4
Performance analysis for Gavin data with CYC2008 and NewMIPS

	# complex	Recall	Precession	F-measure	CR
Gavin with CYC2008					
SPiCi	91	0.36	0.76	0.491	0.504
ClusterONE	258	0.508	0.419	0.459	0.633
NCMine	621	0.513	0.393	0.445	0.64
PEWCC	656	0.517	0.402	0.453	0.596
IPCA	464	0.53	0.457	0.491	0.626
MCODE	101	0.021	0.05	0.03	0.118
SETS	246	0.475	0.602	0.531^{1st}	0.656^{1st}
Gavin with NewMIPS					
SPiCi	91	0.372	0.736	0.494	0.248
ClusterONE	258	0.53	0.419	0.468	0.417
NCMine	621	0.549	0.39	0.456	0.422
PEWCC	656	0.552	0.433	0.485	0.392
IPCA	464	0.573	0.47	0.516	0.413
MCODE	101	0.021	0.059	0.031	0.045
SETS	246	0.524	0.607	0.563^{1st}	0.43^{1st}

significant in BioGRID and human datasets (Figure 2) and ranks second with regard to the rest of the datasets, competing with SPICi, IPCA and ClusterONE algorithms (Appendix 1).

SETS predict overlapping complexes as explained in Table 7 that reports some of these complexes that have high OS scores with benchmark complexes and share some of their

Table 5
Performance analysis for Krogan data with CYC2008 and NewMIPS

	# complex	Recall	Precession	F-measure	CR
Krogan with CYC2008					
SPICi	131	0.458	0.641	0.534	0.583
ClusterONE	240	0.492	0.512	0.502	0.598
NCMine	578	0.458	0.433	0.445	0.593
PEWCC	708	0.525	0.496	0.51	0.593
IPCA	472	0.517	0.595	0.553	0.599
MCODE	60	0.03	0.117	0.047	0.111
SETS	220	0.479	0.764	0.589^{1st}	0.68^{1st}
Krogan with NewMIPS					
SPICi	131	0.479	0.618	0.54	0.352
ClusterONE	240	0.442	0.458	0.45	0.323
NCMine	578	0.479	0.427	0.452	0.362
PEWCC	708	0.534	0.476	0.503	0.368
IPCA	472	0.515	0.574	0.543	0.36
MCODE	60	0.021	0.1	0.035	0.038
SETS	220	0.485	0.732	0.583^{1st}	0.391^{1st}

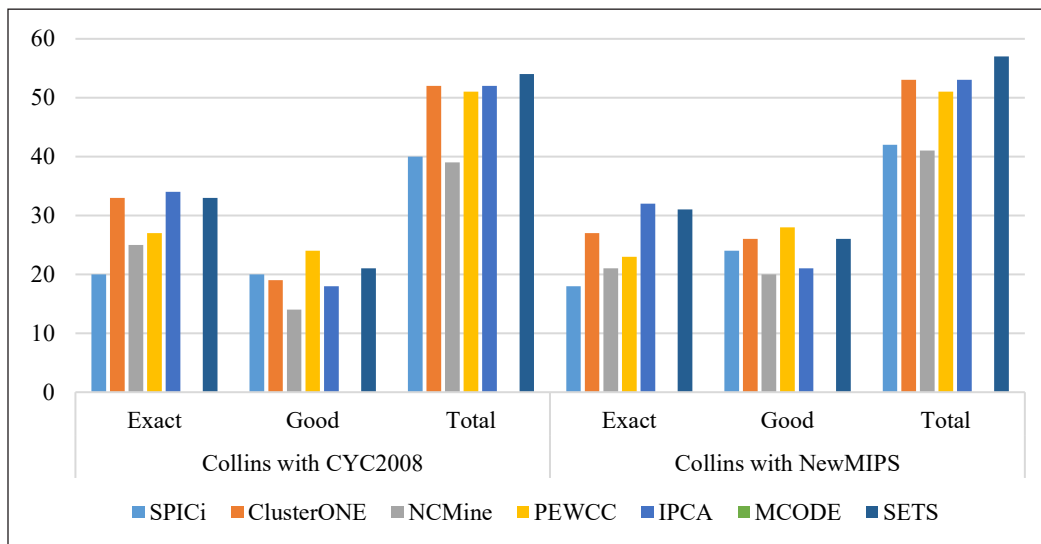


Figure 1. Number of exact and well-predicted complexes in Collins dataset

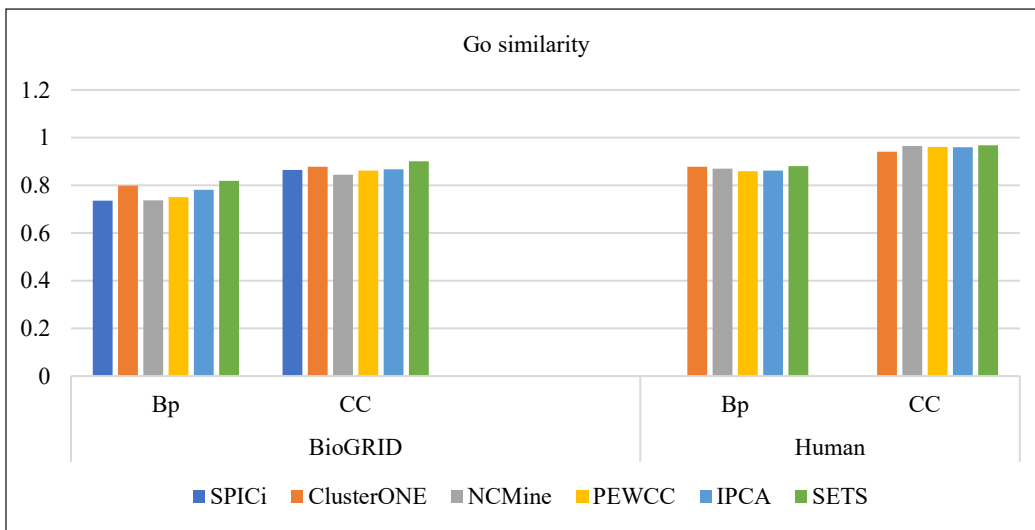


Figure 2. Biological significance of predicted complexes in BioGRID and Human

Table 6
DIP with Newmips reports low density (D.) complexes with a high OS

Real complex		Predicted complex		Inter.	D.	OS	
YKR068C	YBR254C	YDR472W	YLR342W	YKR068C	YBR254C	10 0.36 0.91	
YDR108W	YMR218C	YDR246W	YDR472W	YDR108W	YMR218C		
YML077W	YOR115C	YGR166W	YDR246W	YML077W	YOR115C		
YDR407C			YGR166W	YDR407C			
Length = 10		Length = 11					
YDL005C	YNL236W	YPR070W	YDL005C	YNL236W	YPR070W	21 0.44 0.84	
YOL135C	YNR010W	YDR308C	YOL135C	YNR010W	YDR308C		
YBR193C	YBR253W	YNL025C	YBR193C	YBR253W	YMR112C		
YPR168W	YMR112C	YGL025C	YGL025C	YCR081W	YGL151W		
YCR081W	YGL151W	YOL051W	YOL051W	YOR174W	YLR071C		
YOR174W	YLR071C	YGR104C	YGR104C	YHR041C	YHR058C		
YGL127C	YHR041C	YHR058C	YER022W	YBL093C	YDR443C		
YPL042C	YER022W	YBL093C	Length = 21				
YDR443C							
Length = 25							
Q0080	YDR322C-A	YPL271W	Q0080	Q0130	YDR322C-A	YBL099W	17 0.29 0.8
YPR020W	YBL099W	YML081C-A	YNL315C	YPL271W	YBR039W		
YDR377W	YOL077W-A	YDL004W	YPL078C	YLR295C	YML081C-A		
YKL016C	YDR298C	YGR008C	YPR020W	YDR377W	YDR298C		
YBR039W	YLR295C	Q0085	Q0130	YDL004W	YKL016C	Q0085	
YDL130W-A	YDL181W	YPL078C	YDL181W	YJR121W			
YJR121W			Length = 18				
Length = 20							

Table 6 (continue)

Real complex			Predicted complex			Inter.	D.	OS
YIL084C	YMR075W	YOL004W	YIL084C	YMR075W	YOL004W	10	0.36	0.7
YMR128W	YPL181W	YDL076C	YPL181W	YIL035C	YDL076C			
YPR023C	YMR263W	YNL330C	YNL330C	YMR263W	YLR103C			
YPL139C	YNL097C		YPL139C	YNL097C	YBR095C			
Length = 11			Length = 13					
YKR068C	YBR254C	YDR472W	YLR342W	YKR068C	YBR254C	7	0.36	0.64
YDR108W	YDR246W	YML077W	YDR472W	YDR108W	YMR218C			
YOR115C			YDR246W	YML077W	YOR115C			
Length = 7			YGR166W YDR407C					
			Length = 11					

Note. Inter. is the interaction between predicted and real complexes

Table 7

Predicted overlapping complexes with high OS score from Collins using NewMIPS

Predicted complex	Real complex	OS	Predicted complex	Real complex	OS	Overlapping proteins
YFL039C	YFL039C	1	YFL039C	YFL039C	0.92	YNL107W
YJL081C	YJL081C		YDR334W	YDR334W		YFL039C
YOR244W	YOR244W		YJL081C	YJL081C		YJL081C
YHR090C	YHR090C		YML041C	YLR385C		YGR002C
YNL107W	YNL107W		YNL107W	YML041C		
YNL136W	YNL136W		YBR231C	YNL107W		
YEL018W	YEL018W		YLR085C	YBR231C		
YHR099W	YHR099W		YDR485C	YDR485C		
YPR023C	YPR023C		YDR190C	YLR085C		
YFL024C	YFL024C		YAL011W	YDR190C		
YJR082C	YJR082C		YPL235W	YAL011W		
YDR359C	YDR359C		YGR002C	YPL235W		
YGR002C	YGR002C			YGR002C		
YKL144C	YHR143W-A	0.94	YHR143W-A	YHR143W-A	0.93	YOR210W
YOR210W	YKL144C		YOR341W	YOR341W		YPR110C
YOR116C	YOR210W		YOR210W	YOR210W		YNL113W
YPR190C	YOR116C		YPR110C	YPR110C		YBR154C
YPR110C	YPR190C		YNL113W	YNL113W		YPR187W
YNL113W	YPR110C		YOR340C	YOR340C		YOR224C
YDL150W	YNL113W		YJR063W	YJR063W		
YBR154C	YDL150W		YOR151C	YBR154C		
YPR187W	YBR154C		YBR154C	YNL248C		
YKR025W	YPR187W		YNL248C	YDR156W		
YDR045C	YKR025W		YDR156W	YPR187W		
YNR003C	YDR045C		YPR187W	YPR010C		
YNL151C	YNR003C		YPR010C	YOR224C		
YOR207C	YNL151C		YOR224C	YJL148W		
YJL011C	YOR207C		YJL148W			
YOR224C	YJL011C					
	YOR224C					

proteins. The complexes predicted by SETS are of various densities and not restricted to dense ones as is the case with other algorithms that use the topological structure of PPI. SETS can achieve higher F-measure with different densities of PPI network in contrast with other algorithms whose F-measure decreases when the PPI network density does. Table 6 reports some of the low-density complexes that have high OS scores with benchmark complexes.

