# IIUM ENGINEERING JOURNAL

Volume 20 Number 2

December 2019

IIUM Press

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

ISSN: 1511-788X E-ISSN: 2289-7860

http://journals.iium.edu.my/ejournal

# IIUM ENGINEERING JOURNAL

**CHIEF EDITOR**
Ahmad Faris Ismail, IIUM, Malaysia

**EXECUTIVE EDITOR**
AHM Zahirul Alam, IIUM, Malaysia

**ASSOCIATE EDITOR**
Nor Farahidah Za'bah, IIUM, Malaysia

**LANGUAGE EDITOR**
Lynn Mason, Malaysia

**COPY EDITOR**
Hamzah Mohd. Salleh, IIUM, Malaysia

**EDITORIAL BOARD MEMBERS**

Abdullah Al-Mamun, IIUM, Malaysia
Abdumalik Rakhimov, IIUM, Malaysia
Amir Akramin Shafie, IIUM, Malaysia
Anis Nurashikin Nordin, IIUM, Malaysia
Erry Yulian Triblas Adesta, IIUM, Malaysia
Erwin Sulaeman, IIUM, Malaysia
Hanafy Omar, Saudi Arabia
Hazleen Anuar, IIUM, Malaysia
Konstantin Khanin, University of Toronto, Canada
Ma'an Al-Khatib, IIUM, Malaysia
Md Zahangir Alam, IIUM, Malaysia
Meftah Hrairi, IIUM, Malaysia
Mohamed B. Trabia, United States
Mohammad S. Alam, Texas A&M University-Kingsville, United States
Mustafizur Rahman, National University Singapore, Singapore
Ossama Abdulkhalik, Michigan Technological University, United States
Razi Nalim, IUPUI, Indianapolis, Indiana, United States
Rosminazuin AB. Rahim, IIUM, Malaysia
Waqar Asrar, IIUM, Malaysia

**AIMS & SCOPE OF IIUMENGINEERING JOURNAL**

The **IIUM Engineering Journal**, published biannually, is a carefully refereed international publication of International Islamic University Malaysia (IIUM). Contributions of high technical merit within the span of engineering disciplines; covering the main areas of engineering: Electrical and Computer Engineering; Mechanical and Manufacturing Engineering; Automation and Mechatronics Engineering; Material and Chemical Engineering; Environmental and Civil Engineering; Biotechnology and Bioengineering; Engineering Mathematics and Physics; and Computer Science and Information Technology are considered for publication in this journal. Contributions from other areas of Engineering and Applied Science are also welcomed. The IIUM Engineering Journal publishes contributions under *Regular papers and Invited review papers*. It also welcomes contributions that address solutions to the specific challenges of the developing world, and address science and technology issues from an Islamic and multidisciplinary perspective.

**REFEREES' NETWORK**

All papers submitted to IIUM Engineering Journal will be subjected to a rigorous reviewing process through a worldwide network of specialized and competent referees. Each accepted paper should have at least two positive referees' assessments.

**SUBMISSION OF A MANUSCRIPT**

A manuscript should be submitted online to the IIUM-Engineering Journal website: http://journals.iium.edu.my/ejournal. Further correspondence on the status of the paper could be done through the journal website and the e-mail addresses of the Executive Editor: zahirulalam*@iium.edu.my* Faculty of Engineering, International Islamic University Malaysia (IIUM), Jan Gombak, 53100, Kuala Lumpur, Malaysia. *Phone: (603) 6421 4529, Fax:(603) 6196 4488.*

## INTERNATIONAL ADVISORY COMMITTEE

A. Anwar, United States
Abdul Latif Bin Ahmad, Malaysia
Farzad Ismail, USM, Pulau Pinang, Malaysia
Hanafy Omar, Saudi Arabia
Hany Ammar, United States
Idris Mohammed Bugaje, Nigeria
K.B. Ramachandran, India
Kunzu Abdella, Canada
Luis Le Moyne, ISAT, University of Burgundy, France
M Mujtaba, United Kingdom
Mohamed AI-Rubei, Ireland
Mohamed B Trabia, United States
Syed Kamrul Islam, United States
Tibor Czigany, Budapest University of Technology and Economics, Hungary
Yiu-Wing Mai, The University of Sydney, Australia.

Published by:
IIUM Press,
International Islamic University Malaysia
Jalan Gombak, 53100 Kuala Lumpur, Malaysia
Phone (+603) 6421-5014, Fax: (+603) 6196-6298

Whilst every effort is made by the publisher and editorial board to see that no inaccurate or misleading data, opinion or statement appears in this Journal, they wish to make it clear that the data and opinions appearing in the articles and advertisement herein are the responsibility of the contributor or advertiser concerned. Accordingly, the publisher and the editorial committee accept no liability whatsoever for the consequence of any such inaccurate or misleading data, opinion or statement.

ISSN 1511 - 788X
9 771511 788008
**IIUM Engineering Journal**
**ISSN: 1511-788X E-ISSN: 2289-7860**

**IIUM ENGINEERING JOURNAL**

Volume 20, Issue 2, December 2019

https://doi.org/10.31436/iiumej.v20i2

# Table of Content

**MATERIALS AND MANUFACTURING ENGINEERING**



**MECHANICAL AND AEROSPACE ENGINEERING**

https://doi.org/10.31436/iiumej.v20i2.1030

# COMPARISON STUDY OF STAINLESS STEEL CYCLIC VOLTAMMOGRAMS IN VARIOUS NATURAL MEDIA ADDITION: PRODUCT AND VOLTAGE EFFICIENCY

**ISANA SUPIAH YOSEPHINE LOUISE*, ENDANG WIDJAJANTI LAKSONO, DEWI YUANITA LESTARI[1]**

*Chemical Education Department, Mathematics and Science Faculty, Yogyakarta State University, Yogyakarta, Indonesia*

**Corresponding author: isana_supiah@uny.ac.id*

*(Received: 4th December 2018; Accepted: 4th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** The cyclic voltammogram is widely used to analyse the electrolysis process. The use of various media, namely rambutan seed flour (*Nephelium lappaceum*), mango seed flour (*Mangifera indica*), breadfruit leaf powder (*Artocarpus altilis*), peria extract (*Momordi charantia*) and aloe vera extract (*Aloe vera*), in this study is to deliberately mimic the occurrence of those media in wastewater produced by several industries. The electrolysis analysis on the presence of several media would minimize the preparation steps on wastewater utilization for generating hydrogen as renewable energy. The research looks at the potential of wastewater as the raw material for the electrolysis process. In this research, stainless steel cyclic voltammograms were studied on water electrolysis. The electrolysis was done in base solution and adding various media, such as rambutan seed flour (*Nephelium lappaceum*), mango seed flour (*Mangifera indica*), breadfruit leaf powder (*Artocarpus altilis*), peria extract (*Momordi charantia*) and aloe vera extract (*Aloe vera*) in various concentrations, 0 - 10 g per litter of water. By reviewing the activity of a stainless steel electrode to decompose water molecules, the media generally caused the occurrence of covering by relatively large molecules around the electrode surface, resulting in decreased activity of the stainless steel electrodes. The optimum condition occurred with the addition of breadfruit leaf powder in all treatments with similar electrode activity as much as 1.68. The result could be implemented in a wastewater electrolysis processes containing the media to generate hydrogen gas.

***ABSTRAK:*** Voltammogram berkitar banyak digunakan bagi menganalisis proses elektrolisis. Dalam kajian ini pelbagai jenis media telah sengaja digunakan seperti tepung biji rambutan (Nephelium lappaceum), tepung biji mangga (Mangifera indica), serbuk daun sukun (Artocarpus altilis), ekstrak peria (Momordi charantia) dan ekstrak lidah buaya (Aloe vera), bagi memimik terjadinya media-media tersebut dalam sisa air buangan yang terhasil daripada beberapa industri. Analisis elektrolisis pada pelbagai media dapat mengurangkan langkah-langkah persiapan pada penggunaan sisa air bagi menghasilkan hidrogen sebagai tenaga boleh baharu. Penilitian kajian ini bertumpu pada potensi sisa air sebagai bahan kasar bagi proses elektrolisis. Kajian ini adalah tentang elektrolisis air voltammogram berkitar pada keluli tahan karat. Proses elektrolisis telah dilakukan pada larutan dasar dan dengan menambah pelbagai media, seperti tepung biji rambutan (Nephelium lappaceum), tepung biji mangga (Mangifera indica), serbuk daun sukun (Artocarpus altilis), ekstrak petola (Momordi charantia) dan ekstrak lidah buaya (Aloe vera) dalam pelbagai kepekatan, 0 - 10 gram pada setiap liter air. Penurunan aktiviti elektrod keluli tahan karat telah disebabkan oleh aktiviti elektrod keluli tahan karat yang mengurai molekul air dan diliputi molekul-molekul besar pada permukaan

elektrod. Keadaan optimum telah berlaku dengan penambahan serbuk daun sukun pada semua rawatan dengan aktiviti elektrod serupa sebanyak 1.68. Dapatan kajian dapat digunakan dalam proses elektrolisis sisa air yang mengandungi media bagi menghasilkan gas hidrogen.

***KEYWORDS*****:** *voltammogram; stainless steel; media; electrode activity; efficiency*

## 1. INTRODUCTION

The voltammetry method is favourable for synthesis, analysis, and characterization or application, namely linear and cyclic voltammetry. In cyclic voltammetry, the same potential was applied on continuous time measurement for the electric current recording [1-5]. Furthermore, an electric current is plotted against a potential (Voltammogram) which can be used as a basis for analysing samples. For the same sample measured under similar conditions, it will have a similar voltammogram. Otherwise, when the modified state is applied, it will give a different voltammogram. The electrolysis of water will provide a typical cyclic voltammogram with differences occurring for electrodes and media changes. Hence, it could be used to study the efficiency of the electrolysis process.

Water molecules can be decomposed to hydrogen and oxygen gases by electrolysis. The electrolysis process depends on the type of electrolyte, the type of electrode, current, and time used [6-10]. Hydrogen is a well-known renewable energy source with its environmentally friendly property. Hydrogen gas production by electrolysis of water has relatively low-efficiency [11]. Therefore, it is necessary to strive for a condition of electrolysis of water with high efficiency in hydrogen gas production and voltage efficiency. To produce safe, environmentally friendly, and cheap hydrogen it needs an effort to meet energy demand, given the increasing rarity of fossil fuels and their non-renewability.

One of the methods to increase efficiency is through modification of the electrodes. The preparation, characterization, and application of stainless steel electro-catalyst electrodes for producing hydrogen gas had been reported by Isana et al. [12]. The voltammogram pattern of the stainless steel/Fe-Co-Ni electrolysis in base solution had been analysed by [13] and the reaction of hydrogen evolution in the stainless steel/Fe-Co-Ni electrode [14] showed that the ternary electrode has relatively better catalytic activity than binary and single electrodes. The reaction of hydrogen evolution in Fe-Co/s, Fe-Ni/s and Co-Ni/s electrodes [15] showed that stainless steel coated electrodes with binary metals have relatively better catalytic activity than stainless steel electrode. Meanwhile, Islam [16] evaluated the surfaces of Fe-Ni films on various coating bath media and concluded the morphology that is characterized by a coarse-grained and non-smooth surfaces with the presence of micro-cracks on it is superior. Other studies by Hairin et al. [17] and Salim [18] showed that high surface area is produced from the coating using the electrodeposition method.

Another method to increase efficiency is conducted on the batch solution manipulation of water electrolysis. The temperature thermogram against the time of electrolysis of various brands of salt solution using carbon electrodes showed the diversity of thermograms, which differed with each brand that had been studied by Isana [19]. Another study on well water from various places with stainless steel electrodes indicated that each well water has different electrolysis behaviour [20]. Furthermore, the splitting of water molecules in a cornstarch medium had been reported, as well as in the medium of mocaf [21], in the *Ipomoea batatas* L flour medium [22], in dahlia bulbs flour [23, 24], in the bitter melon medium [25], and in mango

https://doi.org/10.31436/iiumej.v20i2.1030

seed flour [26] showing that there was covering at the surface of the electrode by the flour. Both the production of hydrogen and catalytic activity were relatively dependent on the media. The modification of batch solution can be beneficial on either improving the quality of water electrolysis or utilizing the wastewater of several industries containing several media.

To study the effect of media addition in the water electrolysis process, research was carried out on the electrolysis of water in base solutions by using stainless steel electrodes in various media. The media used in this research were rambutan seed flour (*Nephelium lappaceum*), mango seed flour (*Mangifera indica*), breadfruit leaves powder (*Artocarpus altilis*), peria extract (*Momordi charantia*) and aloe vera extract (*Aloe vera*). The major chemical compositions in the media from various studies are described in Table 1 and illustrated in Fig. 1. The problem formulation is (1) how does the activity of the stainless steel electrode affect the process of decomposition of water in various media based on the cyclic voltammogram? (2) How does media concentration affect the decomposition of a water molecule? (3) What is the value of the decomposing water molecule to produce hydrogen and oxygen gases in various media at the optimum condition?

Table 1: Major composition of media in water electrolysis using stainless steel

| Media | Carbohydrate | Fatty acids | Crude protein | Amylose | Reference |
|---|---|---|---|---|---|
| Rambutan Seed Flour | 87.04% | - | 10.07% | 32.14% | [27] |
| Mango Seed Flour | 64.23% | 12.00% | 4.95% | - | [28] |
| Breadfruit leaves powder | 50.00% | - | 22.45% | - | [29] |
| Peria extract | - | - | 18.02% | - | [30] |
| *Aloe vera* extract | 61.68% | - | 11.21% | - | [31] |

(a) (b) (c) (d) (e)

Fig. 1: (a) Rambutan and seed, (b) mango and seed, (c) breadfruit leave, (d) peria and (e) *Aloe vera*.

## 2. MATERIALS AND METHODOLOGY

The equipment used here included glassware for preparation, an electrolysis tube, and a eDAQ EChem voltammeter. The materials used were nitric acid, acetone, and $NaHCO_3$ (Merck). Other materials such as stainless steel grades S-430 (thickness 1.2 mm, width 3 mm,

and length 110 mm), platinum electrodes, and Ag/AgCl electrodes were equipped as part of the electrolysis apparatus. The solvents, such as distilled water and aquabides, were produced locally in laboratory, while the media of rambutan seed flour, mango flour, breadfruit powder, peria extract, and *Aloe vera* extract were prepared and extracted from the local market. Several processes such as rinsing, drying, size reduction, and 100-mesh screening were used in the flour production of the media. Meanwhile, aloe vera extract was obtained by crushing the raw material using a blender followed by a filtration process.

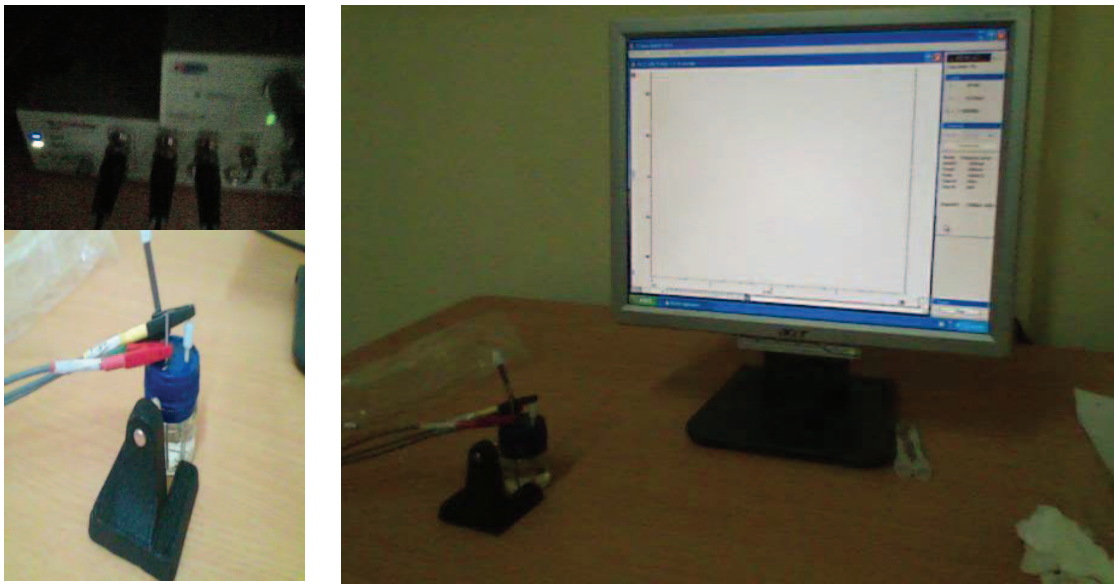


Fig. 2: eDAQ EChem voltammeter equipment in this study.

$NaHCO_3$ (5 g) was poured into one litre of water to make a base solution of mother liquor. Mother solution was added with each media with various concentrations under constant stirring for 1 hour. The concentration of the media varied from 0 to 10 g per 1 liter of water. Electrolysis of water was carried out by using stainless steel electrodes and various media. Meanwhile, the eDAQ EChem voltammeter (Fig. 2) was used to analyse the voltammetry during the electrolysis of water using multiple media. Furthermore, cyclic voltammetry was conducted twice for each sample to ensure data reliability. The eDAQ EChem have recorded the value of anodic-cathodic peak current and anodic-cathodic peak potential of the cyclic voltammogram data file experiment, which was further analysed to gain production efficiency and voltage efficiency. Production efficiency was calculated based on a comparison of the cathodic current peak with or without media, as shown in Eq. 1. Meanwhile, voltage efficiency was compared; observed as the voltage at the cathodic peak condition, for every treatment and the treatment without media as shown in Eq. 2. Furthermore, the electrolysis activity is expressed in Eq. 3 as comparing product efficiency and voltage efficiency.

$$Product\ Efficiency = \frac{I_{with\ media}}{I_{without\ media}} \times 100\% \quad (1)$$

$$Voltage\ efficiency = \frac{Operating\ voltage}{Thermodynami\ \ voltage} \quad (2)$$

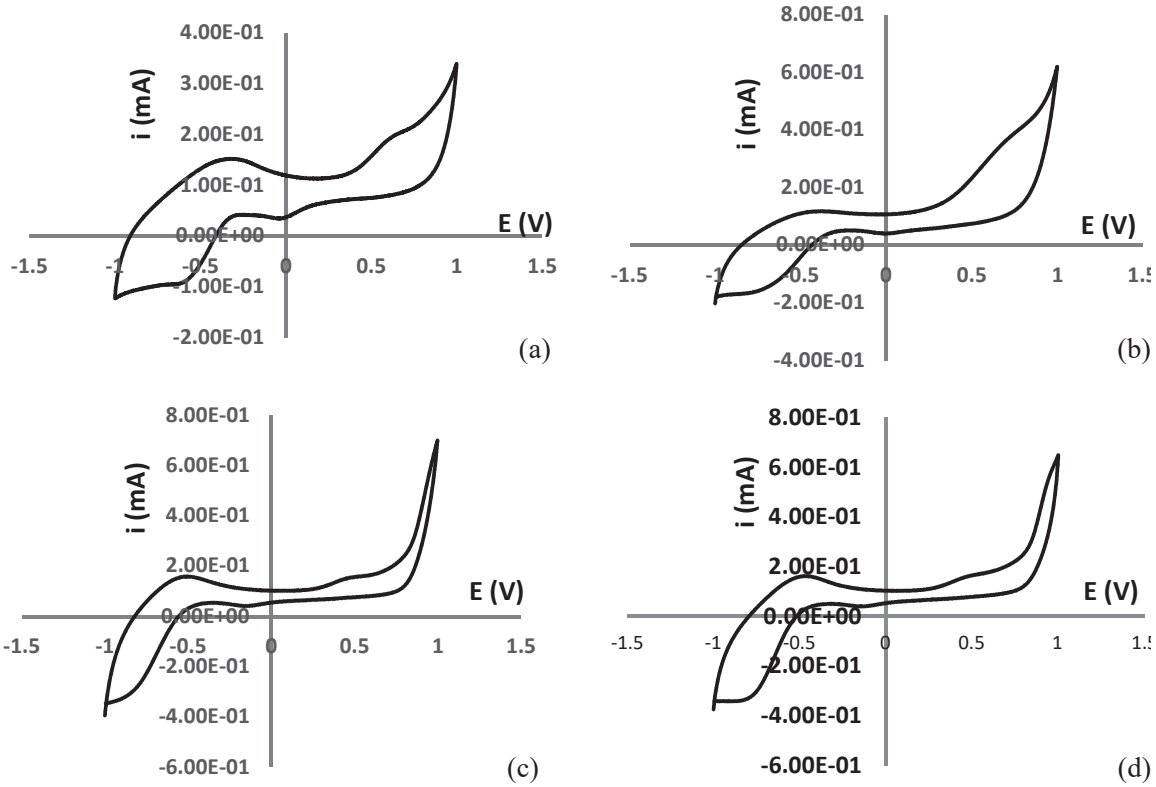
$$Activity\ of\ electrolysis = \frac{Product\ efficiency}{Voltage\ Efficiency} \quad (3)$$

https://doi.org/10.31436/iiumej.v20i2.1030

where thermodynamic voltage is 1.229 V at standard temperature and pressure.

## 3. RESULTS AND DISCUSSION

The stainless steel voltammogram in various media can be seen in Fig. 3. Based on the stainless steel cyclic voltammogram, the cathodic current peak and potential for each system of mass of media versus the efficiency of products are presented in Fig. 4 and voltage efficiency in Fig. 5 in various media such as rambutan seed flour, mango seed flour, breadfruit leaves powder, peria extract, and *Aloe vera* extract. The efficiency of hydrogen product and the voltage efficiency in various media from the cyclic voltammogram are summarized in Table 2. Meanwhile, product efficiency and voltage efficiency also were drawn in Figs. 4 and 5 to compare influences with each other. Furthermore, the activity of electrolysis was shown in Table 3 with the comparison of product efficiency and voltage efficiency in various media. From voltage efficiency, we can calculate the product efficiency using a comparison of voltage efficiency with or without media in every treatment.



https://doi.org/10.31436/iiumej.v20i2.1030

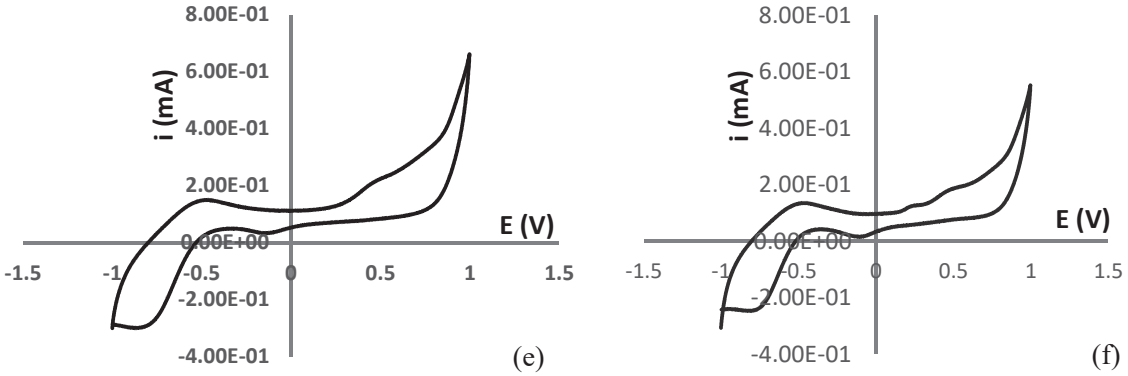



Fig. 3: Cyclic Voltammogram of stainless steel in various media: (a) rambutan seed flour, (b) mango seed flour, (c) breadfruit leaves powder, (d) peria extract, (e) *Aloe vera* extract, and (f) without media.

Tabel 2: The efficiency of hydrogen product and voltage efficiency in various media

| No. | System | Product efficiency (%) | | | | | Voltage efficiency (%) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | RSF | MSF | BLP | PE | AVE | RSF | MSF | BLP | PE | AVE |
| 1 | 0 g | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 67.4 | 59.2 | 59.6 | 59.3 | 59.8 |
| 2 | 1 g | 74.3 | 46.6 | 97.4 | 61.6 | 62.3 | 55.2 | 48.7 | 58.0 | 58.6 | 60.0 |
| 3 | 2 g | 69.7 | 34.8 | 96.3 | 38.1 | 58.3 | 53.5 | 44.7 | 57.4 | 57.7 | 59.1 |
| 4 | 3 g | 49.6 | 35.5 | 100.8 | 53.2 | 51.0 | 54.4 | 43.8 | 60.0 | 58.7 | 59.7 |
| 5 | 4 g | 59.8 | 30.8 | 102.2 | 56.5 | 48.4 | 53.5 | 43.9 | 60.9 | 58.7 | 59.5 |
| 6 | 5 g | 72.7 | 36.1 | 98.9 | 55.7 | 40.2 | 53.5 | 47.0 | 58.9 | 56.1 | 59.6 |
| 7 | 6 g | 67.4 | 34.0 | 95.0 | 55.1 | 36.2 | 55.2 | 46.6 | 56.6 | 56.6 | 59.2 |
| 8 | 7 g | 68.7 | 26.7 | 98.1 | 21.3 | 36.0 | 54.4 | 50.7 | 58.4 | 56.1 | 59.2 |
| 9 | 8 g | 53.4 | 30.4 | 95.6 | 19.5 | 22.8 | 53.5 | 47.2 | 57.0 | 54.4 | 59.1 |
| 10 | 9 g | 40.7 | 36.7 | 99.5 | 48.5 | 25.5 | 54.4 | 46.3 | 59.2 | 53.6 | 59.2 |
| 11 | 10 g | 40.6 | 35.6 | 96.5 | 35.8 | 27.7 | 53.5 | 47.4 | 57.4 | 49.9 | 59.3 |

RSF: rambutan eed flour, MSF: mango seed flour, BLP: breadfruit leaves powder, PE: peria extract, and AVE: *Aloe vera* extract

Cyclic stainless steel voltammograms in various media can be used to study the activity of the stainless steel electrode on decomposition of water molecules to form hydrogen and oxygen gases. The comparison between product gained, and the voltage efficiency is the best method to represent the activity of the stainless steel electrode on decomposition of water molecules to form hydrogen and oxygen gases. Randles-Sevcik's equation describes that the peak current for a reversible system is affected by electron stoichiometry (n), electrode area (A, $cm^2$), diffusion coefficient (D, $cm^2/s$), concentration ($mol/cm^3$), and scan rate (ν, V/s) as shown at Eq. 5. Thus, the concentration of the product is affecting the peak current of cyclic voltammetry. Voltage efficiency is the value of energy used to produce hydrogen and oxygen gas compared to the initial condition without any addition of the media.

$$I_p = (2.69 \times 10^5) n^{3/2} A D^{1/2} C \nu^{1/2} \quad (5)$$

https://doi.org/10.31436/iiumej.v20i2.1030

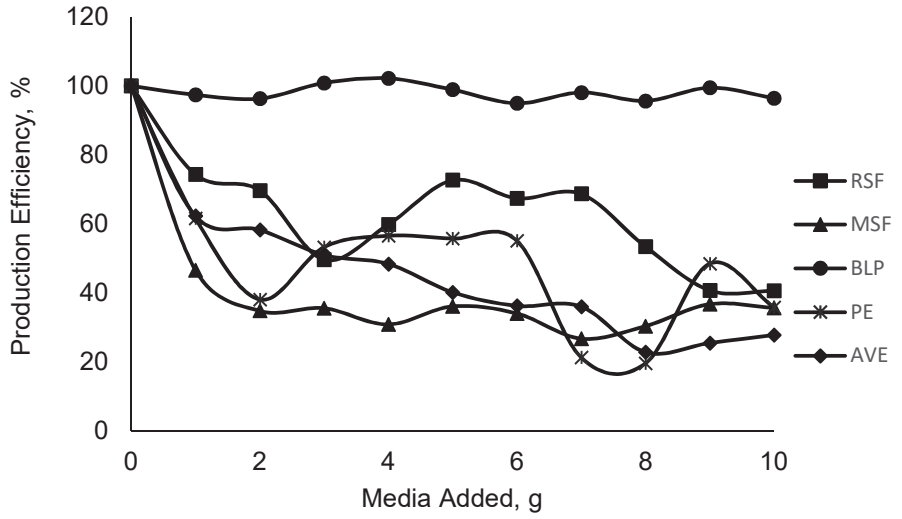



Fig. 4: The efficiency of hydrogen product of various media.

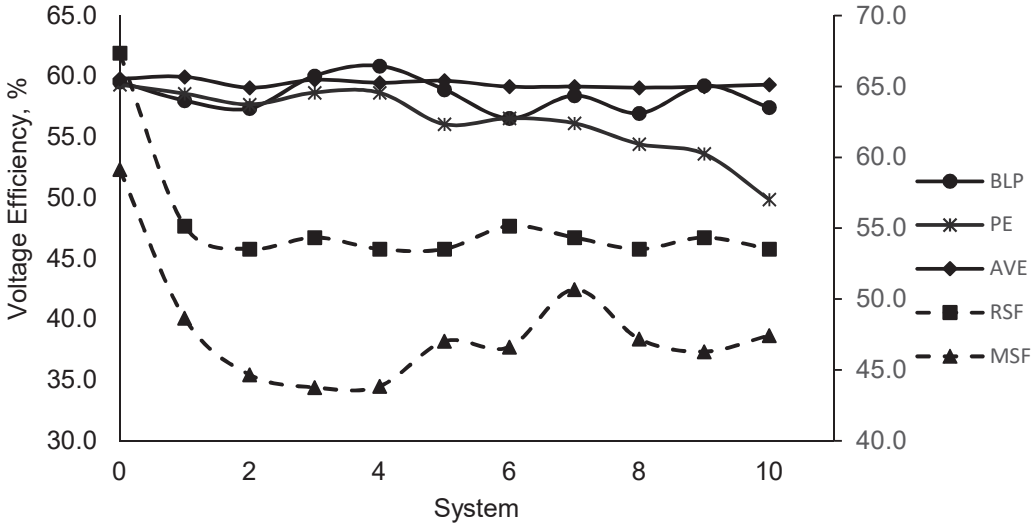



Fig. 5: The voltage efficiency of various media, solid and dotted line indicated major and minor y-axis respectively.

Regarding voltage efficiency, the addition of rambutan seed flour and mango seed flour showed decreasing result. Meanwhile, the other media did not significantly impact voltage efficiency. The result of the hydrogen product yield, ranked from relatively high efficiency to relatively small efficiency, is as follows: breadfruit leaves powder, rambutan seed flour, peria extract, *Aloe vera* extract, and mango seed flour. When viewed from energy consumption, illustrated with voltage efficiency, the sequence becomes: mango seed flour, rambutan seed flour, peria extract, breadfruit leaf powder, and *Aloe vera* extract. The most unfavourable

https://doi.org/10.31436/iiumej.v20i2.1030

medium to decompose water molecules into hydrogen and oxygen gases was the *Aloe vera* extract.

From Table 3, it can be concluded that the bigger the value, the better the activity of the stainless steel electrode is. The calculation using least significant difference was conducted to analyse the significance of the activity of electrolysis value. In general, the media caused the covering surface of the electrode because of its relatively large molecules, resulting in decreased activity of the electrode, as presented in Fig. 6 (b). From Eq. 4, it can be shown that the presence of media can disturb the diffusivity (D, $cm^2/s$) because the media are covering the electrode. Interestingly, the addition of Breadfruit Leaves Powder gave an unusual trend, as shown in Fig. 6 (a). Those graph tendencies are strongly believed to be the result of the polymer structure in the media that had not covered the electrode. The polymer occurrence in every media affects the interaction between the surface of the electrode and the polymer itself influences the electrode activity.

Tabel 3: Activity of electrolysis in various media

| No | System | Activity of Electrolysis | | | | |
|---|---|---|---|---|---|---|
| | | RSF | MSF | BLP | PE | AVE |
| 1 | 0 g | $1.49^b$ | $1.70^a$ | $1.68^a$ | $1.69^a$ | $1.68^a$ |
| 2 | 1 g | $1.35^b$ | $0.96^b$ | $1.68^a$ | $1.06^b$ | $1.04^b$ |
| 3 | 2 g | $1.31^b$ | $0.78^c$ | $1.68^a$ | $0.66^c$ | $0.99^b$ |
| 4 | 3 g | $0.92^c$ | $0.82^c$ | $1.68^a$ | $0.91^c$ | $0.86^c$ |
| 5 | 4 g | $1.12^b$ | $0.71^c$ | $1.68^a$ | $0.97^b$ | $0.82^c$ |
| 6 | 5 g | $1.36^b$ | $0.77^c$ | $1.68^a$ | $1.00^b$ | $0.68^c$ |
| 7 | 6 g | $1.23^b$ | $0.73^c$ | $1.68^a$ | $0.98^b$ | $0.62^c$ |
| 8 | 7 g | $1.27^b$ | $0.53^c$ | $1.68^a$ | $0.38^c$ | $0.61^c$ |
| 9 | 8 g | $1.00^b$ | $0.65^c$ | $1.68^a$ | $0.36^c$ | $0.39^c$ |
| 10 | 9 g | $0.75^c$ | $0.80^c$ | $1.68^a$ | $0.91^c$ | $0.44^c$ |
| 11 | 10 g | $0.76^c$ | $0.76^c$ | $1.68^a$ | $0.72^c$ | $0.47^c$ |

RSF: rambutan seed flour, MSF: mango seed flour, BLP: breadfruit leaves powder, PE: peria extract, and AVE: *Aloe vera* extract
a,b, and c show significance group.

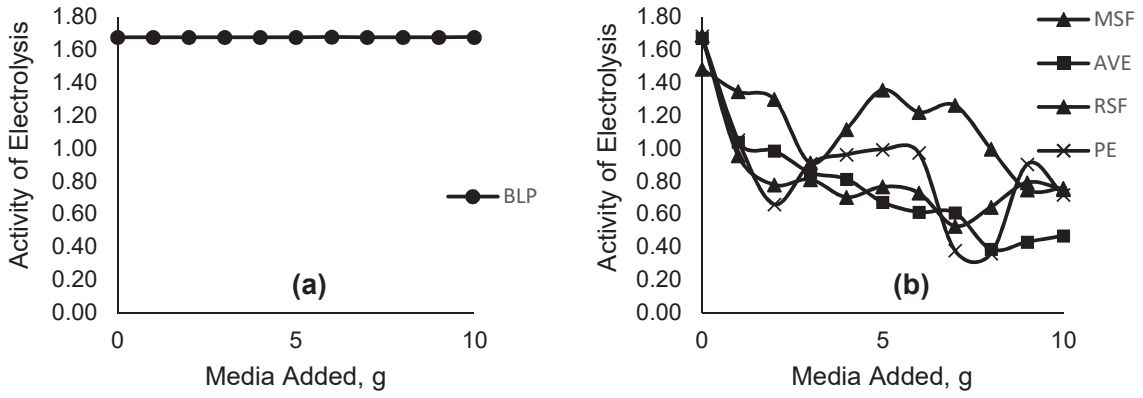



Fig. 6: The activity of electrolysis in various media; (a) unusual trend and (b) usual trend.

The optimum condition occurred when the electrolysis activity is highest. It is obtained when the production efficiency is high, and the voltage efficiency is low. The higher the cathodic current peak, the higher hydrogen production is, so that the higher the activity of the electrode; and the lower the potential, the lower the voltage efficiency. That is the best condition to decompose water molecules into hydrogen and oxygen gases. Therefore, the optimum condition from this research occurred with the addition of breadfruit leaves powder for which the electrolysis activity was constant around 1.68.

## 4. CONCLUSION

The activity of stainless steel electrodes in the decomposition of a water molecule into hydrogen and oxygen gases in various media varied greatly. The optimum conditions for water molecule decomposition into hydrogen and oxygen gases were achieved by adding breadfruit leaves powder at every system with similar activity of electrolysis around 1.68. The addition of various media, regularly, could not improve the electrode activity, but the addition of breadfruit leaves powder could stabilize the electrode activity. The results could give an illustration and initial data to the generalization of hydrogen gas using the electrolysis process in wastewater mainly containing flour media.

## ACKNOWLEDGEMENT

The authors send gratitude to the University for the facilities and support during the research.

## REFERENCES

[1] Fricke, K., Harnisch, F., and Schröder, U., 2008, On the use of cyclic voltammetry for the study of anodic electron transfer in microbial fuel cells, *Energy & Environmental Science*, 1(1): 144-147.

[2] Karlberg, G., Jaramillo, T., Skulason, E., Rossmeisl, J., Bligaard, T., and Nørskov, J.K., 2007, Cyclic voltammograms for H on Pt (111) and Pt (100) from first principles, *Physical Review Letters*, 99(12): 126101.

[3] Kissinger, P.T. and Heineman, W.R., 1983, Cyclic voltammetry, *Journal of Chemical Education*, 60(9): 702.

[4] Van Benschoten, J.J., Lewis, J.Y., Heineman, W.R., Roston, D.A., and Kissinger, P.T., 1983, Cyclic voltammetry experiment, *J. Chem. Educ.*, 60(9): 772.

[5] Rountree, E.S., McCarthy, B.D., Eisenhart, T.T., and Dempsey, J.L., Evaluation of homogeneous electrocatalysts by cyclic voltammetry, ACS Publications, 2014.

[6] Bard, A.J., Faulkner, L.R., Leddy, J., and Zoski, C.G., 1980, *Electrochemical methods: fundamentals and applications*, wiley New York,

[7] De Nora, O., De Nora, V., and Spaziante, P.M., Electrolysis cells, Google Patents, 1977.

[8] DeNora, O., Electrolysis cell, Google Patents, 1986.

[9] Olivares-Ramirez, J., Campos-Cornelio, M., Godínez, J.U., Borja-Arco, E., and Castellanos, R., 2007, Studies on the hydrogen evolution reaction on different stainless steels, *International Journal of Hydrogen Energy*, 32(15): 3170-3173.

[10] Stojić, D.L., Marčeta, M.P., Sovilj, S.P., and Miljanić, Š.S., 2003, Hydrogen generation from water electrolysis—possibilities of energy saving, *Journal of Power Sources*, 118(1-2): 315-319.
[11] Zeng, K. and Zhang, D., 2010, Recent progress in alkaline water electrolysis for hydrogen production and applications, *Progress in Energy and Combustion Science*, 36(3): 307-326.
[12] Isana SYL, W.T., Agus Kuncaka and Triyono, 2016, Preparation, Characterization and Application of Stainless Steel/Fe, Co, Ni Electrocatalyst Electrodes, *J. Chem. & Cheml. Sci.*, 6 (11)1151-1160.
[13] Isana, S., Trisunaryanti, W., and Kuncaka, A., 2012, Studies On The Hydrogen Evolution Reaction On Fe-Co-Ni, *IOSR Journal of Applied Chemistry*, 6-10.
[14] Louise, I.S.Y., 2014, Stainless steel/Fe-Co-Ni sebagai Elektro Katalis pada Reaksi Evolusi Hidrogen,
[15] Isana S. Y. L., W.T., Agus K., Triyono, Studies on the hydrogen evolution reaction on Fe-Co/s, Fe-Ni/s and Co-Ni/s electrodes, The International Post Graduate Conference on Science and Mathematics 2013, Convention Hall, e-Learning Building, Universiti Pendidikan Sultan Idris, Malaysia, 2013.
[16] Islam, M., 2009, Anomalous Electrodeposition Of Fe-Ni Alloy Coating From Simple And Complex Baths and its Magnetic Property, *IIUM Engineering Journal*, 10(2): 108-122.
[17] Hairin, A.N., Othman, R., Ani, M.H., and Saputra, H., 2011, High Discharge Rate Electrodeposited Zinc Electrode for Use in Alkaline Microbattery, *IIUM Engineering Journal*, 12(5): 115-122.
[18] Salim, W.W.A.W., Benoudjit, A., Guthoos, H.F.A., and Arris, F.A., 2017, PEDOT: PSS-Modified Platinum Microelectrodes for Measurement in Aqueous Media: Effect of Polymer Surface Area on Long-term ANodic Peak Current Stability, *IIUM Engineering Journal*, 18(2): 11-15.
[19] Isana, S., Variasi Temperatur Dan Waktu Pada Elektrolisis Larutan Garam Dapur Berbagai Merk, Seminar Nasional Kimia, 2009.
[20] Isana, S., Perilaku sel elektrolisis air dengan elektroda stainless steel, Prosiding Seminar Nasional Kimia dan Pendidikan Kimia, Yogyakarta, 2010, pp. 1-9.
[21] Louise, Y.A.A.a.I.S.Y., Reaksi Evolusi Hidrogen Menggunakan Media Tepung Mocaf dengan Elektroda Stainless Steel/Fe-Co-Ni, Jurusan Pendidikan Kimia, Universitas Negeri Yogyakarta, 2016.
[22] Isana S. Y. L., S., Heru Pratomo, Voltamogram siklik stainless steel dalam media tepung ubi jalar (Ipomoea batatas L) (Cyclic voltammogram of stainless steel in Ipomoea batatas L flour), Seminar Nasional Kimia, Jurusan Pendidikan Kimia FMIPA Universitas Negeri Yogyakarta, Yogyakarta, Indonesia, 2016.
[23] Isana, S., Yuanita, D., and Pratomo, H., 2015, Pemecahan Molekul Air dengan Menggunakan Media Tepung Umbi Dahlia (Dahlia Pinnata), *Jurnal Sains Dasar*, 4(2): 173-178.
[24] Isana, S., Yuanita, D., Sulistyani, A., and Pratomo, H., Voltammogram of stainless steel/Fe-Co-Ni electrode on water electrolysis in base condition with dahlia pinnata tuber starch media, American Institute of Physics Conference Series, 2017.
[25] Isana S. Y. L., H.P.A., Sulistyani, dan Lathifa H., The efficiency of hydrogen production from water electrolysis with bitter melon as a media using stainless steel/Fe-Co-Ni

electrode, Basic and Applied Sciences Interdisciplinary Conference (BASIC), Universitas Indonesia, Jakarta, Indonesia, 2017.

[26] Isana S. Y. L., S., dan Nur A. R., Breaking of H2O in mango seed flour medium with stainless steel/Fe-Co-Ni electrode, Basic and Applied Sciences Interdisciplinary Conference (BASIC), Universitas Indonesia, Jakarta, Indonesia, 2017.

[27] Eiamwat, J., Wanlapa, S., Sematong, T., Reungpatthanapong, S., Phanthanapatet, W., Hankhuntod, N., and Kampruengdet, S., 2015, Rambutan (Nephelium lappaceum) seed flour prepared by fat extraction of rambutan seeds with SC-CO2, *Isan Journal of Pharmaceutical Sciences*, 10138-146.

[28] Okpala, L. and Gibson-Umeh, G., 2013, Physicochemical properties of mango seed flour, *Nigerian Food Journal*, 31(1): 23-27.

[29] Oladipupo, A.A. and Abiodun, T.G., 2014, Phytochemical analysis and antimicrobial effect of Chrysophillum albidum leave extract on gastrointestinal tract pathogenic bacteria and fungi in human, *Journal of Applied Chemistry*, 71-5.

[30] Yuwai, K.E., Rao, K.S., Kaluwin, C., Jones, G.P., and Rivett, D.E., 1991, Chemical composition of Momordica charantia L. fruits, *J. Agric. Food Chem.*, 39(10): 1762-1763.

[31] Scala, K.D., Vega-Gálvez, A., Ah-Hen, K., Nuñez-Mancilla, Y., Tabilo-Munizaga, G., Pérez-Won, M., and Giovagnoli, C., 2013, Chemical and physical properties of aloe vera (Aloe barbadensis Miller) gel stored after high hydrostatic pressure processing, *Food Science and Technology*, 33(1): 52-59.

https://doi.org/10.31436/iiumej.v20i2.1062

# PRODUCTION OF CARRAGEENAN BY DIFFERENT STRAINS OF *Kappaphycus alvarezii* CULTIVATED IN SERANG, INDONESIA

**ANAS ZAKARIA[1], EGI NURUL RAHMAN[1], ULIN NOOR RAHMANI[1], ROBERT MANURUNG[1], NOOR ILLI MOHAMAD PUAD[2], MUHAMMAD YUSUF ABDUH[1*]**

*[1] School of Life Sciences and Technology, Institut Teknologi Bandung, Jalan Ganesha No. 10 Bandung 40132 Indonesia*
*[2]Bioprocess and Molecular Engineering Research Unit, Department of Biotechnology Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, P.O. Box 10, 50728 Kuala Lumpur, Malaysia*

**Corresponding author: yusuf@sith.itb.ac.id*

*(Received: 22nd January 2019; Accepted: 26th June 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** *Kappaphycus alvarezii* is one of the main seaweeds cultivated in Indonesia. Its use as a raw material for the kappa-carrageenan industry has increased its cultivation activities. However, standard cultivation protocol for *Kappaphycus alvarezii* has yet to be well-documented, particularly on strain selection for kappa-carrageenan production. There are various strains of *Kappaphycus alvarezii* grown in Indonesia, specifically in Serang, Banten such as green, red, and yellowish green strains. In this study, growth rate, carrageenan yield, and gel strength from different strains of *Kappaphycus alvarezii* were investigated. It was observed that the specific growth rate of green type (4.14% /day) differed significantly from the other two types (red; 3.41% /day and yellowish green; 3.47% /day). The red type had the highest yield of carrageenan (45.16%) followed by yellowish green (36.66%) and green strain (33.67%). Gel strength was not significantly affected by the strain variations where the observed values for yellowish green, green, and red strain were 344.46, 316.91 and 313.10 g/cm$^2$, respectively. Green *Kappaphycus alvarezii* had the highest biomass (0.27 g/cm/d) and carrageenan productivity (0.012 g/cm/d). It can be concluded that the green strain of *Kappaphycus alvarezii* is the most suitable for cultivation in Serang, particularly for the kappa-carrageenan industry.

***ABSTRAK:*** *Kappaphycus alvarezii* adalah sebahagian daripada rumpai air yang dituai di Indonesia. Penggunaannya sebagai bahan dasar kepada industri kappa-carrageenan telah menambah kepada aktiviti penuaian. Walau bagaimanapun, protokol penuaian biasa bagi *Kappaphycus alvarezii* adalah tertakluk pada dokumentasi penuh terutama dalam pemilihan strain pada penghasilan kappa-carrageenan. Terdapat pelbagai jenis strain *Kappaphycus alvarezii* yang tumbuh di Indonesia, terutama di Serang, Banten seperti hijau, merah, dan strain hijau kekuningan. Kajian ini, kadar pertumbuhan, hasil karaginan dan kekuatan gel daripada pelbagai strain *Kappaphycus alvarezii* telah dikaji. Didapati bahawa kadar pertumbuhan tertentu pada jenis hijau (4.14% /hari) berbeza ketara berbanding dari dua jenis lain (merah; 3.41% /hari dan hijau kekuningan; 3.47% /hari). Jenis merah mempunyai hasil karaginan tertinggi (45.16%) diikuti oleh hijau kekuningan (36.66%) dan strain hijau (33.67%). Kekuatan gel adalah tidak ketara mempengaruhi kepelbagaian variasi di mana nilai yang dipantau pada strain hijau kekuningan, hijau dan merah adalah 344.46, 316.91 dan 313.10 g/cm$^2$, masing-masing. *Kappaphycus alvarezii* hijau mempunyai biojisim tertinggi (0.27 g/cm/d) dan penghasilan karaginan (0.012 g/cm/d). Kesimpulannya strain hijau *Kappaphycus*

*alvarezii* adalah lebih sesuai untuk disemai di Serang terutamanya pada industri kappa-karaginan.

***KEYWORDS:*** *Kappaphycus alvarezii; carrageenan yield; gel strength; productivity; specific growth rate*

# 1. INTRODUCTION

*Kappaphycus alvarezii* belongs to the Rhodophyta group and its growing habitat is generally located at the top of the sub-territorial zone beneath the seawater tidal line on the rock or sand with moderate or medium speed ocean waves [1]. *K. alvarezii* has several strains that are distinguished mainly because of the difference in their color. It is the main source of kappa-carrageenan, a hydrocolloid that has been used extensively as an additive in various industries such as food, cosmetics, and pharmaceutical industries [2,3]. Carrageenan is a linear sulfated polysaccharide made up of a repeating disaccharide sequence of α-D-galactopyranose linked 1,3 called the A residue and β-D-galactopyranose residues linked through positions 1,4 (B residues) that can be extracted from *K. alvarezii* [4].

*K. alvarezii* was firstly reported to be cultivated in the Philippines back in 1970 [5]. Then in the year 1985, with the increasing global demand for carrageenan, the cultivation of *K. alvarezii* made a debut in Indonesia [6]. Several factors, such as climate and weather conditions, ease, and low costs required for its cultivation have contributed to the vast amount of cultivation activities in Indonesia [7]. The three strains commonly cultivated in Serang, Indonesia are red, green, and yellowish green strains. However, there is still a lack of reports regarding the characteristics of carrageenan produced from these strains. This study aimed to determine the effect of strain variations towards the yield and gel strength of the extracted carrageenan. The results would assist the local community in selecting a strain yielding a good quality of carrageenan to be cultivated in Serang, Indonesia.

# 2. MATERIALS AND METHODS

## 2.1 Cultivation of *K. alvarezii*

The cultivation of *K. alvarezii* was carried out at Desa Lontar, Kecamatan Tirtayasa, Kabupaten Serang, Banten, Indonesia using the off-bottom method. The cultivation period in this study was set similar to the common practice in the field, namely 45 days. The cultivation parameters were: salinity of sea water ranging between 30 ppt - 33 ppt; average water temperature of 32°C; sea water pH of 7.9-8.3; depth of sunlight penetration of 13.75 cm and received light intensity ranging between 59,100 – 62,400 lux. The shoots of vegetative growth of green, red, and yellowish green strains were obtained from Desa Lontar, Serang. About 30 g of seaweed fragments were tied on polyethylene (PE) ropes and the spacing between those fragments were 5, 10, and 15 cm, respectively (Fig. 1). The rope was then mounted on the pile horizontally.

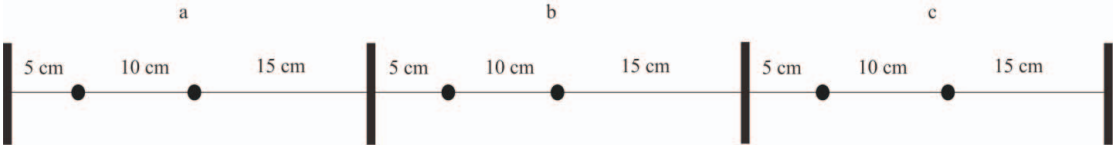



Fig. 1: Schematic diagram of *K. alvarezii* strains cultivation (a) green (b) yellowish green and (c) red with different spacing of seaweed fragments (5 cm, 10 cm and 15 cm).

https://doi.org/10.31436/iiumej.v20i2.1062

The fresh weight of each *K. alvarezii* strain was recorded every 7 days during the 45 days period of cultivation. The seaweed was then harvested after 45 days and the fresh weight was again measured. The growth rate (Eq. (1)) was then calculated according to Lignell and Pedersen [8].

$$Growth\ Rate\ (\%) = \left[\left(\frac{W_t}{W_i}\right)^{\frac{1}{n}} - 1\right] \times 100 \quad (1)$$

where $W_t$ = fresh weight after 45 days (g); $W_i$ = initial fresh weight (g); $n$ = cultivation period (45 days).

### 2.2 Carrageenan Extraction

Each strain of 45 days old *K. alvarezii* was sun-dried for 3 days until the water content reached 20-25%. Extraction was carried out by adding 3 g of dried *K. alvarezii* to 4% NaOH solution [9]. The process was carried out for 2 hours at 80 °C with continuous stirring. The extract was vacuum-filtered and the carrageenan was recovered from the solution by precipitation with 5% KCl solution. The precipitate containing carrageenan was then again immersed in a 5% KCl solution prior to washing with water. The extracted carrageenan was dried in an oven at 105 °C until a constant dry weight was achieved. The yield of carrageenan extract was calculated using Eq. (2).

$$Yield\ (\%) = \frac{Dry\ weight\ of\ extracted\ carrageenan\ (g)}{Dry\ weight\ of\ K.alvarezii} \times 100\% \quad (2)$$

### 2.3 Carrageenan Extract Analysis

The gel strength of extracted carrageenan was determined by dissolving 1.5% on a weight basis (w/w) of carrageenan in an aqueous solution at 90 °C. The solution was then poured into a plastic tube (15 mm in diameter) and left at the room temperature for 3 hours [10]. The gel strength was analyzed using a TA-XT Plus Texture Analyzer (Stable Micro Systems, England) by applying a 2 kg maximum force using a 1 cm (diameter) cylindrical probe. The probe pressed the gel at a rate of 2 mm/s. Maximal penetration strength was considered as the gel strength, which was registered by the Texture Expert software (Stable Micro Systems Ltd., England). All analyses were conducted in triplicates.

### 2.4 Estimation of Carrageenan Productivity

The productivity of the cultivation process and carrageenan yield were determined based on the data of 10 cm cultivation spacing, 45 days cultivation time, as well as the growth rate and carrageenan yield for each *K. alvarezii* strain (Eq. (3)).

$$P = (W_t - W_i)/\ (L\ x\ t) \quad (3)$$

where $P$ = biomass productivity (g/cm/d); $W_t$ = the final weight of biomass (g), $W_i$ = weight of initial biomass (g), $L$ = length of cultivation distance (cm), $t$ is the cultivation time (d).

The estimation of carrageenan productivity was calculated using Eq. (4).

$$C_P = Y\ x\ M_d/t \quad (4)$$

where $C_P$ = carrageenan productivity (g/cm/d); $Y$ = carrageenan yield (%) and $M_d$ = mass of dried *K. alvarezii* after extraction with 25% of water (g), $t$ is the cultivation time (d).

https://doi.org/10.31436/iiumej.v20i2.1062

### 2.5 Statistical Analysis

All experiments in this study were carried out using a Completely Randomized Design with one main factor: variation of the *K. alvarezii* strain (green, red, and yellowish green). Each treatment was repeated three times (randomly). The data was processed using SPSS 23 for MacBook Pro at a 95% confidence level.

## 3. RESULTS AND DISCUSSION

### 3.1 The Growth of Different Strains of *K. alvarezii*

The rate of biomass accumulation for *K. alvarezii* green strain for various cultivation spacing as shown in Fig. 2 implies that the spacing did not significantly affect its growth rate ($p > 0.05$). Similar observations were recorded for the red and yellowish green strains. Hence, 10 cm spacing was further used and discussed in this study since it is the common distance used by the seaweed farmers in the Serang area. The growth rate calculated using Eq. (1) was then compared with the strains grown in real conditions at Serang (Fig. 3). The calculated and observed values were in a good agreement with each other and it indicates that Eq. (1) can be used to estimate the rate of biomass accumulation of *K. alvarezii*.

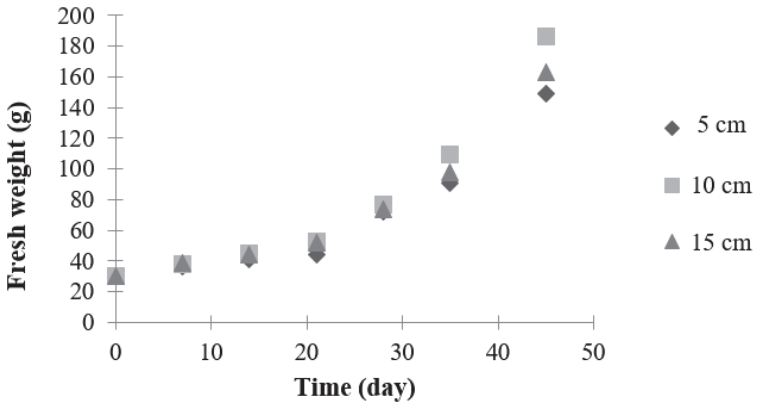



Fig. 2: Growth profile of *K. alvarezii* green strain at different cultivation spacing between seaweed fragments.

The green strain of *K. alvarezii* had the highest growth rate (4.14 %/day) compared to red (3.41 %/day) and yellowish green (3.47 %/day) strains (Fig. 4). The difference in terms of growth rate between the green strain and the other two strains were significant ($p < 0.05$). This result is similar with what has been reported by Munoz et al. [11] that green strain of *K. alvarezii* demonstrated a higher growth rate compared to the red strain. However, it was also reported that the red strain (3.8 %/day) achieved a slightly higher growth rate when compared to the green strain (3.4 %/day) [12]. Meanwhile for the red strain in this study, the growth rate was lower than reported by Munoz et al. [12], which ranged from 1.1 %/day to 3.8 %/ day but below the range that has been reported by Hayashi et al. [2] (5.2-7.2 %/day). Overall, the values of *K. alvarezii* growth rate indicated that the Lontar village area, Serang is a suitable place to cultivate *K. alvarezii*. This is because its growth rate falls within the range of 3-5 %/day which is the standard value to specify a suitable cultivation areas for *K. alvarezii* [1].

### 3.2 Carrageenan Yield and Gel Strength

Carrageenan yield is one of the parameters observed in this study. This is because the yield of carrageenan from the *K. alvarezii* extraction process would directly affect its productivity and economic value. The yield of carrageenan extract for three strains of *K. alvarezii* is depicted in Fig. 5. The highest yield of carrageenan was produced by the red strain (45.16%) whereas the carrageenan yield from green and yellowish green strains were 33.67% and 36.66%, respectively. These values were higher than reported by Hayashi et al. [2] which ranged between 20-32% and were comparable with the range (8-30%) reported by Valderrama et al. [13]. The difference in the yield of carrageenan for each strain was obvious, especially between the green and red strains (Fig. 5). This is possibly due to their pigment compositions. Each pigment has a different absorption rate for the color spectrum, which will eventually affect the main product of photosynthesis which is carrageenan [14]. The results of the ANOVA analysis suggest that strain variations had a significant effect on the carrageenan yield (($p < 0.05$).

Carrageenan properties, such as gel strength, have made carrageenan into an additive in various industries. In this study, the gel strength varied from 313.10 - 344.46 g/cm$^2$ (Fig. 6). The minimum gel strength was found in the carrageenan extracted from the red strain (313.10 g/cm$^2$) while the maximum one was produced by the yellowish green strain (344.46 g/cm$^2$). The results of the ANOVA analysis suggest that strain variations did not have a significant effect on the gel strength (($p > 0.05$).

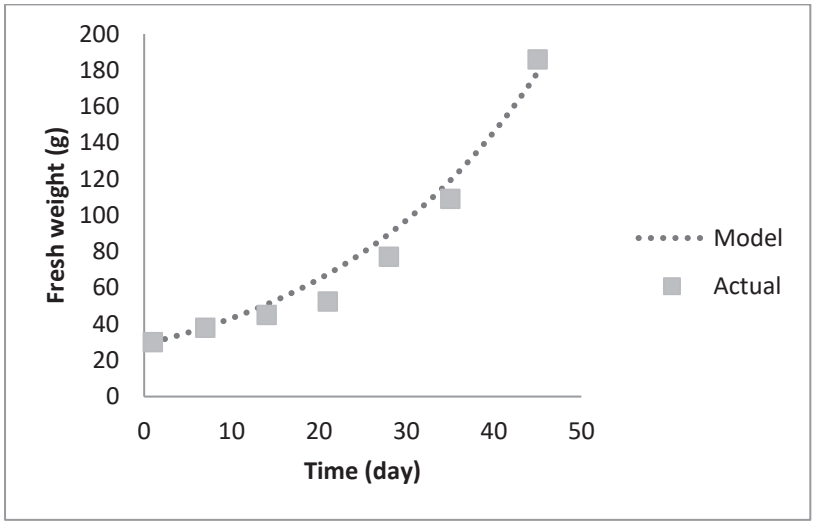
The gel strength values were higher compared to the one reported previously which ranged from 3 to 112 g/cm$^2$ [15]. However, carrageenan gel strength in this study was lower compared to the carrageen quality standard set by the Food and Agriculture Organization (FAO) which lies in the range of 900 and 1200 g/cm$^2$ [16, 17] as well as from the values reported by Basmal et al. [18] which lie in the range of 1029-1243 g/cm$^2$. Variations in the reported gel strength values might be due to different carrageenan extraction processes.

The yield and quality of carrageenan produced is not only influenced by the strains of seaweed that are being cultivated but also by several biological and environmental parameters such as plant age, light, nutrient, temperature, and salinity of the seawater [2, 19-24]. There could be other factors that contribute to the differences in the results achieved in this study with other reported results from studies conducted in different places.

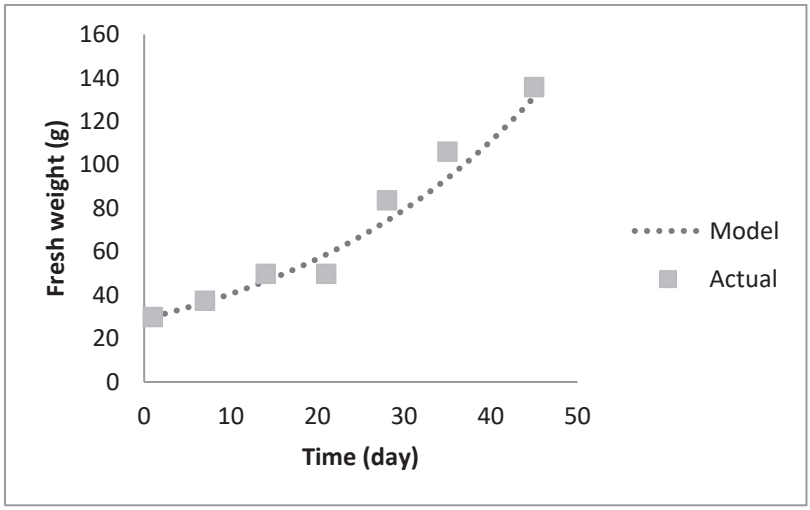
### 3.4 Comparison of Biomass and Carrageenan Productivity

The biomass productivity was estimated based on the difference between the final and initial weight of *K. alvarezii* during 45 days of cultivation on polyethylene rope with a 10 m spacing between seaweed fragments. The cultivation was carried out for 45 days following the common practice in the field that the final weight of the biomass may have reached up to 400% of its initial weight [25]. It has been reported in a previous study that the highest biomass productivity of *K. alvarezii* was recorded after 44 days of cultivation, after which it started to decline when the cultivation period was prolonged to 59 days [2]. In another study, it is also reported that *K. alvarezii* that was cultivated for 45 days had a higher neutrophic activity than *K. alvarezii* that was harvested sooner (15 days) [26]. In this study, the percentage of biomass increase varies for all the three strains with the green strain having the highest percentage of biomass increase (496%) followed by the yellowish green strain (349%) and the red strain (337%).

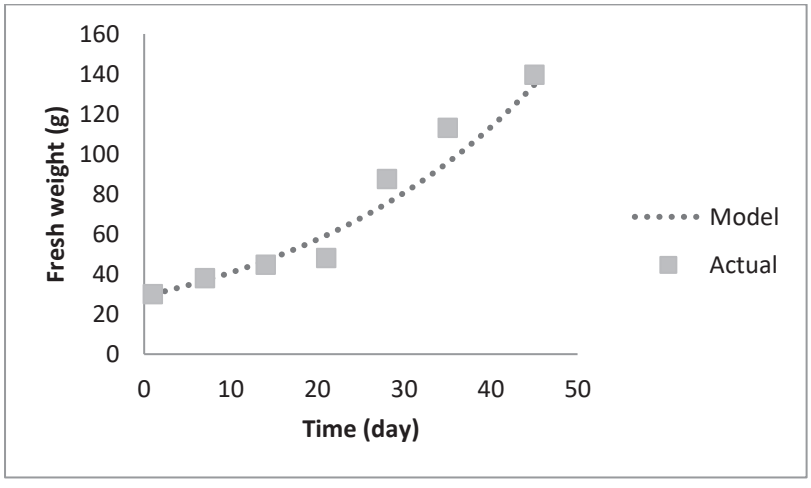
https://doi.org/10.31436/iiumej.v20i2.1062




(a)




(b)




(c)

Fig. 3: Actual and modelled biomass accumulation rate for various *K. alvarezii* strain (a) green, (b) red, (c) yellowish green strains.

https://doi.org/10.31436/iiumej.v20i2.1062

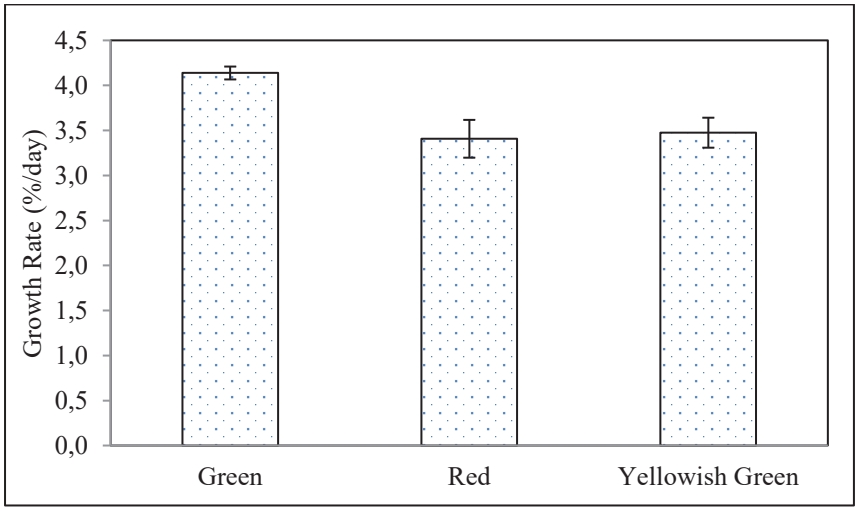



Fig. 4: Growth rate of *K. alvarezii* for different strain variations.

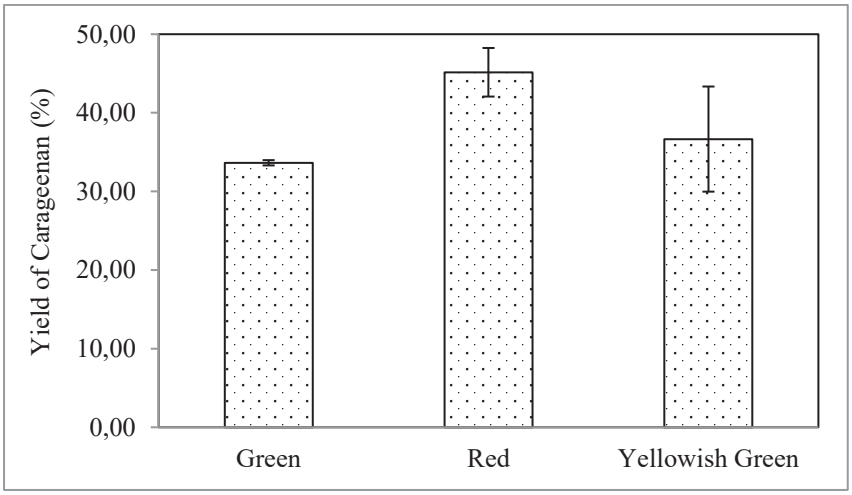



Fig. 5: Comparison of carrageenan yield for different strains of *K. alvarezii.*

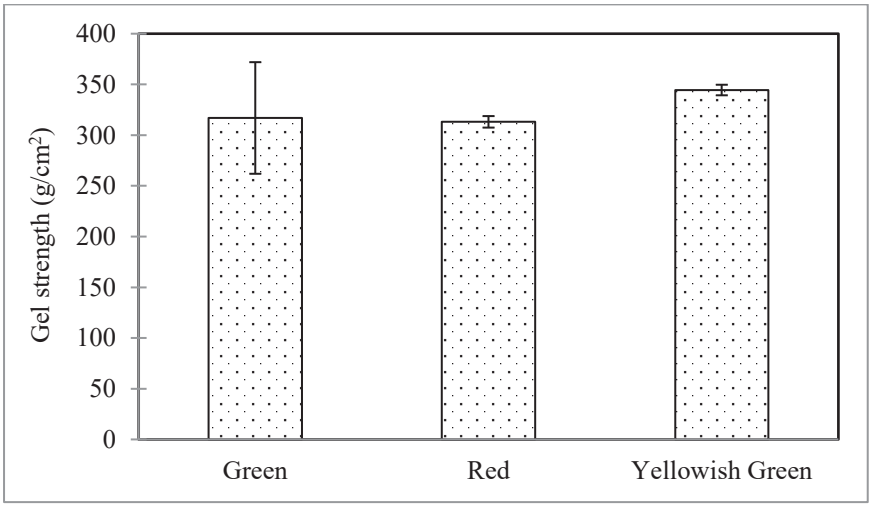



Fig. 6. Comparison of carrageenan gel strength for different *K. alvarezii* strains. There is no significant difference for each strain type.

In a previous study on the cultivation of *K. alvarezii* using a longline method in Southwest Sulawesi, the biomass was cultivated for 40 days and recorded a percentage of biomass increase of 350% [27]. The study also compared a different cultivation technique, namely floating cage, and reported a higher percentage of biomass increase of 676%. Several studies have reported the advantages of the floating cage method as compared to the longline method because it provides a protection to the seaweed from various herbivores and has a higher biomass productivity [27,28]. However, a longline method was used in this study due to the low depth of the cultivation area in Serang. Moreover, the longline method has been widely used by seaweed farmers in the Serang area for cultivating *K. alvarezii*.

Overall, the productivity of the green strain in terms of biomass and carrageenan was the highest among the strains (Table 1). The biomass productivity of the green strain was 0.27 g/cm/d while the red and yellowish green strains were 0.16 g/cm/d and 0.17 g/cm/d. The results resemble the biomass productivity of 0.22 g/cm/d reported in a previous study for the cultivation of *K. alvarezii* using a longline method in Southwest Sulawesi [27]. Based on the yield of carrageenan obtained in the previous section, the estimated productivity of carrageenan in ascending order for yellowish green, red and green are 0.008, 0.01 and 0.012 g/cm/d, respectively. Although several studies have reported the yield of carrageenan extraction process, studies that provide estimation on carrageenan productivity are still very scarce. Hence, the estimated carrageenan productivity reported in this study may provide useful information to produce carrageen from *K. alvarezii*.

Table 1: Productivity of *K. alvarezii* strains

| Strain | Productivity (g/cm/d) | |
|---|---|---|
| | Biomass | Carrageenan |
| Green | 0.27 | 0.012 |
| Red | 0.16 | 0.010 |
| Yellowish green | 0.17 | 0.008 |

## 4. CONCLUSION

In brief, variations of cultivation spacing between seaweed fragments did not affect the growth rate of *K. alvarezii*. On the other hand, different strains of *K. alvarezii* showed different growth rates and carrageenan yields. The green strain recorded the highest growth rate (4.14% /day) while the highest carrageen yield (45.16%) was produced by the red *K. alvarezii* strain. Green *K. alvarezii* had the highest biomass productivity (0.27 g/cm/d) as well as carrageenan productivity (0.012 g/cm/d) compared to the other two strains. These observations can be used as a reference in selecting *K. alvarezii* strain to be cultivated in the area of Serang, Banten.

## REFERENCES

[1] McHugh DJ. (2003) A Guide to the Seaweed Industry. Food and Agriculture Organization of the United Nations.

[2] Hayashi L, Oliveira EC, Bleicher-Lhonneur G, Boulenguer P, Pereira RTL, von Seckendorff R, Shimoda VT, Leflamand A, Vallée P, Critchley AT. (2007) The effects of selected cultivation conditions on the carrageenan characteristics of Kappaphycus alvarezii (Rhodophyta, Solieriaceae) in Ubatuba Bay, São Paulo, Brazil. Journal of Applied Phycology, 19: 505-511.

[3] Masarin F, Cedeno FR, Chavez EG, de Oliveira LE, Gelli VC, Monti R. (2016) Chemical analysis and biorefinery of red algae Kappaphycus alvarezii for efficient production of glucose from residue of carrageenan extraction process. Biotechnology for biofuels, 9:122.
[4] Campo VL, Kawano DF, Silva DBd, Carvalho I. (2009) Carrageenans: Biological properties, chemical modifications and structural analysis – A review. Carbohydrate Polymers, 77(2): 167-180.
[5] Azanza-Corrales R, Sa-a P. (1990) The farmed Eucheuma species (Gigartinales, Rhodophyta) in Danajon Reef, Philippines: carrageenan properties. Hydrobiologia, 204(1): 521-525.
[6] Luxton DM. (1993) Aspects of the farming and processing of Kappaphycus and Eucheuma in Indonesia. Hydrobiologia, 260(1): 365-371.
[7] WWF-Indonesia TP. (2014) Better Management Practices (BMP)- Budidaya Rumput Laut Jenis Kotoni (Kappaphycus alvarezii), Sacol (Kappaphycus striatum), dan Spinosum (Eucheuma denticulatum). Indonesia: WWF-Indonesia.
[8] Lignell Å, Pedersén M. (1989) Agar Composition as a Function of Morphology and Growth Rate. Studies on Some Morphological Strains of Gracilaria secundata and Gracilaria verrucosa (Rhodophyta). In: Botanica Marina; pp 219.
[9] Minghou J. (1990) Processing and extraction of phycocolloids. In: Report on the regional workshop on the culture and utilization of seaweeds (Vol. II) Bangkok, Thiland: NACA.
[10] Thrimawithana TR, Young S, Dunstan DE, Alany RG. (2010) Texture and rheological characterization of kappa and iota carrageenan in the presence of counter ions. Carbohydrate Polymers, 82(1): 69-77.
[11] Munoz J, Freile-Pelegrin Y, Robledo D. (2004) Mariculture of Kappaphycus alvarezii (Rhodophyta, Solieriaceae) color strains in tropical waters of Yucatan, Mexico Aquaculture, 239: 161-177.
[12] Hurtado-Ponce AQ. (1995) Carrageenan Properties and Proximate Composition of Three Morphotypes of Kappaphycus alvarezii Doty (Gigartinales, Rhodophyta) Grown at Two Depths. In: Botanica Marina, pp 215.
[13] Valderrama D, Cai J, Hishamunda N, Ridler N. (2013) Social and economic dimensions of carrageenan seaweed farming. In: Rome: Food and Agricultural Organizarion (FAO) Fisheries and Aquaculture Technical Paper, pp 217.
[14] Sunarto. (2008) Karakteristik biologi dan peranan plankton bagi ekosistem laut. Bandung: Universitas Padjajaran.
[15] Bono A, Anisuzzaman SM, Ding OW. (2014) Effect of process conditions on the gel viscosity and gel strength of semi-refined carrageenan (SRC) produced from seaweed (Kappaphycus alvarezii). Journal of King Saud University - Engineering Sciences, 26(1): 3-9.
[16] FAO. (2003) A Guide to The Seaweed Industry. Rome: FAO Fisheries Technical Paper.
[17] FAO. (2007) Carrageenan. FAO JECFA Monographs 4 2007; Prepared at the 68th JECFA.
[18] Basmal J, Ikasari D. (2014) Production of Semi Refine Carrageenan (SRC) from Fresh Kappaphycus alvarezii using Modified Technique with Minimum Use of Fuel Squalen Bulletin of Marine & Fisheries Postharvest & Biotechnology, 9(1): 17-24.
[19] Freile-Pelegrín Y, Robledo D. (2009) Carrageenan of Eucheuma isiforme (Solieriaceae, Rhodophyta) from Nicaragua. In: Nineteenth International Seaweed Symposium: Proceedings of the 19th International Seaweed Symposium, held in Kobe, Japan, 26-31 March, 2007. Edited by Borowitzka MA, Critchley AT, Kraan S *et al.* Dordrecht: Springer Netherlands, pp 87-91.
[20] Glenn EP, Doty MS. (2001) Growth of the seaweeds Kappaphycus alvarezii, K. striatum and Eucheuma denticulatum as affected by environment in Hawaii. Aquaculture, 84(3):245-255.
[21] Hurtado AQ, Agbayani RF, Sanares R, de Castro-Mallare MTR. (2001) The seasonality and economic feasibility of cultivating Kappaphycus alvarezii in Panagatan Cays, Caluya, Antique, Philippines. Aquaculture, 199(3): 295-310.
[22] de Paula EJ, Pereira RTL, Ohno M. (1999) Strain selection in Kappaphycus alvarezii var. alvarezii (Solieriaceae, Rhodophyta) using tetraspore progeny. Journal of Applied Phycology, 11(1): 111-121.

[23] Mendoza Wilson G, Ganzon-Fortes Edna T, Villanueva Ronald D, Romero Jumelita B, Montaño Marco Nemesio E. (2006) Tissue age as a factor affecting carrageenan quantity and quality in farmed Kappaphycus striatum (Schmitz) Doty ex Silva. In Botanica Marina, pp 57.

[24] Mulyani S, Tuwo A, Syamsudin R, Jompa J. (2018) Effect of seaweed *Kappaphycus alvarezii* aquaculture on growth and survival of coral *Acropora muricata*. AACL Bioflux, 11(6): 1792-1799.

[25] Kamlasi Y. (2008) Kajian Kajian Ekologis dan Biologi Untuk Pengembangan Budidaya Rumput Laut (*Euchuema Cottonii*) di Kecamatan Kupang Barat Kabupaten Kupang Nusa Tenggara Timur (Tesis). Sekolah Pascasarjana Institut Pertanian Bogor.

[26] Tirtawijaya G, Meinita MDN, Marhaeni B, Hazque MN, Moon IS, Hong YK. (2018) Neutrophic activity of the carrageenophyte *Kapphycus alvarezii* cultivated at different growth period in various areas of Indonesia. Evidence-Based Complementary and Alternative Medicine, 1-7.

[27] Kasim M, Mustafa A. (2017) Comparison growth of Kappaphycus alvarezii (*Rhodophyta, Solieriaceae*) cultivation in floating cage and longline in Indonesia. Aquaculture Reports, 6: 49-55.

[28] Kasim M, Mustafa A, Munier T. (2016) The growth rate of seaweed (*Eucheuma denticulatum*) cultivated in longline and floating cage. AACL Bioflux, 9(2): 291-299.

https://doi.org/10.31436/iiumej.v20i2.1109

# NUTRIENTS AND CHEMICAL OXYGEN DEMAND (COD) REMOVALS BY MICROALGAE – BACTERIA CO-CULTURE SYSTEM IN PALM OIL MILL EFFLUENT (POME)

**AMIRAH SAMSUDIN, AZLIN SUHAIDA AZMI, MOHD NAZRI MOHD NAWI, AMANATUZZAKIAH ABDUL HALIM***

*Department of Biotechnology Engineering, Kulliyyah of Engineering, International Islamic University Malaysia. P.O. Box 10, 50728 Kuala Lumpur, Malaysia,*

*Corresponding author: amana@iium.edu.my*

*(Received: 3rd April 2019; Accepted: 11th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** In Malaysia, large amounts of waste known as palm oil mill effluent (POME) are generated during the production process of crude palm oil. Conventionally, POME is treated using biological treatment that involves two processes; aerobic and anaerobic. These processes however, require long hydraulic retention time and produce methane and carbon dioxide ($CO_2$) that can cause environmental problems. Alternatively, POME can be treated by a combination of microalgae and bacterial co-culture that requires a shorter treatment time and is environmentally friendly. In this study, a microalgae strain, *Chlorella vulgaris* was co-cultured with a bacteria strain *Azospirillum brasilense* in POME with an initial concentration of $1.9\times10^6$ cells/mL and $10^4$ CFU/mL, respectively. The removal of chemical oxygen demand (COD) and nutrients (phosphorus and ammonium) were analyzed using Standard Methods, APHA 1999. The effectiveness of the co-culture system in POME treatment under agitation and aeration conditions for nutrients and COD removals were studied. Results show that the removal of ammonium by microalgae was much higher under the aeration condition (73.5%) compared to that of the agitation condition (34.4%) in POME. Moreover, co-culture system exhibits better removal of ammonium, phosphorus, and COD (84%, 87.3% and 51.8%, respectively) compared to that in an axenic microalgae system (67%, 84.2% and 41.1%, respectively). The kinetic studies on the co-culture system and the nutrients removal were also conducted. The kinetic coefficients of maximum specific growth rate ($\mu_{max}$) and half-saturation coefficient ($k_s$) obtained from Lineweaver-Burk plot were 0.192 $d^{-1}$ and 27.32 mg/L, respectively. Based on the findings obtained, the co-culture system could be implemented as an efficient and inexpensive alternative method for POME treatment.

***ABSTRACT****:* Di Malaysia, banyak bahan buangan kilang minyak kelapa sawit yang dikenali sebagai (POME) telah terhasil ketika proses penghasilan minyak kelapa sawit mentah. Dahulu, POME dirawat menggunakan rawatan biologi yang terdiri daripada dua proses; erob dan anaerob. Walau bagaimanapun, proses-proses ini memerlukan masa yang panjang bagi pengekalan hidraulik dan gas metana dan karbon dioksida ($CO_2$) telah terhasil yang menyebabkan masalah alam. Sebagai alternatif, POME dapat dirawat dengan kombinasi mikroalga dan sistem bakteria ko-kultur melalui masa rawatan yang lebih pendek dan mesra alam. Kajian ini, strain mikroalga, *Chlorella vulgaris* telah di ko-kultur dengan strain bakteria *Azospirillum brasilense* dalam POME dengan ketumpatan awal $1.9\times10^6$ sel/mL dan $10^4$ CFU/mL masing-masing. Penyingkiran kehendak oksigen secara kimia (COD) dan nutrisi (fosferus dan ammonium) telah dikaji menggunakan Kaedah Biasa, APHA 1999. Keberkesanan system ko-kultur dalam rawatan

POME di bawah keadaan kisaran dan pengudaraan bagi nutrisi dan penyingkiran COD telah diselidiki. Dapatan kajian menunjukkan penyingkiran ammonium menggunakan mikroalga lebih banyak melalui keadaan pengudaraan (73.5%) berbanding keadaan kisaran (34.4%) dalam POME. Tambahan, system ko-kultur menunjukkan lebih bagus dalam penyingkiran ammonium, fosferus dan COD (84%, 87.3% dan 51.8%, masing-masing) dibandingkan dengan sistem mikroalga aksenik (67%, 84.2% dan 41.1%, masing-masing). Kajian kinetik pada sistem ko-kultur dan penyingkiran nutrisi turut dijalankan. Pekali kinetik kadar maksimum pertumbuhan sebenar ($\mu_{max}$) dan pekali separuh-penepuan ($k_s$) telah diperolehi melalui plot Lineweaver-Burk iaitu 0.192 $d^{-1}$ dan 27.32 mg/L, masing-masing. Berdasarkan penemuan ini, sistem ko-kultur boleh dijalankan dengan cekap dan murah sebagai pilihan alternatif kepada rawatan POME.

***KEYWORDS:*** *co-culture; microalgae; palm oil mill effluent; symbiotic; wastewater treatment*

# 1. INTRODUCTION

Malaysia is one of the largest producers and exporters of palm oil in the world. During the palm oil production process, huge amounts of wastewater, known as palm oil mill effluent (POME) consisting of various pollutants and nutrients, are generated [1]. POME contains a high concentration of phosphorus, organic nitrogen, and other nutrients. Moreover, the biochemical oxygen demand (BOD) and chemical oxygen demand (COD) content in POME are also high, reaching 100 times higher than municipal sewage [2]. Hence, POME treatment is very crucial prior to discharge into natural water resources, such as rivers, which can cause water pollution and negatively affect aquatic life.

POME has been treated using conventional methods (e.g. anaerobic and aerobic digestion). However, these methods produce $CO_2$ and methane which can lead to enviromental problems. Furthermore, these conventional methods require long hydraulic retention time (HRT) and can take at least 20 days for the aerobic digestion process. As an alternative for POME treatment, utilization of microalgae has been introduced due to its ability to grow in robust environments and still remove various pollutants in wastewater such as phosphorus and nitrogen. It was also reported that bacteria known as microalgae growth promoting bacteria (MGPB) can be co-cultured with microalgae to improve the removal of nutrients [3-6]. Moreover, this co-culture system shares a symbiotic relationship. The oxygen ($O_2$) produced by microalgae during photosynthesis will be used by heterotrophic bacteria and the bacteria in return release $CO_2$ that can be consumed by microalgae for growth [7].

The aim of this study is to evaluate the efficiency of microalgae and bacteria species, *Chlorella vulgaris* and *Azospirillum brasilense* for nutrients (ammonium and phosphorus) and COD removals in partially treated POME under different cultivation conditions to replace the aerobic digestion process in the conventional POME treatment. The kinetics studies of the co-culture system were also determined.

# 2. METHODOLOGY

## 2.1 Raw Materials and Inoculum Preparation

A species of microalgae (*Chlorella vulgaris)* and a bacteria strain *(Azospirillum brasilense)* were used in this study. The microalgal strain and bacteria were purchased from Universiti Malaya, Kuala Lumpur and Leibniz-Institute DSMZ, Germany respectively. Microalgae was grown in Tris-Acetate-Phosphate (TAP) medium at 25 $^{o}$C and 150 rpm in a rotary shaker and bacteria was maintained in nutrient media at 30±2 $^{o}$C and 120 rpm [3]. Partially treated palm oil mill effluent (POME) (sample after anaerobic process) was

https://doi.org/10.31436/iiumej.v20i2.1109

collected from West Mill Sdn Bhd, Sime Darby Research Centre, Carey Island, Malaysia. Partially treated POME was used because this study aims to replace the aerobic process and the amount of nutrients available were acceptable for microalgae to growth. Sample obtained was stored in 5 L plastic containers and kept at 4 $^{o}$C to limit the activity of biodegradation process and prevent any contamination [7]. The ammonium, phosphorus and COD contents in partially treated POME are 162, 18, and 874 mg/L, respectively.

### 2.2 Experimental Set Up

Microalgae and bacteria were cultivated in POME at 10% concentration ($V_{microalgae-bacteria}/V_{POME}$) in a 250 mL flask for 6 days and 25±2 $^{o}$C under 45 – 50 µmol photon $m^{-2}s^{-1}$ of illumination of white fluorescent light (alternate 12 h/12 h of light/dark periods) [7]. The initial concentrations of microalgae and bacteria were $1.9\times10^{6}$ cells/mL and $10^{4}$ CFU/mL respectively [7].

Two different cultivation conditions of microalgae – bacteria for POME treatment, agitation and aeration conditions, were tested. For the agitation condition, POME treatment was incubated in incubator shaker at 25 $^{o}$C and 150 rpm. For the aeration condition, a shake flask containing POME and microbes was connected to an air pump (1.8 L/min) to provide aeration to the treatment system. A sterilized 0.22 µm nylon filter was used to avoid contamination. During cultivation, 2 mL of algae suspension was collected for every 2 days and filtered using 0.45 µm nylon filter before it was analyzed for ammonium, phosphorus and COD using their respective method. The growth of microalgae and bacteria were also recorded.

### 2.3 Analytical Methods

Microalgae cell number was calculated by Neubauer haemocytometer counting and bacterial concentrations were determined as colony forming units (CFU) on nutrient media agar plates. Besides that, the Standard Methods for the Examination of Water and Wastewater were used to analyse ammonium, phosphorus and COD removal [8]. Ammonium content was analysed using phenate method while phosphorus content (ortophosphate) was analysed using ascorbic acid method and COD was analysed using closed reflux, colorimetric method [7].

### 2.4 Kinetics Study

Based on the best growth conditions (agitation or aeration) in terms of COD and nutrient removal, batch kinetics studies were carried out by using approximately $1.9\times10^{6}$ cells/mL and $10^{4}$ CFU/mL of microalgae and bacteria initial concentration, respectively. They were cultivated in partially treated POME at room temperature (25±2 $^{o}$C). POME was diluted to study the kinetics of different N/P ratios. The removals and the growth of microalgae were recorded and analyzed for six days of treatment. Kinetics coefficient, maximum specific growth rate, $\mu_{max}$ ($d^{-1}$) and half-saturation coefficient, $k_s$ (mg/L) were determined using Lineweaver–Burk plot of $1/\mu$ vs. $1/S$.

## 3. RESULTS AND DISCUSSION

### 3.1 POME Treatment in Agitation and Aeration Conditions

The characteristic of the partially treated POME sample collected is tabulated in Table 1.

Table 1: Characteristic of partially treated POME

| Parameters | Concentration (mg/L), except for pH and turbidity | Average Concentration (mg/L) |
|---|---|---|
| pH | 7.5 – 8.5 | 8 |
| COD | 600 – 900 | 750 |
| Ammonium | 160 – 250 | 205 |
| Phosphorus | 10 – 30 | 20 |
| Total Suspended Solid | 1120 – 1160 | 1140 |
| Total Volatile Solid | 880 – 1170 | 1025 |
| Turbidity (NTU) | 2755 – 2785 | 2775 |
| Oil and Grease Content | 118 - 234 | 176 |

In this study, microalgae were grown in non-sterilized partially treated POME under agitation and aeration cultivation conditions. Throughout the treatment, the pH of POME was observed and measured. At day 0, the pH of POME samples was 7.5 and it increased to 8.8 after six days of treatment in both conditions. The percentage of ammonium, phosphorus, and COD removal by agitation and aeration conditions in non-autoclaved partially treated POME are illustrated in Fig. 1.

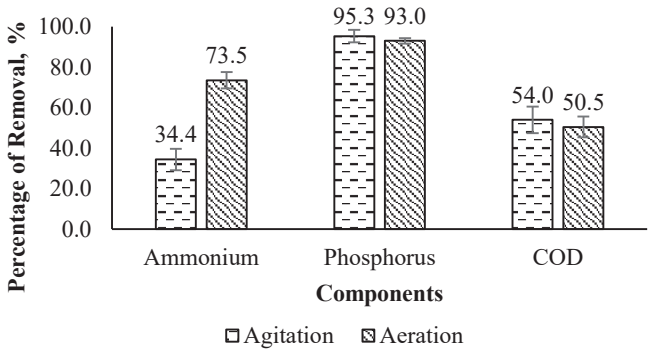



Fig. 1: Percentage of ammonium, phosphorus, and COD removal in agitation and aeration conditions by axenic microalgae in a 250 mL flask containing 150 mL non-sterilized partially treated POME with 10% concentration ($V_{microalgae\ –\ bacteria}$ /$V_{POME}$) for 6 days at 25±2 $^{o}$C and under 45 – 50 µmol photon $m^{-2}s^{-1}$ of illumination of white fluorescent light (alternate 12 h/12 h periods of light/dark). Initial concentration of Ammonium = 162 mg/L, Phosphorus = 18 mg/L and COD = 874 mg/L. Error bars represent one standard deviation about the mean ($n$ = 3).