ECC vs. CN with SETS

The CN calculation in Algorithm 1 (Appendix 1) is replaced with an edge clustering coefficient (ECC) to compare the F-measure in both cases (Figure 3). SETS with ECC is high only with Collins, which has the highest network density. It, on the other hand, achieved a lower F-measure than SETS with CN in other datasets that have different densities. Radicchi et al. (2004) realized that ECC might not suite for PPI networks as it was disassortative. This was proven with SETS that performed better using CN with different network densities.

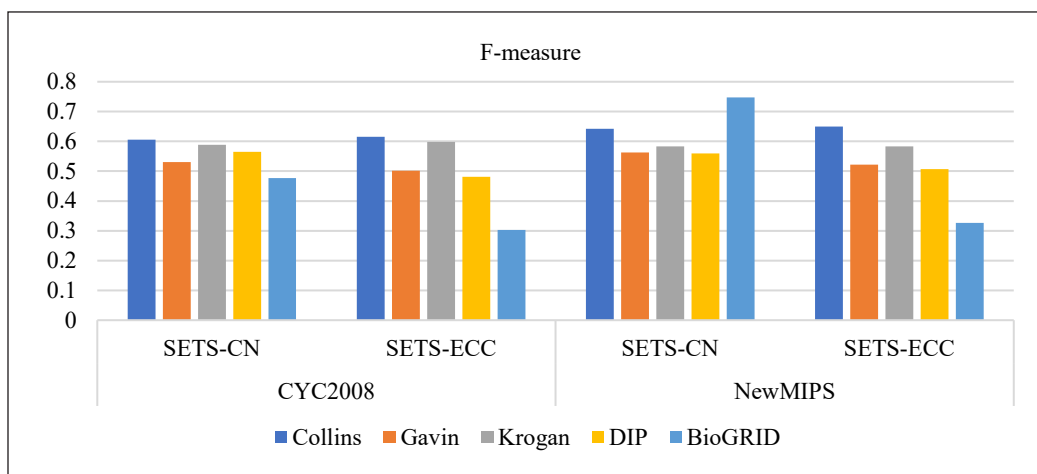


Figure 3. F-measure of ECC vs. CN with SETS

CONCLUSION

In this paper, the seed-expansion model has been proposed based on the topological structure of PPI networks to predict overlapping protein complexes with various densities. The main idea behind this algorithm is (i) choosing the first node in Q that is not visited before as a seed, (ii) adding the seed's neighbour that shares a specific percentage of common neighbours and accepting the complex if its density is more than or equal to the density threshold (DT) and (iii) expanding each complex by adding the proteins that are close to the complex's proteins. SETS could achieve high accuracy in all datasets that have

different densities with good biological significance of predicted complexes compared to other methods. SETS can be further improved by using biological information as gene expression or gene ontology.

ACKNOWLEDGEMENT

We would like to thank the University of Kerbala and the University of Babylon, Iraq for partially sponsoring this research.

REFERENCES

- Adamcsek, B., Palla, G., Farkas, I. J., Derényi, I., & Vicsek, T. (2006). CFinder: locating cliques and overlapping modules in biological networks. *Bioinformatics*, 22(8), 1021-1023. <https://doi.org/10.1093/bioinformatics/btl039>
- Aloy, P., Böttcher, B., Ceulemans, H., Leutwein, C., Mellwig, C., Fischer, S., Gavin, A.-C., Bork, P., Superti-Furga, G., & Serrano, L. (2004). Structure-based assembly of protein complexes in yeast. *Science*, 303(5666), 2026-2029. <https://doi.org/10.1126/science.1092645>
- Altaf-Ul-Amin, M., Shinbo, Y., Mihara, K., Kurokawa, K., & Kanaya, S. (2006). Development and implementation of an algorithm for detection of protein complexes in large interaction networks. *BMC Bioinformatics*, 7, Article 207. <https://doi.org/10.1186/1471-2105-7-207>
- Bader, G. D., & Hogue, C. W. (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4, Article 2. <https://doi.org/10.1186/1471-2105-4-2>
- Brohée, S., & van Helden, J. (2006). Evaluation of clustering algorithms for protein-protein interaction networks. *BMC Bioinformatics*, 7, Article 488. <https://doi.org/10.1186/1471-2105-7-488>
- Feng, J., Jiang, R., & Jiang, T. (2010). A max-flow-based approach to the identification of protein complexes using protein interaction and microarray data. *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, 8(3), 621-634. <https://doi.org/10.1109/TCBB.2010.78>
- Friedel, C. C., Krumsiek, J., & Zimmer, R. (2008). Bootstrapping the interactome: unsupervised identification of protein complexes in yeast. In M. Vingron & L. Wong (Eds.), *Lecture notes in computer science: Research in computational molecular biology* (Vol. 4955, pp. 3-16). Springer. https://doi.org/10.1007/978-3-540-78839-3_2.
- Goldberg, D. S., & Roth, F. P. (2003). Assessing experimentally derived interactions in a small world. *Proceedings of the National Academy of Sciences, USA*, 100(8), 4372-4376. <https://doi.org/10.1073/pnas.0735871100>
- Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, 402(6761), C47-C52. <https://doi.org/10.1038/35011540>
- Jiang, P., & Singh, M. (2010). SPICi: A fast clustering algorithm for large biological networks. *Bioinformatics*, 26(8), 1105-1111. <https://doi.org/10.1093/bioinformatics/btq078>

- Krogan, N. J., Cagney, G., Yu, H., Zhong, G., Guo, X., Ignatchenko, A., Li, J., Pu, S., Datta, N., & Tikuisis, A. P. (2006). Global landscape of protein complexes in the yeast *Saccharomyces cerevisiae*. *Nature*, *440*(7084), 637-643. <https://doi.org/10.1038/nature04670>
- Li, M., Chen, J. E., Wang, J. X., Hu, B., & Chen, G. (2008). Modifying the DPCLUS algorithm for identifying protein complexes based on new topological structures. *BMC Bioinformatics*, *9*, Article 398. <https://doi.org/10.1186/1471-2105-9-398>
- Li, M., Chen, W., Wang, J., Wu, F. X., & Pan, Y. (2014). Identifying dynamic protein complexes based on gene expression profiles and PPI networks. *BioMed Research International*, *2014*, Article 375262. <https://doi.org/10.1155/2014/375262>
- Li, X. L., Foo, C. S., Tan, S. H., & Ng, S. K. (2005). Interaction graph mining for protein complexes using local clique merging. *Genome Informatics*, *16*(2), 260-269. https://doi.org/10.11234/gi1990.16.2_260
- Liu, G., Wong, L., & Chua, H. N. (2009). Complex discovery from weighted PPI networks. *Bioinformatics*, *25*(15), 1891-1897. <https://doi.org/10.1093/bioinformatics/btp311>
- Liu, G., Yong, C. H., Wong, L., & Chua, H. N. (2010, December 18-21). *Decomposing PPI networks for complex discovery* [Paper presentation]. 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Hong Kong, China. <https://doi.org/10.1109/BIBM.2010.5706577>.
- Ma, C. Y., Chen, Y. P. P., Berger, B., & Liao, C. S. (2017). Identification of protein complexes by integrating multiple alignment of protein interaction networks. *Bioinformatics*, *33*(11), 1681-1688. <https://doi.org/10.1093/bioinformatics/btx043>
- Maraziotis, I. A., Dimitrakopoulou, K., & Bezerianos, A. (2007). Growing functional modules from a seed protein via integration of protein interaction and gene expression data. *BMC Bioinformatics*, *8*, Article 408. <https://doi.org/10.1186/1471-2105-8-408>
- Mewes, H. W., Amid, C., Arnold, R., Frishman, D., Güldener, U., Mannhaupt, G., Münsterkötter, M., Pagel, P., Strack, N., & Stümpflen, V. (2004). MIPS: Analysis and annotation of proteins from whole genomes. *Nucleic Acids Research*, *32*(suppl_1), D41-D44. <https://doi.org/10.1093/nar/gkh092>
- Nepusz, T., Yu, H., & Paccanaro, A. (2012). Detecting overlapping protein complexes in protein-protein interaction networks. *Nature Methods*, *9*(5), 471-472. <https://doi.org/10.1038/nmeth.1938>
- Palla, G., Derényi, I., Farkas, I., & Vicsek, T. (2005). Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, *435*(7043), 814-818. <https://doi.org/10.1038/nature03607>
- Peng, X., Wang, J., Peng, W., Wu, F. X., & Pan, Y. (2017). Protein-protein interactions: Detection, reliability assessment and applications. *Briefings in Bioinformatics*, *18*(5), 798-819. <https://doi.org/10.1093/bib/bbw066>
- Pizzuti, C., & Rombo, S. E. (2014). Algorithms and tools for protein-protein interaction networks clustering, with a special focus on population-based stochastic methods. *Bioinformatics*, *30*(10), 1343-1352. <https://doi.org/10.1093/bioinformatics/btu034>
- Pu, S., Wong, J., Turner, B., Cho, E., & Wodak, S. J. (2009). Up-to-date catalogues of yeast protein complexes. *Nucleic Acids Research*, *37*(3), 825-831. <https://doi.org/10.1093/nar/gkn1005>