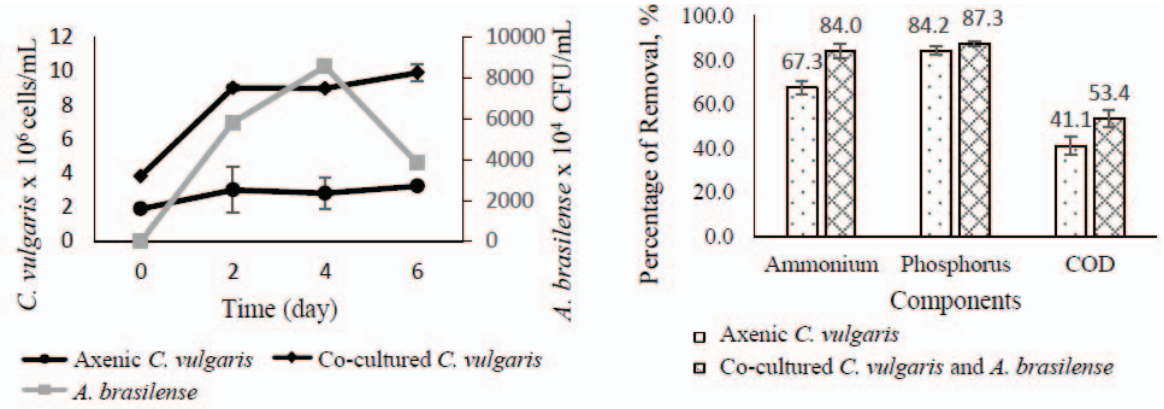
Based on Fig. 1, the removal of ammonium in aeration condition is 73.5% which is higher compared to that in agitation condition which is 34.4%. The ammonium removal percentage was in line with the growth rate of *C. vulgaris* which were 0.356 $d^{-1}$ and 0.176 $d^{-1}$ for aeration and agitation system respectively. Similar results were obtained whereby the microalgae cultivated with air supply (aerated) recorded higher growth rate compared to the microalgae grown in mixing shaker (non-aerated system) [9]. The optimum turbulence effect created by aeration system promotes better cell growth which consequently helps in the removal of the nutrients. Moreover, under equal light intensity, air flux is produced as a result of the turbulence and creates different movement regimes of the cells within the media. Adequate turbulence enhances mass transfer of nutrients to the microalgae cell and

promotes gas exchange that overcomes the photo-oxidative effect as well as minimizing thick boundary layer between microalgae cell and the surrounding condition of unstirred suspension. In addition to that, sufficient system turbulence enhances the microalgae reproduction compared to the non-aerated medium, which can cause biomass accumulation [9].

Although the removals of phosphorus and COD in the agitation system were higher compared to those in the aeration system, the difference was very small, i.e. less than 5%. Results indicated that the aeration system offered better nutrient removal especially for ammonium removal. Therefore, the aeration system was chosen to be used in the subsequent experiment.

### 3.2 Axenic and Co-culture Systems in Aeration Condition

A co-culture system (*C. vulgaris* and *A. brasilense*) was utilized in the partially treated POME treatment. In this case, sterilized partially treated POME was used with initial concentration of 874 mg/L, 162 mg/L and 18 mg/L of COD ammonium and phosphorus, respectively. The microalgae – bacteria growth conditions were at 0.77 g/L ($1.9\times10^6$ cells/mL), $10^4$ CFU/mL and day 1.52 for microalgae initial concentration, bacteria initial concentration and bacteria inoculation time, respectively. POME treatment was conducted in the aeration condition. The growth profiles of microalgae – bacteria and percentage of nutrients and COD removals in axenic and microalgae – bacteria co-culture systems are shown in Fig. 2 (a) and 2 (b), respectively.




(a) (b)

Fig. 2: (a) Growth profiles of bacteria (*A. brasilense*), microalgae (*C. vulgaris*) in axenic system and microalgae (*C. vulgaris* and *A. brasilense*) in co-culture system (b) Percentage of ammonium, phosphorus and COD removals in aeration condition by axenic microalgae and co-culture microalgae – bacteria in a 250 ml flask containing 150 mL sterilized partially treated POME with 10% concentration ($V_{microalgae-bacteria}/V_{POME}$) for 6 days at 25±2 $^{o}$C, under 45 – 50 µmol photon $m^{-2}s^{-1}$ of illumination of white fluorescent light (alternate 12 h/12 h periods of light/dark) and aerated. Initial concentration of Ammonium = 162 mg/L, Phosphorus = 18 mg/L and COD = 874 mg/L. Error bars represent one standard deviation about the mean ($n$ = 3).

Figure 2 (a) shows that, co-culture of microalgae and bacteria enhanced the microalgae population. The specific growth rate of co-cultured microalgae was higher (0.570 $d^{-1}$), compared to that of axenic microalgae (0.228 $d^{-1}$). Similar results were obtained by a study that reported that the growth of *C. vulgaris* was significantly enhanced when co-

immobilized with *A. brasilense* bacteria under aeration conditions (1.8 L/min) [3]. It was also reported that some bacteria species enhance microalgae cells growth [10]. However microalgae and bacteria have a complex relationship. It depends on the species of microlagae and bacteria as well as the growth conditions themselves [5,11]. Mutual symbiosis between microalgae and bacteria can be achieved if the species and conditions were optimum for both microorganisms. The results showed that bacteria promoted the growth of microalgae when the initial concentration was suitable ($10^4$ CFU/mL) and cultivated under aeration conditions. A study reported that, at low bacteria concentration ($10^2$ CFU/mL), the effect of bacteria in the co-culture system was not significant towards the removal of the nutrients. On the other hand, low nutrient removal was also recorded at maximum bacteria concentration ($10^6$ CFU/mL) due to the competition of the nutrients between the bacteria and microalgae. At these conditions, the bacteria growth rate was higher than the microalgae growth rate, hence the utilization of nutrients was dominated by the bacteria cells [7].

In addition, *A. brasilense* bacterium was able to grow well with microalgae and partially treated POME with specific growth rates of 2.264 $d^{-1}$. Besides getting oxygen from the microalgae's photosynthesis process, the aeration process helps to promote gaseous exchange and enhances mass transfer of nutrients in the system. Moreover, excess oxygen from microalgae photosynthesis can be released from the media in the aeration system [10].

Based on Fig. 2 (b), the co-culture system has significantly increased the removals of COD and nutrients after six days of treatment, especially for ammonium and COD removal. The presence of bacteria delivers positive effect to the system. Higher percentage removals of ammonium, phosphorus and COD were obtained for co-cultured *C. vulgaris* system (84%, 87.3% and 53.4%, respectively) compared to that from axenic *C. vulgaris* system (67.3%, 84.2%, and 41.1%, respectively). These results show better removals of COD and nutrients compared to another study that reported that microalgae from *Chlamydomonas incerta* species removed 67.35% of COD with initial concentration of 250 mg/L in POME in 28 days [1]. Ammonium removal in this study was sligthly lower (84%) compared to the previous study, which used *Spirulina plantesis* to achieve 96.5% removal from approximately 296 mg/L of initial concentration [2]. Comparison of DOE standard and results of removals achieved by this study is presented in Table 2.

Table 2: DOE standard B and final results of removals achieved by this study

| | **DOE Standard B** | **This Study** | | |
|---|---|---|---|---|
| | | *Initial Concentration* | *Final Concentration* | |
| | | *Axenic and Co-cultured System* | *Axenic System* | *Co-cultured System* |
| COD (mg/L) | 200 | 874 | 514.8 ± 37.32 | 407.3 ± 32.95 |
| Ammonium (mg/L) | 20 | 162 | 53 ± 4.97 | 25.9 ± 5.3 |
| Phosphorus (mg/L) | 10* | 18 | 2.84 ± 0.39 | 2.29 ± 0.15 |

* DOE Standard B for sewage discharge into enclosed water body

Results obtained in this study showed that, microalgae – bacteria co-cultured system exhibits a good potential to be applied in POME treatment within six days of cultivation compared to aerobic digestion process that required at least 20 days of treatment with final concentration 200, 52 and 12 mg/L of COD, nitrogen and phosphorus respectively [1].The co-cultured system also proves that, it enhances the COD and nutrient removal in the

treatment compared to the axenic system. However, further investigation has to be conducted to improve COD removal in order to achieve the standard set by DOE, Malaysia, prior to discharge treated POME into the river. Heterotrophic bacteria is one of the microorganisms that have an important role in the reduction of COD content by breaking down the organic matter in the wastewater. Isolation and growth characterization of heterotrophic bacteria in partially treated POME sample could be investigated further in order to find more efficient bacteria to improve COD removals in POME.

### 3.3 Kinetics Studies of Microalgae Growth and Nutrients Removal

There are two important nutrients for microalgae growth: phosphorus (P) and ammonium (N). Kinetics constants were determined to study the rate of nutrients uptake by the microalgae. In this study, nutrients removal rates were calculated by Eq. (1) [12].

$$R = \frac{C_0 - C_i}{t_i - t_0} \quad (1)$$

where $R$ represents the removal rate of nutrients (N or P), $C_0$ and $C_i$ represent initial concentration of nutrients and nutrient concentration at time $t_i$ respectively and $t_i$ is the time when there is no significant change of nutrient concentration. The specific nutrient removal rate, $R_i$ was determined by dividing $R$ by the initial cell density (*C. vulgaris*). In these kinetic studies, two POME samples with different N/P ratio were utilized. The N/P ratio of media used for POME 1 and POME 2 were 8.96 and 2.27 respectively. The specific nutrient removal rate, $R_i$ of ammonium and phosphorus for both samples are shown in Table 3.

Table 3: Specific nutrient removal rate of ammonium and phosphorus in POME 1 and POME 2

| Samples | N/P Ratio | $R_N$ (mg $g^{-1}d^{-1}$) | $R_P$ (mg $g^{-1}d^{-1}$) |
|---|---|---|---|
| POME 1 | 8.96 | 29.42 | 3.41 |
| POME 2 | 2.27 | 4.64 | 1.18 |

The specific ammonium removal rates, $R_N$ for *C. vulgaris* grown in POME 1 and POME 2 were 29.42 and 4.64 mg $g^{-1}d^{-1}$ with initial ammonium concentration (POME 1 = 167 mg/L, POME 2 = 22.6 mg/L). These results showed that high $R_N$ was achieved when initial concentration of ammonium introduced was high. These results were in agreement with a previous study whereby the specific N removal rates of *Chlorella* sp. grown in media with 9 and 56 N/P were 1.8 and 3.8 mg $mg^{-1}$ Chl a $d^{-1}$ respectively [12]. Furthermore, the other study observed that the specific N removal rates increased from 0.5 to 3 mg $mg^{-1}$ Chl a $d^{-1}$ when the initial concentration of $NH_4^+$-N increased from 13.2 mg/L to 410 mg/L in synthetic wastewater [12].

In this study, a similar pattern of results was obtained for specific phosphorus removal rates, $R_P$. The $R_P$ increased from 1.18 to 3.41 mg $g^{-1}d^{-1}$ with the increase of phosphorus initial concentration from 9.96 to 18.06 mg/L. Slightly higher P concentration in the primary effluent also led to higher $R_P$. Chlorella *sp.* obtained 0.25 and 0.15 mg $mg^{-1}$Chl a $d^{-1}$ for the diluted sludge centrate and the primary effluent with initial P, 4.0 and 3.5 mg/L respectively [12].

First-order and second-order reactions were tested for both nutrients removal to find the reaction constant. The reaction constant for first-order and second-order was calculated by Eq (2) and (3), respectively [13].

https://doi.org/10.31436/iiumej.v20i2.1109

$$-ln\frac{C_N}{C_{NO}} = kt \quad (2)$$

$$-ln\frac{C_P C_{NO}}{C_N C_{PO}} = (C_{PO} - C_{NO})kt \quad (3)$$

where $C_N$ or $C_P$ and $C_{NO}$ or $C_{PO}$ represent the concentration of ammonium or phosphorus at time $t$ and $t_o$, respectively and $k$ is the reaction constant. Graphs were plotted for both order reactions and in this study, the removal of phosphorus and ammonium fitted most to the first-order reaction since $R^2$ in first-order reaction for both POME were higher than $R^2$ in second-order reaction, as shown in Table 4. This result was in line with a study that reported that the phosphorus and nitrogen removal in their study fitted the first-order reaction very well [12]. Based on the nutrient consumption, the yield of biomass was determined by Eq. (4), [11].

$$Y = \frac{X_i - X_0}{C_0 - C_i} \quad (4)$$

where $Y$ represents the yield of biomass linked to the consumption of nutrients (phosphorus, P or ammonium, N), $C_0$ is the initial nutrient concentration, $C_i$ is the concentration of nutrients when there was no significant decrease of nutrients, $X_i$ represents *C. vulgaris* cell density with respect to the nutrient $C_i$ and $X_0$ is the initial *C. vulgaris* cell density [11]. The yield of biomass based on specific nutrient consumption and first order reaction rate constants are shown in Table 5.

Table 4: Value of $R^2$ for first and second order reaction

| Sample | First-Order | | Second-Order |
|---|---|---|---|
| | $R^2$, $_P$ | $R^2$, $_N$ | $R^2$, $_{NP}$ |
| **POME 1** | 0.9305 | 0.9565 | 0.9231 |
| **POME 2** | 0.8741 | 0.9127 | 0.8052 |

Table 5: Kinetics parameters of nutrients removal for *C. vulgaris* in POME 1 and POME 2

| Sample | $k_P$ | $k_N$ | $Y_N$ | $Y_P$ | $Y_P/Y_N$ |
|---|---|---|---|---|---|
| **POME 1** | 0.3685 | 0.3272 | 0.01406 | 0.1213 | 8.6231 |
| **POME 2** | 0.1271 | 0.1383 | 0.1751 | 0.6886 | 3.9327 |

The biomass yield based on N consumption, $Y_N$, for *C. vulgaris* cultivated in POME 1 was lower than the yield in POME 2 which were 0.01406 g $mg^{-1}$ N and 0.1751 g $mg^{-1}$ N respectively. A similar result was obtained for $Y_P$, 0.1213 g $mg^{-1}$ P and 0.6886 g $mg^{-1}$ P for *C. vulgaris* grown in POME 1 and POME 2, respectively. The ratio of $Y_P/Y_N$, indicates the amount of the ratio of N/P required (mg N consumed/mg P consumed) to produce the same unit amount of biomass. In this study, N/P ratio of POME 1 was almost double the N/P ratio of POME 2. Theoretically, the chemical formula of microalgae is $C_{106}H_{181}O_{45}N_{16}P$ where N/P ratio is 7.2 [14]. The $Y_P/Y_N$ for *C. vulgaris* cultivated in POME 1 was 8.62, which is close to this theoretical value.

In this study, the final N and P concentration in POME 1 were 25.9 and 2.3 mg/L respectively. This case was similar to the result obtained by a study that stated that P was the limiting nutrient for microalgae grown in POME 1. Thus, the kinetic coefficients of

microalgae growing in POME 1 were determined using Lineweaver-Burk equation (Eq. (5)), where $\mu$ is specific growth rate ($d^{-1}$), $\mu_{max}$ is the maximum specific growth rate ($d^{-1}$), S is the concentration of limiting nutrient (mg/L) and $k_s$ is half-saturation coefficient (mg/L). Experimental data were plotted as illustrated in Fig. 3 [11].

$$\frac{1}{\mu} = \frac{k_S}{\mu_{max}} \frac{1}{S} + \frac{1}{\mu_{max}} \quad (5)$$

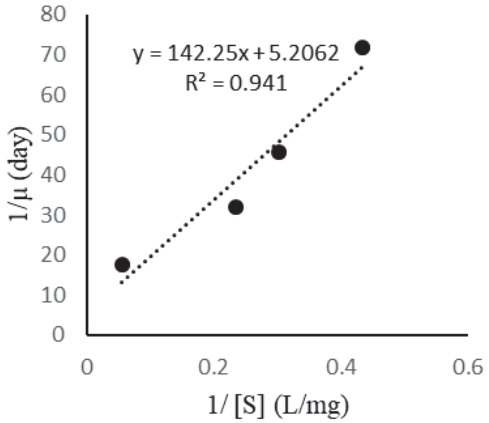



Fig. 3: Lineweaver-Burk plot used to estimate kinetic coefficient.

Based on the linear fit line for this plot (Fig. 3), the kinetic coefficients $\mu_{max}$ and $k_s$ of P removal by *C. vulgaris* were determined to be $\mu_{max}$ = 0.192 $d^{-1}$ and $k_s$ = 27.32 mg/L ($R^2$ = 0.973).

## 4. CONCLUSION

As a conclusion, higher removal of ammonium from POME was obtained under aeration conditions (73.5%) compared to that of agitation conditions (34.4%). In addition, the microalgae – bacteria co-culture system showed higher removal of ammonium, phosphorus, and COD in POME (84%, 87.3% and 51.8%, respectively) compared to that in the axenic microalgae system (67%, 84.2% and 41.1%, respectively). The kinetic coefficients obtained were $\mu_{max}$ = 0.192$d^{-1}$ and $k_s$ = 27.32 mg/L. The obtained data showed that this co-culture system has the potential to improve the current aerobic digestion process in the wastewater treatment. The long HRT required by the aerobic treatment can be shortened by replacing or integrating with this co-culture system. Moreover, this system also promotes green technology and can be considered for future commercial use. However, for COD removal, further investigation on POME treatment is required in order to achieve the standard set by DOE, Malaysia, prior to discharge into the river.

## ACKNOWLEDGMENT

This research was supported by Research Initiative Grant Scheme (RIGS15-144-0144) from the International Islamic University Malaysia and Ministry of Higher Education, Malaysia.

## REFERENCES

[1] Kamyab H, Fadhil M, Keyvanfar A, Abd MZ. (2015) Efficiency of Microalgae Chlamydomonas on the Removal of Pollutants from Palm Oil Mill Effluent (POME). Energy Procedia, 75: 2400-2408.

[2] Hadiyanto, Marcelinus C, Danny S. (2013) Phytoremediation of Palm Oil Mill Effluent

(POME) by Using Aquatic Plants and Microalgae for Biomass Production, J. Environ. Sci. Technol., 6(2): 79-90.

[3] De-Bashan LE, Hernandez JP, Morey T, Bashan Y. (2004) Microalgae growth-promoting bacteria as 'helpers' for microalgae: A novel approach for removing ammonium and phosphorus from municipal wastewater. Water Res., 38(2): 466-474.

[4] Hernandez J, Luz E, Bashan Y. (2006) Starvation enhances phosphorus removal from wastewater by the with *Azospirillum brasilense*. Enzyme Microb. Technol, 38: 190-198.

[5] Mujtaba G, Rizwan M, Lee K. (2015) Simultaneous removal of inorganic nutrients and organic carbon by symbiotic co-culture of *Chlorella vulgaris* and *Pseudomonas putida.* Biotechnol. Bioprocess Eng., 20(6): 1114-1122.

[6] Delgadillo-mirquez L, Lopes F, Taidi B, Pareau D. (2016) Nitrogen and phosphate removal from wastewater with a mixed microalgae and bacteria culture. Biotechnol. Reports, 11: 18–26.

[7] Samsudin A, Azmi AS, Mohd Nawi MN, Abdul Halim A. (2018) Wastewater treatment by microalgae-bacteria co-culture system. Malays. J. Microbiol., 14(2): 131-136.

[8] APHA. (1999) Standard Methods for the Examination of Water and Wastewater Part 4000 Inorganic Nonmetallic Constituents Standard Methods for the Examination of Water and Wastewater.

[9] Gonzalez LE, Bashan Y. (2000) Increased growth of the microalga Chlorella vulgaris when coimmobilized and cocultured in alginate beads with the plant-growth-promoting bacterium *Azospirillum brasilense.* Appl. Environ. Microbiol., 66(4): 1527-1531.

[10] Hariz HB, Takriff MS. (2017) Growth and Biomass Production of Native Microalgae *Chlorella* sp ., *Chlamydomonas* sp . and *Scenedesmus* sp . Cultivated in Palm Oil Mill Effluent (POME) at Different Cultivation Conditions. Transactions on Science and Technology, 4(3): 298-311.

[11] Wang M, Kuo-Dahab WC, Dolan S, Park C. (2014) Kinetics of nutrient removal and expression of extracellular polymeric substances of the microalgae , *Chlorella* sp. in wastewater treatment. Bioresour. Technol., 154: 131-137.

[12] Aslan S , Kapdan IK. (2006) Batch kinetics of nitrogen and phosphorus removal from synthetic wastewater by algae. J. Ecol. Eng., 8: 64-70.

[13] Levenspiel, O. (1999). Interpretation of Batch Reactor Data. (3rd ed.) In : Chemical Reaction Engineering (pp.38-63). New York, USA: John Wiley & Sons.

[14] Grobbelaar, JU. (2007). Algal Nutrition - Mineral Nutrition, Amos, R. (1st ed.) In : Handbook of Microalgae Culture (pp. 95-115). New Jersey, USA: Blackwell Publishing Ltd.

https://doi.org/10.31436/iiumej.v20i2.996

# A REVIEW OF ROAD EMBANKMENT STABILITY ON SOFT GROUND: PROBLEMS AND FUTURE PERSPECTIVE

**RUFAIZAL CHE MAMAT[1,3*], ANUAR KASA[1], SITI FATIN MOHD RAZALI[2]**

*[1]Centre of Engineering and Built Environment Education Research (PeKA)*
*[2]Smart and Sustainable Township Research Centre (SUTRA)*
*Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.*
*[3]Department of Civil Engineering, Politeknik Ungku Omar, Jalan Raja Musa Mahadi, 31400 Ipoh, Perak, Malaysia.*

*[*]Corresponding author: rufaizal.cm@gmail.com*

*(Received: 4th December 2018; Accepted: 30th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** This paper presents an exhaustive review of the challenges faced in the construction of road embankments on soft ground and proposes a direction for future development. Frequently used techniques for soft ground improvement are discussed. The factors that contribute to the stability of the road embankment are reviewed by approach, results of past studies, and historical cases. The findings show that settlement, slope stability, and soil bearing capacity are all challenges to constructing the road embankment. Additionally, it is found that geometric data is a key factor in embankment design. Pre-loading with prefabricated vertical drain (PVDs) methods and lightweight fill were found to be widely used techniques in soft ground improvement. The information from this study can be used to develop design guidance systems, numerical modelling, and to give an overview and knowledge to other researchers who are or will conduct research in this field. Finally, future perspectives for research are related to predictions of factors that affect the stability of road embankment with an artificial intelligence approach.

***ABSTRAK:*** Kertas ini membentangkan ulasan kajian menyeluruh mengenai cabaran yang dihadapi dalam pembinaan benteng jalanraya di atas tanah lembut dan mencadangkan ke arah pembangunan kajian masa depan. Teknik-teknik penambahbaikan tanah lembut yang sering digunakan turut dibincangkan. Faktor- faktor yang menyumbang kepada kestabilan benteng jalanraya diulas dengan pendekatan kepada kajian lepas dan sejarah kes. Hasil kajian ini didapati bahawa enapan, kestabilan cerun dan keupayaan galas tanah merupakan cabaran dalam pembinaan benteng jalanraya. Selain itu, ia didapati bahawa data geometri merupakan faktor penting kepada rekabentuk benteng. Kaedah pra pembebanan dengan prefabrikasi saliran menegak (PVDs) dan isian ringan didapati teknik yang popular digunakan dalam pembaikkan tanah lembut masa kini. Maklumat dari kajian ini boleh digunakan untuk membangunkan sistem panduan reka bentuk, pemodelan berangka serta memberi gambaran dan ilmu kepada penyelidik lain yang sedang atau akan menjalankan kajian dalam bidang ini. Akhir sekali, perspektif masa depan untuk penyelidikan berkaitan ramalan faktor-faktor yang mempengaruhi kestabilan embankment jalanraya dengan pendekatan kepintaran buatan.

***KEYWORDS:*** *road embankment; soft grounds; stability factors; ground improvement*

# 1. INTRODUCTION

In road construction, the embankment is used to increase the height of the road compared to the height of the surrounding area. It is a large earth structure and is often used in civil engineering applications related to infrastructure projects. The construction involves two essential construction components namely fill and foundation [1, 2]. These two components act as definitive parameters to analyse the embankment to find failure slope. Hence, a deep understanding and interpretation of the behaviours of these components such as deformation mechanisms can produce the collapse mechanism of the embankment.

Embankment fill is defined as selected fill placed and compacted with proper specifications so that it can display the required engineering performance [3]. The characteristics of fill material, the degree of compaction and the factors affecting the workability are considered for best results in the construction of embankment fill. Common fill materials are soil [4], natural aggregates [5] and lightweight filling material [6]. These fill materials must possess the required characteristics such as proper drainage to prevent saturation [7], good performance in compaction properties [8], compressibility [9], and shear strength [10]. Even if the embankment fill has the best material properties, instability can occur if it is built on unstable soft ground or soil with properties needing improvement.

The foundation serves as a placement for embankment structure and it may to be stable or unstable. The embankment foundation containing soft grounds is often a challenge for all engineers to ensure stability depending on the nature of the soil. Soft ground with fine particles such as silt, clay, and peat have high moisture content and are located near or below the groundwater level [11]. These soils also have the characteristics of high compressibility, high plasticity, high sensitivity, low shear strength, and low permeability [11,12]. Such features will lead to the problem of high settlement and low bearing capacity during post-construction work. This will reduce the strength of workability and shorten the embankment's lifespan [13]. It will also lead to bumpy road surface and damage, subsequently causing accidents to road users.

Embankment stability is critical in road construction. Hence, before construction, the engineer must provide all geotechnical design reports and include drawings to be submitted to the relevant parties. The report shall state clearly the assumptions, parameter justifications and the methods used in the design to mitigate all issues or concerns. Geometric data is a general requirement in the preparation of design drawings while data materials contain the engineering parameters of fill and foundation of the embankment. Data materials are used in the engineering design of the embankment. The factor of safety (FoS) calculation is used to assess the performance of the embankment to ensure constant stability. The design report contains stability analyses, settlement analyses, the design of porewater pressure and technical recommendations for ground improvement. Embankments built on soft ground require the installation of monitoring instruments on site such as the inclinometer, settlement marker or rod settlement gauge and piezometer. The instrument will be monitored from time to time for the aid in the demonstration of compliance with minimum FoS during construction.

This paper reviews the literature on the challenges of embankment construction on soft ground and the latest techniques that are often used in improving soft ground. To achieve this objective, this paper has been divided into three sections. The first part deals with the factors that affect the embankment stability. The second part is about the factors affecting the design of the embankment, while the third is the latest soft soil improvement technique widely used with considerable aspects of sustainability, cost savings, and construction time. Finally, the main conclusions and future perspectives will be presented.

## 2. THE CHALLENGE OF CONSTRUCTING EMBANKMENTS ON SOFT GROUND

Soft ground has a high groundwater level and typically, it can be found near coastal areas. The soil profile consists of a layer of crust at the top followed by a soft soil layer of either limited or unlimited thickness underneath. This soil is brown from light to dark [11]. In general, soft soil such as clay is dominated by particles of less than 0.002 mm in size and is pliable. It is formed from a combination of clay and mud and with almost 50% clay content [11]. The structure of this soil is identified as being dispersed, flocculated, bookhouse, turbostratic, and cohesive [14].

Soils are characterized by their physical, mechanical, chemical, and biological properties. Soil with poor geotechnical properties, such as soft clay, requires improvement technique application. Yalcin [15] investigated the effect of clay on a landslide in Kanlica Village, Turkey. As a result of laboratory tests, clay was found to contain illite and montmorillonite, which produced low shear strength and high swelling potentials making it prone to landslide problems. This conclusion is also supported by Ohlmacher [16], where the soil comprising mineral content such as illite has a high liquid limit contributing to very low strength parameters. Instability problems such as lateral deformations and excessive settlement often occur in structural and geotechnical construction on soft soils [17]. The load generated by the embankment fill drives the increase in soil stress and pore water pressure leading to deformation mechanisms. This causes the load to be undrained, where the behaviour is similar to the viscous fluid that can lead to critical conditions against structures such as bearing capacity problems. However, a thick hard layer built on the top can absorb the load produced by the fill of embankment.

An example of active embankment failure is the 1 km-long highway corridor located at the Eisenhower Tunnel, Colorado, USA. A survey by the Colorado Transport Department (CDOT) found that frequent slope settlement and failure problems have occurred over the last four decades [18]. The instability of the embankment due to movements of the side ground should be carefully considered during the design process and during engineering projects around the embankment. The road damage incident that occurred in September 2010 at the Port of Tanjung Priok, Jakarta was caused by embankment instability [19]. The embankment was built on clay with high plasticity. The movement of side ground caused the road surface to crack. Road embankment failure also happened in the Northland region of New Zealand [20], where the embankment was built on a soft soil characterized by a low strength with a highly sheared fabric. Other reasons for the failure are settlement and continuous soil movement that led to slope failure. An embankment failure occurred during the construction of a four-lane expressway in St. Stephen, New Brunswick, Canada across the USA border in July 2006, where failure was due to the rapid rate of construction and the intensity of loading for low-strength foundation soils [21]. The land in the foundation was composed of a soft marine clay layer which is a grey clay with a thickness of up to 15m.

Irsyam et al. [22] reported that a failure of the embankment slope occurred in Cipularang Toll Road located in West Java, Indonesia. Based on the record of slope monitoring indicators, the failure plane was due to low ground shear strength. Parameters of soil shear strength at failure conditions, i.e., cohesion and friction angle, were 5 kPa and $13^0$ respectively from the toe of the embankment to the top of the embankment at the median of the highway. Additionally, a road connecting two villages, namely Nata and Pentalia in southwest Cyprus suffered 340 m of severe cracks even though the road was still under construction. The incident began in January 2002 where the foundation ground was

composed of soft soil with high plasticity and low shear resistance [23]. The slope failure also occurred at the embankment due to high rates of rainfall.

Based on the case histories, it is clear that soil properties and rainfall are the cause of road embankment construction failures. To ensure embankment stability, the geotechnical design typically requires consideration of three factors namely: strength, deformation, and permeability or hydraulic conductivity. Bearing capacity and slope stability are among the important elements in the strength factor. Deformation is influenced by differential settlement, surface settlement, and lateral displacement of the ground while pore water pressure and seepage are elements affecting permeability.

### 2.1 Bearing Capacity

Bearing capacity is one of the necessary steps in the design of the embankment. It is the ability of the soil to bear the structural burden on the ground safely without any shear failure by accommodating a large settlement. This means that the foundation with shear failure involving general, local, or punching shear failure mechanisms will experience structure instability. Davis and Booker [24] have significantly contributed to understanding of the problem of clay-bearing capacity with increased strength with depth. Over the last few decades, much research has emphasized that embankment instability is due to a possible loss of bearing capacity of the foundation soil [25, 26]. During soft soil failure under the embankment, the fill moves downward with part of the foundation soil as a rigid body [25].

Several studies investigating the performance of the embankment on Muar clay [27], Bangkok clay [28], and Matagami lacustrine clay [29] were performed with low bearing capacity. Factors thought to influence bearing capacity have been explored in several studies. Popescu et al. [30] studied the effects of random heterogeneity of soil properties on bearing capacity with various shear strength values. On the other hand, Lehtonen et al. [31] investigated the effect of the trainload and the loading time on the embankment on soft soil. It is clear that careful study of the literature reveals that the soil shear strength, depth, and surcharges are the factors that influence the bearing capacity. A clear understanding of these factors is essential in the design of embankments on soft soils.

Over the past decade, some laboratory research projects have been conducted to investigate the soft soil performance based on the factors affecting bearing capacity. Kim and Lee [32] compared the bearing capacity of various sizes of gravel and sand. Two types of samples were composed of gravel and sand of compaction piles. The results revealed that coarse particles had a high bearing capacity compared to fine particles. It clearly showed that particle size contributes to higher bearing capacity, and it is noted that each type of soil has different strengths. Another study was performed with various strengths of soil stabilized by Portland cement [33]. The results found that the higher the percentage of cement mixture, the higher the maximum dry density and bearing capacity. This implies that the strength and density of the soil are higher, resulting in a high bearing capacity.

In order to improve soft soil bearing capacity, several researchers have suggested the application of basal reinforcement [34], chemical stabilization [11], and pile-supporting [35] in the embankment construction. However, cost, time and sustainable construction are the challenges of engineers choosing ground improvement techniques. A number of methods can be used to overcome this challenge with the 'Limited Life Geotextiles' [36] approach as basal reinforcement of embankments and fly ash [37] as a soft soil foundation stabilization.

### 2.2 Slope Stability

Slopes are closely related to the height of an area, whereby the higher the area, the steeper the slope. The stability of the slope is an important aspect in embankment

construction due to its possible side effect of loss of human life and damage to property. A stable slope guarantees safety but otherwise, unstable slopes increase the risk of failure that can trigger landslides that cause the most damage and produce thousands of deaths every year and material losses of billions of dollars [38]. This is of increasing concern with this trend expected to continue in the next decade due to urbanization and development. Thus, the challenging task for engineers is to make decisions by paying more attention to the importance of public safety. Engineers need to understand the factors that led to the failure in order to assess slope stability.

Over the past decade, most research on slope failure was investigated and discussed based on studies and historical cases. Ballantyne [39] reviewed the literature from the period and identified factors that influenced slope failure in Scotland. A total of 740 incidents of landslides due to intense precipitation, engineering excavations, and mass weakening by fluctuating cleft-water pressures were studied. In 1993, over 800 slope failures were triggered by a rainstorm on Lantau Island, Hong Kong [40]. Ohlmacher [16] reports that the landslide that occurred in 1995 damaged two US$ 400,000 homes in Overland Park, Kansas. Based on the investigation, the slope angle was the main factor contributing to the problem of landslides in the study area. Davies [41] reported more than 50 slope failures in Kenya caused by natural geological and geomorphological conditions, coupled with high rainfall. A landslide that occurred in Wenchuan, China in 2008 was triggered by earthquakes and heavy rainfall [42].

Several studies investigating the factors of stability were performed by experiment. Xu et al. [43] studied gravity erosion on the steep loess slope by conducting a series of experiments in the laboratory to test slope stability against various geometry and rainfall rates. The analysis found that climate changes have significant influence over gravity erosion. Yong et al. [44] described the dynamic processes of grain comminution through grain size distribution of soils. High dynamic properties are found in granular soil exposed to soil evolution and mass movements, which can cause instability. The effect of slope length and slope angle on stability was successfully investigated based on laboratory tests by Kinnell [45].

Based on the evidence presented in this section, it is clear that nature, humans, and slope geometry are factors affecting stability. Landslides triggered by natural factors such as heavy rainfall and earthquake are the most common in the world. High intensity of rainfall is prone to contribute to weathering and rapid erosion of rock mass causing a reduction in the slope stability [45]. Erosion agents such as surface runoff enter the cracks, eroding the slope carrying with it a mud and soil stream while runoff absorbed into the soil weakens the bonds between soil particles on the slope surface. The pore water pressure that exists between the soil particles will affect the strength of the soil structure.

The earthquake phenomenon causes strong ground vibration that increases shear stress. The effect of the earthquake's frequency and magnitude can overcome the force of adhesion to the ground and the retaining structure [42]. Besides, the vibration causes the movement of ground particles to be more likely to move according to gravity. Gravity is a force acting on the surface of the earth by pulling something towards the middle of the earth. Material on a flat surface will not move due to gravitational attraction unless the surface is sloped or angled.

Slope excavation is the human factor required for the construction of building foundations and roads because suitable engineering sites do not exist. This activity triggers the movement of slope and ground vibration will cause disturbance that will slide the mass, which makes the deformation accelerate [46]. Additionally, it will also destroy the

vegetation on the slope surface and consequently, the reinforcing effect of the vegetation will be lost and precipitation will infiltrate the slope crack on the surface.

Geometric factors such as slope angle are the main components of the slope for measuring stability. As the slope angle increases, the shear stress in the soil and erosion will also increase [45]. The low slope angle is expected to be stable due to low shear stresses [47]. Conversely, with the increase in slope steepness, water infiltration decreases [48]. Thus, the slope angle is suspected of playing a major role in the control of infiltration and erosion that affects slope stability.

However, the slope stability problem varies between different geographical regions. For example, in European countries, slopes are exposed to varying temperatures and humidity, while in Malaysia, they are exposed to heavy rainfall with an average annual rainfall of 2500 mm. Lack of information on climate change, rainfall rates, and unpredictable clouding has caused engineers to fail to understand the mechanisms of slope failure. A detailed study is, therefore, needed and must be performed with appropriate technology to predict the parameters associated with slope failure factors.

### 2.2 Settlement

Settlement is the total size of the vertical deformation on the soil surface. In theory, embankment settlement occurs when increased compressive stress is generated by baseload compression. This process will cause vertical and horizontal deformation of the soil foundation. Besides changes in water content, soil mass and temperature can also distort soil. It is often associated with cohesive soil such as soft soil and can produce a large settlement for an extended period even under load [11]. During settlement, the ground transition at stress conditions is of its weight to the new state due to the extra load. The stress change caused by the additional load can produce a time-dependent accumulation of particle rolling, sliding, crashing, and elastic distortions in a limited influence zone beneath the loaded area. Hence, significant and excessive impacts will affect the road surface, where a decrease in height will occur and may cause flooding during heavy rains. Terzaghi [49] began the study of soft soil settlement. In theory, a settlement analysis comprises calculations of total and differential settlement, where total settlement is the magnitude of downward movement while the differential settlement is the vertical movement between two places.

The total settlement consists of three modes, namely elastic, consolidation, and creep deposition. Elastic settlement occurs due to the deformation of soil that is loaded without changes in water content and occurs within a short period [50]. It is also known as an undrained or immediate settlement. This type of settlement occurs instantaneously during the loading process with saturated soil that has low permeability and drainage ability. Consolidation settlement occurs due to the consolidation of soil in which water is slowly expelled from the soil cavity [51]. During the consolidation process, the deformation of the soil occurs with volume reduction over a long period. Creep or secondary compression is the same as compaction. It can occur when soil attempts to reposition itself in the cavity left by the water during consolidation settlement [52]. In the typical case, during embankment design, the computation of settlement involves elastic settlement and consolidation while post-construction involves creep.

Azzam and Basha [53] assessed the impact of improvements in the parameter of cohesive soil shear strength by using a soil nailing technique to reduce settlement. Brown to grey clay was taken from the Middle Delta of Egypt and stabilized using steel flat bars as vertical soil nails. The sample is then compressed with an unconfined compression machine

with a low pressing rate. The study showed that increasing vertical inclusions significantly increases the shear strength and reduces settlement rates. Studies on the effects of creep on the side deformation of soil and pore pressure below the embankment were performed by Grimstad et al. [54]. The numerical method is used to investigate the embankment on the I-95 motorway in North Boston with a height of 12.2 m and built on Boston Blue Clay. The study shows that an increase in creep rate also increases soil settlement, pore pressure, and lateral deformations. Earthquakes can cause the settlement of saturated clay layers. This has been discussed in a study conducted by Sato et al. [55] using a Kaolinite clay sample, i.e. Tokyo bay clay, and Kitakyushu clay tested using uniform and irregular cyclic shear tests. The results of the study concluded that the Atterberg limits affect the settlement rate and pore water pressure accumulation during earthquakes. Accumulation and settlement of water pore pressure triggered by the friction cycle are measured by cumulative shear stress. To conclude, the studies conducted by the researchers clearly show that the type of soil, the underground conditions, the techniques to improve soil properties, and the soil drainage are factors that affect the settlement.

Differential value of settlement on the ground surface and uneven settlements can cause major problems such as a function of the uniformity and stiffness of the soil. As shown in Fig. 1, it may be characterized as the distortion, i.e. the settlement over a horizontal distance. It occurs when the ground underneath the structure is unable to withstand the load, but the limitation to differential settlement is dependent on the application and site condition.

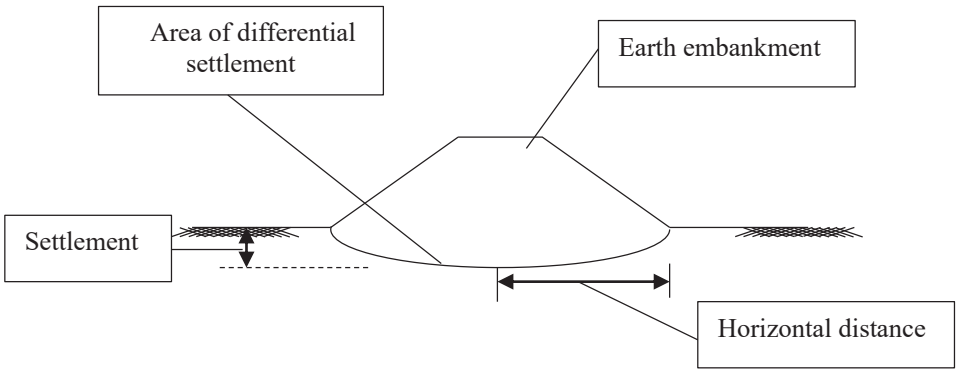



Fig. 1: Definition of the differential of settlement [56].

One of the case histories occurred in Bangkok where significant settlement happened in an area with 12-16 m of soft clay [57]. The 1.5 m to 2.5 m road embankment experienced total settlement of between 500 to 1000 mm in 10 years. Dark grey soft and compressible recent marine clay and frequent flooding from the Chao Phraya river seem to be the major factors. A seven-story building was slanted during the Mexico earthquake that took place in 1985 and continued until 2005, as reported by Moghaddam and Jayawickrama [58]. Based on the results of the review and monitoring it was found that the building underwent a differential settlement leading to vertical inclinations ranging from 241 mm to 310 mm to the north, and 70 mm to 98 mm to the east. This problem is very closely related to the consolidation process of the base soil. Apart from these factors, eccentricity in loading, groundwater extraction, and construction in the vicinity also contributed to this problem.

There are four reasonable and commonly used methods to predict soft clay settlement. These are the conventional one-dimensional (1D), Skempton and Bjerrum [59], stress path and finite element methods. Terzaghi [49] introduced the 1D conventional method aimed at estimating the compression of clay at a certain soil depth. It is now used to predict the magnitude and settlement rates. With the right soil compression and the measurement of a

thin layer of soil in the laboratory under a single load dimension, engineers can successfully and accurately predict performance [60]. According to the theory, stratum compressible is divided into several layers, while the calculation of the initial vertical effective stress and vertical stress increase using the Boussinesq stress [61] is based on the centre of each layer. At the end of the consolidation, when all excess pore water pressure is lost, the stress increase will become effective stress. The vertical strain of each layer can be obtained from the graph plotted between the void ratio with effective vertical stress as a result of the one-dimensional (odometer) compression experiment. This method was used in the study of the settlement of road embankment prediction [62]. However, Duncan [63] highlighted the limitations of this method; among others are the difference in the conditions between the field and the laboratory, the difference in load rates and tensions experienced by the soil between the laboratory and field, and the limitations found in 1D theory for three-dimensional (3D) problems. In addition, there are insufficient resources and documentation on the selection of parameters and the use of theory to predict settlement of silt compared to clay [64].

Skempton et al. [65] used 1D to calculate the total settlement of several buildings but admitted that undrained settlement is a significant parameter in the total settlement. As such, they used the elastic displacement theory to calculate the undrained settlement. The settlement forecasting in 3D has lower consistency compared to 2D due to pore pressure and the lower lateral deformation in 3D. They have thus put together a primary method to calculate the consolidation settlement caused by excessive dissipation of pore pressure. The settlement correction factors proposed in 1957 are dependent on geometric problems and pore pressure. This factor was also supported by Bergado and Teerawattanasuk [66] in their study that geometric effects should be considered as an important factor that could affect numerical simulation results. Bergado et al. [67] used this method to predict the embankment settlement of the Bangna-Bangpakong highway that was constructed on Bangkok soft clay. The study found that these methods provide reasonable estimates of long-term settlement.

The stress path method was proposed by Skempton [59] and developed further by Poulos and Davis [68]. It uses an uninterrupted sample under the appropriate underground depth to perform triaxial experiments under the appropriate initial in situ effective stresses. The purpose of vertical and horizontal stress without allowing drainage to be calculated is to determine vertical strain and excess pore pressure. When necessary, drainage was allowed, and applied stresses were adjusted to determine vertical stress and change in volume. The strain was then measured to obtain undrained settlement and consolidation. Studies on the performance of the screw plate test to determine the properties of deformability and consolidation of Bangkok clay was carried out by Bergado et al. [67]. Comparison of consolidation coefficients, non-flow and flow modulus between SPLIT and stress-path-controlled triaxial consolidation tests was performed. The findings of both experiments are similar to each other, implying the accuracy of the results to be sufficient for use for predictive purposes. Although logical, the method is complicated because the initial vertical and horizontal effective stresses must be known, and it requires advanced experimental skills. Isotropic and homogenous elasticity used to calculate the stress change and uncertainty in the accuracy of horizontal stresses exist. In addition, these experiments are highly costly and require a long time to conduct.

In recent times, numerical methods have been well received for predicting settlements because finite elements are carried out on non-homogeneous anisotropic elastic materials. Besides, it can handle nonlinear stress-strain behaviours of varying complexity. The engineers are more confident in the findings of finite elements because of its strength and

flexibility, without realizing that this package contains inherent idealization and assumptions. Generally, the quality of data inputs plays an important role in determining the accuracy of the forecast. It is therefore important to evaluate and examine the constraints of the experiment as good control ensures high-quality data to be used for numerical analysis.

It was clear that settlement forecasting methods have undergone a sophisticated revolution moving from the classic 1D to numerical methods. The latter is able to solve a variety of complex issues because the program is equipped with advanced parameters and testing ground. The available features with sophisticated computing technology have led to this method becoming popular for application and development.

## 3. EMBANKMENT DESIGN

The design of the embankment is extremely complex because it depends on the purpose of construction, the environment, and the ground foundation conditions. When designing an embankment, engineers must plan, define geometric data, and analyse stability. The road embankment especially consists of a series of compacted layers placed on top of each other until the subgrade surface level is reached. It is also known as structural embankment and is divided into two types: the unreinforced [69] and reinforced, [70] which serves to distribute traffic load to soil foundation. Therefore, appropriate stability assessment and geometric data methods are essential to ensure the desired efficiency of construction. This section discusses three key aspects of geometric data, design, and analysis methods based on the recent literature.

### 3.1 Design and Analysis Methods

The majority of today's engineers and researchers model the embankment design using FEM. It is a numerical method that aims to design and analyse various engineering problems of both practical or research. It is also useful for problems with complicated geometries, loadings and material properties where analytical solutions cannot be obtained. There are three types of finite elements that are often used, namely one-dimensional (line), two-dimensional (plane), and three-dimensional (solid). From the literature, the use of two-dimensional is the choice of many researchers in the study of embankment stability. Plaxis [81], Abaqus [82] and Fast Lagrangian Analysis of Continua (FLAC) [83] are the types of software used for analysis, testing, and design in geotechnical engineering. However, the use of FLAC is less common in the study of the road embankment. Researchers and engineers more popularly use Plaxis as it is friendlier in soil modelling to solve simple problems in complex strata. Beyond that point, it can analyse all cases in geotechnical engineering including pile design [84], ground foundation [85], retaining wall [86], slope stability [87], and dam [88].

The driving force along the slopes that is greater than the resisting force restraining the movement of the slope is the basis of slope stability. Force imbalances such as increased resisting force will fail. Both forces are elements in slope design. The main driving force in most earth movements is gravity while the main resisting force is the material's shear strength. Slope stability is evaluated by calculating and assessing the stability index, or the FoS, defined as the ratio of ground resistance force along the failed surface to the instability force that illustrates the capabilities of a structure, such as an embankment [89]. Duncan [90] also explains that FoS is the ratio of real ground shear strength to the minimum shear strength required to prevent failure. Theoretically, a FoS value greater than 1 indicates a stable slope while a FoS value equal to 1 implies that the slope is in equilibrium.

The slope stability analysis is divided into two, namely global and local stability. Global stability refers to the overall stability of soil mass along the slip surface and in multiple failure mechanisms by providing a lower safety factor. A sloping surface can withstand soil stress. For long-term and short-term flows, FoS ≥ 1.3 [91]. Most slope stability analysis software available, such as Slope/W, provides global FoS calculations. The global FoS calculation using the Phi-C reduction method was performed by Singh et al. [92] in the study of slope stability in Nainital, India. Iñeguez [93] in his study defines universal FoS by using a strength level approach. He also calculates local FoS using enhanced methods as the location of the most critical zone can be evaluated and used to guide the implementation of remedial measures. Local FoS are irregular and different at each point compared to conventional FoS which assumes that it remains constant along the slip surface.

In order to assess the stability of the slope, limit equilibrium (LE) analysis is commonly employed to determine the FoS by researchers and engineers. The LE technique such as the ordinary method of slices [94], Bishop simplified [95], Spencer [96], and Morgenstern-Price [97] methods have been widely used in the slope stability analysis. The slicing technique is known as an uncertain static problem, solved by assuming a distribution of internal forces. As such, different assumptions can result in varied results. Due to this inherent engineering practical weakness, FEM is now a choice. Some researchers have compared the results between LE method and FEM [98, 99], concluding that FEM provides the best and most accurate FoS result, as it is better able to explain the behaviour of soil stress and eliminate stress assumptions used in the LE method.

Strength reduction method (SRM) [100,101] and enhanced limit strength method (ELSM) [102, 103] are popular methods used in FEM although there are also ways to evaluate FoS using stress calculated directly through finite element modelling. The SRM has been used by Zienkiewicz et al. [104] to calculate slope safety factors based on the Mohr-Coulomb model. Most FoS formulas used now are obtained from the Mohr-Coulomb failure criterion [105]. SRM does not require assumptions on the form or location of the failure surface, because failures occur naturally through the zone in the soil mass where soil shear strength is unable to resist the applied shear stress [106]. This means that the critical slip surface is automatically obtained from the increased shear stress when the shear strength decreases. However, this method is unable to determine the position of the local minimum slip. ELSM is a hybrid between the LE method and FEM introduced by Kulhawy [107] as an independent assessment of the distribution of normal and shear stress along the assumed slip surface. Normal and shear stresses from elasticity analysis are used to calculate the overall FoS. In addition, this method uses FEM to estimate the magnitude of the stress as well as to find the critical slip surface with minimum FoS.

ELSM has the advantage over SRM in terms of time, as some finite element analyses need to be conducted separately. Besides, SRM is best suited for global FoS calculations while ELSM is for local FoS calculations. This is because local FoS calculations along the potential failure surface cannot be considered in the SRM procedure.

### 3.2 Geometry

Analysis through FEM requires the establishment of data in constitutive relationships and geometric modelling. Geometric data namely height, crest width, and gradient of side slope are among the important parameters in the production of design drawings. Engineers usually select the parameters based on their own experience. Based on observations, most studies conducted by researchers use embankments that are 1.5 m in height, have a crest width of between 35 m and 40 m, and have a 1:1.5 gradient of side slope (vertical: horizontal). In addition, road design manuals in several countries are described in Table 1.

Most countries recommend embankment height of less than 6m to enable the vertical-horizontal side slope ratio of 1:1.5.

Table: Various sources of standards for side slope and height embankment for design.

| Height | Side slope (V:H) | References |
|---|---|---|
| 0 – 6 m | 1:1.5 | [108] |
| 0 – 6 m | 1:3 | [109] |
| 0 – 1 m | 1:4 | [110] |
| 1 – 3 m | 1:2 | |
| More than 3 m | 1:1.5 | |
| 0 – 1.5 m | 1:8 (flat terrain) | [111] |
| | 1:6 (rolling terrain) | |
| 1.5 m and over | 1:6 (flat or rolling terrain) | |
| 1 m | 1:1.5 | [112] |
| 2 m | 1:2 | |
| 3 m | 1:2.5 | |

Based on literature reviews, geometric data such as crest width and side slope are reported to be too limited. Researchers widely practice the study of critical [113–115], safe [116], and reasonable [117, 118] embankment height as they believe that these parameters are factors that contribute to embankment stability, which is supported by study findings. The study conducted by Cui et al. [119] further supports this statement as it found that the height of the embankment affects the cumulative settlement of the subsoil. Embankments built at an altitude of less than 1.5 m will lead to an increase in the speed of cumulative settlement of the subsoil. However, the determination of height depends on the soil foundation improvement and traffic load method or technique. Studies conducted by Almeida et al. [120] found that the result of FEM analysis showed the difference in settlement of soil surface decreased when the height of the embankment increased. This proves that height is an important determinant of embankment stability.

### 3.3 Height of Embankment

The height of the embankment is a problem often encountered during design and construction work. This is because the embankment built below the critical height can control the deformation and stability of the foundation. Critical height is important to determine performance such as the workability of the embankment and the assessment of potential excessive settlement [121]. For example, surface damage is caused by uneven crest surface. This happens when the embankment is lower than the clear column spacing, leading to the differential settlement at the base of the embankment that can spread up to the crest of the embankment. According to McGuire [122], the minimum height for the column diameter and distance where differential settlement in the crest does not occur is known as the critical height. Thus, the diameter and distance from the column are important factors that affect critical height.

Fagundes et al. [113] in their study used the calculations proposed by Filz et al. [121], McGuire [119], and BS 8006 [123] to find that the critical height suitable for the construction of embankments built on soft soils strengthened by piles without reinforcements is 2.5 m. Bhasi and Rajagopal [114] found that the critical height with the effect of reinforcement modulus on maximum settlements is 5.6m. In this study, embankments stabilized using the geosynthetic reinforced pile method built on soft soils were analysed using the three-dimensional (3D) FEM. The findings from Chen Jia et al. [124] on embankments stabilised by plastic tubes cast-in-place concrete piles found that the

critical height is 2.2 times the pile net spacing distance. The piles are driven by a distance of 1.0 to 1.5 m from one another. This means that the critical height is between 2.2 to 3.3 m. The three studies found that the critical height for the embankment built on soft soil and soil stabilized with piles is between 2.2 m and 5.6 m. However, there are very few studies on the critical height for embankments stabilized using other techniques such as vertical drains, preloading, grouting, and reinforcement.

A lot of research regarding reasonable height was conducted on the Qinghai-Tibet Highway (QTH) and the Qinghai-Tibet Railway (QTR). It was introduced by Zhang et al. [125] to protect the underlying permafrost during construction. The freezing and defrosting cycle of soil can lead to a settlement on the embankment. Hence, maintaining the magnitude of the embankment settlement to reasonable limits is key to ensure embankment stability [126]. The higher the embankment, the greater the differential settlement because an embankment that is not sufficiently high will lead to the freezing and defrosting cycle that affects tension and settlement. The foundation soil found along the highway and the railway is mostly silty clay, which is a type of soft ground.

QTH has a length of 1937 km and has a 632 km permafrost area. Among the studies to determine the reasonable elevation of the embankment in QTH were those carried out by Wang et al. [117] and Zhang et al. [127] who found, through site monitoring, that an embankment height of more than 4 m would lead to a reduction in the permafrost table. Therefore, reasonable elevation shall be determined to a limited extent. Wang et al. [117] have determined a reasonable embankment height for two typical geological sections along the QTH. The findings of the analysis are obtained through the modified numerical method, with a reasonable embankment height implied of 1.63 m. This is a significant study because reasonable height is important for maintaining the thermal and mechanical stability of highways located in the cold areas.

The QTR is 550 km long and has a permafrost area at an altitude of about 4500 m to 5000 m above sea level. Yue et al. [126] and Qi et al. [128] conducted analyses to obtain reasonable embankment of railroads in the permafrost state. The embankment should be built at a reasonable height of 4.0 m to reduce differential settlement [126]. The in-situ monitoring results in Qi et al. [128] showed that the reasonable height is 3.0 m and has a positive impact on frozen soil protection and promotes the increase of the permafrost table under the embankment. A study on the optimum design of crushed stone layers with shading board over the railway embankment on warm perma-frost was done by G. Li et al. [129] and found that the reasonable elevation is between 3.0 m to 5.0 m. The analysis of these findings is obtained through systematic numerical tests.

The safe embankment height is dependent on the gain in strength of soft subsoil and target FoS [130]. The study determines the height of each stage of construction based on increased shear strength to reach the FoS of 1.25. Through the iteration process, they have found the exact height of second-stage loading to be 3.02 m to achieve a total height of 5 m for the embankment. Meanwhile, the study conducted by Kasim et al. [116] concerning the safe height for the construction of the embankment on soft soil is summarised through the simulation results of PLAXIS Version 8.2 software that the maximum height is 4.9 m to ensure its stability against excessive settlement. This is also supported by Jin et al. [131], whereby increasing the height from 1.0 m to 4.0 m reduced the deformation value towards settlement.

Several factors affect the height of an embankment. We have reviewed critical, reasonable, and safe height, which have the same goal of ensuring embankment stability. Overall, height determination depends on several factors. The diameter and distance

between planted pile stalks are factors that determine the critical height. Permafrost is one of the most dangerous elements and affects the stability of the embankment in cold areas. This is due to uncontrolled thaw settlement, and frost heaving processes that can cause settling problems and longitudinal cracks on the embankment. According to Wang et al. [132], embankment height is closely related to the problem. The reasonable height is interlinked with embankments built in areas prone to permafrost. Therefore, the temperature of the permafrost is also a factor in determining a reasonable height. FoS is the determinant of an embankment's safe height. Most studies conducted to determine the safe height use Taylor's stability charts. The road design manual as in Table 1 suggested embankment height of up to 6 m. However, there are also researchers who employ the height in their study as 2.0 m [133], 3.0 m [134], 4.0 m [135] and 5.0 m [102]. We can conclude from the literature that 1.5 m to 6.0 m is the height range in embankment design that is able to ensure stability. The same height range is appropriate for the study or design of embankments. This range is relevant and useful as a guide for young engineers and researchers in the prediction and design of embankments.

## 4. SOFT GROUND IMPROVEMENT

An engineer must possess several strategies that may be adopted to achieve project goals when dealing with troubled ground conditions at the site. Various methods exist to improve compressible soils commonly applied on-site including preloading, lightweight fill, geosynthetic reinforcement, over-excavation and replacement, vertical drains, rigid piles, and injection. As mentioned in the literature review, enhancing density, increasing shear strength, and reducing compressibility can help in promoting stability. On the other hand, influencing permeability also decreases groundwater flow to increase consolidation rates. Despite the fact that existing methods may mitigate stability problems, selection factors need to be considered, such as sustainability, cost, and time of construction. All have become the topic of intense debate within the scholar community. Thus, this section will review the application of preloading with PVDs and geofoam as lightweight fill materials in the road embankment construction on soft ground by considering three selection factors.

### 4.1 Preloading and Prefabricated Vertical Drains

Preloading methods are categorized as static consolidation and are cost-effective and practical. They impose a temporary load to increase soil stress. This causes the substratum to densify and form stiffer support that will help reduce compressibility when the actual structure is built. This method commonly uses earth filled with a pre-determined height, and subsequent monitoring of settlement is conducted during a specific period until the deformation stops.

Dafalla [136] studied the performance of a stabilized soft ground settlement using the preloading and prefabricated vertical drains (PVDs) method along the coast of the Red Sea of Saudi Arabia and found that medium hard silty sand was beneath the site surface. Buildings and pedestrian pathways around the site have been damaged due to excessive settlement. The Saudi Ministry of Municipal and Rural Affairs has proposed that this method is applied for the project. The site uses preloaded sand of 4.5 m to 6 m high for six months and was monitored every week. The study concluded that this method reduces 20 % to 50 % of total settlement with sufficient time to prevent deformation. It can increase the strength of soft soils by accelerating the consolidation process so that the ground is able to bear the load. Therefore, sites with soft grounds are suitable for this method as they can be compressed in a relatively short time when loaded. Pressures exceeding what is required are placed above the soil and allowed to settle until the predicted settlement is reached. When

the settlement has been reached, the excessive load is removed, and the service load is allowed on a strengthened foundation.

Recently, various innovations in this method have been created using modern technology. Among the successful modernizations of this method is combining the use of vertical drains or PVDs. Variation of this method has been successfully applied in many construction projects worldwide requiring ground improvements and reclamation [137]. Therefore, the theory, design, and construction methods for PVDs are significant issues in the preloading or consolidation method. The practice of using horizontal drainage began to emerge with drainage and then evolved into PVDs. Kjellman [138] introduced a combination of two materials, namely wood, and cardboard in 1947. Most PVDs have a width of about 100 mm with a thickness of 3.0 m to 6.0 m made of corrugated plastic cores surrounded by geotextile filters or a layer of natural fibres. It serves as artificial drain lines to accelerate the draining of water in the ground so that the time taken for the soft soil consolidation and the removal of excess pressure in the water pore can be significantly reduced. The elimination of pore water pressure increases soil strength, which would enable it to sustain the load. Da Silva et al. [81] studied the impact of PVDs installation on stability and embankment settlement. Using back-analysis to find sufficient material model and calculation of values using elastic-viscoplastic (EVP) through FEM to four dikes on very soft clay, it was found that PVDs very significantly accelerated settlement and increased FoS. Chu et al. [139] also conducted a study on the use of PVDs on soft clay soil which has a high moisture content for land reclamation projects. Evaluation of the consolidation effectiveness and suitability of the selected PVDs was performed through a large-scale laboratory model. The results showed that the installation of PVDs is very effective to accelerate the consolidation process. However, PVDs performance may decline as the consolidation process takes place due to the quality of the material that is unable to sustain the biological and chemical reaction between the filter and the soil surface. Besides, the soil undergoes a significant increase in undrained shear strength.

It is clear from the above study findings that the installation of the preloading and PVDs methods has advantages. A combination of both methods is efficient to be used in areas with soft ground. This method does not meet the green technology needs that emphasize the environmental impact of construction. This is because the surcharge fill usually involves cutting of high places such as hills, leading to various degrees of destruction and disasters. Although there are alternatives undertaken and introduced by Yang et al. [140], which use water as a surcharge, they are not suitable for areas with hot climates because the evaporation process will reduce the quantity of water and thus can affect the settlement monitoring readings. Besides, there are preloading methods using vacuum pressure [141, 142], which only use surcharges such as minimal soil but still require soil transportation from sources elsewhere. This method is also highly costly because it involves the pumping and plastic pipes that are often exposed to leakage. Therefore, environment-conscious methods should be developed based on the environmental and construction conditions of a particular location.

### 4.2 Geofoam as Lightweight Fill Material

Construction challenges on soft grounds and lack of suitable soil for reclamation work and environmental preservation led to a massive demand for innovative methods. The use of lightweight materials to replace fill material such as clay is an alternative. According to Schaefer et al. [143], lightweight fills are divided into two categories: materials that behave and have the same properties as soil particles, and materials with unconfined compressive strength that possess properties that are similar to the cohesive soil as shown in Table 2. The

unconfined compressive strength is directly related to the cast density of the mixture. In the report by Stark et al. [144], geofoam unconfined compressive strength range is 160-173 kPa. Cellular concrete consists of cement, aggregate, water, and air voids. Its production needs to mix materials such as foaming agents that can produce 10-70% of the air content [145]. Therefore, the unconfined compressive strength is dependent on the mixed material used. The material mixture used in cellular concrete as foam content + bottom ash [146], flue ash [147] and aluminium powder [148] produces a compressive strength of 0.8-5.2 MPa, 1.7-2.7 MPa and 0.9-7.9 MPa, respectively.

Table 2: Lightweight fill materials categories [143].

| Granular lightweight fill materials | Lightweight fill materials with an unconfined compressive strength |
|---|---|
| Wood fiber | Geofoam |
| Blast furnace slag | Cellular concrete |
| Boiler slag | |
| Fly ash | |
| Expanded shale, clay and slate | |
| Shredded tires | |

The use of geofoam is well received because it is more practical as a readily-available source, and there are many manufacturers and suppliers around the world [149]. It has a density value of between 12 to 35 kg/m$^3$ and was first used in the construction of road embankment in 1972 by the Norwegian Road authorities [150]. Based on density, expanded polystyrene (EPS) is divided into seven categories with the index oxygen volume of 24% as shown in Table 3 [151]. It is produced in large light blocks and can be obtained in various sizes and resistance characteristics. It is made of expanded styrene formed through the process of expansion using pentane and steam [152]. The mechanical [153] and dynamic properties [154] of geofoam blocks were investigated in detail. The EPS' light strength properties between surfaces were also investigated and found to be significant across two surfaces of the block [155].

In general, EPS blocks are used to reduce overburden pressure and settlement problems. Umashankar et al. [76] studied the settlement rates of landfill built on soft soils using two types of EPS. The comparison of settlement concentration was done on EPS with an elastic modulus of 1 and 10 MPa respectively. The analysis through FLAC on the embankment loaded following AASHTO specifications found that the maximum embankment settlement with EPS elastic modulus of 10 MPa is between 1.8 cm and 2.0 cm while the elastic modulus EPS 1 MPa is in the range of 15 cm to 16 cm. This demonstrates that the EPS with a hard feature has the advantage of reducing the settlement problem.

Table 3. Types of EPS geofoam [151].

| Types | Density (kg/m$^3$) | Flexural Strength (kPa) |
|---|---|---|
| EPS12 | 11.2 | 69 |
| EPS15 | 14.4 | 172 |
| EPS19 | 18.4 | 207 |
| EPS22 | 21.6 | 240 |

The Indiana Department of Transportation has recommended the use of the EPS blocks for the Borman Highway reconstruction project near Gary, Indiana to reduce the load on soft grounds [156]. The ground in this area consists entirely of soft soils, as it is close to Lake Michigan. Thirty-two truckloads of EPS blocks were used in this project. Michalowski

et al. [157] investigated the road embankment performance which is partially stabilized using the pile-supported method, that is controlled modulus columns of 400 mm in diameter and installed with square grid space of between 1.4 m to 2.0 m from one another and were installed in ground by up to 22.8 m deep [148]. The bulk of the embankment fill used the EPS29 block and mounted on a 9.5 m thick organic clay foundation. The top layer of the EPS block is protected with a 0.4 mm thick synthetic moisture resistant coating, while the end of the block is mounted with at 0.6 mm thick geomembrane. The results obtained through the calculation of settlement were consistent with the settlement of the pile group due to the use of EPS which reduced the burden of the load. In another project, the Utah Transportation Department in 2001 undertook the reconstruction of a 27-km I-15 highway over four years costing $1.4 billion in Salt Lake City [158]. A total of 100,000 $m^3$ EPS15 blocks are used in this project as a road embankment filling material. The widened I-15 alignment requires a large embankment of 8.0 m to 10.0 m above the soft soils that could potentially lead a primary consolidation settlement of more than 1.0 m. Some sections of the highway experienced settlement due to the construction of the embankment and have damaged the utility lines across the road. This has led the design team to decide on the use of EPS blocks in the project to maintain service without the need for relocation or work suspension, which could affect construction costs. After the project was completed, Bartlett et al. [159] produced a design review report and performance assessment of the embankment for 10 years. The results obtained from the site show that the elastic compression and gaps between the two EPS block faces yield about 1% vertical gradients as load distribution slabs. Additionally, a total of 15 mm foundation ground settlement and 0.2 to 0.4% of creep cramps occur on the base soil during the post-construction period of 10 years under pressure levels of 20 to 35 kPa. Long-term monitoring of creep deformation of the criteria of 50 years found that maximum pressure is about 1.5 to 1.7 %.

The case history described above showed that the EPS blocks functioned as a solution to the problem of excess settlement. Its light weight reduces side loads or bearings borne by the base soil. Although EPS with various density characteristics is available to meet different strength requirements, cost-based selection and suitability of use should be considered. EPS with high hardness or elastic modulus is useful for large loads. Many researchers believe that using this method can reduce costs [76, 153, 157], but Kim et al. [160] argued that the use of EPS blocks requires a high cost in terms of installation and construction operations in road construction projects compared to plastic board drains and PVD. This is because even with the reduction in the quantity of soil used for the embankment, it still requires additional soil due to the same longitudinal grade as the original design. Hence, the development of sustainable materials such as expanded cork granules and granite sludge, comparable to EPS features and functions should be carried out.

## 5. DISCUSSION AND FUTURE RESEARCH DEVELOPMENT

The initial objective of this study is to discuss the factors that contribute to the stability of road embankment on soft ground. Based on surveys through literature and some historical cases, current studies have found that settlement, slope stability, and bearing capacity are important factors for the stability of the road embankment. Therefore, it is clear that these factors are contributors to road damage as evidenced by the findings of a historical case investigation that has been reported by researchers [19-23]. However, an unexpected finding was that soil properties and rainfall rates also contribute to this problem. The rainfall rate depends on the weather conditions of the area or country. These two factors are the cause of the occurrence of sediment, failure of bearing capacity, and slope failure. One of the

issues emerging from this finding is that soft soil properties require a high cost for improvement.

Furthermore, these soil properties also require the installation of instruments and frequent lab and field tests during the design process. With the latest technology approaches such as artificial intelligence, it no longer requires excessive testing in the field or laboratory as it can predict the soil properties quickly and accurately. Artificial neural network (ANN), adaptive neuro-fuzzy inference system (ANFIS) and support vector machines (SVM) are artificial intelligence methods that are widely used today. Further work is required to develop a road embankment stability prediction system that includes the key aspects of settlement, soil bearing capacity, and slope stability based on soil properties.

The second objective of this study is to discuss the factors affecting the design of road embankments. In this study, geometric data is found to have an important role in the design. This is because geometric data such as embankment height, side slope, and crest width should be determined before the analysis of stability is done. What is interesting from the literature is that most researchers are more interested in investigating the height of the embankment. Three types of embankment height have been previously reported, which are critical [113, 124], safe [116], and reasonable height [117, 128]. However, most guidelines issued by enforcement agencies in many countries emphasize that slopes and side loads need to be complied with as stipulated. Most of the existing guidelines are limited regarding side slope determination criteria based on the height of the embankment. This is an important issue for future research to develop new guidance that has critical, safe, and reasonable side slope criteria. Moreover, the relationship between crest width, side slope, and embankment height may be investigated in the future to develop a standardized level or value range with critical, safe, and reasonable criteria.

The final objective of this study is to discuss effective soft soil improvement techniques. Based on literature and observations through some historical cases, preloading and prefabricated vertical drains [136, 138] and geofoam as lightweight fill materials [145, 146] are widely used today. The study found that both techniques are environmentally friendly. However, both of these techniques require a long construction time. The preloading and prefabricated vertical drains technique takes up 6-12 months for the soft ground to consolidate. Installation of geofoam is very complicated and requires the construction of a linked network that takes a long time. This problem can lead to cost implications for construction. Thus, this is an important issue for future research and development.

## 6. CONCLUSION

The main goal of the current study was to review the challenges faced in the construction of highway embankments on soft ground and suggest a direction for future development. This study has found that, generally, there are two major challenges that engineers need to face in the construction of road embankments on soft ground during the stability design and selection of improvement techniques. It also found three factors that engineers need to consider in the embankment stability analysis, i.e., bearing capacity, settlement, and slope stability. However, these factors depend on ground properties, rainfall rate, and earthquakes. Additionally, more significant findings in current studies have found that research that has been reported on a side slope and crest width was limited compared to the height of the embankment. Thus, future development studies need to be broadly comprehensible on these three geometric data by considering criteria such as critical, safe, and reasonable. In order to save construction time, this study found that the FEM approach in determining settlement rates and slope stability is useful. The most obvious finding to

emerge from this study is that preloading with PVDs and geofoam is widely applied as a soft soil improvement technique as it is environmentally friendly, cost-effective, and saves construction time. The findings from this study make several contributions to the current literature. First, it extends our understanding of the latest issues in the construction of embankments on soft ground. Secondly, it adds to a growing body of literature on road embankment stability. Finally, it identifies the gap in previous studies for the development of future studies.

## ACKNOWLEDGEMENT

The authors wish to thank the Centre of Engineering and Built Environment Education Research (PeKA) and Smart and Sustainable Township Research Centre (SUTRA), Faculty of Engineering at and Built Environment, Universiti Kebangsaan Malaysia for providing the facility for this research work. Besides, the appreciation also goes to the Ministry of Education, Malaysia, for providing scholarships for the project. Moreover, we are grateful to Politeknik Ungku Omar for providing the opportunity to perform this project.

## REFERENCES

[1] Arulrajah A, Maghoolpilehrood F, Disfani MM, Horpibulsuk S. (2014) Spent coffee grounds as a non-structural embankment fill material: Engineering and environmental considerations. Journal of Cleaner Production, 72:181-186.

[2] Yu Z, Jianhui Z, Xu Z, Xiaodong P, Hongwei L, Hao C. (2017) Finite Element Analysis of Embankment with Soft Foundation Reinforced by Geogrids. Modern Civil and Structural Engineering, 1(1):78-83.

[3] Watts K, Charles A. (2015) Building on fill: geotechnical aspects, Third. Garston, Watford: Building Research Establishment.

[4] Xue J feng, Chen J feng. (2015) Reinforcement strength reduction in FEM for mechanically stabilized earth structures. Journal of Central South University, 22(7):2691-2698.

[5] Zhuang Y, Wang K. (2017) Numerical simulation of high-speed railway foundation improved by PVD-DCM method and compared with field measurements. European Journal of Environmental and Civil Engineering, 21(11):1363-1383.

[6] Yu H, Wang Y, Zou C, Wang P, Yan C. (2017) Study on Subgrade Settlement Characteristics After Widening Project of Highway Built on Weak Foundation. Arabian Journal for Science and Engineering, 42(9):3723-3732.

[7] Quang ND., Dang SM. (2013) Settlement calculation and back-analysis of soil properties for a test embankment on a soft clay ground improved by PVD and vacuum-assisted preloading at a site in Vung Tau, Viet Nam. Springer Series in Geomechanics and Geoengineering, 3:317-322.

[8] Li S, Huang X, Zeng C. (2017) Performance of an Embankment Foundation with Sand over Clay : Experimental and Numerical Analyses. International Journal of Geomechanics, 17(6):1-11.

[9] Nazir R, Moayedi H, Subramaniam P, Gue S-S. (2017) Application and Design of Transition Piled Embankment with Surcharged Prefabricated Vertical Drain Intersection over Soft Ground. Arab J Sci Eng. doi: 10.1007/s13369-017-2628-6.

[10] Yean-Chin T, Peir-Tien L, Kuan-Seng K. (2016) Construction Control Chart Developed from Instrumented Trial Embankment on Soft Ground at Tokai of Kedah, Malaysia. Procedia Engineering, 143(Ictg):548-555.

[11] Che Mamat R. (2013) Engineering properties of Batu Pahat soft clay stabilized with lime, cement and bentonite for subgrade in road construction. MS Thesis, Faculty of Civil and Environmental Engineering, Universiti Tun Hussien Onn Malaysia.

[12] Liu SY, Cai GJ, Tong LY, Du GY. (2008) Approach on the Engineering Properties of Lianyungang Marine Clay from Piezocone Penetration Tests. Marine Georesources &

Geotechnology, 26(3):189-210.
[13] Horpibulsuk S, Wijitchot A, Nerimitknornburee A, Shen SL, Suksiripattanapong C. (2014). Factors influencing unit weight and strength of lightweight cemented clay. Quarterly Journal of Engineering Geology and Hydrogeology, 47(1):101-109.
[14] Le TM, Fatahi B, Khabbaz H. (2012) Viscous Behaviour of Soft Clay and Inducing Factors. Geotechnical and Geological Engineering, 30(5):1069-1083.
[15] Yalcin A. (2007) The effects of clay on landslides: A case study. Applied Clay Science, 38(1-2):77-85.
[16] Ohlmacher GC. (2000) The Relationship between geology and landslide hazards of Atchison, Kansas, and vicinity. Current Research in Earth Sciences, 244(3):1-16.
[17] Oser C, Cinicioglu SF. (2017) Embankment Design Method Combining Limit-State Approach with Stress-Path Application. International Journal of Geomechanics ASCE, 17(4):1-16.
[18] Lu N, Wayllace A, Oh S. (2013) Infiltration-induced seasonally reactivated instability of a highway embankment near the Eisenhower Tunnel, Colorado, USA. Engineering Geology, 162:22-32.
[19] Rahadian H, Hendarto, Prasetya B. (2011) The Failure of Road Embankment Over North Java Soft Soil. In: Geotech. Eng. Disaster Mitig. Rehabil. Highw. Eng. pp 224-232.
[20] Tatarniuk C, Bowman ET. (2012) Case Study of a Road Embankment Failure Mitigated Using Deep Soil Mixing. In: Proc. Fourth Int. Conf. Grouting Deep Mix. American Society of Civil Engineers, pp 471-482.
[21] Mills B, McGinn J. (2010) Design, Construction, and Performance of a Highway Embankment Failure Repaired with Tire-Derived Aggregate. Transportation Research Record, (2170):90-99.
[22] Irsyam M, Susila E, Himawan A. (2007) Slope Failure of an Embankment on Clay Shale at km 97+500 of the Cipularang Toll Road and the Selected Solution. In: Int. Symp. Geotech. Eng. Gr. Improv. Geosynth. Hum. Secur. Environ. Preserv. Bangkok, Thailand, pp 531-540.
[23] Hadjigeorgiou J, Kyriakou E, Papanastasiou P. (2006) A Road Embankment Failure Near Pentalia in Southwest Cyprus. In: Int. Symp. Stab. Rock Slopes Open Pit Min. Civ. Eng. The South African Institute of Mining and Metallurgy, Cape Town, pp 343-352.
[24] Davis EH, Booker JR. (1973) The effect of increasing strength with depth on the bearing capacity of clays. Géotechnique, 23(4):551-563.
[25] Michalowski RL. (1992) Bearing Capacity of Nonhomogeneous Cohesive Soils Under Embankments. Journal of Geotechnical Engineering, 118(7):1098-1118.
[26] Michalowski RL. (1993) Bearing Capacity of Nonhomogeneous Clay Layers under Embankments. Journal of Geotechnical Engineering, 119(10):1657-1669.
[27] Indraratna B, Balasubramaniam AS, Balachandran S. (1992) Performance of test embankment constructed to failure on soft marine clay. Journal of Geotechnical Engineering, 118(1):12-33.
[28] Eide O, Holmberg S. (1972) Test fills to failure on soft Bangkok clay. In: Spec. Conf. Perform. Earth Earth-Supported Struct. ASCE, Lafayette, Indiana, United States, pp 159-180.
[29] Dascal O, Tournier JP, Tavenas F, Rochelle P La. (1972) Failure of a test embankment on sensitive clay. In: Spec. Conf. Perform. Earth Earth-Supported Struct. ASCE, Lafayette, Indiana, United States, pp 129-158.
[30] Popescu R, Deodatis G, Nobahar A. (2005) Effects of random heterogeneity of soil properties on bearing capacity. Probabilistic Engineering Mechanics, 20(4):324-341.
[31] Lehtonen VJ, Meehan CL, Länsivaara TT, Mansikkamäki JN. (2015) Full-scale embankment failure test under simulated train loading. Géotechnique, 65(12):961-974.
[32] Kim BB, Lee S. (2005) Comparison of Bearing Capacity Characteristics of Sand and Gravel Compaction Pile Treated Ground. KSCE Journal of Civil Engineering, 9(3):197-203.
[33] Lopez-Querol S, Arias-Trujillo J, GM-Elipe M, Matias-Sanchez A, Cantero B. (2017). Improvement of the bearing capacity of confined and unconfined cement-stabilized aeolian sand. Construction and Building Materials, 153:374-384.
[34] Rowe RK, Li AL. (2005) Geosynthetic-reinforced embankments over soft foundations.

Geosynthetics International, 12(1):50-85.
[35] Hewlett WJ, Randolph MF. (1988) Analysis of piled embankment. Ground Engineering, 21(3):12-18.
[36] Sarsby RW. (2007) Use of 'Limited Life Geotextiles' (LLGs) for basal reinforcement of embankments built on soft clay. Geotextiles and Geomembranes, 25(4-5):302-310.
[37] Ozdemir MA. (2016) Improvement in Bearing Capacity of a Soft Soil by Addition of Fly Ash. Procedia Engineering, 143(Ictg):498-505.
[38] Brabb EE. (1991) The world landslide problem. Episodes, 14(1):52-61.
[39] Ballantyne CK. (1986) Landslides and slope failures in Scotland: A review. Scottish Geographical Magazine, 102(3):134-150.
[40] Dai FC, Lee CF, Li J, Xu ZW. (2001) Assessment of landslide susceptibility on the natural terrain of Lantau Island, Hong Kong. Environmental Geology, 40(3):381-391.
[41] Davies TC. (1996) Landslide research in Kenya. Journal of African Earth Sciences, 23(4):541-545.
[42] Tang C, Zhu J, Qi X, Ding J. (2011) Landslides induced by the Wenchuan earthquake and the subsequent strong rainfall event: A case study in the Beichuan area of China. Engineering Geology, 122(1-2):22-33.
[43] Xu XZ, Liu ZY, Xiao PQ, Guo WZ, Zhang HW, Zhao C, Yan Q. (2015) Gravity erosion on the steep loess slope: Behavior, trigger and sensitivity. Catena, 135:231-239.
[44] Yong L, Chengmin H, Baoliang W, Xiafei T, Jingjing L. (2017) A unified expression for grain size distribution of soils. Geoderma, 288:105-119.
[45] Kinnell PIA. (2000) The effect of slope length on sediment concentrations associated with side-slope erosion. Soil Science Society of America Journal, 64(3):1004-1008.
[46] Hou Q, Wu G, Li H, Fan G, Zhou J. (2019) Large deformation and failure mechanism analyses of Tangba high slope with a high-intensity and complex excavation process. Journal of Mountain Science, 16(2):453-469.
[47] Duncan JM. (1996) Soil slope stability analysis. In: Turer KA, Schuster RI (eds) Landslides, Investig. Mitig. National Research Council, Washington, DC, US., pp 337-371.
[48] Huat BBK, Ali FH, Low TH. (2006) Water infiltration characteristics of unsaturated soil slope and its effect on suction and stability. Geotechnical & Geological Engineering, 24(5):1293-1306.
[49] Terzaghi K. (1925) Erdbaumechanik. Franz Deuticke, Leipzig and Vienna .
[50] Kong X. (2016) Prediction of Subgrade Settlement Based Fuzzy Self Adaptable Method of Artificial Intelligence. 9th International Symposium on Computational Intelligence and Design, (7):144-147.
[51] Zou C, Wang Y, Lin J, Chen Y. (2016) Creep behaviors and constitutive model for high density polyethylene geogrid and its application to reinforced soil retaining wall on soft soil foundation. Construction and Building Materials, 114:763-771.
[52] Zou L, Wang S, Lai X. (2013) Creep model for unsaturated soils in sliding zone of Qianjiangping landslide. Journal of Rock Mechanics and Geotechnical Engineering, 5(2):162–167.
[53] Azzam WR, Basha A. (2017) Utilization of soil nailing technique to increase shear strength of cohesive soil and reduce settlement. Journal of Rock Mechanics and Geotechnical Engineering, 9(6):1104-1111.
[54] Grimstad G, Haji Ashrafi MA, Degago SA, Emdal A, Nordal S. (2016) Discussion of 'Soil creep effects on ground lateral deformation and pore water pressure under embankments'. Geomechanics and Geoengineering, 11(1):86-93.
[55] Sato H, Nhan TT, Matsuda H. (2018) Earthquake-induced settlement of a clay layer. Soil Dynamics and Earthquake Engineering, 104(May 2017):418-431.
[56] LaGatta MD, Boardman BT, Cooley BH, Daniel DE. (1997) Geosynthetic Clay Liners Subjected to Differential Settlement. Journal of Geotechnical and Geoenvironmental Engineering, 123(5):402-410.
[57] Paveenchana T, Arayasiri M. (2009) Solving the Problems of Differential Settlement of Pavement Structures in the Bangkok Area. In: GeoHunan Int. Conf. 2009. American Society of Civil Engineers, Changsha, Hunan, China, pp 180-185.