- Radicchi, F., Castellano, C., Cecconi, F., Loreto, V., & Parisi, D. (2004). Defining and identifying communities in networks. *Proceedings of the National Academy of Sciences, USA*, *101*(9), 2658-2663. <https://doi.org/10.1073/pnas.0400054101>
- Rives, A. W., & Galitski, T. (2003). Modular organization of cellular networks. *Proceedings of the National Academy of Sciences, USA*, *100*(3), 1128-1133. <https://doi.org/10.1073/pnas.0237338100>
- Schlicker, A., Domingues, F. S., Rahnenführer, J., & Lengauer, T. (2006). A new measure for functional similarity of gene products based on Gene Ontology. *BMC Bioinformatics*, *7*, Article 302. <https://doi.org/10.1186/1471-2105-7-302>
- Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*(11), 2498-2504. <https://doi.org/10.1101/gr.1239303>
- Tadaka, S., & Kinoshita, K. (2016). NCMine: Core-peripheral based functional module detection using near-clique mining. *Bioinformatics*, *32*(22), 3454-3460. <https://doi.org/10.1093/bioinformatics/btw488>
- Van Dongen, S. M. (2000). *Graph clustering by flow simulation* [Doctoral dissertation, Utrecht University]. Utrecht University Publication. <https://dspace.library.uu.nl/bitstream/handle/1874/848/full.pdf?sequence=1&isAllowed=y>.
- Wang, J., Liu, B., Li, M., & Pan, Y. (2010). Identifying protein complexes from interaction networks based on clique percolation and distance restriction. *BMC Genomics*, *11*, Article S10. <https://doi.org/10.1186/1471-2164-11-S2-S10>
- Wang, R., Liu, G., Wang, C., Su, L., & Sun, L. (2018). Predicting overlapping protein complexes based on core-attachment and a local modularity structure. *BMC Bioinformatics*, *19*, Article 305. <https://doi.org/10.1186/s12859-018-2309-9>
- Wang, Y., You, Z., Li, X., Chen, X., Jiang, T., & Zhang, J. (2017). PCVMZM: Using the probabilistic classification vector machines model combined with a zernike moments descriptor to predict protein-protein interactions from protein sequences. *International Journal of Molecular Sciences*, *18*(5), Article 1029. <https://doi.org/10.3390/ijms18051029>
- Xenarios, I., Salwinski, L., Duan, X. J., Higney, P., Kim, S. M., & Eisenberg, D. (2002). DIP, the database of interacting proteins: A research tool for studying cellular networks of protein interactions. *Nucleic Acids Research*, *30*(1), 303-305. <https://doi.org/10.1093/nar/30.1.303>
- Zaki, N., Efimov, D., & Berenguères, J. (2013). Protein complex detection using interaction reliability assessment and weighted clustering coefficient. *BMC Bioinformatics*, *14*, Article 163. <https://doi.org/10.1186/1471-2105-14-163>
- Zhao, J., & Lei, X. (2019). Detecting overlapping protein complexes in weighted PPI network based on overlay network chain in quotient space. *BMC Bioinformatics*, *20*, Article 682. <https://doi.org/10.1186/s12859-019-3256-9>

APPENDIX 1

Algorithm 1

Inputs: Q that contains ordered proteins

Output: The sets of predicted protein complexes (COMPLEXES).

1. **For** each protein in Q
2. **IF** visited_label == False
3. Add protein to complex set (COMP)
4. Set visited_label of protein to True
5. **For** each neighbour of protein
6. Find the common neighbours (CN) between protein and neighbours
7. **IF** CN \geq T_{CN}
8. Add neighbour to COMP
9. Set visited_label of neighbour to True
10. **IF** density(COMP) \geq DT **and** COMP IS not in COMPLEXES
11. Add COMP to COMPLEXES
12. **ELSE**
13. **For** each protein in COMP
14. Set visited_label of protein to False
15. **For** each complex in COMPLEXES
16. **For** each round
17. Find neighbours N_{CC} of complex's proteins
18. **For** each protein in N_{CC}
19. **IF** CS(CC, protein) \geq T_{CS}
20. Add protein to complex
21. **Return** complexes

Collins

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
CYC2008										
SPICi	0.419	0.736	0.534	0.69	106	78	70	20	20	40
ClusterONE	0.559	0.547	0.553	0.797	203	111	103	33	19	52
NCMine	0.517	0.475	0.495	0.763	377	179	71	25	14	39
PEWCC	0.53	0.521	0.525	0.738	426	222	89	27	24	51
IPCA	0.542	0.64	0.587	0.751	342	219	68	34	18	52
MCODE	0.051	0.107	0.069	0.121	103	11	79	0	0	0
SETS	0.521	0.725	0.606	0.767	218	158	70	33	21	54
NewMips										
SPICi	0.473	0.726	0.573	0.443	106	77	70	18	24	42
ClusterONE	0.588	0.542	0.564	0.519	203	110	103	27	26	53
NCMine	0.537	0.501	0.518	0.493	377	189	71	21	20	41
PEWCC	0.546	0.533	0.539	0.479	426	227	89	23	28	51
IPCA	0.567	0.705	0.628	0.486	342	241	68	32	21	53
MCODE	0.03	0.087	0.045	0.055	103	9	79	0	0	0
SETS	0.555	0.761	0.642	0.488	218	166	70	31	26	57

Note. R: Recall, P: Precision, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

Gavin

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
CYC2008										
SPICi	0.36	0.76	0.491	0.504	91	70	13	14	10	24
ClusterONE	0.508	0.419	0.459	0.633	258	108	40	11	22	33
NCMine	0.513	0.393	0.445	0.64	621	244	43	9	14	23
PEWCC	0.517	0.402	0.453	0.596	656	264	36	11	20	31
IPCA	0.53	0.457	0.491	0.626	464	212	37	15	19	34
MCODE	0.021	0.05	0.03	0.118	101	5	137	0	0	0
SETS	0.475	0.602	0.531	0.656	246	148	37	12	25	37
NewMips										
SPICi	0.372	0.736	0.494	0.248	91	67	13	11	15	26
ClusterONE	0.53	0.419	0.468	0.417	258	108	40	11	19	30
NCMine	0.549	0.39	0.456	0.422	621	242	43	10	16	26
PEWCC	0.552	0.433	0.485	0.392	656	284	36	13	21	34
IPCA	0.573	0.47	0.516	0.413	464	218	37	17	25	42
MCODE	0.021	0.059	0.031	0.045	101	6	137	0	0	0
SETS	0.524	0.607	0.563	0.43	246	159	37	13	31	44

Note. R: Recall, P: Precession, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

Krogan

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
CYC2008										
SPICi	0.458	0.641	0.534	0.583	131	84	20	17	15	32
ClusterONE	0.492	0.512	0.502	0.598	240	123	23	12	15	27
NCMine	0.458	0.433	0.445	0.593	578	250	25	5	17	22
PEWCC	0.525	0.496	0.51	0.593	708	351	31	15	24	39
IPCA	0.517	0.595	0.553	0.599	472	281	22	19	15	34
MCODE	0.03	0.117	0.047	0.111	60	7	73	0	0	0
SETS	0.479	0.764	0.589	0.68	220	168	62	19	20	39
NewMips										
SPICi	0.479	0.618	0.54	0.352	131	81	20	16	17	33
ClusterONE	0.442	0.458	0.45	0.323	240	110	23	9	12	21
NCMine	0.479	0.427	0.452	0.362	578	247	25	7	13	20
PEWCC	0.534	0.476	0.503	0.368	708	337	31	10	22	32
IPCA	0.515	0.574	0.543	0.36	472	271	22	13	18	31
MCODE	0.021	0.1	0.035	0.038	60	6	73	0	0	0
SETS	0.485	0.732	0.583	0.391	220	161	62	14	21	35

Note. R: Recall, P: Precession, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

DIP

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
CYC2008										
SPICi	0.555	0.507	0.53	0.541	219	111	22	13	8	21
ClusterONE	0.436	0.336	0.38	0.466	342	115	23	7	7	14
NCMine	0.542	0.291	0.378	0.497	1074	312	28	8	10	18
PEWCC	0.678	0.317	0.432	0.582	1544	490	42	22	18	40
IPCA	0.589	0.318	0.413	0.516	826	263	32	17	10	27
MCODE	0.008	0.04	0.014	0.116	50	2	180	0	0	0
SETS	0.653	0.498	0.565	0.593	540	269	41	18	14	32
NewMips										
SPICi	0.573	0.479	0.522	0.334	219	105	22	11	8	19
ClusterONE	0.412	0.304	0.35	0.265	342	104	23	5	5	10
NCMine	0.546	0.287	0.376	0.32	1047	308	28	5	11	16
PEWCC	0.683	0.318	0.434	0.39	1544	491	42	16	17	33
IPCA	0.579	0.311	0.405	0.323	826	257	32	18	8	26
MCODE	0.006	0.04	0.011	0.046	50	2	180	0	0	0
SETS	0.64	0.496	0.559	0.39	540	268	41	16	17	33

Note. R: Recall, P: Precision, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