[58] Moghaddam RB, Jayawickrama PW. (2017) General bearing capacity theory and soil extraction method for the mitigation of differential settlements. In: Geotech. Front. 2017. American Society of Civil Engineers, Orlando, Florida, pp 314-323.
[59] Skempton AW, Bjerrum L. (1957) A Contribution to the settlement analysis of foundations on clay. Géotechnique, 7(4):168-178.
[60] Bo MW, Fabius M, Chu J, A. Arulrajah. (2011) Predicting consolidation settlements using small strain, large-strain and stress path methods. In: Proc. 17th Int. Conf. Soil Mech. Geotech. Eng. pp 3481-3485.
[61] Boussinesq MJ. (1885) Application Des Potentiels. Gauthier-Villars, Paris [in French]: à l'Étude de l'Éqilibre et du Movvement Des Solides Elastiques.
[62] Disfani MM, Arulrajah A, Suthagaran V, Bo MW. (2013) Long-term settlement prediction for wastewater biosolids in road embankments. Resources, Conservation and Recycling, 77:69-77.
[63] Duncan M. (1993) Limitation of conventional analysis of consolidation settlement. Journal of Geotechnical Engineering, 119(9):1333-1359.
[64] Carroll R. (2011) Use of CRS test to predict settlement in an Irish silt. 21st Eur. Young Geotech. Eng. Conf. Rotterdam .
[65] Skempton AW, Peck RB, MacDonald DH. (1955) Settlement analyses of six structures in Chicago and London. Proceedings Institution of Civil Engineers, 4(4):525-542.
[66] Bergado DT, Teerawattanasuk C. (2008) 2D and 3D numerical simulations of reinforced embankments on soft ground. Geotextiles and Geomembranes, 26(1):39-55.
[67] Bergado D., Chong KC, Daria PAM, Alfaro MC. (1990) Deformability and consolidation characteristics of soft Bangkok clay using screw plate tests. Canadian Geotechnical Journal, 27(5): 531-545.
[68] Poulos HG, Davis EH. (1963) Triaxial testing and three dimensional settlement analysis. In: Proc. 4th Aust. New Zeal. Conf. SM FE, Adelaide. pp 233-243.
[69] Chai J chun, Shrestha S, Hino T, Uchikoshi T. (2017) Predicting bending failure of CDM columns under embankment loading. Computers and Geotechnics, 91:169-178.
[70] Smith CC, Tatari A. (2016) Limit analysis of reinforced embankments on soft soil. Geotextiles and Geomembranes, 44(4):504-514.
[71] Buttling S, Cao R, Lau W, Naicker D. (2018) Class A and Class C numerical predictions of the deformation of an embankment on soft ground. Computers and Geotechnics, 93:191-203.
[72] Keller GR. (2016) Application of geosynthetics on low-volume roads. Transportation Geotechnics, 8(April):119-131.
[73] Yapage N, Liyanapathirana S. (2018) Behaviour of geosynthetic reinforced column supported embankments. Journal of Engineering, Design and Technology, 16(1):44-62.
[74] Zhou WH, Lao JY, Huang Y, Chen R. (2017) Group effect on soil arching in geogrid-reinforced pile-supported embankments. Japanese Geotechnical Society Special Publication, 5(2):130-134.
[75] Hegde AM, Sitharam TG. (2015) Effect of infill materials on the performance of geocell reinforced soft clay beds. Geomechanics and Geoengineering, 10(3):163-173.
[76] Umashankar B, Mouli S, Hariprasad C. (2015) Settlement of Embankment Constructed with Geofoam. In: Iskander M, Suleiman MT, Anderson JB, Laefer DF (eds) Int. Found. Congr. Equip. Expo 2015. American Society of Civil Engineers, pp 161-170.
[77] Paolo Di Pietro. (2017) Practical Applications with Geosynthetic Mats Reinforced with Steel Wire Meshes to Prevent Embankment Damage by Burrowing Large Rodents and Beavers. Journal of Civil Engineering and Architecture, 11(1):8-15.
[78] Chaiyaput S, Bergado DT, Artidteang S. (2014) Measured and simulated results of a Kenaf Limited Life Geosynthetics (LLGs) reinforced test embankment on soft clay. Geotextiles and Geomembranes, 42(1):39-47.
[79] Marto A, Othman BA, Kasim F, Bakar I. (2012) Comparison of Field Performance between Bamboo-Geotextile Composite Embankment and High Strength Geotextile Embankment. Advanced Materials Research, 587:77-80.
[80] Lal D, Sankar N, Chandrakaran S. (2017) Effect of reinforcement form on the behaviour of

coir geotextile reinforced sand beds. Soils and Foundations, 57(2):227-236.
[81] Da Silva EM, Justo JL, Durand P, Justo E, Vázquez-Boza M. (2017) The effect of geotextile reinforcement and prefabricated vertical drains on the stability and settlement of embankments. Geotext Geomembranes. doi: 10.1016/j.geotexmem.2017.07.001.
[82] Zhuang Y, Wang KY. (2016) Finite-Element Analysis on the Effect of Subsoil in Reinforced Piled Embankments and Comparison with Theoretical Method Predictions. International Journal of Geomechanics, 16(5):04016011-1–15.
[83] Parsa-Pajouh A, Fatahi B, Vincent P, Khabbaz H. (2014). Trial Embankment Analysis to Predict Smear Zone Characteristics Induced by Prefabricated Vertical Drain Installation. Geotechnical and Geological Engineering, 32(5):1187-1210.
[84] Sheil BB, McCabe BA. (2014) A finite element-based approach for predictions of rigid pile group stiffness efficiency in clays. Acta Geotechnica, 9(3):469-484.
[85] Rezania M, Nguyen H, Zanganeh H, Taiebat M. (2018) Numerical analysis of Ballina test embankment on a soft structured clay foundation. Computers and Geotechnics, 93:61-74.
[86] Xue JF, Chen JF, Liu JX, Shi ZM. (2014) Instability of a geogrid reinforced soil wall on thick soft Shanghai clay with prefabricated vertical drains: A case study. Geotextiles and Geomembranes, 42(4):302-311.
[87] Rabie M. (2014) Comparison study between traditional and finite element methods for slopes under heavy rainfall. HBRC Journal, 10(2):160–168.
[88] Ghafari A, Nikraz HR, Sanaeirad A. (2016) Finite element analysis of deformation and arching inside the core of embankment dams during construction. Australian Journal of Civil Engineering, 14(1):13-22.
[89] Zhao L, Yang F, Zhang Y, Dan H, Liu W. (2015) Effects of shear strength reduction strategies on safety factor of homogeneous slope based on a general nonlinear failure criterion. Computers and Geotechnics, 63:215-228.
[90] Duncan JM. (1996) State of the Art: Limit Equilibrium and Finite-Element Analysis of Slopes. Journal of Geotechnical Engineering, 122(7):577-596.
[91] Kirkby W, Pickett PE. (2018). Geotechnical Manual. Texas Department of Transportation.
[92] Singh TN, Verma AK, Sarkar K. (2010) Static and dynamic analysis of a landslide. Geomatics, Natural Hazards and Risk, 1(4):323-338.
[93] Iñeguez JB. (2016) Enhanced Limit Method for Slope Stability Analysis. Electronic Journal of Geotechnical Engineering, 21(26):10215-10232.
[94] Fellenius W. (1936) Calculation of the stability of earth slope. In: Trans. 2nd Congr. Large Dams. Washington, DC, US., pp 445-462.
[95] Bishop AW. (1955) The use of slip circle in the stability analysis of slopes. Geotechnique, 5(1):7-17.
[96] E. Spencer. (1967) A method of analysis of the stability of embankments assuming parallel interslice forces. Géotechnique, 17(1):11-26.
[97] Morgenstern NR, Price VE. (1965) The Analysis of the Stability of General Slip Surfaces. Géotechnique, 15(1):79–93.
[98] Liu SY, Shao LT, Li HJ. (2015) Slope stability analysis using the limit equilibrium method and two finite element methods. Computers and Geotechnics, 63:291-298.
[99] Guo T, He Z. (2012) Comparison of factor of safety of a roadway slope based on the limit equilibrium method and shear strength reduction method. In: GeoHunan Int. Conf. American Society of Civil Engineers, Hunan, China, pp 34-40.
[100] Luo N, Bathurst RJ, Javankhoshdel S. (2016) Probabilistic stability analysis of simple reinforced slopes by finite element method. Computers and Geotechnics, 77:45–55.
[101] Alemdag S, Kaya A, Karadag M, Gurocak Z, Bulut F. (2015) Utilization of the limit equilibrium and finite element methods for the stability analysis of the slope debris: An example of the Kalebasi District (NE Turkey). Journal of African Earth Sciences, 106:134-146.
[102] Liu KW, Rowe RK. (2015) Numerical modelling of prefabricated vertical drains and surcharge on reinforced floating column-supported embankment behaviour. Geotextiles and Geomembranes, 43(6):493-505.
[103] Ishii Y, Ota K, Kuraoka S, Tsunaki R. (2012) Evaluation of slope stability by finite element

method using observed displacement of landslide. Landslides, 9(3):335-348.
[104] Zienkiewicz OC, Humpheson C, Lewis RW. (1975) Associated and non-associated visco-plasticity and plasticity in soil mechanics. Géotechnique, 25(4):671-689.
[105] Sun C, Chai J, Xu Z, Qin Y, Chen X. (2016) Stability charts for rock mass slopes based on the Hoek-Brown strength reduction technique. Engineering Geology, 214:94-106.
[106] Griffiths DV, Lane P a. (1999) Slope stability analysis by finite elements. Geotechnique, 49(3):387-403.
[107] Kulhawy FH. (1969) Finite element analysis of the behavior of embankments. The University of California at Berkeley, California. U.S.A.
[108] Korea Expressway Corporation. (1996) Road design manual, soil and foundation (in Korean). Korea Expressway Corporation.
[109] Public Works Department Malaysia. (2010) Guideline for slope design..
[110] Ministry of Work Tanzania. (2011) Road geometric design manual..
[111] Texas Department of Transportation. (2014) Roadway Design Manual..
[112] New Jersey Department of Transportation. (2015) Roadway design manual..
[113] Fagundes D de F, Almeida M de SS de, Girout R, Blanc M, Thorel L. (2015) Behaviour of Piled Embankment Without Reinforcement. Proceedings of the Institution of Civil Engineers - Geotechnical Engineering, 168(6):514-525.
[114] Bhasi A, Rajagopal K. (2014) Geosynthetic-Reinforced Piled Embankments: Comparison of Numerical and Analytical Methods. International Journal of Geomechanics, 15(5):1-12.
[115] Chen Y, Qi C, Xu H, Ng CWW. (2013) Field Test Research on Embankment Supported by Plastic Tube Cast-in-place Concrete Piles. Geotechnical and Geological Engineering, 31(4):1359-1368.
[116] Kasim F, Marto A, Othman BA, Bakar I, Othman MF. (2013) Simulation of Safe Height Embankment on Soft Ground Using Plaxis. APCBEE Procedia, 5:152-156.
[117] Wang S, Qi J, Liu F. (2016) Study on the Reasonable Height of Embankment in Qinghai–Tibet Highway. Geotechnical and Geological Engineering, 34(1):1-14.
[118] Lu Z, Wang H, Yao H. (2014) Method to determine height and water content of highway subgrade filled with fine grained soil materials. Materials Research Innovations, 18(sup5):S5-5-S5-8.
[119] Cui X, Zhang N, Li S, Zhang J, Wang L. (2015) Effects of embankment height and vehicle loads on traffic-load-induced cumulative settlement of soft clay subsoil. Arabian Journal of Geosciences, 8(5):2487-2496.
[120] Almeida MSS, Riccio M, Hosseinpour I. (2013) Performance of a geosynthetic-encased column (GEC) in soft ground: numerical and analytical studies. Geosynthetics International, 20(4):252-262.
[121] Filz G, Sloan J, McGuire MP. (2012) Column-Supported Embankments: Settlement and Load Transfer. GeoCongress 2012 Geotech Eng State Art Pract. doi: 10.1007/s13398-014-0173-7.2.
[122] McGuire MP. (2011) Critical Height and Surface Deformation of Column-Supported Embankments. Virginia Polytechnic Institute and State University, Blacksburg, VA.
[123] BS 8006. (2010) Code of practice for strengthened / reinforced soils and other fills. Br Stand Inst. doi: BS 8006:1995.
[124] Chen Jia F, Hai Bin W, Bao Ping A, Zhang P, Yang Peng Z. (2013) Application of BP neural network embankment settlement prediction in seasonal frozen areas. In: 2013 4th Int. Conf. Digit. Manuf. Autom. ICDMA 2013. pp 276-279.
[125] Zhang J, Zhang M, Lui Y. (2006) Study on the reasonable embankment height of Qinghai-Tibet railway in permafrost regions. China Railway Science, 27(5):28-34.
[126] Yue Z, Ge J, Li Z, Liu Y. (2007) Study on settlement of unprotected railway embankment in permafrost. Cold Regions Science and Technology, 48(1):24-33.
[127] Zhang J, Zhang M, Lui Y. (2009) Reasonable Height of Roadway Embankment in Permafrost Regions. In: 14th Conf. Cold Reg. Eng. American Society of Civil Engineers, pp 486-495.
[128] Qi C, Wu Q, Wu J. (2007) Analysis of Thermal State of Permafrost Under High Embankment Along Qinghai–Tibet Railway. Chinese Journal of Rock Mechanics and

Engineering, 26(S2):4518-4524.
[129] Li G, Li N, Kang J, Niu F, Yu W, Shi L, Bi G. (2008) Study on design optimization of a crushed stone layer with shading board placed on a railway embankment on warm permafrost. Cold Regions Science and Technology, 54(1):36-43.
[130] Sinha AK, Havanagi VG, Mathur S. (2007) Inflection point method for predicting settlement of PVD improved soft clay under embankments. Geotextiles and Geomembranes, 25(6):336-345.
[131] Jin L, Wang S, Chen J, Dong Y. (2012) Study on the height effect of highway embankments in permafrost regions. Cold Regions Science and Technology, 83-84:122-130.
[132] Wang S, Chen J, Qi J. (2009) Study on the technology for highway construction and engineering practices in permafrost regions. Sciences in Cold and Arid Regions, 1(5):412-422.
[133] Aljanabi QA, Chik Z, Allawi MF, El-Shafie AH, Ahmed AN, El-Shafie A. (2018) Support vector regression-based model for prediction of behavior stone column parameters in soft clay under highway embankment. Neural Computing and Applications, 30(8): 2459–2469.
[134] Artidteang S, Chaiyaput S, Bergado DT, Tanchaisawat T. (2015) Embankment reinforced with limited life geotextiles on soft clay. Ground Improvement, 168(2):130-143.
[135] Zhuang Y, Wang K. (2017) Finite element analysis on the dynamic behavior of soil arching effect in piled embankment. Transportation Geotechnics, 14:8-21.
[136] Dafalla MA. (2009) Improvement of coastal silty sand of Saudi Arabia using preloading technique. In: GeoHunan Int. Conf. American Society of Civil Engineers, Changsha, Hunan, China, pp 100-105.
[137] Chan KF, Poon BM, Perera D. (2018) Prediction of embankment performance using numerical analyses – Practitioner's approach. Computers and Geotechnics, 93:163-177.
[138] Kjellman W. (1948) Accelerating consolidation of fine grained soils by means of cardboard wicks. In: 2nd Int. Conf. Soil Mech. Found. Eng. Rotterdam, pp 302-305.
[139] Chu J, Bo MW, Choa V. (2006) Improvement of ultra-soft soil using prefabricated vertical drains. Geotextiles and Geomembranes, 24(6):339-348.
[140] Yang H, Xiao J, He Y. (2010) An Economical, Practical, and Environmental Friendly Surcharge Preloading Method to Improve Soft Ground of Municipal Road. In: GeoShanghai Int. Conf. American Society of Civil Engineers, pp 267-272.
[141] Sun L, Gao X, Zhuang D, Guo W, Hou J, Liu X. (2018) Pilot tests on vacuum preloading method combined with short and long PVDs. Geotextiles and Geomembranes, 46(2):243-250.
[142] Zhang Z, Ye GB, Xu Y. (2018) Comparative analysis on performance of vertical drain improved clay deposit under vacuum or surcharge loading. Geotextiles and Geomembranes, 46(2):146-154.
[143] Schaefer VR, Berg RR, Collin JG, Christopher BR, DiMaggio JA, Filz GM, Bruce DA, Ayala D. (2017) Ground modification methods reference manual – Volume I. U.S. Dep. Transp. Fed. Highw. Adm. FHWA-NHI-16-027 I:
[144] Stark TD, Horvath JS, Leshchinsky D. (2004) Guideline and Recommended Standard for Geofoam Applications in Highway Embankments. Washington, D.C.
[145] Panesar DK. (2013) Cellular concrete properties and the effect of synthetic and protein foaming agents. Construction and Building Materials, 44:575-584.
[146] Onprom P, Chaimoon K, Cheerarot R. (2015) Influence of Bottom Ash Replacements as Fine Aggregate on the Property of Cellular Concrete with Various Foam Contents. Advances in Materials Science and Engineering, 2015:381704.
[147] Struhárová A, Rouseková I. (2007) Porous Structure of Cellular Concrete and its Impact on Selected Physical- Mechanical Properties of Cellular Concrete. Slovak Journal of Civil Engineering, 2:35-43.
[148] Hu W, Neufeld RD, Vallejo LE, Kelly C, Latona M. (1997) Strength Properties of Autoclaved Cellular Concrete with High Volume Fly Ash. Journal of Energy Engineering, 123(2):44-54.
[149] Marradi a, Pinori U, Betti G. (2012) The Use of Lightweight Materials in Road Embankment Construction. Procedia - Social and Behavioral Sciences, 53(0):1000-1009.

[150] Elias V, Welsh J, Warren J, Lukas R. (1999) Ground improvement technical summaries. Washington DC, US: Department of Transportation Federal Highway Administration.
[151] ASTM D6817 / D6817M-17. (2017) Standard Specification for Rigid Cellular Polystyrene Geofoam. ASTM Int. doi: 10.1520/D6817_D6817M-17.
[152] Koerner RM. (2012) Designing with geosynthetics, 6th ed. Bloomington, IN: Xlibris Corporation.
[153] Beju YZ, Mandal JN. (2017) Expanded Polystyrene (EPS) Geofoam: Preliminary Characteristic Evaluation. In: Procedia Eng. pp 239-246.
[154] Trandafir AC, Bartlett SF, Lingwall BN. (2010) Behavior of EPS geofoam in stress-controlled cyclic uniaxial tests. Geotextiles and Geomembranes, 28(6):514-524.
[155] Özer AT, Akay O. (2016) Interface Shear Strength Characteristics of Interlocked EPS-Block Geofoam. Journal of Materials in Civil Engineering, 28(4):04015156.
[156] Shong D, Sutmoller N. (2011) Project showcase: Highway reconstruction on sof soils not a problem with geofoam. Geosynthetics, 29(2):14-17.
[157] Michalowski RL, Asce F, Wojtasik A, Duda A, Florkiewicz A, Park D, Asce SM. (2018) Failure and Remedy of Column-Supported Embankment : Case Study. Journal of Geotechnical and Geoenvironmental Engineering, 144(3):1-14.
[158] Newman MP, Bartlett SF, Lawton EC. (2010) Numerical Modeling of Geofoam Embankments. Journal of Geotechnical and Geoenvironmental Engineering, 136:290-298.
[159] Bartlett SF, Lawton EC, Farnsworth CB, Perry MN. (2012) Design and Evaluation of Expanded Polystyrene Geofoam Embankments for the I-15 Reconstruction Project, Salt Lake City, Utah. Salt Lake City,US.
[160] Kim T-H, Lee HW, Hong S-W. (2016) Value engineering for roadway expansion project over deep thick soft soils. Journal of Construction Engineering and Management, 142(2):1-9.

https://doi.org/10.31436/iiumej.v20i2.1075

# LIFE CYCLE COST COMPARISON BETWEEN PERFORMANCE BASED AND TRADITIONAL CONTRACTS FOR ROADS IN INDONESIA

**BETTY SUSANTI[1*], REINI D. WIRAHADIKUSUMAH[2], BIEMO W. SOEMARDI[2], MEI SUTRISNO[3]**

[1]*Dept. of Civil Engineering, Universitas Sriwijaya, Palembang, Indonesia.*
[2]*Dept. of Civil Engineering, Institut Teknologi Bandung, Bandung, Indonesia.*
[3]*Dept. of Civil Engineering, Politeknik Negeri Bandung, Bandung, Indonesia.*

[*]*Corresponding author: bettysusanti0401@gmail.com*

*(Received: 2nd February 2019; Accepted: 30th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** Implementation of Performance-based Contracts (PBC) for road maintenance in Indonesia still requires various studies, especially related to potential long-term cost efficiencies that can be achieved by the road agencies by shifting the application of traditional contracts to PBC. This study assesses the effectiveness of PBC compared to traditional contract implementation based on a case study on one of the PBC pilot projects in Indonesia. The effectiveness of the PBC reviews in terms of project life cycle cost efficiency was calculated. LCC calculations were conducted by considering the influence of the length of contracted road, the initial conditions of contracted road, and the duration of PBC project. The LCC calculation shows that the implementation of PBC generally has the potential to generate LCC efficiency compared to the traditional contract. This study also proved that the implementation of PBC can guarantee the performance of road services during multi-year contracts and also provide other benefits for the road agency and road users.

***ABSTRAK:*** Pelaksanaan Kontrak Berasaskan Prestasi (PBC) untuk penyelenggaraan jalan raya di Indonesia masih memerlukan berbagai kajian, terutama yang berkaitan dengan potensi efisiensi biaya jangka panjang yang dapat dicapai oleh pihak agensi jalan dengan mengalihkan penerapan kontrak tradisional ke PBC. Kajian ini menilai keberkesanan PBC berbanding dengan pelaksanaan kontrak tradisional berdasarkan kajian kes pada salah satu projek perintis PBC di Indonesia. Keberkesanan ulasan PBC dari segi kecekapan kos kitaran hayat projek. Pengiraan LCC dilakukan dengan mempertimbangkan pengaruh panjang jalan yang dikontrak, keadaan awal jalan terkontrak, dan jangka waktu projek PBC. Pengiraan LCC menunjukkan bahawa pelaksanaan PBC secara amnya berpotensi untuk menghasilkan kecekapan LCC berbanding dengan kontrak tradisional. Kajian ini juga membuktikan bahawa pelaksanaan PBC dapat menjamin prestasi perkhidmatan jalan selama kontrak multi-tahun dan juga memberikan manfaat lain untuk agensi jalan dan pengguna jalan raya.

***KEYWORDS:*** *life-cycle cost; maintenance; performance-based contract; road; traditional contract*

## 1. INTRODUCTION

National road maintenance projects in Indonesia are generally delivered using in-house systems and input-based contracts with a Design-Bid-Build (DBB) method. An in-house system is an approach of self-managed routine maintenance project implementation by a

https://doi.org/10.31436/iiumej.v20i2.1075

road agency. This system is only applied for maintenance of roads in good condition. In the in-house system, the road agency plays its role directly in the process of implementing the physical works of maintenance and project supervision. Meanwhile, the DBB pattern is an approach for implementing the maintenance of road that is contracted to the construction service provider. This pattern is applied on non-routine road maintenance work such as road improvement works, periodic maintenance, and road rehabilitation. Both in-house and DBB systems are generally denoted as part of the traditional contract approach. This traditional approach applies a prescriptive or input-based system. It is performed in a single year with a short maintenance warranty period. The mechanism of road maintenance always applies a corrective approach and it is implemented in the packages of the contract with a small value. Many studies have shown that these traditional contracts generate high costs and inadequate road quality [1,2].

Although traditional contracts remain the most commonly used type of contract in Indonesia, there are currently efforts being made that focus on the application of non-traditional contracts to fund and perform road maintenance projects. This non-traditional contract, or an alternative contract, generally focuses on the output or the outcome of the project. It is performed on a multi-year basis with a long maintenance warranty period. Additionally, the contract integrates various maintenance works in the maintenance project package. There are various types of non-traditional contract for road maintenance namely Design-Build, extended warranty contract, lane rental, cost plus time, and performance-based contract. Some previous studies showed that one alternative to traditional contracts that can overcome a problem related to the high amount of cost and the low road quality is Performance-Based Contract (PBC), as suggested by [3-5].

PBC is perceived as *Output and Performance-based Road Contract* [6]. This contract regulates the road service performance that should be achieved during the contract period in a multi-year contract. Consequently, the road maintenance strategy is preventive. Design works, construction, or road maintenance are carried out in an integrated way by one service provider. The scope of work that is contracted generally includes a relatively long road. With those characteristics stated above, PBCs have the potential to reduce maintenance costs and improve road performance compared to traditional contracts [3,4]. However, various efforts to assess the effectiveness of PBC implementation for road maintenance projects are limited to the assessment of the potential cost savings during the contract period [3,4]. Those approaches have not been based on the assessment of long-term PBC potential cost savings. The appropriate approach to determine cost savings of PBC is to evaluate the actual cost before and after PBC implementation or by making estimates based on road deterioration and life-cycle cost models [7].

Several studies have been conducted regarding the implementation of PBC in Indonesia. Reference [8] identified that the development of methods to estimate the cost efficiency of PBC is required, while [9] emphasized the need to develop PBC quantitative evaluation schemes to justify the claims of efficiency and effectiveness of the use of PBC in national road management in Indonesia. However, quantitative studies assessing the effectiveness of PBC implementation have not been conducted. The objective of this study was to assess the effectiveness of the PBC implementation in road maintenance projects compared to the traditional approach based on a case study of a PBC pilot project in Indonesia. The effectiveness of PBC was reviewed in terms of project's life cycle cost efficiency.

https://doi.org/10.31436/iiumej.v20i2.1075

## 2. CASE STUDY: PBC PILOT PROJECT IN INDONESIA

The implementation of PBC for road maintenance projects in Indonesia started in 2011. Currently, six PBC pilot projects in Indonesia are being carried out (Table 1), all of which are national road improvement works. Of the six projects, five are on roads with high traffic volume and overloads located on Java Island, whereas one project is in the Sei Hanyu–Tumbang Lahung segment on Kalimantan Island, a road with a low volume of traffic. The implementation of PBC is currently more applicable to national roads on Java Island because of the readiness of road agencies and availability of adequate historical data to support PBC implementation.

Table 1: PBC pilot projects in Indonesia

| No | Road Segment | Road Length (kilometers) | Contract Period | Project Location |
|---|---|---|---|---|
| 1 | Ciasem-Pamanukan | 18.50 | 2011-2015 | North coast lane of Java Island |
| 2 | Demak-Trengguli | 7.68 | | |
| 3 | Semarang-Bawen | 22 | | |
| 4 | Bojonegoro-Padangan | 11 | 2012-2019 | Central lane of Java Island |
| 5 | Padangan-Ngawi | 10.70 | | |
| 6 | Sei Hanyu-Tumbang Lahung | 50.60 | 2013-2020 | Kalimantan Island |

This study was conducted on one PBC pilot project at one section of a national road on the northern coast of West Java, i.e. Ciasem-Pamanukan section (Fig. 1). The contract was completed in 2015. The project was for road improvement work along 18.5 kilometers of road with a pavement width of 14 meters. The type of pavement consisted of flexible and rigid pavements on the fast and slow lanes, respectively. The road agency is the owner of the Karawang-Cikampek-Pamanukan section, whose territory is 74.24 kilometers of the total 273.31 kilometers of the northern coast of West Java Section.

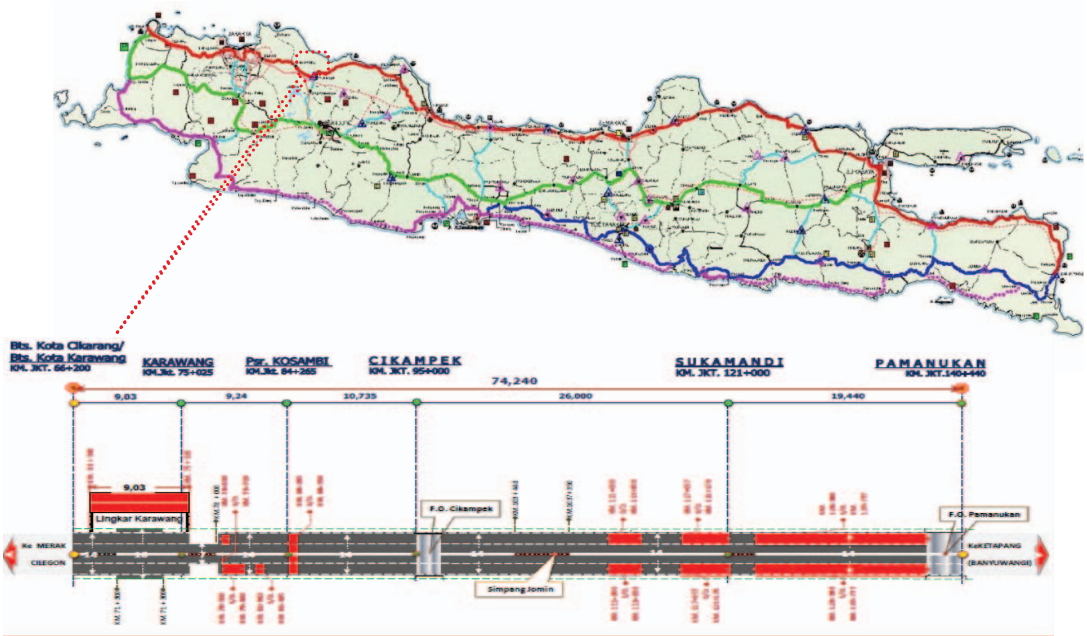



Fig. 1: Location of Ciasem-Pamanukan section.

https://doi.org/10.31436/iiumej.v20i2.1075

National road sections on Java Island have a higher traffic volume and traffic loads compared to roads on other islands, especially on the north coast lane of Java. Reference [10] showed that the average daily traffic on Ciasem-Pamanukan section in 2011 or first year of contract was more than 46 million vehicles in both directions, where the traffic volume is distributed mostly on the fast lane with flexible pavement. Approximately, 60% of heavy vehicles are distributed in the fast lane [11].

## 3. LIFE CYCLE COST (LCC) ANALYSIS

The LCC analysis refers to the models that have been developed [12-14]. In line with the PBC concept that is applied for road maintenance works, the scope of work reviewed in this study was limited to road maintenance works, which consist of routine maintenance, periodic maintenance, and road betterment. The LCC analysis simply considered the cost components of the road agency. They are the total cost that should be spent by the government as the road agency to implement all works related to the road maintenance project [15]. The component of road agency cost for each contract approach that is considered in this study is shown in Table 2. The type of pavements reviewed were flexible pavements as most of the traffic volume was distributed on these.

Table 2: The components of road agency costs

| **Type of road works** | **Road Agency Cost for Traditional Contract Approach** | |
|---|---|---|
| | **Delivery System** | **Road agency costs components** |
| Routine maintenance | In-house | Routine maintenance cost |
| Periodic maintenance | D-B-B | Periodic maintenance cost |
| | | Contractor procurement cost |
| | | Design cost |
| | | Designer procurement cost |
| | | Supervision cost |
| | | Supervisor procurement cost |
| Road improvement | D-B-B | Road improvement cost |
| | | Contractor procurement cost |
| | | Design cost |
| | | Designer procurement cost |
| | | Supervision cost |
| | | Supervisor procurement cost |
| Type of road works | | Road Agency Cost for PBC |
| | Delivery System | Road agency costs components |
| Design and road maintenance (integrated) | D-B-O-M | Integrated procurement costs (for designer and contractor) |
| | | Integrated costs for design and road works |
| | | Procurement cost for external supervision |
| | | Supervision cost |

This study applied the LCC method with a deterministic approach. To overcome the uncertainty associated with an estimated present value of LCC, we applied deterministic sensitivity analysis techniques in order that the estimation indicate the variability of LCC for road maintenance projects performed using different types of contracts. According to [3] and [4], the characteristics of a project applying a PBC scheme—such as the length and initial condition of the road, the duration of the contract, the number and type of activities

https://doi.org/10.31436/iiumej.v20i2.1075

comprised in the contract package, the extension or addition of contract period, and the performance indicators applied—affects its cost-efficiency.

In this study, a deterministic sensitivity analysis associated with the length and initial condition of the road contracted and duration of the PBC project was performed on 48 scenarios, 36 of which were for PBC projects, and 12 for traditional projects. Three overall scenarios were designed, each with varying road lengths: the first scenario used a road length complying with the scope of the case study project; the contracted road in the second scenario was extended to road networks under the authority of the owner of Karawang-Cikampek-Pamanukan Sections; and the contracted road in the third scenario was extended to the northern coast lane of West Java. For each of these scenarios of different road lengths, the initial road conditions were simulated as being good, fair, minor damage, and heavy damage. For each initial road condition, the PBC project durations were simulated to follow the existing practical implementations of PBC in Indonesia—4 and 7 years—and forthcoming PBC projects—10 years. A summary of all the scenarios used in the LCC calculations are shown in Table 3.

Table 3: Scenarios used in the LCC Analysis

| Road Length (kilometers) | Initial Road Condition | LCC of Traditional Projects | Duration of PBC Projects | LCC of PBC Projects |
|---|---|---|---|---|
| 18.5 | Good | LCC TC-1 | 4-years | LCC PBC-1 |
| | | | 7-years | LCC PBC-2 |
| | | | 10-years | LCC PBC-3 |
| | Fair | LCC TC-2 | 4-years | LCC PBC-4 |
| | | | 7-years | LCC PBC-5 |
| | | | 10-years | LCC PBC-6 |
| | Minor Damage | LCC TC-3 | 4-years | LCC PBC-7 |
| | | | 7-years | LCC PBC-8 |
| | | | 10-years | LCC PBC-9 |
| | Heavy Damage | LCC TC-4 | 4-years | LCC PBC-10 |
| | | | 7-years | LCC PBC-11 |
| | | | 10-years | LCC PBC-12 |
| 74.24 | Good | LCC TC-5 | 4-years | LCC PBC-13 |
| | | | 7-years | LCC PBC-14 |
| | | | 10-years | LCC PBC-15 |
| | Fair | LCC TC-6 | 4-years | LCC PBC-16 |
| | | | 7-years | LCC PBC-17 |
| | | | 10-years | LCC PBC-18 |
| | Minor Damage | LCC TC-7 | 4-years | LCC PBC-19 |
| | | | 7-years | LCC PBC-20 |
| | | | 10-years | LCC PBC-21 |
| | Heavy Damage | LCC TC-8 | 4-years | LCC PBC-22 |
| | | | 7-years | LCC PBC-23 |
| | | | 10-years | LCC PBC-24 |
| 273.31 | Good | LCC TC-9 | 4-years | LCC PBC-25 |
| | | | 7-years | LCC PBC-26 |
| | | | 10-years | LCC PBC-27 |
| | Fair | LCC TC-10 | 4-years | LCC PBC-28 |
| | | | 7-years | LCC PBC-29 |
| | | | 10-years | LCC PBC-30 |
| | Minor Damage | LCC TC-11 | 4-years | LCC PBC-31 |
| | | | 7-years | LCC PBC-32 |
| | | | 10-years | LCC PBC-33 |
| | Heavy Damage | LCC TC-12 | 4-years | LCC PBC-34 |
| | | | 7-years | LCC PBC-35 |
| | | | 10-years | LCC PBC-36 |

https://doi.org/10.31436/iiumej.v20i2.1075

The LCC considers all initial and future costs. All costs included in this LCCA are adjusted to the cost of project implementation for both contract approaches, as shown in Table 2. This LCCA input consists of the historical data of road maintenance costs; surface roughness at the beginning of contract year, number of modified structures on the road, and the cumulative equivalent single axle load (CESAL) value for predicting the road pavement performance; design and construction supervision costs; and administration costs for procuring planning service provider, contractor, and project supervisor.

A schematic diagram of the LCC analysis procedure is shown in Fig. 2. The initial step to calculate the project's LCC is to determine the road condition and maintenance works required (Fig. 2). Initial road conditions were varied to represent good, fair, minor damaged and heavily damaged conditions.

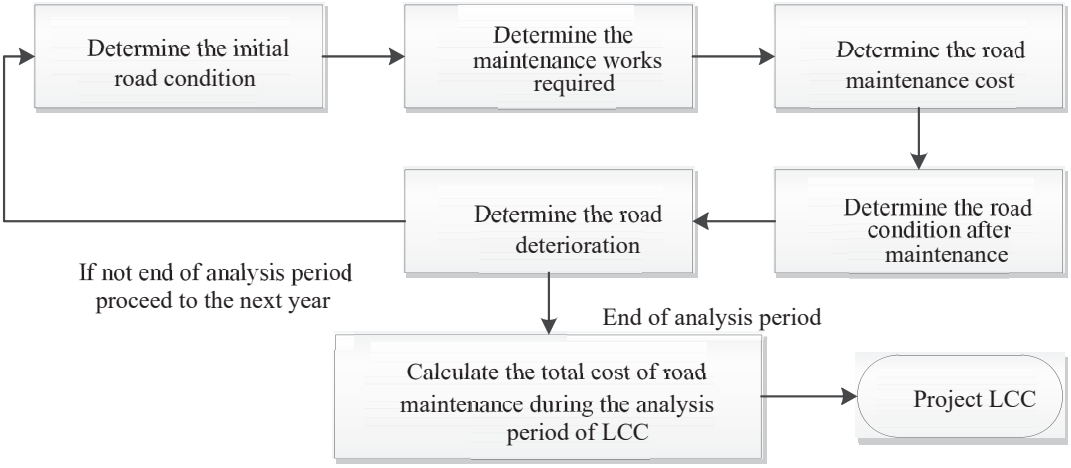



Fig. 2: LCC calculation procedure.

In line with the performance indicators set in the PBC project, the road conditions were measured based on the road's functional performance using a roughness indicator (International Roughness Index, IRI). The road conditions were assumed to be uniform and were expressed using a single IRI value. The Public Works Ministry Regulation on Procedures for Road Maintenance and Surveillance [16] was used to determine the type of maintenance works suitable for the road condition (Table 4). Table 4 is also used to determine the road's conditions after maintenance based on IRI values.

Table 4. Guidelines for determination of road maintenance works

| Road Condition | IRI Value Range | IRI Value Used | Required Works | IRI Value after Maintenance |
|---|---|---|---|---|
| Good | $0 \leq IRI \leq 3.5$ | 1.75 | Routine (preventive) maintenance | Same as initial IRI |
| Fair | $3,5 \leq IRI \leq 6$ | 4.75 | Routine (corrective) maintenance | Previous IRI – 0,5 |
| Minor Damage | $6 \leq IRI \leq 8$ | 7 | Periodic maintenance | 3,00 |
| Heavy Damage | $IRI \geq b$ | 9 | Betterment | 3,00 |

PBCs and traditional approaches have different road maintenance strategies. The maintenance works under a PBC scheme are conducted throughout the year by considering the road conditions and the project's performance indicators, which in this project case study was IRI ≤ 4 m/km. The maintenance work applied to the PBC project may be higher than

https://doi.org/10.31436/iiumej.v20i2.1075

what is required by regulation (Table 4). For example, good road conditions require only routine preventive maintenance, but if this action is implemented, the deterioration prediction for the coming year would exceed the performance indicators set out in the PBC project; and higher maintenance works in this case routine corrective maintenance should be applied. In traditional projects, road maintenance is conducted every year and only considers the road conditions. In general, preventive maintenance strategies are applied to the PBC scheme, which delays and prevents widespread damage to the roads; whereas in traditional projects, corrective maintenance strategies are applied in an attempt to repair damages that have already occurred. In this study, those two contract approaches are set by applying the same performance indicator for the road that will be contracted. Thus, the implication of the LCC difference between both types of contract is simply on the mechanism of project implementation.

The next step in the LCC calculation procedure is to determine the road maintenance cost. The cost for each type of road maintenance work was estimated based on historical cost data of road projects contracted using traditional contracts during the years 2003-2011. The historical data was collected from the same road segment and vicinity as the one reviewed in this study. The unit cost estimation was carried out for the year 2011 and was used as a baseline of the LCC calculation (Table 5). Unit cost of PBC project not including the cost of service provider procurement and supervision works.

Table 5: Unit cost estimate of road maintenance works

| **Type of Works** | **Project Unit Cost (Rp/km)** |
|---|---|
| Routine (preventive) maintenance | 98,054,957 |
| Routine (corrective) maintenance | 11,738,118 |
| Periodic maintenance | 2,116,526,108 |
| Betterment | 8,735,724,931 |

Note: 1 USD = Rp. 14,500.

In traditional approach, routine maintenance works are always carried out under an in-house system in which the planning and supervision are carried out by in-house staff and the maintenance work is conducted using equipment available at *Balai Pelaksanaan Jalan Nasional*. The total cost of in-house works consists only of material costs and wages of field workers. Periodic maintenance and betterment on the traditional approach are carried out under the DBB delivery system. The road agency expenditures include costs for design, construction, and supervision, as well as the cost of procuring service providers for each of those works.

PBC projects differ from traditional projects in that PBC projects are organized under a Design-Build-Operate-Maintain (DBOM) project delivery system, as stated in [4,5,7]. The road agency expenses encompass the costs for design, construction, and maintenance, all in an integrated manner [4]. In this case study project, there was also a cost for the activities of an independent supervisor that were meant to ensure that the project quality assurance was performing adequately. Procurement costs in PBC projects consist of service provider procurement of both the contractor and supervisor.

This study determined the LCC project based on deterioration of the road. This study uses the Indonesian Integrated Road Management System (IIRMS) method to estimate the deterioration of flexible pavements, which is expressed in the following equation:

$$IRI_t = e^{0.023.t}[IRI_o + 263\ (1+SNC)^{-5}.NE_t] \quad \text{Eq. 1}$$

https://doi.org/10.31436/iiumej.v20i2.1075

where $IRI_t$ is the road deterioration prediction after 1 year of maintenance, $IRI_o$ is the IRI value immediately after maintenance, t is the time of evaluation (t=1 as the maintenance work is to be conducted every year), SNC is the structural number capacity of the pavements, and $NE_t$ is the cumulative equivalent single axle load (CESAL) at time t (per 1 million ESAL). SNC and CESAL data are obtained from [15]. The CESAL value in 2011 was 18,785,729 with a growth rate set to 7.5% per year. In this study, the SNC values were assumed to be influenced by the roads' functional performance where SNC values for roads with good and fair conditions were each set at 6.10 and 3.32, with a condition factor of 0.92 and 0.5, respectively.

Results of the deterioration evaluation of the current year were used as input to assess the initial road conditions for the following year and, based on these predictions, the required maintenance works for the following years were identified. This process was carried out throughout the period of analysis.

The project LCC is the present value of the total cost of road maintenance during the analysis period. Based on the IIRMS guideline that becomes the pavement management system for the road network in Indonesia, the service life of the road for improving the existing road is established for 10 years. The IIRMS guideline also the flexible pavement is designed to be able to serve traffic for 10 years. Based on the considerations above, this study decided that the service life of roads was 10 years. Considering that the longest service life of roads reviewed was 10 years, the LCC analysis period in this study was also set for 10 years. However, as the contract duration of the PBC projects is 4 or 7 years, identification of the required maintenance works and their cost during the 4 or 7 years was conducted in accordance with the PBC maintenance strategy. After the end of the PBC contract period, road maintenance is assumed to be delivered using a traditional approach until the end of the LCC analysis period. Therefore, identification of the required maintenance works and calculation of road maintenance costs for the rest of the analysis period was conducted using a traditional scheme.

As the PBC project reviewed in this study was implemented in 2011, the baseline of the LCC calculation and estimation of the LCC present value were set for the year 2011. LCC calculation was done using an average discount rate during 2005-2014. The discount rate used here was 7.97%.

## 4. PROJECT LCC EFFICIENCY

In general, the results of this calculation show that PBC projects generate a lower LCC compared to traditional projects, as shown in Fig. 3. The LCC calculation for all scenarios shows that the PBC application has potential in generating an LCC efficiency of 9.4% compared to the application of the traditional contract. For all contract duration scenarios and the initial conditions of the road to be contracted, the implementation of PBC generates lower LCC compared to the implementation of traditional contracts. The longer the duration of PBC contracts, there is a tendency for the LCC project to decline. This is due to the low administrative costs for multi-year contract management as in PBC projects compared to single-year contracts on traditional projects, as indicated by [17].

https://doi.org/10.31436/iiumej.v20i2.1075

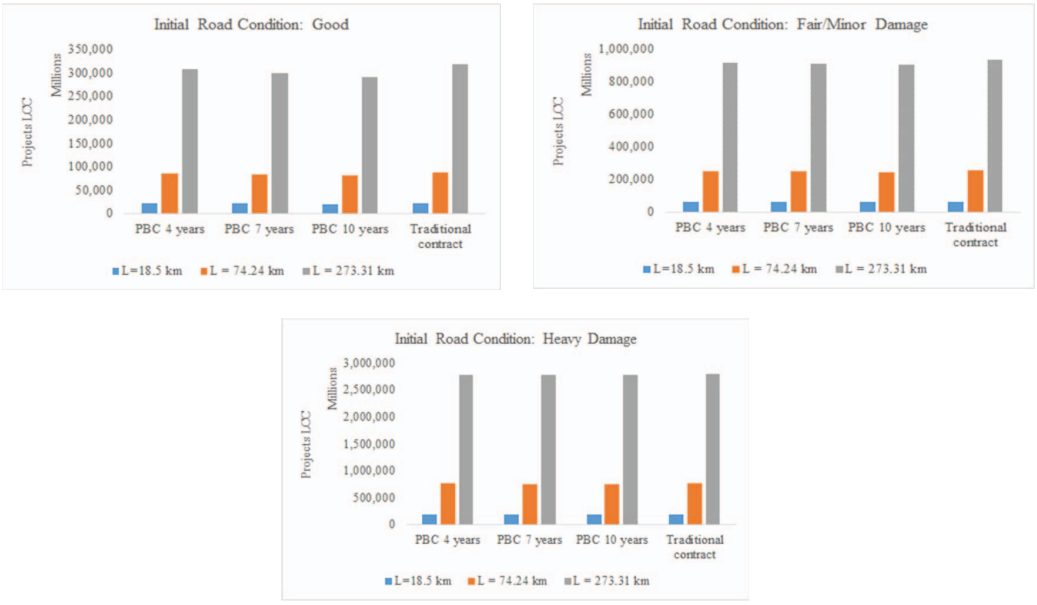



Fig. 3: LCC of PBC and traditional projects.

LCC projects increase as the length of contracted roads increases; however project costs per kilometer of road decrease by 1.9%, as shown in Fig. 4. This is consistent with [18] and [19], thus, working on longer road lengths has the potential to generate lower LCC. Implementation of PBC for the maintenance of long roads can potentially generate a higher LCC efficiency due to the economies of scale in unit cost of road maintenance. There is also a benefit in the form of ensuring road performance on a long road network.

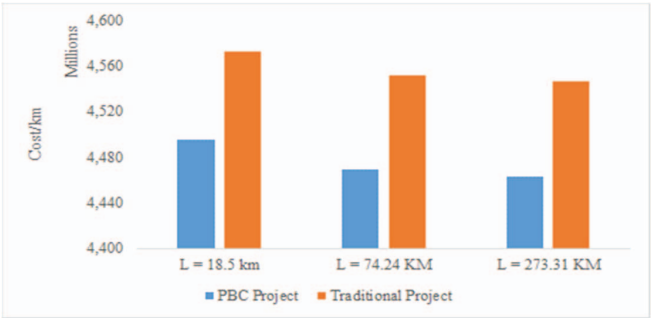



Fig. 4: Project cost comparison per kilometer of road.

The LCC calculation results also show that the implementation of PBC has the potential to generate a lower LCC by as much as 6.4% under the traditional contract for each of the same road conditions, as shown in Fig. 5. Implementation of PBC also has the potential to generate a low LCC if the roads contracted are in good condition. The lower the initial road conditions, the higher the LCC of PBC projects due to the high cost of road maintenance needed in the early years of the contract to meet the required performance indicators, and vice versa. This is consistent with [20-22]. Implementation of PBC on roads with a heavily damaged condition and high traffic loads will encourage contractors to perform road betterment works in the early years of the contract because of the high levels of uncertainty associated with the pavement structural condition.

https://doi.org/10.31436/iiumej.v20i2.1075

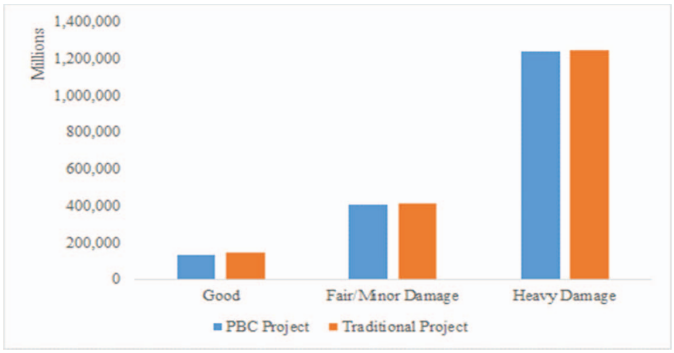



Fig. 5: Project cost comparison for different road condition.

The LCC analysis also showed that the implementation of PBC with a duration of 10 years generated an LCC efficiency of 2.5% compared to the shorter contract duration, as shown in Fig. 6. This result is in line with the findings of [3] and [4]. Although long contract durations require a significantly higher initial cost, they can ensure road performance over a long period of time. Other benefits of a long contract duration include less bureaucracy for the road agencies due to the assurance of available funding of road maintenance for several years, allowing the agency to focus on efforts to increase their role and expertise in the areas of asset management. Roads maintained at good conditions provide benefits for road users, such as reducing the vehicle operating costs, accident rate, and travel time, as also demonstrated by [7].


Fig. 6: Project cost comparison for different contract duration.

Implementation of PBC should be based on the consideration of its effectiveness rather than LCC efficiency alone. To determine a viable PBC project, a PBC effectiveness assessment is required on various project scenarios. PBC effectiveness is expressed as the ratio between the traditional LCC project to the PBC LCC project for the same project scenario. A project scenario with a LCC ratio of more than 1 is considered feasible to be contracted under PBC scheme. The results of the PBC effectiveness assessment on this case study project are shown in Fig. 7. The effectiveness of PBC increases when implemented to projects with longer contracted road, good road conditions, and long contract duration. This is in line with the study conducted by [3] and [4].

https://doi.org/10.31436/iiumej.v20i2.1075


Fig.7: PBC effectiveness for different contract scenarios.

## 5. CONCLUSION

This study conducted LCC calculation and analysis of a PBC pilot project on national roads in Indonesia with high traffic volumes and overloads. The LCC calculation showed that implementation of PBC for road maintenance has the potential to generate lower LCC compared to the implementation of traditional approaches. Preventive treatment on PBC projects could delay and restrict road deterioration, thereby reducing the long-term cost compared to implementation of corrective treatments with traditional contracts.

This study contributes to the determination of an accurate PBC scheme based on the indicator of LCC efficiency. The results showed that PBC projects are highly effective on roads with good conditions and long-term contract duration. These findings are specifically for projects that are similar to this case study; however, the principles of LCC assessment and PBC effectiveness assessment presented in this study can be applied in general. Furthermore, this study has proven that PBC implementation may decrease the road maintenance cost and also provide other benefits that may not be easily reflected in monetary value.

The LCC calculations in this study used a deterministic approach and did not take into consideration the uncertainty of LCCA input variables that can affect the results of LCC efficiency potential. The sensitivity of LCC in this study was limited to the initial conditions and length of the road contracted, and the contract duration of the PBC project. There are various other variables that can affect the variability of LCC, such as the structural condition of roads and vehicle load that might influence the performance deterioration significantly.

## REFERENCES

[1] Carpenter B, Fekpe E, Gopalakrishna D. (2003) Performance Based Contracting for the Highway Construction Industry- an Evaluation of the Use of Innovative Contracting and Performance Specification in Highway Construction. http://www.ncppp.org/resources/papers/battellereport.pdf, Final Report. Koch Industries Inc. Washington, D.C. accessed on September 9, 2011.

[2] World Bank. (2012) Investing in Indonesia's Roads - Improving Efficiency and Closing the Financing Gap. Road Sector Public Expenditure Review 2012. Jakarta.

[3] Anastasopoulous P, McCullouch BG, Gkritza K, Mannering FL, Sinha KC. (2010) Cost Savings Analysis of Performance-Based Contracts for Highway Maintenance Operations. Journal of Infrastructure Systems-ASCE, 16 (4): 251-263.

[4] Tamin RZ, Tamin AZ, Marzuki PF. (2011) Performance Based Contract Application Opportunity and Challenges in Indonesian National Roads Management. *Proceedings: The*

Twelfth East Asia-Pacific Conference on Structural Engineering and Construction, Procedia Engineering 14 (2011): 851-858.
[5] Gajurel, A. (2014). Performance-based Contracts for Road Projects - Comparative Analysis of Different Types. ISBN 978-81-322-1302-4, Springer, India.
[6] World Bank. (2005). Sample bidding documents-procurement of works and services under output- and performance-based road contracts and sample specifications.
[7] Hyman WA. (2009) Performance-Based Contracting for Maintenance. NCHRP Synthesis 389, http://onlinepubs.trb.org/onlinepubs/nchrp/nchrp_syn_389.pdf, accessed on July 20, 2011.
[8] Susanti B, Wirahadikusumah RD. (2011) Identifikasi Tantangan dan Prasyarat Penerapan Kontrak Berbasis Kinerja untuk Proyek Konstruksi Jalan di Indonesia, Prosiding Konferensi Nasional Pascasarjana Teknik Sipil, ISSN. 2089-3051.
[9] Wirahadikusumah RD, Susanti B, Soemardi,B, Sutrisno M. (2014) Drivers for Increased Benefits in Performance-Based Contracts of Road Projects. Proceedings of Conference for Civil Engineering Research Networks, ISSN: 2407-1374.
[10] Care F. (2012) Evaluation of Functional and Structural for flexible Pavement Using Non-Destructive Method – Case Study: Ciasem- Pamanukan Section at Northern Coast of West Java. Master's Thesis. Institut Teknologi Bandung.
[11] Subagio BS, RAMC F, Rahman H, Kusumawati, A. (2013) Structural and Functional Evaluation of Flexible Pavement in Indonesia, Case Study: Ciasem-Pamanukan Section. Proceedings: The Eastern Asia Society for Transportation Studies, 9: 1-17.
[12] Harrison R, Waalkes S, Wilde WJ. (1999) A life cycle cost analysis of rigid pavement. Project Summary Report 1739-S, Center for Transportation Research - The University of Texas at Austin. https://www.utexas.edu/research/ctr/pdf_reports/1739_S.pdf, accessed on October 31, 2012.
[13] Wilde WJ, Waalkes S, Harrison R. (2001) Life cycle cost analysis of Portland Cement Concrete Pavements, Center for Transportation Research. University of Texas at Austin. Research Report 167205-1. http://d2dtl5nnlpfr0r.cloudfront.net/swutc.tamu.edu/publications/technicalreports/167205-1.pdf, accessed on April 24, 2014.
[14] Goh KC, Yang J. (2009) Extending Life-Cycle Costing (LCC) Analysis for Sustainability Considerations in Road Infrastructure Projects. Proceedings: 3rd CIB International Conference on Smart and Sustainable Built Environment, QUT Digital Repository: http://eprints.qut.edu.au/ accessed on November 12, 2012.
[15] Federal Highway Administration. (2002) Life-Cycle Cost Analysis Primer, https://www.fhwa.dot.gov/asset/lcca/010621.pdf, accessed on November 5, 2015
[16] Ministry of Public Works. (2011) Public Works Minister Regulation No. 13/2011 on Procedures for Road Maintenance and Surveillance.
[17] Sultana M, Rahman A, Chowdhury S. (2012) An Overview of Issues to Consider Before Introducing Performance-Based Road Maintenance Contracting, Proceedings of World Academy of Science, Engineering and Technology, 62: 350-355.
[18] Wilmot CG, Cheng G. (2003) Estimating Future Highway Construction Cost. Journal of Construction Engineering and Management, 129(3): 272-279.
[19] Washington State Department of Transportation. (2015). Cost Estimating Manual for Projects. http://www.wsdot.wa.gov/publications/manuals/fulltext/M3034/EstimatingGuidelines.pdf, accessed on December 10, 2015.
[20] Burningham S, Stankevich N. (2005) Why Road Maintenance is Important and How to Get it Done. Transport Note No. TRN-4. The World Bank. Washington, DC.http://siteresources.worldbank.org/INTTRANSPORT/Resources/336291-1227561426235/5611053-1231943010251/TRN4_Road_Maintenance.pdf, accessed on April 4, 2013.
[21] Kong JS, Frangopol DM. (2003) Evaluation of Expected Life-Cycle Maintenance Cost of Deteriorating Structures. Journal of Structural Engineering, 129(5): 682-691.

[22] Litzka J, Weninger-Vycudil A. (2012) The Effect of Restricted Budgets for Road Maintenance. Proceeding: Transport Research Arena – Europe 2012, Procedia – Social and Behavioral Science, 48: 484-494.

https://doi.org/10.31436/iiumej.v20i2.1006

# NO REFERENCE QUALITY OF HAZY IMAGES DEPENDING ON TRANSMISSION COMPONENT ESTIMATION

**HANA HASAN . KAREEM[1], ESRAA GATEA. DAWAY[2], HAZIM GATEA. DAWAY[3]**

[1]*Department of Physics, College of Education, Mustansiriyah University, Baghdad, Iraq.*
[2]*Department of Medical Instrumentation Techniques Engineering, Bilad AlRafidain University College, Baghdad, Iraq.*
[3]*Department of Physics, College of Science, Mustansiriyah University, Baghdad, Iraq.*

[*]*Corresponding author: hazimdo@uomustansiriyah.edu.iq*

*(Received: 6th October 2018; Accepted: 30th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** The research aim is to measure the quality of hazy images using a no-reference scale based on the Transmission Component and Wavelet Transform (TCWT) by calculating the histogram in the High and Low (HL) component. The system is designed to capture several images at different levels of distortion from little to medium to high and the quality is studied in the transmission component. This measure is compared with the other no-reference measurements as a Haze Distribution Map based Haze Assessment (HDMHA) and Entropy by calculating the correlation coefficient between the no reference measurements and the reference scale Universal Quality Index (UQI). The results show that the proposed algorithm TCWT is a good measure of the quality of hazy images.

***ABSTRAK:*** Kajian ini bertujuan bagi mengukur kualiti imej berjerebu dengan menggunakan skala tiada-rujukan berdasarkan Komponen Transmisi dan Penukaran Signal Gelombang (TCWT) dengan mengira komponen Tinggi dan Rendah (HL) histogram. Sistem ini dicipta bagi mengumpul imej pada tahap berbeza dari takat selerakan paling rendah kepada paling tinggi dan kualiti imej diselidik dalam komponen transmisi. Ukuran ini dibandingkan dengan ukuran tiada-rujukan lain sebagai Peta Selerak Berjerebu (UQI). Keputusan menunjukkan algoritma kualiti imej berjerebu TCWT yang dicadangkan adalah berkualiti baik..

***KEYWORDS:*** *no reference quality; wavelet transform; hazy images; transmission*

## 1. INTRODUCTION

The atmosphere reacts to light due to the atmosphere's composition under steady conditions. The atmosphere includes a variety of different molecules and small types of suspended particles, called aerosols [1]. As such, the primary problem presented is the distortion of aerial imaging by these particles. Aerosols, by definition, are very fine liquid or solid particles suspended (in the air) in the atmosphere with a very low fall speed. Their size mostly lies between $10^{-2}$ m and 100 μm. Furthermore, dust, fog, haze, drizzle, smoke, rain, and snow are some of the examples of aerosols [2]. The majority of aerial images have been taken under bad conditions and thus, need a quality scale. The quality scale provides a pointer on the degree of deformation as well as the degree of improvement. Correspondingly, the image quality assessment is a hot spot in the field of image processing.

https://doi.org/10.31436/iiumej.v20i2.1006

In the classified Image Quality Assessment (IQA) [3], the two main categories are fidelity and intelligibility of the image. The fidelity is described as deviation level between the evaluated image and a standard image. The intelligibility is the image's ability to provide related information to people or a computer. The image quality measure can be estimated by quantitative measurement of the combination of the two classes. Currently, IQA is segregated into (i) the subjective estimation method, and (ii) the objective estimation method. The previous method uses people's subjective sense in order to evaluate the quality level of the images while the latter simulates perception technique of Human Visual System (HVS) in order to measure it. Several corresponding existing approaches are known as Full-Reference (FR), Reduced-Reference (RR), and No-Reference (NR).

Among these methods, the Full Reference method is based on HVS [4,5], Structural Similarity Measurement (SSIM) [6] ,and Universal Quality Index (UQI) [7], which were developed a long time ago and were subsequently applied in several areas. In general, the No Reference IQA method can be divided into two categories [8]: (i) Algorithms developed for specific types of distortion, such as blur [9], JPEG and JPEG2000 compression [10,11], and noise [12] and (ii) Non-distortion specific algorithms. Moorthy and Bovik [13] proposed a two-step framework (BIQI) for No Reference IQA based on natural scene statistics, which did not require any prior knowledge of the distorting process, once trained, as evidenced by Tang et al. [14]. Another research [15] proposed the method to determine No Reference image quality by depending on the entropy (ENT) of the first derivative of the lightness. Furthermore, an evaluation of the histogram of the edge regions in [16] suggested that the No Reference measure, which depends on the Range Channel (RGB) of the image, is defined and the Haze Distribution Map (HDM) that is extracted from the hazy image and then the Haze Assessment Metric HDM-based Haze Assessment (HDMHA) is designed according to the HDM using Google earth maps.

## 2. PROPOSED ALGORITHM TO MEASURE THE QUALITY OF THE HAZY IMAGES

In real visual systems, the effects of the haze in the scene rely on the haze density level and coverage area of the haze. Increasing the density of the haze and increasing the area covered by haze, leads to an increased distortion in the resulting image. The proposed method, thus, depends on the estimation the transmittance components from the image. Following which, a wavelet transform of this component is calculated. The wavelet transform is further analyzed by calculating the histogram.

### 2.1 Estimating the Transmission

In a single image distorted by haze, an optical transmission (T) can only be evaluated depending on the hazy image [17-19]. In the calculation, possibly, an Atmospheric Light (A) value may be given. Thus, by superimposition, the transmission in a local patch R(x) is constant. We denote the patch's transmission $\hat{t}(x)$ to have the min value in the local patch R(x) on the haze image. The transmission is, thereby, given by [20]:

$$\hat{t}(x) = 1 - d_{\circ} \min_{c}(\min_{y \in R(x)} \left(\frac{I^{c}(y)}{A^{c}}\right)) \quad (1)$$

In this equation, the term $\min_{y \in R(x)} \left(\frac{I^{c}(y)}{A^{c}}\right)$ is dark channel of the haze image. An optional is to keep a very small amount of haze for distant objects, depending on the constant $d_{\circ}$ ($0<d_{\circ}\leq 1$). In this study, we have used $d_{\circ} = 0.95$, $A^{c} = 0.62$ as in [18]. Figure 1 illustrates the transmission for the flower images using equation (1).

https://doi.org/10.31436/iiumej.v20i2.1006

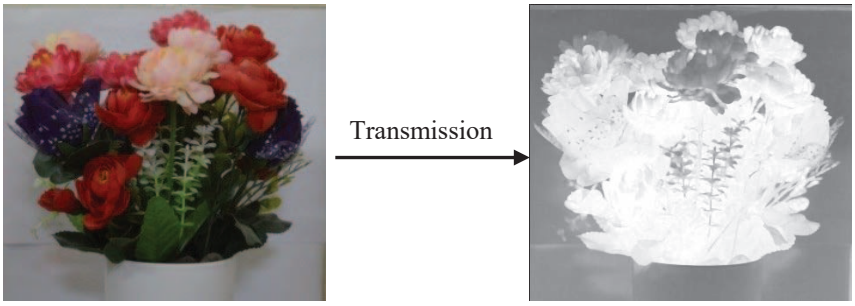



Fig. 1: The flower image and its transmission.

## 2.2 Histogram of the HL Component in the Wavelet Transform

Wavelet transforms are significantly beneficial tools for image processing; they have been used extensively in procedures involving image processing, data compression, and signal processing. Figure 2 illustrates the wavelet transform for one level decomposition using a haar transform [21]. Wherein, the quality scale is based on a conversion of transmission image rather than the original image, and the histogram is calculated for the HL component. Figure 3 shows this transform for the original transmission image (free haze) and its histogram of the HL component.

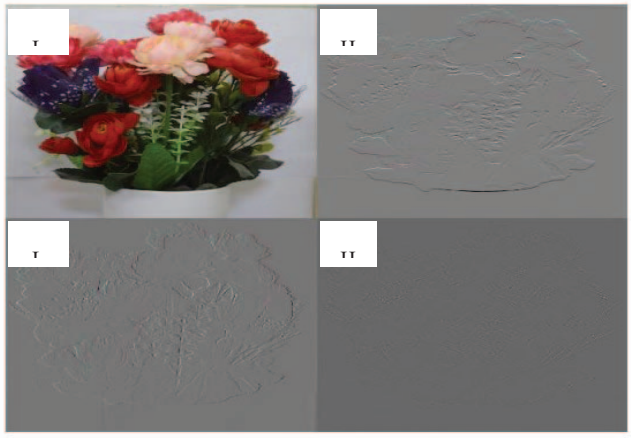


Fig. 2: The wavelet transform for the flower image and its components.

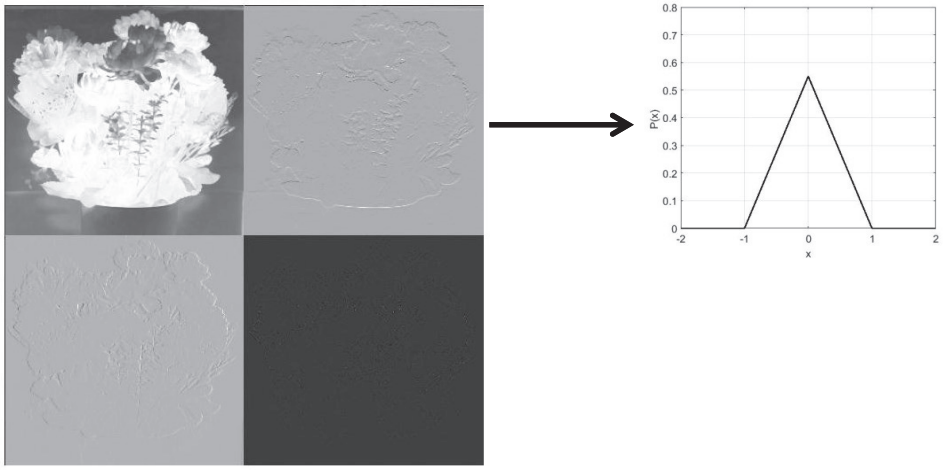



Fig. 3: The wavelet transform for the transmission of the flower image with free haze and histogram of the HL component.

https://doi.org/10.31436/iiumej.v20i2.1006

Figure 4 illustrates this transform for the transmission image (for haze images) and its histogram of the HL component. In the histogram distributions, we can see that the maximum value is inversely proportional to the haze ratio in the image, where this feature can be investigated in the creation of a blind-reference quality measure by using:

$$Q\_TCWT = 1/max\ (hist\_HL) \tag{2}$$

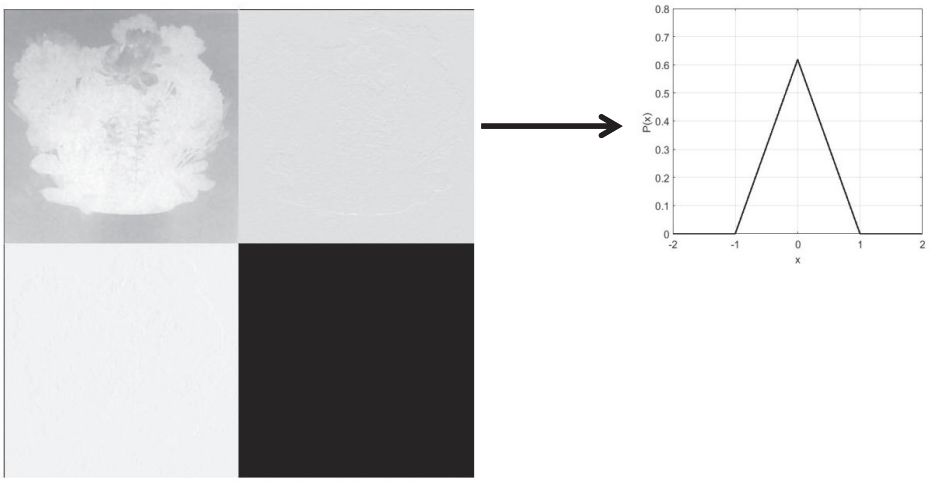



Fig. 4: The wavelet transform for the transmission flower image with distortion by haze and histogram of a HL component.

The increase in the haze ratio leads to a decrease in quality by reducing the Q_TCWT factor while an increase in quality, leads to increased Q_TCWT. The TCWT algorithm can be performed with the following procedure:

1. Input hazy images Ih(x, y).
2. Estimate the Transmission using Eq. (1) to get Th(x,y).
3. Transform Th(x, y) image by finding the wavelet transform for one level.
4. Estimate the HL component from the wavelet transform to get ThHL(x, y).
5. Calculate the histogram for ThHL(x,y).
6. Compute the quality based on TCWT by using Eq. (2).

## 3. UNIVERSAL QUALITY INDEX

Instead of using traditional error summation measurements, we can use a method referred to as Universal Quality Index (UQI) to model any image distortion. [7]:

$$\mathrm{UQI} = \frac{4\bar{I}\,\bar{I}_n\sigma_{II_n}}{(\bar{I}^2+\bar{I}_n{}^2)(\sigma_I{}^2+\sigma_{I_n}{}^2)} \quad , \tag{3}$$

where $\bar{I}$ and $\bar{I}_n$ are the mean of the original and processing (or hazy) images, $\sigma_I$ and $\sigma_{I_n}$ are the standard deviation between the original and processed (or noisy) images. $\sigma_{II_n}$ is the covariance, which can be defined as[7]:

$$\sigma_{II_n} = \frac{1}{N-1}\sum_{i=1}^{N}(I-\bar{I})(I_n-\bar{I}_n) \tag{4}$$

or

https://doi.org/10.31436/iiumej.v20i2.1006

$$\mathrm{UQI} = \left(\frac{\sigma_{\mathrm{II_n}}}{\sigma_{\mathrm{I_n}}\sigma_{\mathrm{I}}}\right)\left(\frac{2\bar{\mathrm{I}}\,\bar{\mathrm{I}}_{\mathrm{n}}}{\left(\bar{\mathrm{I}}^2+\bar{\mathrm{I}}_{\mathrm{n}}{}^2\right)}\right)\left(\frac{2\sigma_{\mathrm{I_n}}\sigma_{\mathrm{I}}}{\sigma_{\mathrm{I}}{}^2+\sigma_{\mathrm{I_n}}{}^2}\right) \quad . \tag{5}$$

## 4. RESULTS AND DISCUSSION

In this study, we have designed a system to measure the level of the haze (dust form), which distorts the color image. Figure 5 shows the complete system, which contains a firm glass box, HeNe laser, fan, lux meter, camera stand, and camera. Dust is raised by the fan in the box, whereby the lux meter immediately measures the laser intensity. While more dust is added into the box, the dust in the bottom of the box is stirred. After a short span of time, the dust settles at the bottom of the box, and once the dust has settled, images are taken from the camera a second time. The lux meter reads the intensity of the laser light, which decreases as the intensity of dust increases and increases according to the reduced density at the same time as the camera captures the images. During the course of the experiment, in the current study, 39 images were taken, each group contained 13 images (distorted from high to low by hazy) with size (1068*1040 pixel). Subsequently, all the images were processed using a MATLAB program version 2013a.

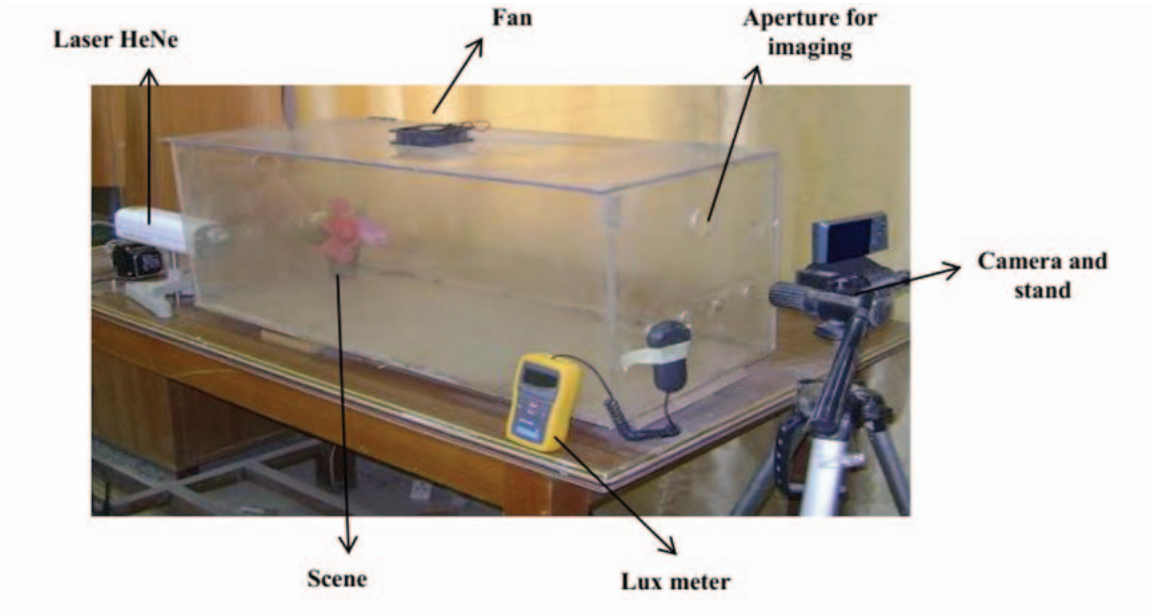



Fig. 5: The system used to capture the hazy image with deformation levels from high to low.

In this research, the reference measure values (UQI) of the hazy images was calculated as well as no-reference measures (ENT, HDMHS, and TCWT) were computed for all measurements. These relationships are illustrated in Figs. 7, 8 and 9 for all the group images. Note, as apparent from all these relationships, the non-reference measure TCWT is closer to reference measure UQI, followed by HDMHS then ENT. This behaviour is reflected in the correlation coefficients between the reference scale and non-reference measures. Note that the greatest value of the correlation is 0.9864 because it is close to one for the TCWT as shown in Table 1, and then correlation coefficients are lower in HDMHS and ENT.

https://doi.org/10.31436/iiumej.v20i2.1006

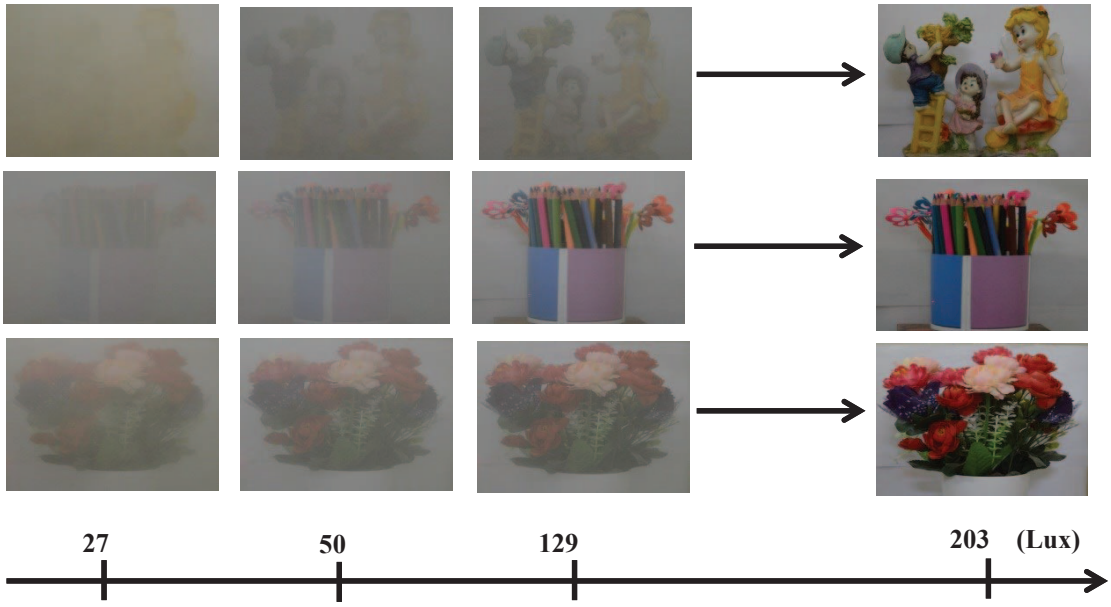



Fig. 6: The first row represents some images in the first group, i.e., artwork images, the second row represents the second group, i.e., the pens images and the third row represents the second group, i.e., the flower images. All images were taken in about 27 to 203 lux.

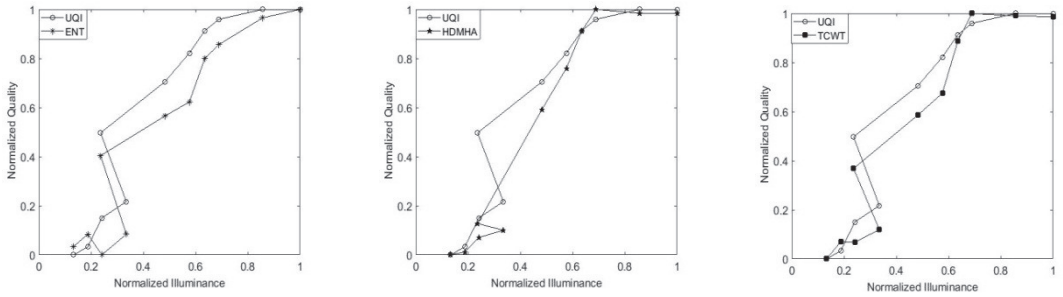



Fig.7: The relationship between normalized quality measurements and normalized illuminance for the first group images.

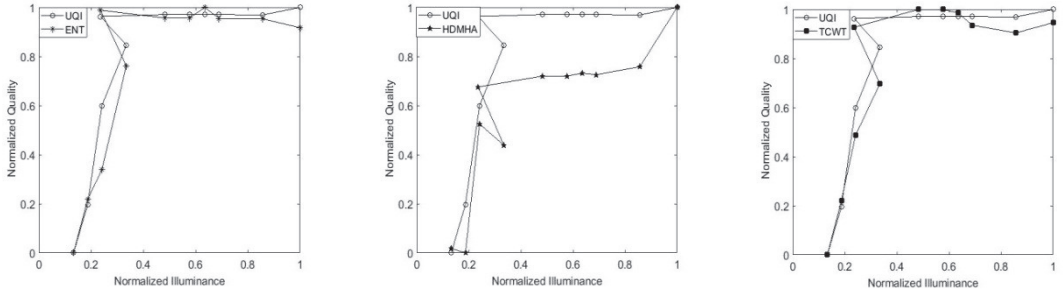



Fig. 8: The relationship between normalized quality measurements and normalized illuminance for the second group images.

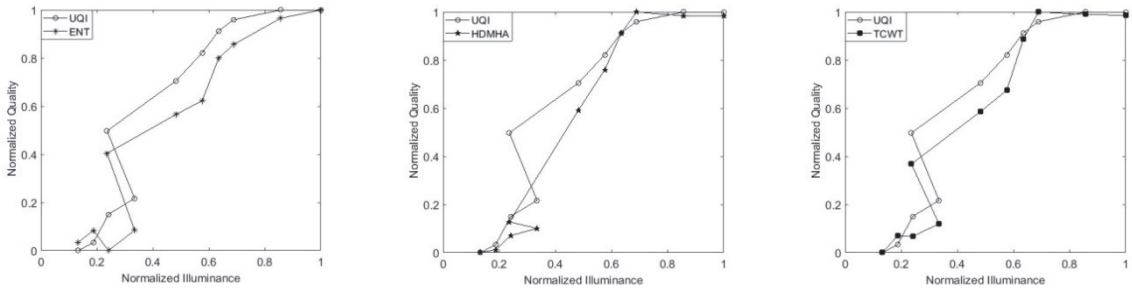



Fig. 9: The relationship between normalized quality measurements and normalized illuminance for the third group images.

Table 1: Correlation Coefficient between the reference measure (UQI) and non-reference measure for all image groups

| Method | First G. | Second G. | Third G. |
|---|---|---|---|
| *TCWT* | 0.9864 | 0.9857 | 0.9864 |
| *HDMHW* | 0.9798 | 0.9735 | 0.9798 |
| *ENT* | 0.9686 | 0.9339 | 0.9686 |

## 5. CONCLUSION

In this research a non-reference scale TCWT was proposed to measure the quality of hazy images based on Transmission. This measure is compared with the HDMHW and ENT measurements by calculating the correlation coefficient. Through the analysis of the results, we note that the TCWT is a suitable performance metric for measuring the quality of hazy images and a good proportion if it is comparable with other non-reference measures in the ENT and HDMHW, because the proposed method has higher correlation (0.9864) compared to other methods.

## REFERENCES

[1] Aber, J. S., Marzolff, I., & Ries, J. B. (2010). Small-format aerial photography. Boston: Elsevier. 266.
[2] Murphy, B. L., and Morrison, R. D. (2002). Introduction to Environ-mental Forensics . San Diego, CA: Academic Press.
[3] Abdou IE, Dusaussoy NJ. (1986) Survey of image quality measurements. Paper presented at Proceedings of 1986 ACM Fall joint computer conference, IEEE Computer Society Press. pp71-78.
[4] Yu Z, Wu HR, Winkler S, Chen T. (2002) Vision-model-based impairment metric to evaluate blocking artifacts in digital video. Proceedings of the IEEE, 90(1): 154-169.
[5] Nill NB, Bouzas B. (1992) Objective image quality measure derived from digital image power spectra. Optical Engineering, 31(4): 813-826.
[6] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. (2004) Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing, 13(4): 600-612.
[7] Wang Z, Bovik AC. (2002) A universal image quality index. IEEE Signal Processing Letters, 9(3): 81-84.
[8] Lu Y, Xie F, Wu Y, Jiang Z, Meng R. (2015) No reference uneven illumination assessment for dermoscopy images. IEEE Signal Processing Letters, 22(5): 534-538.

[9] Cao Z, Wei Z, Zhang G. (2014) A no-reference sharpness metric based on the notion of relative blur for Gaussian blurred image. Journal of Visual Communication and Image Representation, 25(7): 1763-1773.
[10] Corchs S, Gasparini F, Schettini R. (2014) No reference image quality classification for JPEG-distorted images. Digital Signal Processing, 30: 86-100.
[11] Liang L, Wang S, Chen J, Ma S, Zhao D, Gao W. (2010) No-reference perceptual image quality metric using gradient profiles for JPEG2000. Signal Processing: Image Communication, 25(7): 502-516.
[12] Liu M, Zhai G, Zhang Z, Sun Y, Gu K, Yang X. (2014) Blind image quality assessment for noise, paper presented at Broadband Multimedia Systems and Broadcasting (BMSB), 2014 IEEE International Symposium on, IEEE, 1–5.
[13] Moorthy AK, Bovik AC. (2010) A two-step framework for constructing blind image quality indices. IEEE Signal Processing Letters, 17(5): 513-516.
[14] Tang H, Joshi N, Kapoor A. (2011) Learning a blind measure of perceptual image quality, paper presented at Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on, IEEE.
[15] Thuy Tuong Nguyen, Xuan Dai Pham, Dongkyun Kim and Jae Wook Jeon, (2008,"Automatic Exposure Compensation for Line Detection Applications",IEEE International Conference on Multisensor Fusion and Integration for Intelligent SystemsSeoul, Korea.
[16] Pan X, Xie F, Jiang Z, Shi Z, Luo X. (2016) No-reference assessment on haze for remote-sensing images. IEEE Geoscience and Remote Sensing Letters, 13(12): 1855-1859.
[17] Daway, H.G.; Mohammed, F.S.; Abdulabbas, D.A. (2016) Aerial Image Enhancement Using Modified Fast Visibility Restoration Based on Sigmoid Function. Adv. Nat. Appl. Sci. 2016, 10 (11), 16–22
[18] He K, Sun J, Tang X. (2011) Single image haze removal using dark channel prior. IEEE transactions on pattern analysis and machine intelligence, 33(12): 2341-2353.
[19] Tan RT. (2008) Visibility in bad weather from a single image, paper presented at Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on, IEEE.
[20] Wang Z, Feng Y. (2014) Fast single haze image enhancement. Computers & Electrical Engineering, 40(3): 785-795.
[21] Gonzalez RC, Woods RE. (2002) Digital image processing, edited, Prentice hall New Jersey

https://doi.org/10.31436/iiumej.v20i2.1096

# USING SIMILARITY DEGREES TO IMPROVE FUZZY MINING ASSOCIATION RULE BASED MODEL FOR ANALYZING IT ENTREPRENEURIAL TENDENCY

**ENDANG SUPRIYATI[1*], MOHAMAD IQBAL[2], TUTIK KHOTIMAH[2]**

[1]*Informatic Department, Universitas Muria Kudus, Central Java, Indonesia.*
[2]*Electrical Engineering Department, Universitas Muria Kudus, Central Java, Indonesia .*

[*]*Corresponding author: endang.supriyati@umk.ac.id*

*(Received: 27th February 2019; Accepted: 18th August 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** Higher education has great potential in producing new startups in the IT (Information Technology) field. Many choices influence students to become IT-entrepreneurs. Association Rule can be used to obtain a model by analysing data so that it can be used to make a rule to the IT entrepreneurship-student model, but the association algorithm has disadvantages in handling large datasets. We propose reducing candidate itemsets using degrees of fuzzy similarity. The membership function in fuzzy sets can be used to measure the quality of rules obtained. The purpose of this study is to improve the algorithm by evaluating the similarity of candidate itemsets to get a good quality rule. This research method has 2 phases, namely (1) calculating the membership function with similarity itemset and (2) applying fuzzy mining association rule. Phase 1 has several steps, including: preparation of a transaction database, the taxonomy process, and identification of similar itemset. Phase 2 has several steps as well. The first is defining membership functions, and the last is a fuzzy mining fuzzy association rule. In this study, a questionnaire was distributed to 1225 students who were members of the IT entrepreneurship program. The results of this study were reduced into 823 itemsets and produced an IT entrepreneurship rule model.

***ABSTRAK:*** Pendidikan tinggi mempunyai potensi besar dalam menghasilkan permulaan baru dalam bidang IT. Banyak pilihan mempengaruhi pelajar bagi menjadi usahawan-IT. Kaedah Bersekutu boleh digunakan bagi mendapatkan model dengan menganalisa data supaya ianya dapat digunakan menjadi model kepada pelajar keusahawanan-IT, namun algoritma bersekutu mempunyai kelemahan dalam mengendalikan dataset yang besar. Kami mencadangkan pengurangan bilangan set item menggunakan tahapan persamaan kabur. Fungsi ahli dalam set kabur dapat digunakan bagi mengukur kualiti aturan yang diperoleh. Tujuan kajian ini adalah bagi meningkatkan algoritma dengan menilai persamaan set item calon bagi mendapatkan aturan kualiti yang baik. Kaedah penyelidikan ini mempunyai 2 peringkat, iaitu (1) mengira fungsi ahli dengan set item persamaan dan (2) menerapkan aturan perlombongan bersekutu kabur. Peringkat 1 mempunyai beberapa langkah, iaitu: urus niaga pangkalan data, proses taksonomi, identifikasi set item yang sama. Tahap 2 mempunyai beberapa langkah, iaitu: menentukan fungsi keahlian, dan akhirnya, aturan perlombongan bersekutu. Dalam kajian ini, soal selidik telah diedarkan kepada 1225 pelajar yang menjadi ahli program keusahawanan IT. Dapatan kajian menunjukkan pengurangan nombor dataset kepada 823 set item dan menghasilkan model aturan teknologi keusahawanan IT.

***KEYWORDS:*** *fuzzy mining; association rule; similarity degrees; IT entrepreneurial*

https://doi.org/10.31436/iiumej.v20i2.1096

# 1. INTRODUCTION

Data mining technology is widely used to analyse large amounts of data, including classification, clustering, and association. The association algorithm is a classic association algorithm and widely used. Classic association algorithms look for relationship patterns between one or more itemsets in the dataset. This algorithm is famous for finding high-frequency patterns. This frequent itemset pattern is used to compile associative rules [1]. Frequent itemsets play an important role in determining the pattern of associations. A frequent itemset is able to detect problems to determine patterns in the dataset [2]. The main key to the association algorithm is in the iteration stage in the database. Each iteration produces a frequency pattern to get the support value of each item. The association algorithm can reduce the number of candidates whose support must be calculated by removing unnecessary objects and attributes to get minimum attributes, and this algorithm has a better performance than the classic association algorithm [3].

Research on improvement of association algorithms has been carried out, all aimed at improving algorithm performance, i.e. to reduce the calculation time of candidate items by duplicating the frequent itemset to build the next candidate item [4]. Some paths are used to improve the association algorithm [5-8]: (1) Reduction of transactions, transactions that do not contain the itemset are often deleted from the itemset candidates, (2) hashing tables to calculate the occurrence of set items, (3) partitioning each itemset often must be in one partition. Fuzzy theory can be used to improve the association algorithm. The basic idea comes from dividing quantitative values into a crisp set that will weaken or overestimate the specified limit. Fuzzy sets can solve problems by allowing membership to be in different sets. Reference [9] applied the fuzzy association rule mining algorithm to find associated relationship information. Mining fuzzy rules may lead to the discovery of more general and important knowledge from the data [10-12].

The focus of this study is the reduction of data by measuring the similarity of items in the transaction. This measurement of similarity uses the average distance between items. Distance measurement for similarity is used for decision making based on certain criteria. Therefore the measure of similarity must be able to distinguish similarities from fuzzy sets so as not to overlap [13,14]. After the process of measuring similarity, $\alpha - cut$ is used to reduce the dataset. The $\alpha - cut\ sets$ contain members whose membership value is not less than $\alpha$ value. The selection of $\alpha - cut$ will have an effect on the results. The selection of the appropriate $\alpha - cut$ is based on the previous experience of the decision maker and/or the case under study [15-17]

Association algorithms are used to analyse data to obtain a model, but association algorithms have disadvantages in handling large datasets. We propose reducing candidate set items using degrees of fuzzy similarity. The focus of this study is the reduction of datasets by measuring the similarity of items based on distance in the dataset with Fuzzy similarity degrees.

# 2. METHOD

There are 2 phases in this research, namely the phase of seeking membership functions, then the phase of the fuzzy mining association rule. The first phase consists of (1) transaction database, (2) taxonomy process, (3) identification of similarity items, the second phase consists of (4) defining membership functions (MFs), (5) Fuzzy mining and (5) fuzzy association rule. Figure 1 describes the research method.

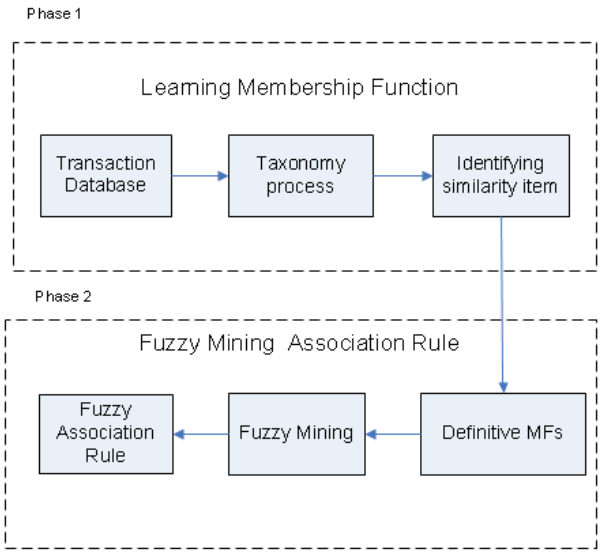



Fig. 1: Research method.

## 2.1 Transaction Database

In this process, the activities carried out include collecting and using historical data to find order, patterns, or relationships in the dataset. The dataset used in this study were 1225 itemsets. The dataset used for this research came from questionnaire survey data of students who took part in entrepreneurship programs at Muria Kudus University.

Table 1: The dataset of participant

| ID | Case |
|---|---|
| T1 | MD, AI, EB |
| T2 | MD, E |
| T3 | DG, EB, EL, EM |
| --- | --- |
| --- | --- |
| --- | --- |
| T873 | SE, D,WD,MD,DG,P,EB |
| T874 | AI,EB |
| T875 | WD,DG,P,EB |

In Table 1, the values T1, T2, T3, ..., T1225 are the IDs of students participating in the program, students are given a survey to choose the knowledge that will be used to pioneer into independent IT entrepreneurs. These choices are Mobile Development (MD), Software Engineering (SE), Database (D), Web Development (WD), Artificial Intelligence (AI), Robotics (R), Mechanics (M), Electric (E), e-marketing (EM), Graphics Design (DG), Photography (P), e-learning (EL), e-business (EB).

## 2.2 Taxonomy Process

Relevant taxonomy items are usually predetermined and can be represented as tree hierarchies. Figure 2 is a picture of the IT entrepreneurship taxonomy. This taxonomy was built based on a combination of computer science taxonomy [18-21] and the concept of

entrepreneurship [22,23]. In this Taxonomy, engineering students can see the opportunity to use IT and engineering to be more innovative, be able to share knowledge, and utilize technology to develop their businesses.

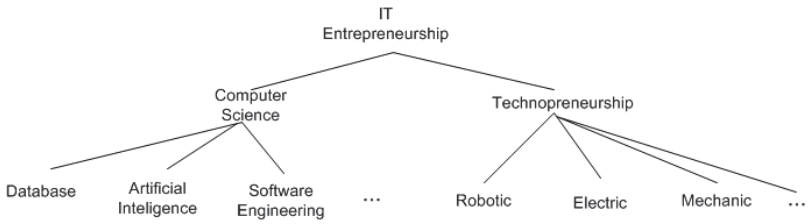



Fig. 2: Taxonomy of IT-Entrepreneurship.

### 2.3 Similarity Identification

The Euclidean distance formula (1) is used to measure the distance between items. Euclidean Distance is the most common use of distance. Euclidean distance examines the root of square differences between coordinates of a pair objects. If $x = (x_1, x_2)$ and $y = (y_1, y_2)$ then the distance is given by:

$$d = \sqrt{(x_1 - x_2) + (y_1 - y_2)^2}. \quad (1)$$

Finding distance between several points can be reached by the average formula (2). If distance $d_1, d_2, d_2$ and sample $n$ then the distance is given by

$$\overline{d} = \frac{\Sigma d}{n} \quad (2)$$

For example, to calculate the similarity item in Table 2, T1 is measured against cases MD, AI, and EB. The measurement of the distance of MD to AI, MD to EB, AI to the EB is carried out using the Euclidean formula. The third result of the distance is then calculated on average.