BioGRID

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
CYC2008										
SPICi	0.432	0.186	0.26	0.613	440	82	141	4	2	6
ClusterONE	0.487	0.265	0.343	0.697	476	126	83	1	7	8
NCMine	0.737	0.123	0.211	0.807	3671	451	95	4	10	14
PEWCC	0.873	0.196	0.32	0.872	4048	792	750	11	21	33
IPCA	0.576	0.14	0.226	0.758	2718	381	80	2	14	16
MCODE	0.008	0.036	0.014	0.073	56	2	192	0	0	0
SETS	0.644	0.379	0.477	0.816	633	240	93	6	23	29
NewMips										
SPICi	0.436	0.18	0.254	0.442	440	79	141	2	5	7
ClusterONE	0.488	0.25	0.331	0.496	476	119	83	1	6	7
NCMine	0.695	0.13	0.219	0.551	3671	478	95	5	8	13
PEWCC	0.826	0.21	0.335	0.627	4048	850	750	10	20	30
IPCA	0.591	0.138	0.223	0.538	2718	374	80	2	13	15
MCODE	0.006	0.036	0.01	0.029	56	2	192	0	0	0
SETS	0.622	0.382	0.474	0.561	633	242	93	6	21	27

Note. R: Recall, P: Precision, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

SETS Algorithm for Prediction Overlapping Protein Complexes

Human

	R	P	F	CR	# Complexes	# matched complexes	Max	Exact	Good	Total
SPICi	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
ClusterONE	0.223	0.235	0.229	0.33	1037	252	96	11	9	20
NCMine	0.552	0.221	0.315	0.459	7776	1716	111	7	12	19
PEWCC	0.68	0.276	0.393	0.559	9036	2495	394	21	34	55
IPCA	0.463	0.266	0.338	0.455	6533	1736	93	5	9	14
MCODE	0.001	0.014	0.002	0.04	74	1	377	0	0	0
SETS	0.498	0.405	0.447	0.484	2026	822	223	18	26	44

Note. R: Recall, P: Precision, F: F-measure, CR: Coverage Rate, Max: Maximum size of the complex.

Biological significance

	Collins		Gavin		Krogan		DIP		BioGRID		Human	
	Bp	CC	Bp	CC	Bp	CC	Bp	CC	Bp	CC	Bp	CC
SPICi	0.953	0.976	0.954	0.977	0.957	0.982	0.947	0.962	0.736	0.865	NA	NA
ClusterONE	0.902	0.958	0.782	0.893	0.849	0.91	0.809	0.876	0.799	0.878	0.879	0.941
NCMine	0.924	0.961	0.858	0.917	0.838	0.917	0.788	0.926	0.737	0.845	0.871	0.966
PEWCC	0.942	0.968	0.88	0.933	0.861	0.928	0.804	0.935	0.751	0.862	0.86	0.961
IPCA	0.967	0.981	0.882	0.937	0.883	0.938	0.802	0.946	0.782	0.868	0.862	0.96
SETS	0.962	0.976	0.897	0.944	0.897	0.947	0.833	0.951	0.82	0.901	0.881	0.968

ECC vs. CN in SETS

	CYC2008		NewMIPS		Human - CORUN	
	SETS-CN F-measure	SETS-ECC F-measure	SETS-CN F-measure	SETS-ECC F-measure	SETS-CN F-measure	SETS-ECC F-measure
Collins	0.606	0.615	0.642	0.65	0.447	0.333
Gavin	0.53	0.502	0.563	0.522		
Krogan	0.589	0.598	0.583	0.583		
DIP	0.565	0.481	0.559	0.507		
BioGRID	0.477	0.303	0.474	0.327		

APPENDIX 2

Analysis of Benchmark Complexes

The benchmark dataset is analysed using PPI networks. Tables contain the number of proteins in each PPI as well as the number of proteins that are in benchmark complexes but are not in PPIs. The number of complexes in the benchmark dataset is reported, the benchmark complexes from proteins that are not in PPI are filtered out and only the complexes that have a length of more than two proteins are retained. The benchmark complexes are filtered again and only those complexes that have all its proteins in PPI are retained. The CN is calculated between the proteins of the same complex for different thresholds. The number of complexes where at least two of its proteins are satisfied at the threshold is reported and according to the number of complexes that satisfied T_{CN} to the number of complexes from the second filter, almost 25% of complexes from the second filter, the threshold T_{CN} is set to each PPI. The F-measure of each dataset with a different threshold proved the accuracy of the selected threshold.

Collins

	# proteins In Collins	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CYC2008	1662	382	236	145	102
NewMIPS	1662	695	328	221	106

T_{CN}	CYC2008	NewMIPS
0.1	1	2
0.2	14	10
0.3	24	24
0.4	36	38
0.5	47	48

T_{CN}	F-measure (CYC2008)	F-measure (NewMIPS)
0.1	0.602	0.638
0.2	0.606	0.638
0.3	0.606	0.642
0.4	0.588	0.623
0.5	0.571	0.604

Gavin

	# proteins In Collins	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CYC2008	1855	439	236	143	86
NewMIPS	1855	724	328	218	90

T _{CN}	CYC2008	NewMIPS
0.1	5	6
0.2	17	17
0.3	25	29
0.4	33	36
0.5	47	51

TCN	F-measure (CYC2008)	F-measure (NewMIPS)
0.1	0.483	0.51
0.2	0.499	0.526
0.3	0.53	0.563
0.4	0.537	0.566
0.5	0.544	0.565

Krogan

	# proteins In Krogan	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CYC2008	2675	389	236	169	119
NewMIPS	2675	604	328	249	123

T _{CN}	CYC2008	NewMIPS
0.1	12	17
0.2	40	43
0.3	70	68
0.4	88	90
0.5	99	103

TCN	F-measure (CYC2008)	F-measure (NewMIPS)
0.1	0.6	0.575
0.2	0.589	0.583
0.3	0.531	0.548
0.4	0.468	0.494
0.5	0.421	0.442

DIP

	# proteins In DIP	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CYC2008	4930	138	236	226	191
NewMIPS	4930	194	328	313	231

T_{CN}	CYC2008	NewMIPS
0.1	61	100
0.2	133	173
0.3	158	201
0.4	168	215
0.5	170	218

TCN	F-measure (CYC2008)	F-measure (NewMIPS)
0.1	0.565	0.559
0.2	0.516	0.518
0.3	0.464	0.507
0.4	0.35	0.407
0.5	0.194	0.407

BioGRID

	# proteins In BioGRID	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CYC2008	5361	6	236	236	231
NewMIPS	5361	31	328	322	301

T_{CN}	CYC2008	NewMIPS
0.1	75	165
0.2	159	242
0.3	199	277
0.4	217	292
0.5	222	296

TCN	F-measure (CYC2008)	F-measure (NewMIPS)
0.1	0.406	0.416
0.2	0.477	0.474
0.3	0.45	0.485
0.4	0.358	0.408
0.5	0.247	0.297

Human

	# proteins In Human dataset	# proteins in benchmark but not in PPI	# benchmark complexes	First filter	Second filter
CORUM	15459	157	2351	2340	2196

T_{CN}	CORUM
0.1	1483
0.2	1948
0.3	2071
0.4	2123
0.5	2145

TCN	F-measure (CORUM)
0.1	0.447
0.2	0.417
0.3	0.325
0.4	0.242
0.5	0.137



REFEREES FOR THE PERTANIKA JOURNAL OF SCIENCE & TECHNOLOGY

VOL. 29 (2) APR. 2021

The Editorial Board of Pertanika Journal of Science & Technology wishes to thank the following:

Adeyi Abel Adekanmi
(ABUAD, Nigeria)

Agus Arsad
(UTM, Malaysia)

Ahmad Hilmy Abdul Hamid
(USM, Malaysia)

Ahmad Kamil Arshad
(UiTM, Malaysia)

Ahmad Shukri Yahaya
(USM, Malaysia)

Ahmad Taufek Abdul Rahman
(UiTM, Malaysia)

Ahmed Jalal Khan Chowdhury
(IIUM, Malaysia)

Ahmed Lateef Khalaf
(Al-Ma'moon University, Iraq)

Arien Heryansyah
(UTM, Malaysia)

Arifah Bahar
(UTM, Malaysia)

Aslina Baharum
(UMS, Malaysia)

Azizul Azhar Ramli
(UTHM, Malaysia)

Azni Zain Ahmed
(UiTM, Malaysia)

Azura Hamzah
(UTM, Malaysia)

Bimo Ario Tejo
(UPM, Malaysia)

Bing Zhou
(Khalifa University, UAE)

Chaw May May
(UM, Malaysia)

Che Mohd Imran Che Taib
(UMT, Malaysia)

Chuah Joon Huang
(UM, Malaysia)

Chuan Zun Liang
(UMP, Malaysia)

Darius El Pebrian
(UiTM, Malaysia)

Dat Tran Van
(IWEM, Vietnam)

Dian Darina Indah Daruis
(UPNM, Malaysia)

Effariza Hanafi
(UM, Malaysia)
Elia Godoong
(UMS, Malaysia)

Hafiz Rashidi Harun @ Ramli
(UPM, Malaysia)

Hedzlin Zainuddin
(UiTM, Malaysia)

Jalilah Abdul Jalil
(UniMAP, Malaysia)

Joel Than Chia Ming
(Swinburne University of Technology,
Malaysia)

Kek Sie Long
(UTHM, Malaysia)

Khairul Hamimah Abas
(UTM, Malaysia)

Leong Wah June
(UPM, Malaysia)

Lim Yoke Mui
(USM, Malaysia)

Maizatul Akmar Ismail
(UM, Malaysia)

Marzanah Jabar
(UPM, Malaysia)

Md. Noordin Abu Bakar
(USM, Malaysia)

Md. Suffian Idris
(UMT, Malaysia)

Meric Aziz Berge
(DEU, Turkey)

Mirta Widia
(UMP, Malaysia)

Mohamad Roslan Mohamad
Kasim
(UPM, Malaysia)

Mohamed Ansari Nainar
(UNITEN, Malaysia)

Mohammad Firuz Ramli
(UPM, Malaysia)

Mohan Reddy Moola
(Curtin University, Malaysia)

Mohd Amrallah Mustafa
(UPM, Malaysia)

Mohd Azri Mohd Izhar
(UTM, Malaysia)

Mohd Hafiz Mohd Zin
(USM, Malaysia)

Mohd Hanafi Idris
(UMT, Malaysia)

Mohd Hariri Arifin
(UKM, Malaysia)

Mohd Hudzari Haji Razali
(UiTM, Malaysia)

Mohd Shahrudin Abd Manan
(UPM, Malaysia)

Mohd Shahrul Nizam Mohd
Danuri
(KUIS, Malaysia)

Mohd Tahir Ismail
(USM, Malaysia)

Muhamad Azhar Ghazali
(USM, Malaysia)

Muhamad Fazly Abdul Patah <i>(UM, Malaysia)</i>	Norizan Mohamed <i>(UMT, Malaysia)</i>	Suzana Ab Hamid <i>(UPM, Malaysia)</i>
Muhammad Nur Salihin Yusoff <i>(USM, Malaysia)</i>	Nur Fadhilah Ibrahim <i>(UMT, Malaysia)</i>	Syazreen Niza Shair <i>(UiTM, Malaysia)</i>
Muhammed Nafis Osman <i>Zahid (UMP, Malaysia)</i>	Nur Tantiyani Ali Othman <i>(UKM, Malaysia)</i>	Teoh Sian Hoon <i>(UiTM, Malaysia)</i>
Muzalwana Abdul Talib <i>(UM, Malaysia)</i>	Phang Chang <i>(UTHM, Malaysia)</i>	Thaned Rojsiraphisal <i>(CMU, Thailand)</i>
Nadiatul Adilah Ahmad Abdul Ghani <i>(UMP, Malaysia)</i>	Rahida Wati Sharudin <i>(UiTM, Malaysia)</i>	Tye Ching Thian <i>(USM, Malaysia)</i>
Nasri Sulaiman <i>(UPM, Malaysia)</i>	Razali Ngah <i>(UTM, Malaysia)</i>	Wan Fatma Zuharah Wan Mustapha <i>(USM, Malaysia)</i>
Nazre Saleh <i>(UPM, Malaysia)</i>	Riza Wirawan <i>(ITB, Indonesia)</i>	Wan Mohd Nazmee Wan Zainon <i>(USM, Malaysia)</i>
Ng Kok Haur <i>(UM, Malaysia)</i>	Rossita Mohamad Yunus <i>(UM, Malaysia)</i>	Wan Zakiyatussariroh Wan Husin <i>(UiTM, Malaysia)</i>
Ng Seng Beng <i>(UPM, Malaysia)</i>	Ruhaizin Sulaiman <i>(UPM, Malaysia)</i>	Yu Ke Xin <i>(MSU, Malaysia)</i>
Nik Norsyahariati Nik Daud <i>(UPM, Malaysia)</i>	Siti Mahani Marjugi <i>(UPM, Malaysia)</i>	Zainuddin Md Yusoff <i>(UPM, Malaysia)</i>
Nooriati Taib <i>(USM, Malaysia)</i>	Siti Nurbaya Oslan <i>(UPM, Malaysia)</i>	Zeinab Abbas Jawad <i>(Curtin University, Malaysia)</i>
	Suraya Mohd Tahir <i>(UPM, Malaysia)</i>	

ABUAD – Afe Babalola University
 CMU – Chiangmai University
 DEU – Dokuz Eylül University
 IIUM – International Islamic University Malaysia
 ITB – Bandung Institute of Technology
 IWEM – Institute for Water Resources Economics and Management
 KUIS – Kolej Universiti Islam Antarabangsa Selangor
 MSU – Management & Science University
 UiTM – Universiti Teknologi MARA
 UKM – Universiti Kebangsaan Malaysia
 UM – Universiti Malaya

UMP – Universiti Malaysia Pahang
 UMS – Universiti Malaysia Sabah
 UMT – Universiti Malaysia Terengganu
 UniMAP – Universiti Malaysia Perlis
 UNITEN – Universiti Tenaga Nasional
 UPM – Universiti Putra Malaysia
 UPNM – Universiti Pertahanan Nasional Malaysia
 USM – Universiti Sains Malaysia
 UTHM – Universiti Tun Hussein Onn Malaysia
 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

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.

Pertanika Journal of Science & Technology

Our goal is to bring high-quality research to the widest possible audience

INSTRUCTIONS TO AUTHORS

(REGULAR ISSUE)

(Manuscript Preparation & Submission Guide)

Revised: November 2020

Please read the *Pertanika* guidelines and follow these instructions carefully. The Chief Executive Editor reserves the right to return manuscripts that are not prepared in accordance with these guidelines.

MANUSCRIPT PREPARATION Manuscript Types

Pertanika accepts submission of mainly 4 types of manuscripts

- that have not been published elsewhere (including proceedings)
- that are not currently being submitted to other journals

1. Regular article

Regular article is a full-length original empirical investigation, consisting of introduction, methods, results, and discussion. Original research work should present new and significant findings that contribute to the advancement of the research area. *Analysis and Discussion* must be supported with relevant references.

Size: Generally, each manuscript is **not to exceed 6000 words** (excluding the abstract, references, tables, and/or figures), a maximum of **80 references**, and **an abstract of less than 250 words**.

2. Review article

A review article reports a critical evaluation of materials about current research that has already been published by organising, integrating, and evaluating previously published materials. It summarises the status of knowledge and outlines future directions of research within the journal scope. A review article should aim to provide systemic overviews, evaluations, and interpretations of research in a given field. Re-analyses as meta-analysis and systemic reviews are encouraged.

Size: Generally, it is expected **not to exceed 6000 words** (excluding the abstract, references, tables, and/or figures), a maximum of **80 references**, and **an abstract of less than 250 words**.

3. Short communications

Each article should be timely and brief. It is suitable for the publication of significant technical advances and maybe used to:

- (a) reports new developments, significant advances and novel aspects of experimental and theoretical methods and techniques which are relevant for scientific investigations within the journal scope;
- (b) reports/discuss on significant matters of policy and perspective related to the science of the journal, including 'personal' commentary;
- (c) disseminates information and data on topical events of significant scientific and/or social interest within the scope of the journal.

Size: It is limited to **3000 words** and have a maximum of **3 figures and/or tables, from 8 to 20 references, and an abstract length not exceeding 100 words**. The information must be in short but complete form and it is not intended to publish preliminary results or to be a reduced version of a regular paper.

4. Others

Brief reports, case studies, comments, concept papers, letters to the editor, and replies on previously published articles may be considered.

Language Accuracy

Pertanika emphasises on the linguistic accuracy of every manuscript published. Articles can be written in **English** or **Bahasa Malaysia** and they must be competently written and presented in clear and concise grammatical English/Bahasa Malaysia. Contributors are strongly advised to have the manuscript checked by a colleague with ample experience in writing English manuscripts or a competent English language editor. For articles in Bahasa Malaysia, the title, abstract and keywords should be written in both English and Bahasa Malaysia.

Author(s) **may be required to provide a certificate** confirming that their manuscripts have been adequately edited. **All editing costs must be borne by the authors.**

Linguistically hopeless manuscripts will be rejected straightaway (e.g., when the language is so poor that one cannot be sure of what the authors are really trying to say). This process, taken by authors before submission, will greatly facilitate reviewing, and thus, publication.

MANUSCRIPT FORMAT

The paper should be submitted in **one-column format** with 1.5 line spacing throughout. Authors are advised to use Times New Roman 12-point font and *MS Word* format.

1. Manuscript Structure

The manuscripts, in general, should be organised in the following order:

Page 1: Running title

This page should **only** contain the running title of your paper. The running title is an abbreviated title used as the running head on every page of the manuscript. The running title **should not exceed 60 characters, counting letters and spaces.**

Page 2: Author(s) and Corresponding author's information

General information: This page should contain the **full title** of your paper **not exceeding 25 words**, with the name of all the authors, institutions and corresponding author's name, institution and full address (Street address, telephone number (including extension), handphone number, and e-mail address) for editorial correspondence. **The corresponding author must be clearly indicated with a superscripted asterisk symbol (*).**

Authors' name: The names of the authors should be named **in full without academic titles**. For Asian (Chinese, Korean, Japanese, Vietnamese), please write first name and middle name before surname (family name). The last name in the sequence is considered the surname.

Authors' addresses: Multiple authors with different addresses must indicate their respective addresses separately by superscript numbers.

Tables/figures list: A list of the number of **black and white/colour figures and tables** should also be indicated on this page. See **"5. Figures & Photographs"** for details.

Example (page 2):

Fast and Robust Diagnostic Technique for the Detection of High Leverage Points

Habshah Midi^{1,2*}, Hasan Talib Hendi¹, Jayanthi Arasan² and Hassan Uraibi³

¹*Institute for Mathematical Research, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

²*Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, 43400 UPM, Serdang, Selangor, Malaysia*

³*Department of Statistics, University of Al-Qadisiyah, 88 -Al-Qadisiyah -Al-Diwaniyah, Iraq*

E-mail addresses

habshah@upm.edu.my (Habshah Midi)

h.applied.t88@gmail.com (Hasan Talib Hendi)

jayanthi@upm.edu.my (Jayanthi Arasan)

hssn.sami1@gmail.com (Hassan Uraibi)

*Corresponding author

List of Table/Figure: Table 1.

Figure 1.

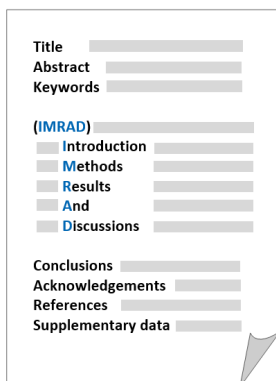
Page 3: Abstract

This page should **repeat** the **full title** of your paper with only the **Abstract**, usually in one paragraph and **Keywords**.

Keywords: *Not more than 8 keywords in alphabetical order must be provided to describe the content of the manuscript.*

Page 4: Text

A regular paper should be prepared with the headings *Introduction, Materials and Methods, Results and Discussions, Conclusions, Acknowledgements, References, and Supplementary data* (if any) in this order. The literature review may be part of or separated from the *Introduction*.