### 2.4 Fuzzy Membership Function

Once the similarity between elements is known, the next step is to calculate the degree of fuzzy membership in similarity. A fuzzy set is a set containing elements that have varying degrees of membership in the set. Fuzziness is characterized by its membership function. The membership function is a curve that describes the mapping from input to membership level between 0 and 1. Through the membership function that has been compiled, the input values become fuzzy information that is useful later for fuzzy processing. The membership function $\mu_A$ (x), is the usual fuzzy set A defined as:

$$\mu_A: X \rightarrow [0.1] \quad (3)$$

where [0,1] shows real value intervals from 0 to 1. While notations $x_1, x_2, \ldots, x_n$ are the elements of fuzzy set A, notation $\mu_1, \mu_2, \ldots, \mu_n$ is the value of the degree of membership in A. A is usually represented as follows:

$$A = {}^{\mu_1}/_{x_1} + {}^{\mu_2}/_{x_2} + \cdots + {}^{\mu_n}/_{x_n} \quad (4)$$

The membership function used is a triangle curve and a linear curve. Membership in fuzzy sets has different curve shapes [24]. Fuzzy sets can also be defined by assigning a continuous function to describe the membership either analytically or graphically. Some commonly used membership functions are shown in Fig. 3.

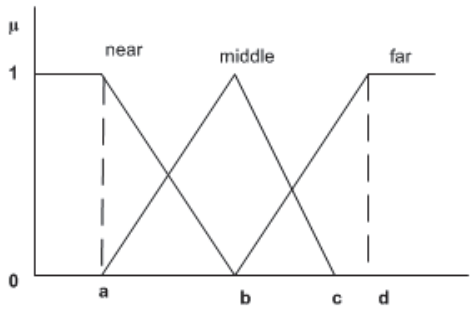



Fig. 3: The membership function of distance similarity.

The triangular membership function in Fig. 4 be expressed as:

$$\mu_x = \begin{cases} 0; & x \le a \mid x \ge c \\ \frac{x-a}{b-a} & a \le x \le b \\ \frac{c-x}{c-b} & b \le x \le c \end{cases} \quad (5)$$

In the linear curve, the fuzzy set consists of 2 states, namely linear up and linear down. For the linear down, the straight line starts from the value of the domain with the highest degree of membership on the left side, then moves down to the value of the domain that has a lower degree of membership with the membership function:

$$\mu_x = \begin{cases} \frac{b-x}{b-a} & ; \quad a \le x \le \mathrm{b} \\ 0\,; & x \ge b \end{cases} \quad (6)$$

In linear rise, the increase in the set starts at the domain value that has zero membership degree [0] moves to the right towards the value of the domain that has a higher degree of membership with the membership function:

$$\mu_x = \begin{cases} 0; & x \le c \\ \frac{x-c}{d-c}; & c \le x \le d \\ 1; & x \ge b \end{cases} \quad (7)$$

## 2.5 Fuzzy Mining

Mining is a technique for extracting process models from an execution log. Fuzzy mining can be applied to very large and unstructured logs. [24,25] In this stage, data on fuzzy distance similarity is collected. In Table 2, the similarity item represents near, middle, and far similarity.

## 2.6 Fuzzy Association Rule

Fuzzy logic is used as data mining which involves calculations based on predictions and clustering. Fuzzy logic data mining algorithms are not only able to analyse data but also to develop accurate results that are easy to implement. Threshold value $\alpha - cut$ is a domain threshold value based on the membership value for each domain, where α-cut has 2 conditions: weak α-cut can be expressed as: $\mu_{(x)} \ge \alpha$, $\alpha - cut$ strong can be stated as: μ (x) <α.

In this paper the Fuzzy Association Rule algorithm is:

1. The main candidates are generated based on similarity
2. Fuzzy operations are carried out to obtain fuzzy mining
3. Reduction of candidates with $\alpha - cut$
4. Fuzzy rules are built using fuzzy support and confidence.

Fuzzy set operations are used to combine and modify fuzzy sets. The value of membership as a result of the operation of two sets is often known as fire strength or $\alpha - cut$. Several fuzzy operations [26] are used in this study to get associative linguistic rules, namely:

(i) Intersection of two fuzzy sets A and B are denoted A∩B, the membership function is as follows:

$$\mu_{A\cap B}(x) = min\{\mu_A(x), \mu_B(x)\}, \forall\, x \in X \quad (8)$$

(ii) Union two fuzzy sets A and B are denoted A ∪ B, the membership function is as follows:

$$\mu_{A\cup B}(x) = max\{\mu_A(x), \mu_B(x)\}, \forall x \in X \quad (9)$$

Table 2: Fuzzy mining of itemsets

| ID | Case | Distance Similarity | Near | Middle | Far |
|---|---|---|---|---|---|
| T1 | MD, AI, EB | 6.30 | 0.00 | 0.90 | 0.10 |
| T2 | MD, E | 2.24 | 1.00 | 0 | 0 |
| T3 | DG, EB, EL, EM | 3.82 | 0.73 | 0.27 | 0 |
| | --- | | | | |
| | --- | | | | |
| | --- | | | | |
| T48 | SE, D | 2.45 | 1.00 | 0.00 | 0.00 |
| T49 | E, R, AI | 2.49 | 1.00 | 0.00 | 0.00 |
| T50 | WD, MD, AI | 4.41 | 0.53 | 0.47 | 0.00 |

The minimum criterion is used to find the intersection of these sets and take the minimum membership value. However, different decision agents may have different perceptions for the fuzzy operation. The min operation is used for near and middle similarity, and does not use far similar values which are of course not very similar. It is determined that α–cut is 0.3, hence only ID with a value of α –cut ≥ 0.3 in the near set will be processed into the Association algorithm. α-cut determination decision, based on the author's decision. If the α-cut is large or too small, the data entering the association algorithm becomes large or too small, and there is increased probability that the results do not represent optimal results. The near and middle sets are chosen because the level of proximity is stronger than the other. Table 3 contains the results of the intersection operation on the fuzzy set.

Table 3: Intersection fuzzy set

| ID | Case | Distance Similarity | Near | Middle | Far | Near ∩ Middle |
|---|---|---|---|---|---|---|
| T1 | MD, AI, EB | 6.30 | 0.00 | 0.90 | 0.10 | 0.00 |
| T2 | MD, E | 2.24 | 1.00 | 0.00 | 0.00 | 0.00 |
| T3 | DG, EB, EL, EM | 3.82 | 0.73 | 0.27 | 0.00 | 0.27 |
| | --- | --- | --- | --- | --- | --- |
| | --- | --- | --- | --- | --- | --- |
| | --- | --- | --- | --- | --- | --- |
| T48 | SE, D | 2.45 | 1.00 | 0.00 | 0.00 | 0.00 |
| T49 | E, R, AI | 2.49 | 1.00 | 0.00 | 0.00 | 0.00 |
| T50 | WD, MD, AI | 4.41 | 0.53 | 0.47 | 0.00 | 0.47 |

# 3. RESULTS

## 3.1 Processing Data

The object of the study was 1225 engineering faculty students consisting of 560 women and 665 men, aged 19-22 years. Students are required to give a value between 1-9 on the questionnaire items used in entrepreneurship. Students are asked to write material that is mastered and provide an assessment of their interests, level of difficulty, and material needs on the market. Questionnaire results were then averaged. The students have participated in an entrepreneurship program held by the university for 1 year. The questionnaire results are averaged as shown in Table 4.

Table 4: Results in the average value of questionnaire

| Attribute | Difficulty | Interest | Market |
|---|---|---|---|
| Mobile Develpoment (MD) | 9 | 9 | 7 |
| Sofware Engineering (SE) | 7 | 9 | 8 |
| Database (D) | 5 | 8 | 9 |
| Web Development (WD) | 7 | 8 | 8 |
| Artificial Inteligent (AI) | 9 | 4 | 5 |
| Robotic (R ) | 9 | 4 | 5 |
| Mechanic (M) | 8 | 8 | 7 |
| Electric ( E ) | 8 | 7 | 7 |
| e-marketing (EM) | 3 | 9 | 9 |
| Design Graphics (DG) | 3 | 4 | 7 |
| Photography (P) | 2 | 4 | 5 |
| e-learning (EL) | 4 | 6 | 6 |
| e-business (EB) | 4 | 9 | 9 |

## 3.2 Similarities by Calculating Distance

In this step, each data is calculated by the Euclidean distance formula. Table 5 describes the results of calculating the similarity of distance between data.

https://doi.org/10.31436/iiumej.v20i2.1096

Table 5: Results of calculation of the similarity of distance between data

| ID | Case | Distance Similarity |
|---|---|---|
| T1 | MD, AI, EB | 6.30 |
| T2 | MD, E | 2.24 |
| T5 | MD | - |
| T8 | MD, EM | 6.00 |
| T10 | AI, EB | 8.12 |
| T12 | SE, D, WD, EB | 2.26 |
| T13 | SE, D, WD, MD, DG, EM, P, EB | 1.87 |
| T16 | DG, AI, EB | 8.60 |
| T18 | MD, P | 7.40 |
| T25 | SE, WD, AI, EB | 8.40 |
| T26 | R, AI, M, E | 2.94 |
| T31 | DG, P | 2.24 |

### 3.3 Degree of Membership

From the calculation of distance similarity, the value of membership is sought with fuzzy theory. Distance variables are divided into three sets, namely near, middle, and far. Table 6 is the result of calculating fuzzy values, representing the itemset similarity.

Table 6: Results of calculation of fuzzy values

| ID | Case | Distance Similarity | Near | Middle | Far |
|---|---|---|---|---|---|
| T1 | MD, AI, EB | 6.30 | 0.00 | 0.90 | 0.10 |
| T2 | MD, E | 2.24 | 1.00 | 0 | 0.00 |
| T3 | DG, EB, EL, EM | 3.82 | 0.73 | 0.27 | 0.00 |
| T4 | MD, EM, EB | 4.13 | 0.62 | 0.38 | 0.00 |
| T5 | MD | - | 0.00 | 0 | 0.00 |
| T6 | SE, D, WD, AI, P | 4.51 | 0.50 | 0.50 | 0.00 |
| T7 | WD, MD, DG, P, EB | 5.83 | 0.06 | 0.94 | 0.00 |
| T8 | MD, EM | 6.00 | 0.00 | 1.00 | 0.00 |
| T9 | SE, D, WD, MD, DG, P, EB | 5.21 | 0.26 | 0.74 | 0.00 |
| T10 | AI, EB | 8.12 | 0.00 | 0.29 | 0.71 |
| T11 | AI, E | 3.74 | 0.75 | 0.25 | 0.00 |
| T12 | SE, D, WD, EB | 2.26 | 1.00 | 0.00 | 0.00 |

### 3.4 Fuzzy Mining Process

After the process of calculating the degree of membership from the distance of similarity. the next process performs an intersection operation such as step 2.6. Distance is represented in terms of near, middle and far. In this case the intersection operation is near and middle. Far term is not used, because the distance between items is far apart. Table 7 is

the result of fuzzy operations on Fuzzy Mining. Fuzzy intersection operations are used to find the minimum value between itemset membership degrees.

Table 7: The result of fuzzy operations on Fuzzy Mining

| ID | Case | Distance Similarity | Near | Middle | Far | Near ∩ Middle |
|---|---|---|---|---|---|---|
| T1 | MD, AI, EB | 6.30 | 0.00 | 0.90 | 0.10 | 0.00 |
| T2 | MD, E | 2.24 | 1.00 | 0 | 0.00 | 0.00 |
| T3 | DG, EB, EL, EM | 3.82 | 0.73 | 0.27 | 0.00 | 0.27 |
| T4 | MD, EM, EB | 4.13 | 0.62 | 0.38 | 0.00 | 0.38 |
| T5 | MD | - | 0.00 | 0 | 0.00 | 0.00 |
| T6 | SE, D, WD, AI, P | 4.51 | 0.50 | 0.50 | 0.00 | 0.50 |
| T7 | WD, MD, DG, P, EB | 5.83 | 0.06 | 0.94 | 0.00 | 0.06 |
| T8 | MD, EM | 6.00 | 0.00 | 1.00 | 0.00 | 0.00 |
| T9 | SE, D, WD, MD, DG, P, EB | 5.21 | 0.26 | 0.74 | 0.00 | 0.26 |
| T10 | AI, EB | 8.12 | 0.00 | 0.29 | 0.71 | 0.00 |
| T11 | AI, E | 3.74 | 0.75 | 0.25 | 0.00 | 0.25 |
| T12 | SE, D, WD, EB | 2.26 | 1.00 | 0.00 | 0.00 | 0.00 |

## 3.5 Fuzzy Association Rule

Before running this algorithm, the minimum support parameter or the percentage of occurrence of a rule must be set first. Minimum confidence or percentage of occurrence of rules for the appearance of the antecedent must also be set. The value depends on the researcher. Rules are presented in antecedent ⇒ consequent.

### *3.5.1 First Implementation*

In the first experiment, all of the 1225 itemsets were used, after several tests, the values of support = 0.5 and min confidence = 0.4 were obtained. The rules discovered where :

**Fuzzy Mining Association Rule Model**

**Complete Model**

Minimum support: 0.4
Minimum onfidence : 0.5
Iteration : 11

**Best Rules :**

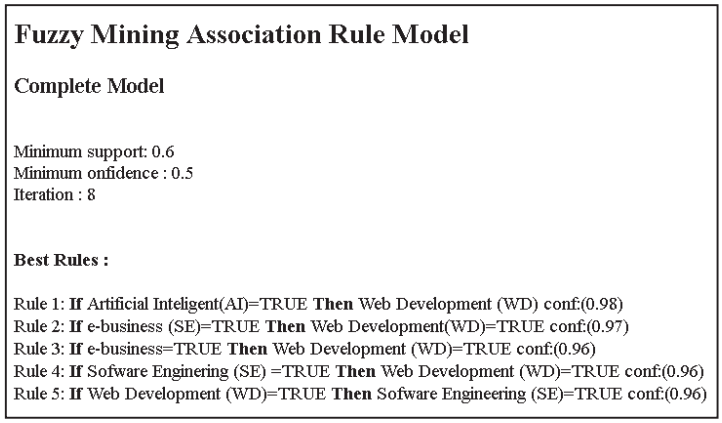
Rule 1: **If** Design Graphics (DG)=TRUE **Then** Robotic (SE)=FALSE conf:(0.78)
Rule 2: **If** e-business (SE)=TRUE **Then** Robotic (EB)=FALSE conf:(0.77)
Rule 3: **If** Sofware Engineering (SE)=TRUE **Then** Web Development (WD)=TRUE conf:(0.75)
Rule 4: **If** Sofware Enginering (SE) =TRUE **Then** Robotic (R)=FALSE conf:(0.75)
Rule 5: **If** Sofware Engineering (SE)=TRUE **Then** e-business (M)=TRUE conf:(0.73)

Fig. 4: Results of the fuzzy mining association rule algorithm.

In Fig. 4, three rules have an antecedent Software Engineering and consequent Robotic. They are associations (think correlation) not necessarily causally related, so be very careful about interpreting association rules.

*3.5.2 Second Implementation*

This experiment determined a strong alpha α-cut ≥0.3, after several test, the value of min support was 0.6 and min confidence = 0.5. From the initial 1225 itemsets, it was reduced to 823 itemsets with an α-cut. In this study, we focused on reducing the number of items included in the a priori algorithm using α-cut. The rules discovered were:


**Fuzzy Mining Association Rule Model**

**Complete Model**

Minimum support: 0.6
Minimum onfidence : 0.5
Iteration : 8

**Best Rules :**

Rule 1: **If** Artificial Inteligent(AI)=TRUE **Then** Web Development (WD) conf:(0.98)
Rule 2: **If** e-business (SE)=TRUE **Then** Web Development(WD)=TRUE conf:(0.97)
Rule 3: **If** e-business=TRUE **Then** Web Development (WD)=TRUE conf:(0.96)
Rule 4: **If** Sofware Enginering (SE) =TRUE **Then** Web Development (WD)=TRUE conf:(0.96)
Rule 5: **If** Web Development (WD)=TRUE **Then** Sofware Engineering (SE)=TRUE conf:(0.96)

Fig. 5: Results of the fuzzy mining association rule algorithm.

In Fig. 5 that algorithm presented 5 rules learned from the dataset. The number associated with the antecedent is absolute coverage in the dataset (in this case a number out of a possible total of 823). The number next to consequent is the absolute number of instance that matches the antecedent and the consequent. In this case, no rule has confidence (number of antecedent divided by number matching consequent) less than 0.96.

## 4. CONCLUSION

Knowledge extraction in databases is a way of extracting knowledge in the form of rules. This rule expresses association relationships between various data, certain data items associated with other data items. In this study Euclidean distance is used to measure the proximity or similarity itemset. Fuzzy membership functions are used to obtain fuzzy mining. Fuzzy logic operation was performed to reduce the size of the dataset, with α-cut. The results of this study are the reduction of the number of dataset from 1225 itemsets to 823 itemsets and to produce an IT entrepreneurship rule model.

## ACKNOWLEDGEMENT

This research received funding from Ministry of Research, Technology and Higher Education of the Republic of Indonesia, through Entrepreneurship Development Program in 2018.

## REFERENCES

[1] Lin, X. (2014) Mr-apriori: Association rules algorithm based on mapreduce. In Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on., pp. 141-144.
[2] Kumar KS, Chezian RM. (2012) A survey on association rule mining using apriori algorithm. International Journal of Computer Applications (IJCA), pp. 47-50.

[3] Patel B, Chaudhari VK, Karan RK, Rana YK. (2011) Optimization of association rule mining apriori algorithm using ACO. International Journal of Soft Computing and Engineering, Vol.1, No.1. pp 24-26.
[4] Al-Maolegi M, Arkok B. (2014) An improved apriori algorithm for association rules. International Journal on Natural Language Computing (IJNLC) Vol.3, No.1. pp 21-29.
[5] Kaur J, Madan N. (2015) Association rule mining: A survey. International Journal of Hybrid Information Technology, Vol.8, No.7. pp 239-242.
[6] Kumar CP, Anjaiah P, Patil S, Lingappa E, Rakesh M. (2017) Mining Association Rules from No-SQL data bases using Map-Reduce Fuzzy Association Rule Mining Algorithm. International Journal of Applied Engineering Research, Vol.12, No.2. pp. 10472-10476.
[7] Yuan, X. (2017). An improved Apriori algorithm for mining association rules. In AIP Conference Proceedings, pp 1-6.
[8] Helm BL, Hahsler PDDM. (2007) Fuzzy Association Rules. Vienna University of Economics and Business Administration.
[9] Verma SK, Thakur RS. (2017) Fuzzy Association Rule Mining based Model to Predict Students' Performance. International Journal of Electrical & Computer Engineering. Vol.3, No.4. pp.2223-2231
[10] Moustafa A, Abuelnasr B, Abougabal MS. (2015) Efficient mining fuzzy association rules from ubiquitous data streams. Alexandria Engineering Journal, 54(2): 163-174.
[11] Imran A, Vladimir K, Viktor L, Olga M. (2017) Fuzzy methods and algorithms in data mining and formation of digital plan-schemes in earth remote sensing. Procedia computer science, 120:120-125.
[12] Johanyak ZC, Kovács S. (2005) Distance based similarity measures of fuzzy sets. Proceedings of SAMI,pp 1-12
[13] Liao H, Xu Z, Zeng XJ. (2014) Distance and similarity measures for hesitant fuzzy linguistic term sets and their application in multi-criteria decision making. Information Sciences, 271: 125-142.
[14] Liu HW. (2005) New similarity measures between intuitionistic fuzzy sets and between elements. Mathematical and Computer Modelling, 42(1-2): 61-70.
[15] Wong KW, Gedeon T, Tikk D. (2000) An improved multidimensional alpha-cut based fuzzy interpolation technique, International Conference on Artificial in Science and Technology, pp 1-6.
[16] Mendel JM. (2017) Type-2 fuzzy sets. In Uncertain Rule-Based Fuzzy Systems. Springer, pp. 259-306.
[17] Wrapper H, Meijer H. (1997) A taxonomy for computer science. Technical Report CSI-R9713
[18] Sujatha R, Bandaru R, Rao R (2011) Taxonomy construction techniques –issues and challenges. Indian Journal of Computer Science and Engineering,Vol.2, No.5. pp 661-671.
[19] Ibrahim A. (2004) Fuzzy logic for embedded systems applications. Elsevier Science.
[20] Salatino AA, Thanapalasingam T, Mannocci A, Osborne F, Motta E. (2018) The computer science ontology: a large-scale taxonomy of research areas. In International Semantic Web Conference, Springer, pp. 187-205.
[21] Masapanta-Carrión S, Velázquez-Iturbide J Á. (2018) A Systematic Review of the Use of Bloom's Taxonomy in Computer Science Education. In Proceedings of the 49th ACM Technical Symposium on Computer Science Education, pp. 441-446.
[22] Dimapilis H. (2013) Are we ready for Technopreneurhip? A Study on selected local Entrepreneur within the city. Research Congress, De La Salle, Universty Manila, pp:1-7.
[23] Amante AD, Ronquillo TA. (2017) Technopreneurship as an outcomes-based education tool applied in some engineering and computing science programme. Australasian Journal of Engineering Education, 22(1): 32-38
[24] Hong TP, Lee YC, Wu MT. (2014) An Effective Parallel Approach for Genetic-Fuzzy Data Mining. Expert Systems with Applications, 41(2): 655-662.
[25] Michael H. (2005) Applied Fuzzy Aritmatic An Introdution With Engineering Applications. Springer.

[26] Cho M, Song M, Yoo S. (2014) A Systematic Methodology for Outpatient Process Analysis Based on Process Mining. In Asia-Pacific Conference on Business Process Management, Springer, pp. 31-42.

https://doi.org/10.31436/iiumej.v20i2.1097

# DESIGN OF WLAN AND WIMAX BAND REJECTION UTILIZING UWB PLANAR ANTENNA COMPRISING SLIT IN THE CONDUCTOR PLANES

**SITI FATIMAH JAINAL[1], NORLIZA MOHAMED[2] AND AZURA HAMZAH[3]**

[1]*Faculty of Engineering, Lincoln University College (LUC), Kelana Jaya, Selangor, Malaysia.*
[1,3]*Malaysia Japan International Institute of Technology, UTM Kuala Lumpur, Malaysia.*
[2]*Razak Faculty of Technology and Informatics, UTM Kuala Lumpur, Malaysia.*

**Corresponding author: sitifatimah@lincoln.edu.my*

*(Received: 27th February 2019; Accepted: 30th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** A compact and low profile ultra wideband planar antenna comprises dual notched-band characteristics for WIMAX and WLAN are presented. UWB communication system is allocated between 3.1 and 10.6GHz, which coexisted with the WLAN and WIMAX frequency bandwidths at 3.3 to 3.6GHz, and 5 to 6GHz, repsectively. The coexistence between multiple frequency bandwidths possibly can cause interference into the communication systems such as data loss and signal disruption. Thus, it is essential to eliminate the coexisted frequency bandwidhs from UWB spectrum. The UWB planar antenna is costructed with a radiator of an elliptical-shaped, and half-ground element which is subjected to suppress the frequency bandwidth for 3.3 to 3.7 and 5 to 6 GHz. Slits are engraved in the elliptical radiator and ground element by etching the conductor elements. Slit shapes are designed in simple and optimized to realize the maximum band notch characteristics. Slit placements are scrutinized and the band notch characteristics are determined. It is considered that the slit in the ground element and the elliptical radiator have stimulated the band notches frequency bandwidths for 3.3 to 3.7 and 5 to 6 GHz, respectively. The UWB planar antennas are compared with the reference antenna and the results are verified. Measured reflection coefficient $S_{11}$ for band notch peaks at the WLAN and WIMAX frequency bandwidths are about -3.0 and -4.0 dB, respectively. Radiation pattern co-polarizations in the H- and E-plane are in omni- and bi-directional, respectively. Maximum gain $G$ is located in the $-z$ –axis and $-x$ –axis in H- and E-plane in the frequency of interest. Surface currents are distributed in the slit areas. Slits in the elliptical radiator and the ground element are not substantially affect the UWB planar antenna overall performances.

***ABSTRAK:*** Antena jalur lebar paling satah yang padat dan bersusuk rendah telah diperkenalkan dan terdiri daripada dua ciri lebar-takik bagi WIMAX dan WLAN. Sistem komunikasi UWB berada pada 3.1 dan 10.6GHz, bertindan dengan jalur lebar frekuensi WLAN dan WIMAX yang berada pada 3.3 hingga 3.6GHz, dan 5 hingga 6GHz, masing-masing. Sifat bertindan antara beberapa jalur lebar frekuensi mungkin akan menyebabkan gangguan pada sistem komunikasi seperti kehilangan data dan gangguan isyarat. Oleh itu, adalah penting bagi membuang jalur lebar frekuensi yang bertindan dengan spektrum UWB. Antena satah UWB telah dibina dengan radiator pemancar berbentuk elips, dan unsur separuh-bumi (lapisan asas) dibawah jalur lebar frekuensi pada 3.3 hingga 3.7 dan 5 hingga 6 GHz. Jalur celahan telah diukir pada radiator elips dan lapisan asas dengan mengukir unsur konduktor. Jalur celahan telah direka mudah dan dioptimumkan bagi mencapai jalur takik maksimum. Kedudukan jalur celahan diteliti dan ciri-diri jalur takik diperolehi. Jalur celahan pada lapisan asas dan radiator elips diperhatikan menyebabkan frekuensi jalur lebar takik sebanyak 3.3 hingga 3.7 dan 5 hingga 6 GHz, masing-masing.

https://doi.org/10.31436/iiumej.v20i2.1097

Antena satah UWB dibandingkan dengan antena rujukan dan dapatan kajian telah disahkan tidak mempengaruhi keputusan antena planar UWB dengan ketara. Ukuran pantulan pekali $S_{11}$ yand diukur pada frekuensi jalur lebar takik WLAN and WIMAX adalah -3.0 dan -4.0 dB, masing-masing. Corak pancaran radiasi ko-polar pada satah H- dan E- adalah omni- dan bi-arah, masing-masing. Kekuatan isyarat maksima *G* berada di paksi *–z* dan *–x* pada satah H- dan E- pada frekuensi yang dipilih. Elektrik pada permukaan tersebar dalam kawasan jalur celahan. Celah radiator elips dan lapisan asas tidak mempengaruhi prestasi keseluruhan antena satah UWB.

***KEY WORDS:*** *ultra wideband, planar antenna, WLAN, WIMAX, and slit*

# 1 INTRODUCTION

Ultra wideband (UWB) frequency bandwidth is assigned from 3.1 to 10.6 GHz and has been commonly applied for communications and sensors, position location and tracking [1], wireless monitoring of transplanted organs [2], monitoring human's activities such as sport and quality of sleep by using wireless body network (WBAN) [3], sensor to detects positioning and tracking of a target by a quad-copter and target [4], detection of human's body behind wall [5], and indoor positioning, radar/medical imaging and target sensor data collection [6-7]. Ultra wideband communication system benefits such as high data rate transfer, less path loss and better immunity to multipath propagation, availability in low-cost transceivers, and low transmit power and low interference, and the requirements in the means of reducing interference due to the coexistence with other narrowband systems have been discussed in [6]. UWB antenna has been proposed in [8] in the means of improving its performances by using conductive adhesive of carbon composite. UWB frequency bandwidth coexisted with other narrowband system and the UWB communication characteristics has been compared with single co-channel interference for various materials of partitions in real environments such as brick board, cloth office partition, concrete block, dry-wall, plywood, structure wood, and single co-channel, had been investigated in [9] where the study shows that the channel capacity had degraded with interference for mostly on the drywall partition and found lowest for structural wood partition.

UWB interference has been studied in [10] and channel loss estimation using the Okumura channel and tested within the ZigBee circuit with dual-power. Possible interference between an indoor UWB system and an outdoor fixed wireless access (FWA) system operating in the 3.5-5.0 GHz has been discussed in [11]. Multiband-orthogonal frequency-division multiplexing (MB-OFDM) coexistence with time hopping ultra-wideband (TH-UWB) networks has been studied analytically in the means of modeling the interference for TH-UWB [12]. Mutual interference performance has been studied in [13] by performing pulse shaping from UWB to other narrowband system. The wavelet technique and reduction of interference from other narrowband system to UWB system in conjunction with Transmitted-Reference (TR-UWB) signaling scheme with adaptive receiver is utilized. Multiple access interference in the bi-orthogonal modulation for UWB communication systems performances has been investigated in [14]. UWB interference such as UWB activity or UWB density is a critical matter when protecting WiBro performance and analytical method based on a system level simulation of a WiBro (OFDMA) is proposed in [15]. The narrowband interferences which had affected the wideband system has been investigated and characterized in [16-17]. Interference from a UWB transmitter into a narrowband receiver has been studied in [18].

https://doi.org/10.31436/iiumej.v20i2.1097

The impacts of ultra wideband in narrowband interferences have also been investigated in [19-22] and had described the degradation of the system, and although filtering can be used in the system in the means of correcting the narrowband interference, it has failed in at high interference level. The analysis of the effect of narrowband interference on UWB system in the presence of multipath fading had demonstrated narrowband problems at certain condition as stated in [23]. UWB signals had encountered many interference sources primarily from narrowband systems and also had affected many narrowband radios as described in [24].

Thus, studies have been indicating that the coexistence between multiple frequency bandwidths could cause interferences to other communication systems. The effects could be in signal degradation, data loss, buzzing sound and communication error. Furthermore, more severe effects could have been resulted by the interference such as system failure and malfunction. Thus, it is essential to eliminate the coexistence frequency bandwidth from the UWB frequency bandwidth.

There are various methods to suppress the mutual narrowband interference (NBI) in UWB communication system. Chirp waveforms of two non-linear UWB based on the arctrigonometric and archperbolic variable is proposed in the means of alleviating the narrowband interference (NBI) in [25]. The effect of interference to bit error rate performance which used the time hopping pulse modulation for impulse radio ultra-wideband is analyzed in [26]. Method of using chirp waveforms of non-linear in the means of suppressing the interference of narrowband in UWB communication system has been explained in [27]. Band-stop filter is applied to prevent the narrowband interference (NBI) in a receiver for transmitted-reference ultra-wideband (TR-UWB) systems consists of narrowband interference (NBI) and inter-pulse interference (IPI) mitigation [28]. A null phase-shift polarization (NPSP) which combined a linear polarization-vector transformer (PVT), single notch polarization (SNP) filter and an amplitude and phase compensator (APC) has been proposed in [29] to suppress narrowband interference (NBI). Elimination of narrowband interference (NBI) could use antenna design modification as an option. The method is simple and low profile as it requires the etching of the conductor element of the antenna which has been reported [30-38]. Structure for the reference and proposed antenna were discussed in section 2. Results were substantiated in section 3 and then concluded in section 4, respectively.

## 2 STRUCTURES OF UWB PLANAR ANTENNAS

Slit placements in the elliptical radiator and ground element are simulated in the early stage for UWB planar antenna type 2A and 2B. Slit is etched in four points known as point a, b, c, and d, consecutively as depicted in Figure 1. Slit is etched in one-sided of the elliptical element and ground plane as it is in symmetric. Thus, the results would be in identical. The elliptical radiator and ground element is divided into two parts (lower and upper). The origin is placed in the middle point of the elliptical radiator and ground element, respectively. Slit is positioned in the centre of the lower and upper area of the elliptical radiator and ground plane. The point coordinates for the slit placement a, b, c, and d are tabulated in Table 1. Slit placements are simulated consecutively and the reflection coefficients $S_{11}$ are compared. Band notch characteristics are studied in the means of achieving the desired band notch frequency bandwidths. Slit placements are optimized in order to obtain the optimum band notch characteristics for UWB planar antenna type 2A, and 2B, respectively. The finalized slit configurations are listed in Table 2.

https://doi.org/10.31436/iiumej.v20i2.1097

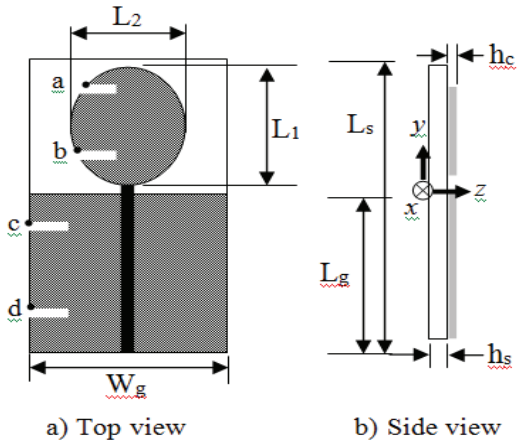



Fig. 1. Slit placements for UWB planar antenna

Table 1: Slit placements in the elliptical element and ground plane

| Point | P(x,y,z) |
|---|---|
| **a** | P(-6, 12, 0.035) |
| **b** | P(-6, 4, 0.035) |
| **c** | P(-11,-6, 0.035) |
| **d** | P(-11,18,0.035) |

Table 2: Parameters for slit $S_1$ and $S_2$

| Antenna type | Slit type | Parameter | Dimension |
|---|---|---|---|
| **2A** | $S_1$ | Length, $l$ | 7.12 mm |
| | | Width, $w$ | 1.10 mm |
| | | Slope angle, $\theta$ | 7.98 deg |
| | $S_2$ | Length, $l$ | 11.0 mm |
| | | Width, $w$ | 0.60 mm |
| | | Slope angle, $\theta$ | 39.43 deg |
| **2B** | $S_1$ | Length, $l$ | 8.28 mm |
| | | Width, $w$ | 0.13 mm |
| | | Slope angle, $\theta$ | 1.38 deg |
| | $S_2$ | Length, $l$ | 11.27 mm |
| | | Width, $w$ | 0.60 mm |
| | | Slope angle, $\theta$ | 41.04 deg |

The reference, UWB planar antenna type 2A, and 2B structures are illustrated in Figure 2 and the dimensions are tabulated in Table 3. UWB planar antenna type 2A and 2B substrates used FR-4 with the permittivity $\varepsilon_r$ of 4.6 and 4.4, respectively. The dielectric electric tangent delta δ is given 0.019. Size of the reference antenna is compact. The conductor planes are placed one-sided on the substrate and half-ground element is used. Thickness of the substrates for UWB planar antenna type 2A and 2B are different due to the value of the substrate permittivity $\varepsilon_r$, thus to achieve impedance matching.

https://doi.org/10.31436/iiumej.v20i2.1097

UWB planar antenna type 2A and 2B used the thickness $h_{s(2A)}$=1.6 and $h_{s(2B)}$=0.76 mm, respectively. However, the conductor planes thickness and the size of the antenna are maintained. Impedance matching is achieved by determining the eccentricity $e$ of the elliptical radiator as Eq. 1.

$$e = \sqrt{\frac{L_2^2}{L_1^2}} \tag{1}$$

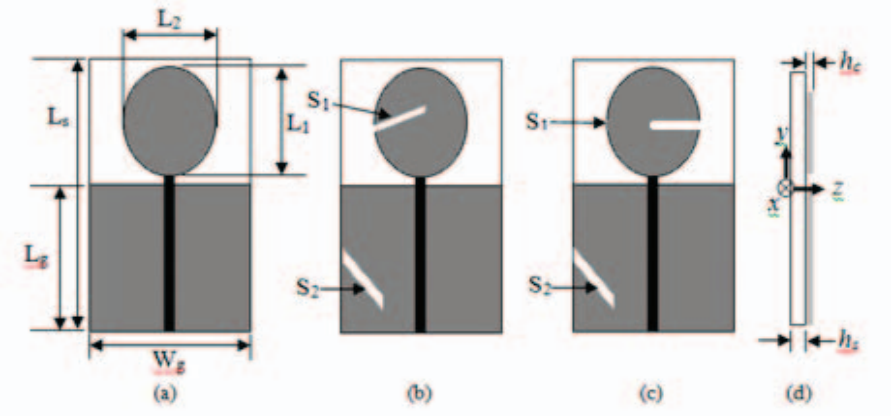



Fig. 2. UWB planar antenna (a) reference (b) type 2A (c) type 2B (d) side view

Table 3: Parameters for UWB planar antenna

| Parameter | Dimension [mm] |
|---|---|
| $L_s$ | 45 |
| $L_g$ | 24 |
| $L_1$ | 16 |
| $L_2$ | 12.8 |
| $w_g$ | 21 |
| $h_c$ | 0.035 |
| $h_{s(2A)}$ | 1.6 |
| $h_{s(2B)}$ | 0.76 |
| $\varepsilon_{r(2A)}$ | 4.6 |
| $\varepsilon_{r(2B)}$ | 4.4 |

Slits are engraved in the elliptical radiator and the ground element, respectively. Slit is etched in the middle side of the elliptical radiator for UWB planar antenna type 2A and 2B. However, slit in the UWB planar antenna type 2A is etched slanted compare with slit in horizontal for type 2B. Slit configurations are determined by the impedance matching and to optimize the reflection coefficient $S_{11}$. Higher value of the reflection coefficient $S_{11}$ is desired for the band notch characteristics. Slit in the ground plane is etched slanted and downwards in about at the centre of the ground plane. Slit configurations in the ground plane are quite similar to each other for UWB planar antenna type 2A and 2B.

However, slits in the elliptical element have major different in slit width $w$ and slope angle, $\theta$. Theoretically, this is owing to the difference in the substrate permittivity $\varepsilon_r$, and slit $S_1$ and $S_2$ configurations are counted in the means of optimizing the band notch characteristics. Generally, surface current distributions are drifted in the elliptical element compared in the ground plane owing to the feed point. Therefore, slit placements in the elliptical element have become critically affected the band notch characteristics compared in the ground plane, where the surface current is disseminated inferiorly.

https://doi.org/10.31436/iiumej.v20i2.1097

Simulation procedure was depicted in Figure 3. First step is to set the simulation setting. Simulation parameters such as boundary, background, and frequency bandwidth are encoded. Optimization is carried out in the means of achieving the optimum band notch characteristics for UWB planar antenna type 2A and 2B. Reference and UWB planar antenna type 2A and 2B are simulated by time domain in Computer Simulation Software (CST) using Antenna-Planar (Microwave/Radio Frequency/Optical) module. Simulation frequency is allocated between 3 and 11GHz. The boundary condition and reflection level are given in open-space and 0.0001, respectively.

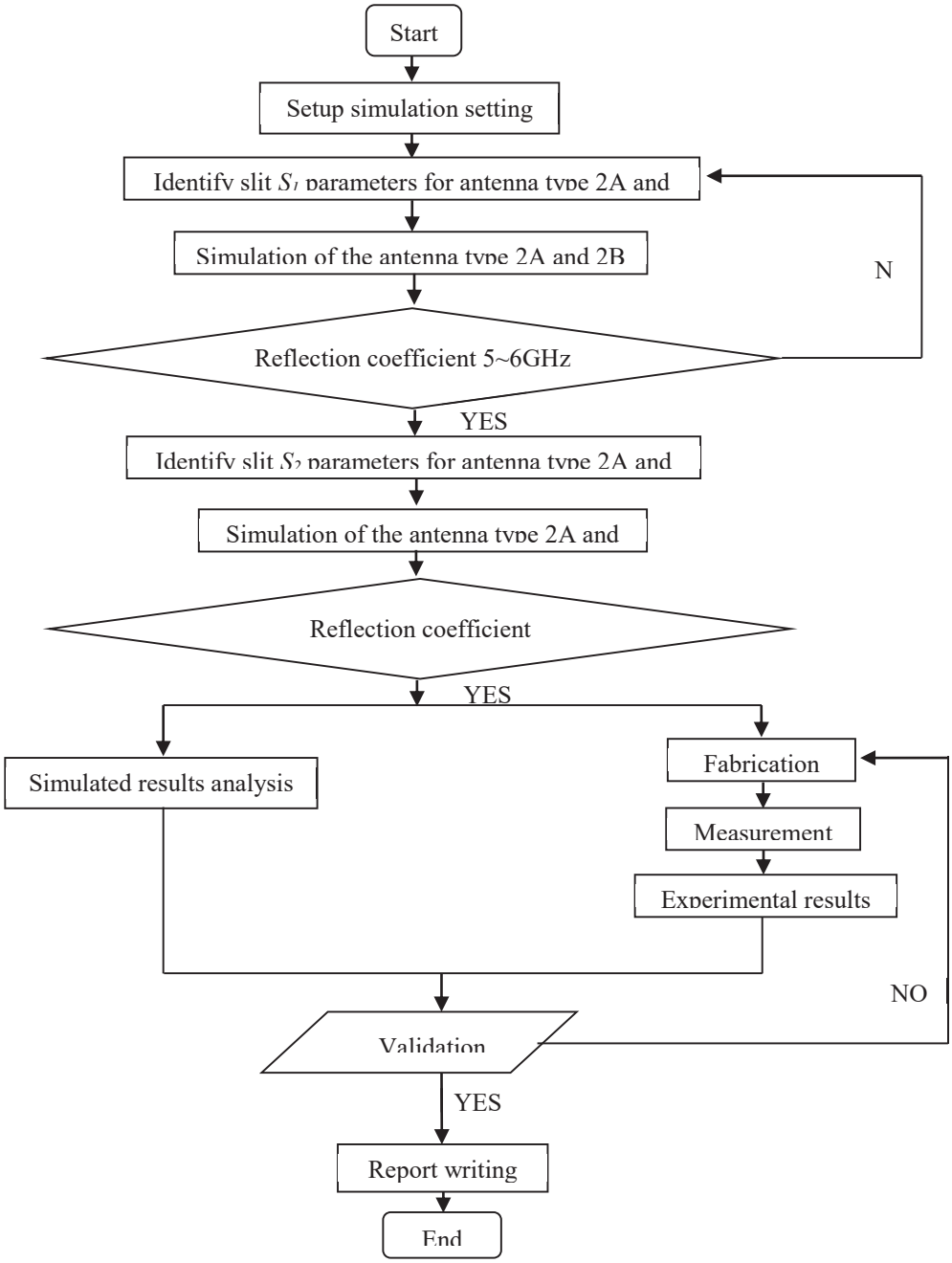



Fig. 3. Workflow for UWB planar antenna

The open boundary is in convolution perfect match layer (CPML) for minimizing the reflection at the boundary during simulation. UWB planar antenna operating frequency is

https://doi.org/10.31436/iiumej.v20i2.1097

between 3.1 and 10.6GHz, thus the reflection coefficient $S_{11}$ must be below -10dB between this frequency bandwidth. However, the reflection coefficients $S_{11}$ for the band notch frequency bandwidths must be above -10dB level in the means of eliminating the band notch frequency bandwidths. The higher reflection coefficients $S_{11}$ level between the band notch frequencies bandwidths the better eliminations of the band notch frequency bandwidths.

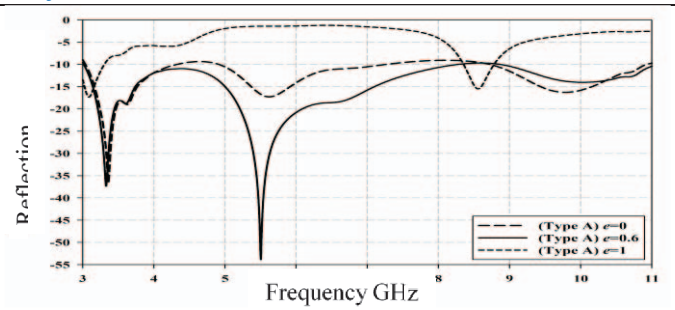
# 3 RESULT AND DISCUSSION

The reference, UWB planar antenna type 2A, and 2B performances are compared after optimization of the band notch characteristics by adjusting the slits parameters. It is desired that the UWB planar antenna type 2A and 2B performances are unchanged when compared with the reference antenna. Thus, it is substantiated that the slits existence on the conductor elements not significantly changes the UWB planar antenna type 2A and 2B performances. The UWB planar antennas are compared with [39-40] as their structures are nearly similar to the UWB planar antenna and etching method also been used for the band notching, and the results are comparable. However, the slits structures are more complicated compared to the UWB planar antennas. The reflection coefficients $S_{11}$ for the band notch peaks are in the same range. The radiation patterns also exhibit in omni-directional and bi-directional for the H- and E-plane, respectively similar to the UWB planar antennas. The numbers of lobes are increased for higher frequency regions. The maximum gains $G$ for the UWB planar antenna is comparable with [39-40] as the maximum gain is achieved at 6.9 dB. However, the overall gains $G$ for [40] is lower than the UWB planar antennas. It is considered that the slit structures in the radiator could affect the overall gain $G$ of the designed antenna due to slit in [40] is considered larger than the UWB planar antenna. Reflection coefficient $S_{11}$, surface current distribution, radiation patterns, gain, and efficiency are discussed in sub-section 3.1, 3.2, and 3.3, consecutively.