MAKE YOUR ARTICLES AS CONCISE AS POSSIBLE

Most scientific papers are prepared according to a format called IMRAD. The term represents the first letters of the words Introduction, Materials and Methods, Results, And, Discussion. It indicates a pattern or format rather than a complete list of headings or components of research papers; the missing parts of a paper are: Title, Authors, Keywords, Abstract, Conclusions, and References. Additionally, some papers include Acknowledgments and Appendices.

The Introduction explains the scope and objective of the study in the light of current knowledge on the subject; the Materials and Methods describes how the study was conducted; the Results section reports what was found in the study; and the Discussion section explains meaning and significance of the results and provides suggestions for future directions of research. The manuscript must be prepared according to the Journal's instructions to authors.

2. Levels of Heading

Level of heading	Format
1 st	LEFT, BOLD, UPPERCASE
2 nd	Flush left, Bold, Capitalise each word
3 rd	Bold, Capitalise each word, ending with .
4 th	Bold italic, Capitalise each word, ending with .

3. Equations and Formulae

These must be set up clearly and should be typed double-spaced. Numbers identifying equations should be in square brackets and placed on the right margin of the text.

4. Tables

- All tables should be prepared in a form consistent with recent issues of *Pertanika* and should be numbered consecutively with Roman numerals (Table 1, Table 2).
- A brief title should be provided, which should be shown at the top of each table (APA format):

Example:

Table 1

PVY infected Nicotiana tabacum plants optical density in ELISA

- Explanatory material should be given in the table legends and footnotes.
- Each table should be prepared on a new page, embedded in the manuscript.
- Authors are advised to keep backup files of all tables.

**** Please submit all tables in Microsoft word format only, because tables submitted as image data cannot be edited for publication and are usually in low-resolution.**

5. Figures & Photographs

- Submit an original figure or photograph.
- Line drawings must be clear, with a high black and white contrast.
- Each figure or photograph should be prepared on a new page, embedded in the manuscript for reviewing to keep the file of the manuscript under 5 MB.
- These should be numbered consecutively with Roman numerals (Figure 1, Figure 2).
- Provide a brief title, which should be shown at the bottom of each table (**APA format**):

Example: *Figure 1. PVY-infected in vitro callus of Nicotiana tabacum*

- If a figure has been previously published, acknowledge the original source, and submit written permission from the copyright holder to reproduce the material.
- Authors are advised to keep backup files of all figures.

**** Figures or photographs must also be submitted separately as TIFF or JPEG, because figures or photographs submitted in low-resolution embedded in the manuscript cannot be accepted for publication. For electronic figures, create your figures using applications that are capable of preparing high-resolution TIFF files.**

6. Acknowledgement

Any individuals and entities who have contributed to the research should be acknowledged appropriately.

7. References

References begin on their own page and are listed in alphabetical order by the first author's last name. Only references cited within the text should be included. All references should be in 12-point font and double-spaced. If a Digital Object Identifier (DOI) is listed on a print or electronic source, it is required to include the DOI in the reference list. Use Crossref to find a DOI using author and title information.

NOTE: When formatting your references, please follow the **APA-reference style** (7th edition) (refer to the examples). Ensure that the references are strictly in the journal's prescribed style, failing which your article will **not be accepted for peer-review**. You may refer to the *Publication Manual of the American Psychological Association* (<https://apastyle.apa.org/>) for further details.

Examples of reference style are given below:

Books		
	Insertion in text	In reference list
Book/E-Book with 1-2 authors	<p>Information prominent' (the author's name is within parentheses):</p> <p>... (Staron, 2020)</p> <p>... (Darus & Rasdi, 2019)</p> <p>... Or</p> <p>'Author prominent' (the author's name is outside the parentheses):</p> <p>(Starron, 2020)...</p> <p>Darus and Rasdi (2019) ...</p>	<p>Staron, M. (2020). <i>Action research in software engineering</i>. Springer International Publishing. https://doi.org/10.1007/978-3-030-32610-4</p> <p>Darus, A., & Rasdi, I. (2019). <i>Introduction to occupational health a workbook</i>. UPM Press.</p>
Book/E-Book with 3 or more authors	<p><i>For all in-text references, list only the first author's family name and followed by 'et al.'</i></p> <p>Information prominent' (the author's name is within parentheses):</p> <p>... (Yusof et al., 2020)</p> <p>... Or</p> <p>'Author prominent' (the author's name is outside the parentheses):</p> <p>Yusof et al. (2020) ...</p>	<p>Yusof, N. A., Azmi, U. Z. M., Ariffin, N., & Rahman, S. F. A. (2020). <i>Biosensors and chemical sensors: A practical approach</i>. UPM Press.</p>
Book/E-Book with more than 20 authors		<p>For books with more than 20 authors, please follow the guidelines for journal articles with more than 20 authors.</p>
Chapter in an edited Book/E-Book	<p>Inform ation pr ominent' (the author 's name is within parentheses):</p> <p>... (Mainzer, 2020) ...</p> <p>... (Tang et al., 2020) ...</p> <p>Or</p> <p>'Author prominent' (the author's name is outside the parentheses):</p> <p>Mainzer (2020) ...</p> <p>Tang et al. (2020) ...</p>	<p>Mainzer, K. (2020). Logical thinking becomes automatic. In K. Mainzer (Ed.), <i>Artificial intelligence-When do machines take over?</i> (pp. 15-45). Springer. https://doi.org/10.1007/978-3-662-59717-0_3</p> <p>Tang, W., Khavarian, M., Yousefi, A., & Cui, H. (2020). Properties of self-compacting concrete with recycled concrete aggregates. In R. Siddique (Ed.), <i>Self-Compacting Concrete: Materials, Properties, and Applications</i> (pp. 219-248). Woodhead Publishing. https://doi.org/10.1016/B978-0-12-817369-5.00009-X</p>

	Insertion in text	In reference list
Editor	<p>Information prominent' (the author's name is within parentheses): ... (Kesharwani, 2020) (Lanza et al., 2020) ... Or 'Author prominent' (the author's name is outside the parentheses): Kesharwani (2020) ... Lanza et al. (2020) ...</p>	<p>Kesharwani, P. (Ed.). (2020). <i>Nanotechnology based approaches for tuberculosis treatment</i>. Academic Press.</p> <p>Lanza, R., Langer, R., Vacanti, J. P., & Atala, A. (Eds.). (2020). <i>Principles of tissue engineering</i>. Academic press. https://doi.org/10.1016/C2018-0-03818-9</p>
Several works by the same author in the same year	<p>Information prominent' (the author's name is within parentheses): ... (Aggarwal & Aggarwal, 2020a, 2020b) ... Or 'Author prominent' (the author's name is outside the parentheses): Aggarwal & Aggarwal (2020a, 2020b) ...</p>	<p>Aggarwal, P., & Aggarwal, Y. (2020a). Strength properties of SCC. In R. Siddique (Ed.), <i>Self-Compacting Concrete: Materials, Properties, and Applications</i> (p. 83-115). Woodhead Publishing. doi: https://doi.org/10.1016/B978-0-12-817369-5.00004-0</p> <p>Aggarwal, P., & Aggarwal, Y. (2020b). Carbonation and corrosion of SCC. In R. Siddique (Ed.), <i>Self-Compacting Concrete: Materials, Properties, and Applications</i> (p. 147-193). Woodhead Publishing. doi: https://doi.org/10.1016/B978-0-12-817369-5.00007-6</p>
Journals		
Journal article with 1-2 authors	<p>Information prominent' (the author's name is within parentheses): ... (Laan & Fox, 2019) ... Or 'Author prominent' (the author's name is outside the parentheses): Laan and Fox (2019) ...</p>	<p>Laan, E., & Fox, J. W. (2019). An experimental test of the effects of dispersal and the paradox of enrichment on metapopulation persistence. <i>Oikos</i>, 129(1), 49-58. https://doi.org/10.1111/oik.06552</p>
Journal article with 3 or more authors	<p><i>For all in-text references, list only the first author's family name and followed by 'et al.'</i> Information prominent' (the author's name is within parentheses): ... (Midi et al., 2020) (Shagufta et al., 2017) ... Or 'Author prominent' (the author's name is outside the parentheses): Midi et al. (2020) ... Shagufta et al. (2017) ...</p>	<p>Midi, H., Hendi, H. T., Arasan, J., & Uraibi, H. (2020). Fast and Robust Diagnostic Technique for the Detection of High Leverage Points. <i>Pertanika Journal of Science & Technology</i>, 28(4), 1203-1220.</p> <p>Shagufta, B., Sivakumar, M., Kumar, S., Agarwal, R. K., Bhilegaonkar, K. N., Kumar, A., & Dubal, Z. B. (2017). Antimicrobial resistance and typing of Salmonella isolated from street vended foods and associated environment. <i>Journal of Food Science and Technology</i>, 54(8), 2532-2539. doi: https://doi.org/10.1007/s13197-017-2698-1</p>
Journal article with more than 20	<p>Information prominent' (the author's name is within parentheses): ... (Wiskunde et al., 2019) ... Or 'Author prominent' (the author's name is outside the parentheses): Wiskunde et al. (2019) ...</p>	<p>Wiskunde, B., Arslan, M., Fischer, P., Nowak, L., Van den Berg, O., Coetzee, L., Juárez, U., Riyaziyyat, E., Wang, C., Zhang, I., Li, P., Yang, R., Kumar, B., Xu, A., Martinez, R., McIntosh, V., Ibáñez, L. M., Mäkinen, G., Virtanen, E., ... Kovács, A. (2019). Indie pop rocks mathematics: Twenty One Pilots, Nicolas Bourbaki, and the empty set. <i>Journal of Improbable Mathematics</i>, 27(1), 1935–1968. https://doi.org/10.0000/3mp7y-537</p>
Journal article with an article number	<p>Information prominent' (the author's name is within parentheses): ... (Roe et al., 2020) ... Or 'Author prominent' (the author's name is outside the parentheses): Roe et al. (2020) ...</p>	<p>Roe, E. T., Bies, A. J., Montgomery, R. D., Watterson, W. J., Parris, B., Boydston, C. R., Sereno, M. E., & Taylor, R. P. (2020). Fractal solar panels: Optimizing aesthetic and electrical performances. <i>Plos One</i>, 15(3), Article e0229945. https://doi.org/10.1371/journal.pone.0229945</p>
Journal article with missing information	<p>Information prominent' (the author's name is within parentheses): ... (Alfirevic et al., 2017) (Hayat et al., 2020) (Fan et al., 2020) ...</p>	<p>Missing volume number Alfirevic, Z., Stampalija, T., & Dowswell, T. (2017). Fetal and umbilical Doppler ultrasound in high-risk pregnancies (review). <i>Cochrane Database of Systematic Reviews</i>, (6), 1-163. https://doi.org/10.1002/14651858.CD007529.pub4. Copyright</p>

	Insertion in text	In reference list
Journal article with missing information	Or 'Author prominent' (the author's name is outside the parentheses): Alfirevic et al. (2017) ... Hayat et al. (2020) ... Fan et al. (2020) ...	Missing issue number Hayat, A., Shaishta, N., Mane, S. K. B., Hayat, A., Khan, J., Rehman, A. U., & Li, T. (2020). Molecular engineering of polymeric carbon nitride based Donor-Acceptor conjugated copolymers for enhanced photocatalytic full water splitting. <i>Journal of colloid and interface science</i> , 560, 743-754. https://doi.org/10.1016/j.jcis.2019.10.088 Missing page or article number Fan, R. G., Wang, Y. B., Luo, M., Zhang, Y. Q., & Zhu, C. P. (2020). SEIR-Based COVID-19 Transmission Model and Inflection Point Prediction Analysis. <i>Dianzi Keji Daxue Xuebao/Journal of the University of Electronic Science and Technology of China</i> , 49(3). https://doi.org/10.12178/1001-0548.9_2020029
Several works by the same author in the same year	Information prominent' (the author's name is within parentheses): ... (Chee et al., 2019a, 2019b) ... Or 'Author prominent' (the author's name is outside the parentheses): Chee et al. (2019a, 2019b) ...	Chee, S. S., Jawaid, M., Sultan, M. T. H., Alothman, O. Y., & Abdullah, L. C. (2019a). Accelerated weathering and soil burial effects on colour, biodegradability and thermal properties of bamboo/kenaf/epoxy hybrid composites. <i>Polymer Testing</i> , 79, Article 106054. https://doi.org/10.1016/j.polymeresting.2019.106054 Chee, S. S., Jawaid, M., Sultan, M. T. H., Alothman, O. Y., & Abdullah, L. C. (2019b). Evaluation of the hybridization effect on the thermal and thermo-oxidative stability of bamboo/kenaf/epoxy hybrid composites. <i>Journal of Thermal Analysis and Calorimetry</i> , 137(1), 55-63. https://doi.org/10.1007/s10973-018-7918-z
Newspaper		
Newspaper article – with an author	... (Shamshuddin, 2019) ... Or ... Shamshuddin (2019) ...	Shamshuddin, J. (2019, September 23). Lynas plant waste residue can be used to boost oil palm growth? <i>New Straits Times</i> . https://www.nst.com.my/opinion/letters/2019/09/523930/lynas-plant-waste-residue-can-be-used-boost-oil-palm-growth
Newspaper article – without an author	("Zoonotic viruses," 2017). OR "Zoonotic viruses" (2017) ... Use a shortened title (or full title if it is short) in Headline Case enclosed in double quotation marks.	Zoonotic viruses like swine flu are ticking time bombs, say experts. (2020, July 4). <i>New Straits Times</i> , 3.
Dissertation/Thesis		
Published Dissertation or Thesis References	... (Rivera, 2016) ... Or ... Rivera (2016) ...	Rivera, C. (2016). <i>Disaster risk management and climate change adaptation in urban contexts: Integration and challenges</i> [Doctoral dissertation, Lund University]. Lund University Publications. https://lup.lub.lu.se/search/ws/files/5471705/8570923.pdf
Unpublished Dissertation or Thesis References	... (Brooks, 2014) ... Or ... Brooks (2014) ...	Brooks, J. D. (2015). <i>Bamboo as a strengthening agent in concrete beams for medium height structures</i> [Unpublished Doctoral dissertation]. The University of Washington.
Conference/Seminar Papers		
Conference proceedings published in a journal	... (Duckworth et al., 2019) ... Or Duckworth et al. (2019) ...	Duckworth, A. L., Quirk, A., Gallop, R., Hoyle, R. H., Kelly, D. R., & Matthews, M. D. (2019). Cognitive and noncognitive predictors of success. <i>Proceedings of the National Academy of Sciences, USA</i> , 116(47), 23499-23504. https://doi.org/10.1073/pnas.1910510116
Conference proceedings published as a book chapter	... (Bedenel et al., 2019) ... Or Bedenel et al. (2019) ...	Bedenel, A. L., Jourdan, L., & Biernacki, C. (2019). Probability estimation by an adapted genetic algorithm in web insurance. In R. Battiti, M. Brunato, I. Kotsireas, & P. Pardalos (Eds.), <i>Lecture notes in computer science: Vol. 11353. Learning and intelligent optimization</i> (pp. 225-240). Springer. https://doi.org/10.1007/978-3-030-05348-2_21

	Insertion in text	In reference list
Online	... (Gu et al., 2018) ... Or Gu et al. (2018) ...	Gu, X., Yu, J., Han, Y., Han, M., & Wei, L. (2019, July 12-14). <i>Vehicle lane change decision model based on random forest</i> . [Paper presentation]. 2019 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS), Shenyang, China. https://doi.org/10.1109/ICPICS47731.2019.8942520
Government Publications		
Government as author	First in-text reference: Spell out the full name with the abbreviation of the body. ... National Cancer Institute (2019) ... Or ... (National Cancer Institute, 2019) ... Subsequent in-text reference: ... NCI (2019) ... Or ... (NCI, 2019) ...	National Cancer Institute. (2019). <i>Taking time: Support for people with cancer</i> (NIH Publication No. 18-2059). U.S. Department of Health and Human Services, National Institutes of Health. https://www.cancer.gov/publications/patient-education/takingtime.pdf

8. General Guidelines

Abbreviations: Define alphabetically, other than abbreviations that can be used without definition. Words or phrases that are abbreviated in the *Introduction* and following text should be written out in full the first time that they appear in the text, with each abbreviated form in parenthesis. Include the common name or scientific name, or both, of animal and plant materials.

Authors' Affiliation: The primary affiliation for each author should be the institution where the majority of their work was done. If an author has subsequently moved to another institution, the current address may also be stated in the footer.

Co-Authors: The commonly accepted guideline for authorship is that one must have substantially contributed to the development of the paper and share accountability for the results. Researchers should decide who will be an author and what order they will be listed depending upon their order of importance to the study. Other contributions should be cited in the manuscript's *Acknowledgements*.

Similarity Index: All articles received must undergo the initial screening for originality before being sent for peer review. *Pertanika* does not accept any article with a similarity index exceeding **20%**.

Copyright Permissions: Authors should seek necessary permissions for quotations, artwork, boxes or tables taken from other publications or other freely available sources on the Internet before submission to *Pertanika*. The *Acknowledgement* must be given to the original source in the illustration legend, in a table footnote, or at the end of the quotation.

Footnotes: Current addresses of authors if different from heading may be inserted here.

Page Numbering: Every page of the manuscript, including the title page, references, and tables should be numbered.

Spelling: The journal uses American or British spelling and authors may follow the latest edition of the Oxford Advanced Learner's Dictionary for British spellings. Each manuscript should follow one type of spelling only.

SUBMISSION OF MANUSCRIPTS

All submissions must be made electronically using the **ScholarOne™ online submission system**, a web-based portal by Clarivate Analytics. For more information, go to our web page and click "**Online Submission (ScholarOne™)**".

Submission Checklist

1. MANUSCRIPT:

Ensure your manuscript has followed the *Pertanika* style particularly the first-4-pages as explained earlier. The article should be written in a good academic style and provide an accurate and succinct description of the contents ensuring that grammar and spelling errors have been corrected before submission. It should also not exceed the suggested length.

2. DECLARATION FORM:

Author has to sign a declaration form. In signing the form, authors declare that the work submitted for publication is original, previously unpublished, and not under consideration for any publication elsewhere.

Author has to agree to pay the publishing fee once the paper is accepted for publication in Pertanika.

3. COVER LETTER:

In Step 6 of the ScholarOne system, author is asked to upload a cover letter in *Pertanika* format. Please ignore this instruction and replace the cover letter with the **Declaration Form**.

Note:

COPYRIGHT FORM: Author will be asked to sign a copyright form when the paper is accepted. In signing the form, it is assumed that authors have obtained permission to use any copyrighted or previously published material. All authors must read and agree to the conditions outlined in the form and must sign the form or agree that the corresponding author can sign on their behalf. Articles cannot be published until a signed form (original pen-to-paper signature) has been received.

Visit our Journal's website for more details at <http://www.pertanika.upm.edu.my/>.

ACCESS TO PUBLISHED MATERIALS

Under the journal's open access initiative, authors can choose to download free material (via PDF link) from any of the journal issues from *Pertanika*'s website. Under "**Browse Journals**" you will see a link, "*Regular Issue*", "*Special Issue*" or "*Archives*". Here you will get access to all current and back-issues from 1978 onwards. No hard copy of journals or offprints are printed.

Visit our Journal's website at:

http://www.pertanika.upm.edu.my/regular_issues.php for "Regular Issue"

http://www.pertanika.upm.edu.my/cspecial_issues.php for "Special Issue"

http://www.pertanika.upm.edu.my/journal_archives.php for "Archives"

PUBLICATION CHARGE

Upon acceptance of a manuscript, a processing fee of RM 750 / USD 250 will be imposed on authors; RM 750 for any corresponding author affiliated to an institution in Malaysia; USD 250 for any corresponding author affiliated to an institution outside Malaysia. Payment must be made online at <https://paygate.upm.edu.my/action.do?do=>

Any queries may be directed to the **Chief Executive Editor's** office via email to executive_editor.pertanika@upm.edu.my

Solution of Collinearity Problem in Two-Dimensional Electrical Resistivity Tomography using Wenner Array <i>Mustapha Adejo Mohammed Nordiana Mohd Muztaza and Rosli Saad</i>	1199
Potential Groundwater Exploration in Use of 2-D Electrical Resistivity Tomography (ERT) Techniques at the Department of Agriculture Kelantan Research and Developmental Platform Padang Raja Kelantan <i>Wan Fazilah Fazlil Ilahi, Nur Hidayu Abu Hassan, Mohd Razi Ismail, Nik Norasma Che'Ya, Zulkarami Berahim, Mohamad Husni Omar, Nurul Idayu Zakaria and Mohamed Azwan Mohamed Zawawi</i>	1219
Environmental Sciences	
Efficacy of Intelligent Mosquito System (I.M.O.S) with Xmos Mini Aerosol Against <i>Aedes</i> in 17 th College, Universiti Putra Malaysia <i>Latifah Saiful Yazan, Banulata Gopalsamy, Siti Najiha Abu Bakar, Khairul Aiman Manan, Noranis Shahida Shahidan and Lee Yean Wang</i>	1229
Modelling <i>In-situ</i> Factors Affecting Bud's Growth of <i>Rafflesia kerrii</i> Meijer in Lojing Highlands, Kelantan, Peninsular Malaysia <i>Abdul Hamid Mar Iman, Nor Hizami Hassin, Muhamad Azahar Abas and Zulhazman Hamzah</i>	1243
Ecological Roles of Benthic Community in Seafloor Northwards of Pulau Indah, Klang, Malaysia <i>Mohd Sophian Mohd Kasihmuddin and Zaidi Che Cob</i>	1267
Medical and Health Sciences	
Radiographic Measurement of Cochlear in Sudanese Population using High Resolution Computed Tomography (HRCT) <i>Lubna Bushara, Mohamed Yousef, Ikhlas Abdelaziz, Mogahid Zidan, Dalia Bilal and Mohammed El Wathig</i>	1289
Prevalence of Vitreous & Retinal Disorders among Sudanese Diabetic Patients: A B-Scan Ultrasonography Study <i>Mohamed Yousef, Safaa Bashir, Awadalla Wagealla, Mogahid Zidan, Mahmoud Salih Babiker and Mona Mohamed</i>	1299
Material Sciences	
Poly lactide and its Composites on Various Scales of Hardness <i>Abraham Kehinde Aworinde, Eyere Emagbetere, Samson Oluropo Adeosun and Esther Titilayo Akinlabi</i>	1313
Chemical Sciences	
SETS: A Seed-Dense-Expanding Model-Based Topological Structure for the Prediction of Overlapping Protein Complexes <i>Soheir Noori, Nabeel Al-A'araji and Eman Al-Shamery</i>	1323

Mathematical Modeling and Availability Analysis of Leaf Spring Manufacturing Plant <i>Sohan Lal Tyagi, Shikha Bansal, Priyanka Agarwal and Ajay Singh Yadav</i>	1041
Penalized LAD-SCAD Estimator Based on Robust Wrapped Correlation Screening Method for High Dimensional Models <i>Ishaq Abdullahi Baba, Habshah Midi, Leong Wah June and Gafurjan Ibragimov</i>	1053
Relative Risk Estimation for Human Leptospirosis Disease in Malaysia Based on Existing Models and Discrete Space-Time Stochastic Sir Model <i>Sufi Hafawati Ideris, Muhammad Rozi Malim and Norshahida Shaadan</i>	1071
Applied Sciences and Technologies	
Factors in Ergonomic Design of 6-to-18-month Baby Carriers for Elderly People <i>Ariya Atthawuttikul and Sorat Khongkharat</i>	1097
Numerical Analysis of the Crack Inspections Using Hybrid Approach for the Application the Circular Cantilever Rods <i>Saddam Hussein Raheemah, Kareem Idan Fadheel, Qais Hussein Hassan, Ashham Mohammed Aneel, Alaa Abdulazeez Turki Al-Taie and Hussein Kadhim Sharaf</i>	1109
Radiation Modified CS/PVA Film with PVP Coating as Cu Adsorbent Material <i>Norhashidah Talip, Amalina Muhammad Afifi, Mohd Yusof Hamzah, Maznah Mahmud and Sarada Idris</i>	1119
An Investigation on Indoor Temperature of Modern Double Storey House with Adapted Common Passive Design Strategies of Malay Traditional House <i>Maryam Qays Oleiwi and Mohd Farid Mohamed</i>	1135
Assessment by Simulation of Different Topological Integration of Solar Photovoltaic Plant in Medium Voltage Distribution Networks <i>Md. Milon Uddin, Mushfiqur Rahman, Md. Tanzid Ridwan Hossain and Md. Habibur Rahman</i>	1159
Earth Sciences	
Volume Transport Variability in the Western Equatorial Pacific and its Relations to Halmahera Throughflow <i>Marlin Chrisye Wattimena, Agus Saleh Atmadipoera, Mulia Purba, I Wayan Nurjaya and Fadli Syamsudin</i>	1171

Information, Computer and Communication Technologies

Review

- A Review Article on Software Effort Estimation in Agile Methodology 837
Pantjawati Sudarmaningtyas and Rozlina Mohamed

- Reducing Cognitive Impairment Among Dementia Users Through Mobile Application 863
Nur Atheera Mohd Hassan, Aslina Baharum, Zaidatul Haslinda Abdullah Sani, Kent Chau and Noorsidi Aizuddin Mat Noor

- Photonics Rib Waveguide Dimension Dependent Charge Distribution and Loss Characterization 885
Angie Teo Chen Chen, Mohammad Rakib Uddin and Foo Kui Law

- Optimizing Placement of Field Experience Program: An Integration of MOORA and Rule-Based Decision Making 895
Okfalisa Okfalisa, Rizka Hafsari, Gusman Nawanir, Saktioto Toto and Novi Yanti

- Assessing Malaysian University English Test (MUET) Essay on Language and Semantic Features Using Intelligent Essay Grader (IEG) 919
Wee Sian Wong and Chih How Bong

- Artificial Neural Network Intelligent System on the Early Warning System of Landslide 943
Aghus Sofwan, Sumardi, Najib and Indrah Wendah Atma Bhirawa

- Performance Evaluation of Different Membership Function in Fuzzy Logic Based Short-Term Load Forecasting 959
Oladimeji Ibrahim, Waheed Olaide Owonikoko, Abubakar Abdulkarim, Abdulrahman Okino Otuoze, Mubarak Akorede Afolayan, Ibrahim Sani Madugu, Mutiu Shola Bakare and Kayode Elijah Adedayo

Mathematical Sciences

- Comparison of Single and MICE Imputation Methods for Missing Values: A Simulation Study 979
Nurul Azifah Mohd Pauzi, Yap Bee Wah, Sayang Mohd Deni, Siti Khatijah Nor Abdul Rahim and Suhartono

- Robust Multivariate Correlation Techniques: A Confirmation Analysis using Covid-19 Data Set 999
Friday Zinzendoff Okwonu, Nor Aishah Ahad, Joshua Sarduana Apanapudor and Festus Irismisose Arunaye

- Generalized Fibonacci Search Method in One-Dimensional Unconstrained Non-Linear Optimization 1017
Chin Yoon Chong, Soo Kar Leow and Hong Seng Sim

Content

Foreword	i
<i>Mohammad Jawaid</i>	
Engineering Sciences	
High Level Synthesis Optimizations of Road Lane Detection Development on Zynq-7000	707
<i>Panadda Solod, Nattha Jindapetch, Kiattisak Sengchuai, Apidet Booranawong, Pakpoom Hoyingcharoen, Surachate Chumpol and Masami Ikura</i>	
Farmers' Participation in Irrigation Management in the Punggur Utara Irrigation Area, Indonesia	725
<i>Dyah Indriana Kusumastuti, Vera Chania Putri, Dwi Jokowinarno and Endro Prasetyo Wahono</i>	
Pseudo-colour with K-means Clustering Algorithm for Acute Ischemic Stroke Lesion Segmentation in Brain MRI	743
<i>Abang Mohd Arif Anaqi Abang Isa, Kuryati Kipli, Ahmad Tirmizi Jobli, Muhammad Hamdi Mahmood, Siti Kudnie Sahari, Aditya Tri Hernowo and Sinin Hamdan</i>	
Improved High Dynamic Range for 3D Shape Measurement based on Saturation of the Coloured Fringe	759
<i>Shanyu Chua, Chee Chin Lim, Swee Kheng Eng, Yen Fook Chong and Chiun Tai Loh</i>	
The Built Environment and its Impact on Transit based Transportation Users Walking Activity in Semarang, Indonesia	771
<i>Anita Ratnasari Rakhmatulloh, Diah Intan Kusumo Dewi and Dinar Mutiara Kusumo Nugraheni</i>	
Ethylene Yield from a Large Scale Naphtha Pyrolysis Cracking Utilizing Response Surface Methodology	791
<i>Mohamad Hafizi Zakria, Mohd Ghazali Mohd Nawawi and Mohd Rizal Abdul Rahman</i>	
Smartphone Application Development for Rice Field Management Through Aerial Imagery and Normalised Difference Vegetation Index (NDVI) Analysis	809
<i>Nor Athirah Roslin, Nik Norasma Che'Ya, Rhushalshafira Rosle and Mohd Razi Ismail</i>	



Pertanika Editorial Office, Journal Division
Putra Science Park
1st Floor, IDEA Tower II
UPM-MTDC Technology Centre
Universiti Putra Malaysia
43400 UPM Serdang
Selangor Darul Ehsan
Malaysia

<http://www.pertanika.upm.edu.my/>
E-mail: executive_editor.pertanika@upm.edu.my
Tel: +603 9769 1622

PENERBIT
UPM
UNIVERSITI PUTRA MALAYSIA
PRESS

<http://penerbit.upm.edu.my>
E-mail: penerbit@upm.edu.my
Tel: +603 9769 8851