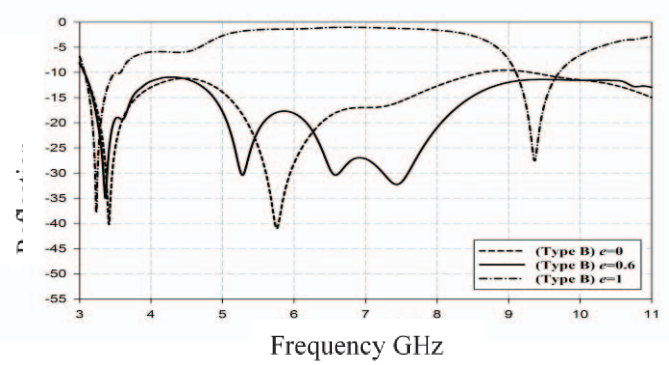
## 3.1 Reflection coefficient

Reflection coefficient $S_{11}$ for eccentricity $e$ and slits placement is depicted in Figure 4 and 5, consecutively. Band notch characteristics for the slit placements are tabulated in Table 4. The reflection coefficient, $S_{11}$ level for operating frequency, $f$ is known below -10 dB. The frequency bandwidths for 4.3~4.9 and 7.3~8.7GHz are located above -10dB when $e=0$, where as for 3.3~8.3 and 8.8~11GHz when $e=1$, for type A. Frequency bandwidths were to be found at 8.7~9.3GHz for $e=0$, and are placed for 3.6~9 and 9.7~11GHz for $e=1$, for type B respectively. The results substantiate that eccentricity $e=0.6$ is ideal for type 2A and 2B as it is within the operating frequency level. Centre frequencies $f_c$ for slit placements at point A for type 2A and 2B is located at 6.7 and 7.3GHz, respectively. Point D is located at 4.4 and 4.3GHz for type 2B, respectively. Thus, point A and D are adjusted in the means of realizing band notch characteristics for 5~6 and 3.3~3.7GHz, respectively. Reference and UWB planar antenna type 2A and 2B were simulated and measured. The comparisons for the reflection coefficient $S_{11}$ are presented in Fig. 6, respectively.

https://doi.org/10.31436/iiumej.v20i2.1097

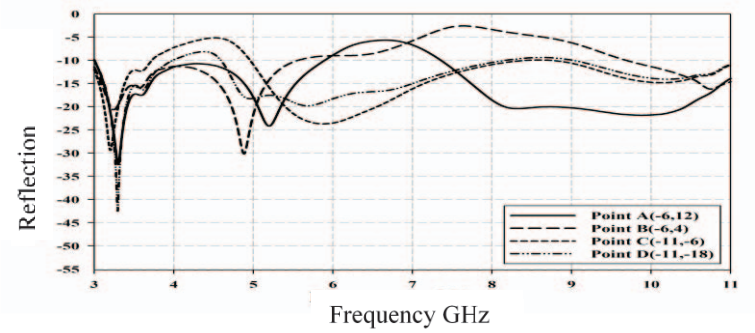



(a) Type 2A

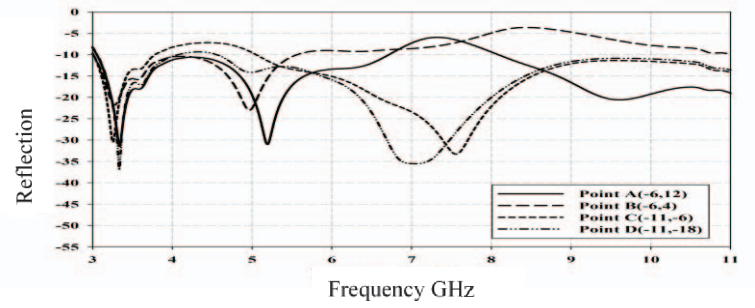



(b) Type 2B

Fig. 4. Reflection coefficient, $S_{11}$ due to eccentricity $e$




(a) Type A



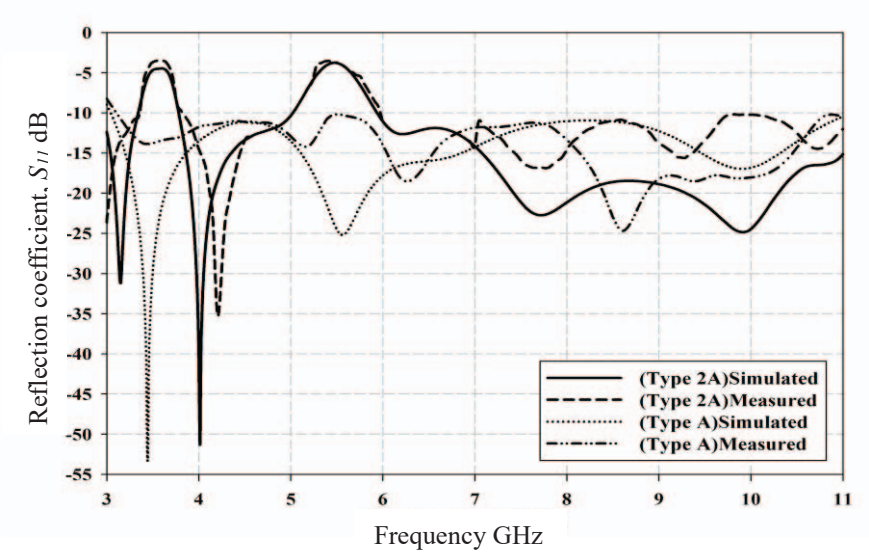

(b) Type B

Fig. 5. Reflection coefficient $S_{11}$ for slit placements

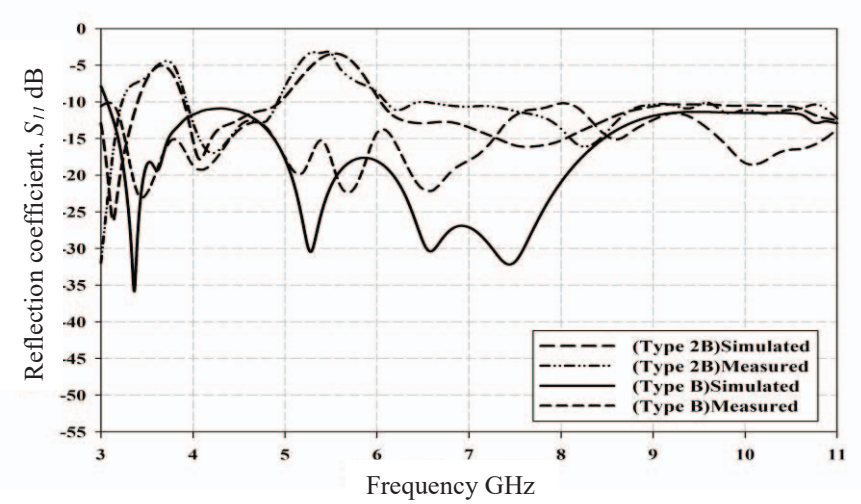
https://doi.org/10.31436/iiumej.v20i2.1097

Table 4: Band notch characteristics due to slit placements

| Antenna type | Point | Centre frequency, $f_c$ [GHz] | Reflection coefficient, $S_{11}$ [dB] |
|---|---|---|---|
| A | a | 6.7 | -5.7 |
| | b | 7.6 | -2.5 |
| | c | 4.5 | -5.1 |
| | d | 4.4 | -8.1 |
| B | a | 7.3 | -6.0 |
| | b | 8.4 | -3.6 |
| | c | 4.4 | -6.9 |
| | d | 4.3 | -9.2 |




(a) Type 2A




(b) Type 2B

Fig. 6. Finalized reflection coefficient, $S_{11}$ for reference and UWB planar antenna type 2A and 2B

BN1 and BN2 are known as the band notch characteristics realized for the frequency bandwidth at 5~6 and 3.3~3.7GHz, correspondingly. The dual notched-band characteristics

https://doi.org/10.31436/iiumej.v20i2.1097

for UWB planar antenna type 2A and 2B are tabulated in Table 5. It is substantiated that dual band notch are generated due to the slits in the elliptical radiator and ground element. The slit in the elliptical radiator and ground element are considered to generate the notched-band at the frequency bandwidth in 5~6GHz and 3.3~3.7GHz, correspondingly. Theoretically, it is considered that the frequency bandwidths for 5~6 and 3.3~3.7GHz are suppressed from the UWB communication system. Simulated and measured results are compared and results have shown reasonable agreement between both.

Table 5: Dual band notch characteristics performances for UWB planar antenna type 2A and 2B

| Antenna type | | | Band notch characteristic | | |
|---|---|---|---|---|---|
| | | | Centre frequency, $f_c$ [GHz] | Reflection coefficient, $S_{11}$ [dB] | Frequency bandwidth, $f_{bw}$ [GHz] |
| **2A** | Simulated | BN1 | 5.5 | -3.7 | 1.0 |
| | | BN2 | 3.6 | -4.4 | 0.4 |
| | Measured | BN1 | 5.4 | -3.4 | 1.0 |
| | | BN2 | 3.6 | -3.3 | 0.4 |
| **2B** | Simulated | BN1 | 5.6 | -3.3 | 1.0 |
| | | BN2 | 3.6 | -5.0 | 0.4 |
| | Measured | BN1 | 5.4 | -3.2 | 1.0 |
| | | BN2 | 3.7 | -3.4 | 0.5 |

### 3.2 Surface current distributions

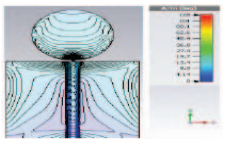
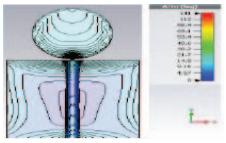
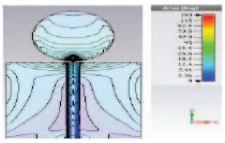
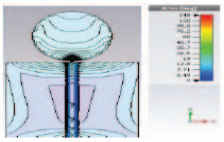
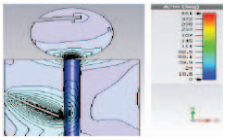
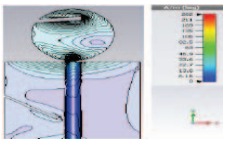
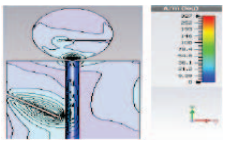
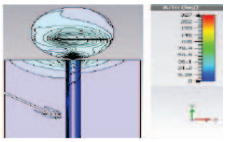
Surface current distributions for the reference and UWB planar antenna type 2A, and 2B for the centre frequency $f_c$, 3.5 and 5.5GHz are exemplified in Table 6. Surface currents are centralized around the slits area. Theoretically, the surface currents surge vertically in the elliptical radiator and ground element. Slits are sited horizontally to intercept the vertically-polarized surface current and consequently mismatched the input impedance in the antenna. Surface currents are surged in the edges of the ground element in the lower frequency area and mainly in the elliptical radiator in the middle frequency region. Surface current is surged in slit in the ground plane and elliptical element for the frequency 3.5 and 5.5GHz, respectively. Hence, it is contemplated that slit in the ground plane and elliptical element have generated the band notch characteristics in the UWB planar antenna type 2A and 2B, accordingly.

### 3.3 Radiation patterns

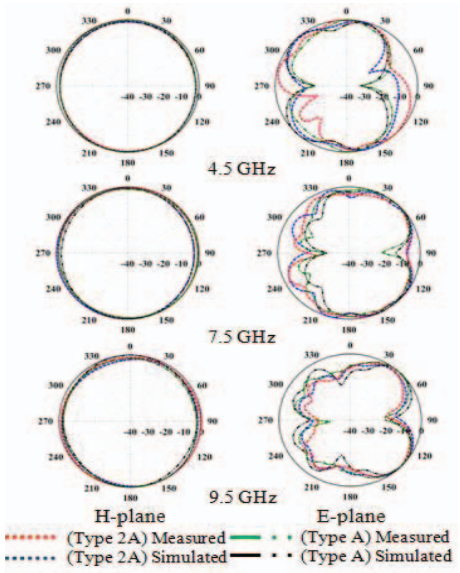
Radiation patterns *RP*, gain *G* and efficiency $e_{ff}$ of the reference, UWB planar antenna type 2A, and 2B are exhibited in Fig. 7 and 8, consecutively. It is demonstrated that the radiation patterns in the H- and E-plane for the UWB planar antenna type 2A and 2B is in omni- and bi-directional, respectively. Radiation patterns in the H-plane of the UWB planar antenna type 2A and 2B is comparable with the monopole antenna. This is as a result of the vertically positioned of the feeding cable on the ground plane. Radiation patterns *RP* performances are tabulated in Table 7. It is corroborated that the higher gain *G* is pointed in the $-z$ –axis and $-x$ –axis in H- and E-plane, correspondingly. However, maximum gain *G* for the frequencies of interest 7.5 and 9.5GHz for UWB planar antenna type 2B in the E-plane is pointed in the $+x$ –axis.

https://doi.org/10.31436/iiumej.v20i2.1097

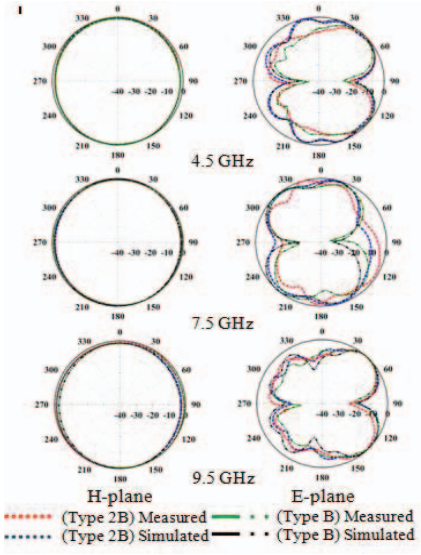
Table 6: Surface current distributions for reference antenna and UWB planar antenna type 2A and 2B

| Antenna type | Surface current distribution / Frequency GHz | |
|---|---|---|
| | 3.5 | 5.5 |
| A |  |  |
| B |  |  |
| 2A |  |  |
| 2B |  |  |

It is considered that the slit in the elliptical radiator is placed in the $+x$ –axis, hence has influenced the radiation patterns of the UWB planar antenna type 2B, accordingly. Radiation patterns of the reference are equated with UWB planar antenna type 2A and 2B, and the results are passable. It is preferred that the performances for UWB planar antenna type 2A and 2B not significantly amended from the reference antenna. Radiation efficiency $e_{ff}$ for the frequency 5~6 and 3.3~3.7GHz have substantiated deterioration for the UWB planar antenna type 2A and 2B, correspondingly. It is demonstrated that the band notch are realized in the respective frequency bandwidths, accordingly.



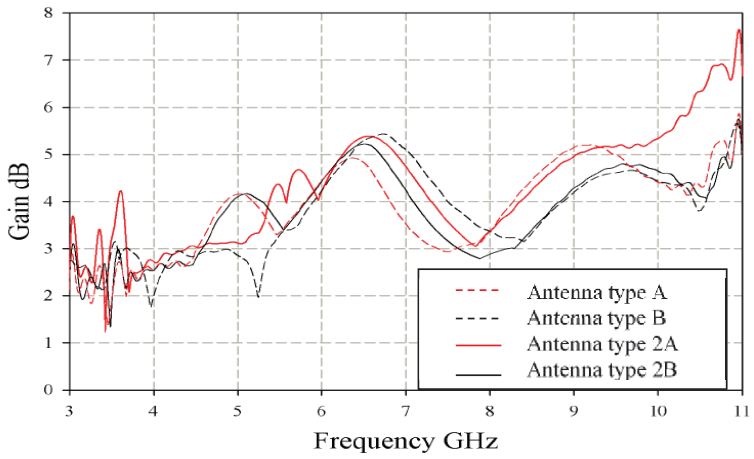

(a) Type 2A



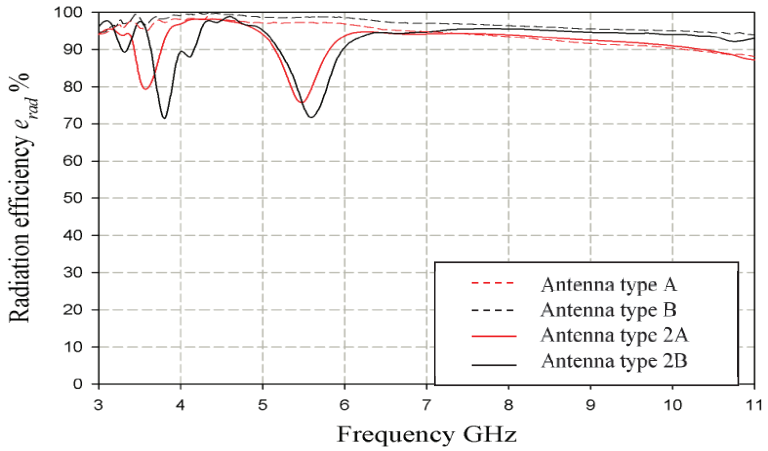

(b) Type 2B

Fig. 7. Radiation patterns for UWB planar antenna type 2A and 2B

https://doi.org/10.31436/iiumej.v20i2.1097




(a) Gain *G*




(b) Efficiency $e_{ff}$

Fig. 8. Gain *G* and efficiency $e_{ff}$ for UWB planar antenna type 2A and 2B

Table 7: Radiation pattern performances for UWB planar antenna type 2A and 2B

| Antenna type | Frequency GHz | Main lobe plane | | Efficiency % | | Gain dB | |
|---|---|---|---|---|---|---|---|
| | | H- | E- | Simulate | Measured | Simulated | Measured |
| **2A** | 4.5 | 180 | 162 | 96.2 | 94.2 | 3.6 | 6.3 |
| | 7.5 | 258 | 196 | 94.0 | 100.0 | 3.6 | 4.8 |
| | 9.5 | 173 | 136 | 91.7 | 83.9 | 5.3 | 4.8 |
| **2B** | 4.5 | 177 | 164 | 94.8 | 81.3 | 2.9 | 5.4 |
| | 7.5 | 172 | 347 | 95.5 | 100.0 | 4.0 | 5.4 |
| | 9.5 | 181 | 42 | 94.4 | 87.5 | 4.7 | 4.9 |

https://doi.org/10.31436/iiumej.v20i2.1097

# 4 CONCLUSION

UWB planar antenna comprises dual notched-band characteristics to decline the frequency bandwidth for WIMAX and WLAN are presented. The method employed in the design is simple and low profile. Slits in the elliptical radiator and ground element apparently have initiated the band notch characteristics in UWB planar antenna type 2A and 2B. The structure of the reference and UWB planar antenna type 2A and 2B uses half-ground plane, and designs are compact. The impedance mismatched is corresponded with the surface currents distributions in the slits area, respectively. The reflection coefficients $S_{11}$ have decreased promptly at the notched-band frequency bandwidths and have achieved about -3.0 and -4.0dB at the center frequency 5.5 and 3.5GHz, respectively. Gain $G$ values are plunged at the band notch center frequencies, 5.5 and 3.5GHz to 4.0 and 3.5dB, respectively. Radiation efficiency $e_{rad}$ for the UWB planar antennas have dropped to 70% from the overall performances. Thus, it is indicated that the frequency bandwidths for 5 to 6 and 3.3 to 3.7GHz are suppressed from the UWB frequency bandwidth, accordingly. However, the radiation patterns for the UWB planar antennas are compared with the reference antenna and have indicated that it is not significantly affected by the slits in the conductor elements. The UWB planar antenna type 2A and 2B performances are evaluated with the reference antenna and the results are verified.

# ACKNOWLEDGEMENT

This research work is funded by Universiti Teknologi Malaysia (UTM) through Razak grant fund vote No. 4J283. Antenna measurements were conducted in Antenna Research Center (ARG), Faculty of Engineering (FOE), Universiti Teknologi Mara (UITM-Shah Alam).

# REFERENCES

[1] Applications on uwb technology. Available: https://arxiv.org/abs/0911.1681.
[2] Pongphan Leelatien, Koichi Ito, Kazuyuki Saito, Manmohan Sharma, Akram Alomainy. (2018) Channel characteristics and wireless telemetry performance of transplanted organ monitoring system using ultra wideband communication. IEEE Journal of Electromagnetics, RF and Microwaves in Medicine and Biology, 2(2):94-101.
[3] Timo Kumpuniemi, Juha Pekka Makela, Matti Hamalaine, Kamya Yekeh Yazdandoost, Jari Iinatti. (2017) Dynamic uwb off-body radio channels-human body shadowing effect. Available: 10.1109/PIMRC.2017.8292584.
[4] Robust target-relative localization with ultra-wideband ranging and communication. Available: https://researchgate.net/publication/324602380.
[5] Jawad Ali, Roshayati Yahya, Noorsaliza Abdullah, Syarfa Zahirah Sapuan. (2017) Ultra-wideband antenna for behind the wall detection. International Journal of Electrical and Computer Engineering (IJECE), 7(6):2936-2941.
[6] Ultra wideband antennas-past and present. Available: https://www.researchgate.net/publication/46093503.
[7] Weihua Zhuang, Xuemin Sheman Shen Qi Bi. (2003) Ultra-wideband wireless communications. Wireless Communications and Mobile Computing, 3:663-685.
[8] Erick Reyes-Vera, Mauricio Arias-Correa, Andreas Giraldo-Munoz, Daniel Catano-Ochoa, Juan Santa-Marin. (2017) Development of an improved response ultra-wideband antenna based on conductive adhesive of carbon composite. Progress in Electromagnetics Research C, 79:199-208.
[9] Chien-Ching Chiu, Chun-Liang Liu, Shu-Han Liao. (2011) Channel characteristics of ultra-wideband systems with single co-channel interference. Wireless Communications and Mobile Computing, 13:864-873.

https://doi.org/10.31436/iiumej.v20i2.1097

[10] Lei Gao, Qun Chen. (2017) Channel loss estimation and test of ultra-wideband propagation from 2 to 10 ghz application. Indonesian Journal of Electrical Engineering and Computer Science, 6(3):663-670.

[11] Romeo Giuliano, Gianluca Guidoni, Ibrahim Habib, Franco Mazzenga. (2004) Evaluation of interference due to uwb hot spot on fixed wireless access systems. J. Wireless Communications and Mobile Computing, 4:947-961.

[12] Farshad Sarabchi, Chahe Nerguizian. (2015) Impact of th-uwb interference on mb-ofdm uwb systems: interference modeling and performance analysis. Wireless Communications and Mobile Computing, 16:960-976.

[13] Bikramaditya Das, Ch. Sasmita Das, Susmita Das. (2010) Interference cancellation schemes in uwb systems used in wireless personal area network based on wavelet based pulse spectral shaping and transmitted reference uwb using awgn channel model. International Journal of Computer Applications, 2(2):88-93.

[14] Simone Morosi, Romano Fantacci, Enrico Del Re, Leornardo Goratti. (2005) Performance of the bi-orthogonal modulation for ultra-wideband communication systems with multiple access interference. Wireless Communications and Mobile Computing, 5:5-14.

[15] Young-Keun Yoon, Heon-Jin Hong, Ik Guen Choi. (2007) Ultra-wideband coexistence with wibro. J. ETRI, 29(2):234-236.

[16] J.R. Foerster. (2002) The performance of a direct-sequence ultra wideband system in the presence of multipath, narrowband interference, and multiuser interference. IEEE Conference on Ultra Wideband Systems and Technologies, 21-23 May. Baltimore (USA), 87-91.

[17] N.G. Ferrara, M.Z.H. Bhuiyan, and S. Soderholm. (2018) A new implementation of narrowband interference detection, characterization, and mitigation technique for a software-defined multi-gnss receiver. Avaible:https//researchgate.net/326784680.

[18] K.Siwiak. (2001) Impact of ultra wide band transmissions on a generic receiver. IEEE VTS 53rd Vehicular Technology Conference, 6-9 May. Rhodes (Greece), 1181-1183.

[19] A.Taha, and K.M.Chugg. (2002) A theoretical study on the effects of interference uwb multiple access impulse radio. Available:https://ieeexplore.org/abstract/document/1197276.

[20] J.Bellorado, S.S.Ghassemzadeh, L.J.Greenstein, T.Sveinsson, and V.Tarokh. (2003) Coexistence of ultra-wideband systems with ieee-802.11a wireless lans. Available:https://ieeexplore.ieee.org/abstract/document/1258271.

[21] Kai Shi, Yi Zhou, Burak Kellici, Timothy Wayne Fischer, Erchin Serpedin, and Aydin Iker Karsilayan. (2007) Impacts of narrowband interference on ofdm-uwb receivers:analysis and mitigation. IEEE Transactions on Signal Processing, 55(3):1118-1128.

[22] A.Giorgetti, M.Chiani, and M.Z.Win. (2005) The effect of narrowband interference on wideband wireless communication systems. IEEE Transactions on Communications, 53(12):2139-2149.

[23] Xiaoli Chu, and R.D.Murch. (2004) The effect of nbi on uwb time-hopping systems. IEEE Transactions on Wireless Communications, 3(5):1431-1436.

[24] Ananthram Swami, B.M.Sadler and J.Turner. (2001) On the coexistence of ultra wideband and narrowband radio systems. Available:https//researchgate.net/publication/3937675.

[25] Zhiquan Bai, Xiatong Li, Dongfeng Yuan, Kyungsup Kwak. (2010) Non-linear chirp based on uwb waveform design for suppression of nbi. Wireless Communications and Mobile Computing, 12:545-552.

[26] Yamen Issa, Iyad Dayoud, Waala Hamouda. (2015) Performance analysis of multiple-input multiple-output relay networks based impulse radio ultra-wideband. Wireless Communications and Mobile Computing, 15:1225-1233.

[27] Hanbin Shen, Weihua Zhang, Kyung Sup Kwak. (2007) Non-linear chirp uwb ranging system with narrowband interference suppression abilities. J. ETRI, 29(4):521-523.

[28] S.Cui, K.C.Teh, K.H.Li, Y.L.Guan, C.L.Law. (2009) Performance analysis of transmitted-reference UWB systems with narrowband interference suppression. Wireless Communications and Mobile Computing, 9:1081-1088.

https://doi.org/10.31436/iiumej.v20i2.1097

[29] Xing Peng Mao, Jon W.Mark. (2008) Polarization filtering for narrowband interference suppression in ultra-wideband communications. Wireless Communications and Mobile Computing, 8:941-952.

[30] A compact fractal uwb antenna with open-ended quarter wavelength slot for band notch characteristics.Available:https://www.researchgate.net/publication/229566986.

[31] A WLAN notched wideband monopole antenna for ultra-wideband communication applications. Available: https://www.researchgate.net/publication/324463199.

[32] Circular slotted monopole printed antenna with grounded stub WLAN band-notch for UWB applications. Available: https://doi.org/10.1080/1448837X.2018.1452331.

[33] Dual-band shared-aperture UHF/UWB RFID reader antenna of circular polarization. Available: https://www.researchgate.net/publication/325327219.

[34] Dual notch band UWB antenna with improved notch characteristics. Available: https://www.researchgate.net/publication/323646136.

[35] Hemachandra R.Gorla, Frances J.Harackiewicz. (2015) Dual trident uwb planar antenna with band notch for wlan. Progress in Electromagnetics Research Letters, 54:115-121.

[36] Novel design of dual band-notched rectangular monopole antenna with bandwidth enhancement for UWB applications. Available: https://www.researchgate.net/ publication/325380975.

[37] Muhibbur Rahman, Jung-Dong Park. (2018) The smallest form factor uwb antenna with quintuple rejection bands for iot applications utilizing rsrr and rcsrr. National Center for Biotechnology Information US National Library of Medicine (NCBI) 18(911):1-16.

[38] Ammar Alhegazi, Zahriladha Zakaria, Noor Azwan Shairi, Sharif Ahmed, Tole Sutikno. (2017) UWB filtenna with electronically reconfigurable band notch using defected misrotrip structure. Indonesian Journal of Electrical Engineering and Computer Science, 8(2):302-307.

[39] Soufian Lakrit, Sudipta Das, Ali El Alami, Debaprasad Barad, Sraddhanjali Mohapatra. (2019) A compact uwb monopole patch antenna with reconfgurable band nocthed characteristics for Wi-MAX and WLAN applications. AEU International Journal of Electronics and Communications. Available: https://www.science direct.com/science/article/pii/S1434841119303048.

[40] Bing Gong, Chuang Ma, Fan Jing, Yaling Hou, Ruibing Shen. (2018) A novel uwb antenna with two ultra narrow and closely space notched bands. Journal of Physics, 1176(2019), 1-7.

https://doi.org/10.31436/iiumej.v20i2.1130

# A REVIEW ON SOFTWARE-DEFINED NETWORKING ENABLED IOT CLOUD COMPUTING

**SUMIT BADOTRA, SURYA NARAYAN PANDA**

*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.*

**Corresponding author: sumit.badotra@chitkara.edu.in*

*(Received: 8th April 2019; Accepted: 24th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** Making use of Internet of Things (IoT) is becoming the necessity of today's life. The data collected and stored through IoT devices comes from the cloud and therefore cloud computing is acting as a backbone for supporting IoT. But it is easy to forget that the cloud is not completely digital in some areas of the world and there is a need for a data centre where data storage can be achieved. Cloud data centres are facing many difficulties and issues because they are using traditional methods of networking. This is where Software Defined Networking (SDN) has come into view; where it has changed the way traditional networks are operated. For example, the separation of the intelligence of the network devices within the data path can be useful in making networks agile and manageable. This paper aims to provide problem areas in the networking used by the cloud data centres and the role of SDN to overcome these issues. Platforms for providing the experimental setup for collaboration between SDN and cloud will ultimately be beneficial in setting up the SDN- enabled IoT cloud, and this is also discussed with the open research problems.

***ABSTRAK:*** Penggunaan Internet Benda (IoT) menjadi keperluan kepada kehendak hari ini. Data yang dikumpul dan disimpan melalui peranti IoT berasal dari awan dan oleh itu pengkomputeran awan bertindak sebagai tulang belakang bagi menyokong IoT. Tetapi, satu perkara yang dilupakan adalah bahawa awan itu tidak sepenuhnya digital, di suatu tempat dalam dunia ini, terdapat keperluan pada pusat data di mana penyimpanan data boleh dicapai. Pusat data awan menghadapi banyak masalah dan isu kerana ianya menggunakan kaedah rangkaian tradisional. Di sinilah Perisian Definisi Perangkaian (SDN) dipandang; ia telah mengubah cara rangkaian tradisional dikendalikan. Sebagai contoh, pemisahan kecerdasan peranti rangkaian dengan laluan data berguna bagi membuat rangkaian pintar dan boleh diurus. Kajian ini bertujuan bagi menyediakan ruang masalah dalam rangkaian yang digunakan oleh pusat Data Awan dan peranan SDN dalam mengatasi masalah ini. Platform persediaan eksperimen bagi kerjasama SDN dan awan akhirnya bermanfaat dalam menubuhkan pengaktifan-SDN awan IOT, juga turut dibincangkan dalam kajian ini berkenaan masalah penyelidikan terbuka.

***KEYWORDS:*** *internet of things (IoT); software defined networking (SDN); data center networks; cloud computing*

## 1. INTRODUCTION

Internet of things (IoT) and cloud computing are strongly related to each other. Data transferred to various IoT devices comes from data centres. IoT is not all about how to connect various devices to the Internet to get the information but how to analyse the data

to get an intelligent answer to a query. Analysis of the data can only be possible if it is being stored at a data centre. Then, the IoT device can make a decision based on a previous data decision. For example, a smart watch is not only smart because it is connected to the Internet but because it is making use of some cloud where the data is stored, and analysing that cumulative data and making a smart decision. Cloud computing has emerged in such a simple way that its acceptance is universal. The computing that has been facilitated by cloud computing has eliminated several problems while providing services to the shoppers like direct investment, minimizing operational expenses, the handiness of on-request computing resources, protractible scaling, and creating pay-per-use [1]. As the utility of this computing model is expanding, it should be matched with a supplier of cloud with a similar brokering system that is dynamic for hosting its application [1]. The application supplier can lease resources used for computing from the cloud supplier by just clicking on a mouse at regular intervals and using their application without direct payment. With a dynamic adjustment (either increasing or decreasing) of the amount of virtualized infrastructure resources such as virtual machines (VMs), the extensibility and flexibility in the cloud atmosphere is achieved [2].

To give versatile support of their shoppers, cloud suppliers work with varied information centres over totally different locations with virtualized quality provisioning. Substantial giant information centres are comprised of thousands of switches interfacing unnumbered servers [3]. Each server will serve varied application demands from varied shoppers by utilizing virtualization advances that have improved readiness and asset capability that were not previously in the network resources [3]. As of now, cloud shoppers will lease virtualized process assets, i.e., VMs (virtual machines), and virtualized storage from different cloud suppliers that are allotted for the shopper. Be that as it may, the system quality virtualization innovation remains an extended method being employed within business public clouds [4].

Servers are allotted to totally different customers if the mass or total requested size from shoppers doesn't exceed server limit. Cloud shoppers accept inquiring for some littler measured resources than the limit of the entire server, and a server will offer varied virtual machines [5]. These virtual machines connect to the customer through varied system switches and switches. So as to transfer the traffic among varied VMs effectively and with information efficiency, network centres are established. Because the Data Centre Network (DCN) interfaces a huge range of servers, the system can have better quality, which makes it exhausting to administrate and scale within the standard systems administration view. In customary systems, every system switch has its own specific management explanation that solely chooses its conduct based on the info obtained from its neighbours [6]. The traditional system approach is wasteful with regards to the information centre of the cloud, where a better thickness provided by varied virtual machine servers that increasingly change over time [5, 6].

In order to overcome the disadvantages of traditional networks like measurability problems, inflexibility, etc. [7] different cloud information centres have started accepting the software-defined networking (SDN) thought in their DCNs [7, 8]. With the help of SDN, provisioning of incorporated centralized management logic with a perspective of the full system at the centralized controller will form dynamic changes on the conduct of the system [9, 10]. The network will likewise be modified by the controller, which is well suited to the more dynamic nature of the service provided by the cloud [11, 12]. Large suppliers of clouds, Google for example, formally received the SDN plan in their information centres [13]. There is a tremendous rise in the rate of IoT devices and to accommodate the growth of this rate, the cloud data centre has to expand its functionality

in order to provide the best services to every user. So, there is a need to implement the new architecture of cloud networking to enable innovation and continue the functionality and services of cloud.

In this paper, each cloud computing and SDN are considered for examination purposes. The main contribution of this paper is to explore the challenges and issues before and after introduction of SDN in IoT and cloud computing for future remedies. The remainder of this research paper is structured as follows:

In section II related literature is discussed along with the issues in the cloud data centre network with a brief introduction of SDN. In section III and IV comprehensive review of the state of art, including on-going projects and research, motivating SDN-Enabled IoT cloud computing is given. Section V is comprised of various evaluation methods that can be used as a platform for performing experiments in SDN-Enabled IoT cloud computing. Section VI explores the spectrum of challenges and obstacles in traditional networks, followed by concluding observations in section VII.

## 2. BACKGROUND

Although several surveys and taxonomies are conferred in SDN and cloud computing contexts, each of them presents their own drawbacks within the space. Toosi et al. [14] did a survey specializing in interconnected cloud computing. This research paper discussed various abilities in the context of multiple information centres and elaborately discussed every approach to work for cloud information centres that were interconnected. This paper had a separate section where it explained various networking challenges for inter-cloud connections, however, the main focus was on the broader problem of desegregation of multiple cloud information centres from the perspective of a cloud broker.

Mastelic et al. [15] additionally conferred a review on the potency of cloud energy for computing. In cloud computing, a scientific class of energy consumption has been instructed with respect to factors like software and hardware systems infrastructure in clouds. The author additionally discussed the networking aspect, accenting the information centre networking (DCN), end-user network, and inter-data-centre network. A detailed survey was done on various aspects of energy potency in networks, but the article was lacking with respect to the context of SDN.

Jararweh et al. [16] discussed in depth about software-defined clouds that specialized in networking, security, systems, and storage. However, his paper emphasized the system design instead of an individual analysis of SDN clouds.

Azodolmolky et al. [17] additionally proposed an innovative software-defined network. SDN can be applied to a number of challenges and will be utilized in future deployments as an economical solution along with cloud computing networking as a possible contributor of a software system outlined network. Some analysis results on the performance of the area unit are presented in this paper.

Yen et al. [18] initially established an association in Nursing SDN that is based on a cloud computing atmosphere via OpenFlow Switches, Controller packages, and Open supply. Then, it explained the practicality of the power saving OpenFlow controller to supply load equalization and observance mechanisms. To prove the feasibility of their planned design, they additionally developed a model of SDN queuing and estimated its analytical resolution for the chances of a steady-state. The result obtained from the experiment showed that the atmosphere based on SDN gave QoS-bonded services for cloud computing.

https://doi.org/10.31436/iiumej.v20i2.1130

Before starting an investigation of SDN-empowered cloud computing, the fundamentals of cloud computing are discussed and that's why the cloud information centres areas unit are represented. During the following section, we try to briefly discuss cloud computing, DCN, and SDN.

### 2.1 Cloud Computing

Cloud computing has come up as a typically acknowledged registering worldview worked around centre ideas [19], as an example, elimination of forthright venture, the diminishment of operational prices, on-request process resources, versatile scaling, and putting in compensation for every utilization arrange action for information innovation and computation administrations [20]. Cloud computing has given economy to ability through financially savvy and versatile IT profit worldview. The specialist additionally outfits its benefits, the pay-as-you-utilize estimating model for example, and flexibility to reinforce their nature of administration and to minimize their price [20]. In cloud condition, skill and flexibility will be accomplished by more decreasing or increasing virtualized foundation assets, e.g. VMs. In fact, virtualization had distended the efficiency and limit of cloud information centres to a degree that ancient system standards cannot offer [21-22]. Figure 1 shows the varied features of cloud computing.

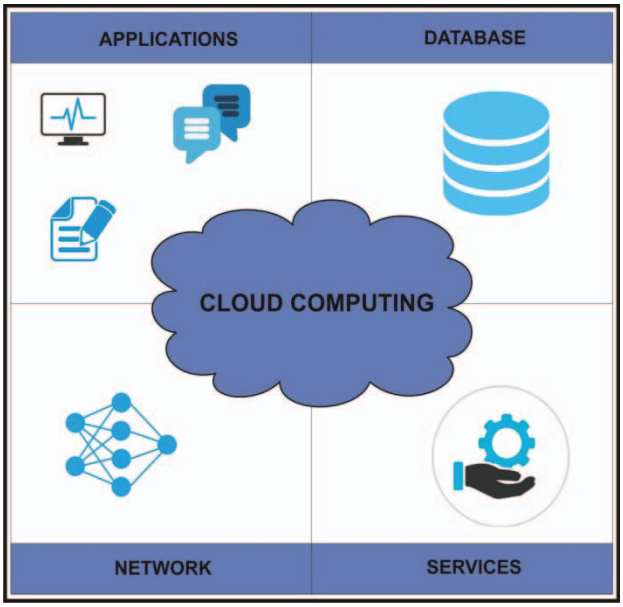



Fig. 1: Features of cloud computing.

The numerous service models of cloud computing enable varied administrations like Platform-as-a-service (PaaS), Software-as-a-Service (SaaS), Infrastructure-as-a-Service(IaaS), and Network-as-a-Service (NaaS) [23]. In SaaS, benefits for the outside provider has various applications that can later be made accessible to the shoppers online [24-25]. The NaaS profit model will incorporate elastic and broadened Virtual Private Network (VPN), transfer speed for the asking, custom steering, multicast conventions, security firewall, Intrusion Detection System, Wide Area Networking (WAN), content monitoring and separation, and antivirus [26-28]. There's no commonplace determination with respect to what's incorporated into NaaS and its executions can shift. Platform-as-a-Service (PaaS) may be a cloud computing model in which an outsider provider conveys instrumentation and programming devices for the main half to those needed for application advancement to shoppers over the web [29]. In PaaS, the supplier has the

instrumentation and programming framework [30-33]. Thus, PaaS allows its shoppers to put in in-house instrumentation and programming to form or run another applications [34]. Figure 2 demonstrates the various administration models of cloud computing.

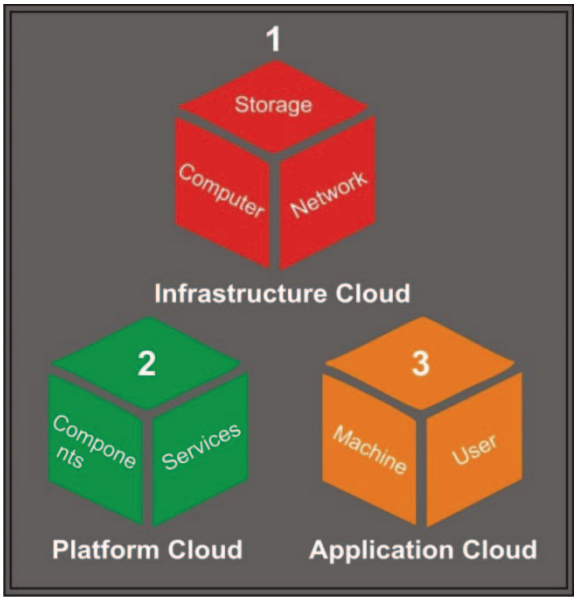



Fig. 2: Models of cloud.

### 2.2 Problem Area in the Cloud Data Centre Network (DCN)

Cloud computing has been well received and generally utilized everywhere. To keep pace with its rapid adoption, service providers must build an ever increasing number of data centres, placing them all around to diminish dormancy to the most extreme conceivable degree for worldwide customers [35]. In cloud situations, flexibility and versatility can be accomplished by expanding or diminishing virtualized framework assets such as virtual machines, yet there are a few issues in overseeing cloud framework on the IaaS level while considering both systems administration asset provisioning and configuring. A large number of networks are making use of a cloud foundation for themselves for dealing with these offices and associating them together. However, the present networking approach utilized in cloud data centre systems is becoming a hurdle to scaling up data centres. In straightforward words, it can neither be effectively organized nor adaptable [36]. With a specific end goal to give cloud benefits, the supplier should construct and keep up at least one vast scale cloud data centre where countless physical hosts are associated through a huge number of system switches [37]. With the high complexity of network connections on a huge scale, the provider needs to think about the DCN with an alternate point of view from the traditional system. Subsequently, it puts an expanding request on the networks and the systems as well [36-37].

### 2.3 Software Defined Networking

SDN is a novel approach in the area of networking [38]. The basic norm behind SDN is the segregation of the Data Plane and the Control Plane. In traditional networking both the intelligence and the forwarding plane are coupled inside proprietary hardware. This increases the complexity and decreases the expandability of the network [39]. With the introduction of SDN in networking, the whole concept of managing the networks has been changed, thereby making them programmable networks that allow the network

https://doi.org/10.31436/iiumej.v20i2.1130

administrator to change the traffic flow from one switch to another with few lines of code. Figure 3 shows the layered architecture of SDN:

a) **Physical Layer**: It is the layer that includes the infrastructure being used in the network like switches, routers, etc.

b) **Control Layer**: It is the main layer where the SDN controller resides and provides various network services and provisioning [40].

c) **Application Layer**: SDN architecture's topmost layer is the application layer. It includes all applications like firewall, load balancer, etc. [38, 39].

d) **API's (Application Programming Interface):** To provide the communication between various layers there are some well-defined API's such as Southbound API (e.g. OpenFlow protocol) and Northbound API (e.g. REST). [40].

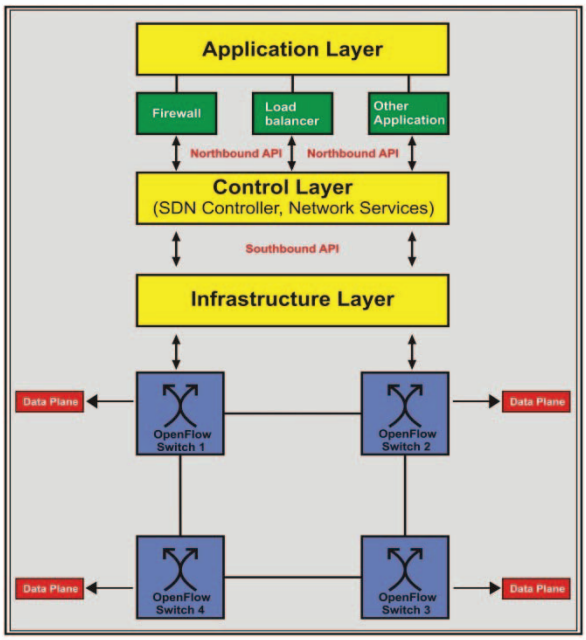



Fig. 3: Architecture of SDN.

Since the bond between sending equipment (or forwarding plane) and control logic has been separated in the case of SDN, so it provides more options to institutions and researchers for establishing new system administration conventions, traffic administration, virtualization, and programming advancements [41]. It prompted the presentation of OpenFlow in which aggregate consequences of various associations are and turn into an accepted standard convention for SDN controllers. Numerous controllers have been presented and effectively created in light of OpenFlow such as RYU, Floodlight, RYU, POX, Open Daylight etc. SDN has likewise opened up inventive open doors in transit of building up the controlling rationale of systems [41, 42]. In traditional systems, the control plane is exclusively created by substantial switch merchants, for example vendors such as Juniper or Cisco, that are combined to equip network hardware and hence make the network devices vendor-specific [41-43]. As the sending equipment is isolated and completely maintained by the SDN controller, programming with respect to SDN based network, various control logic can be combined and executed since it takes after SDN standard, for e.g., OpenFlow. Figure 4 demonstrates the correlation design of traditional network administration and SDN. In this way, SDN has encouraged advancements in the

improvement of different type of system intelligence or the control logic that might have been abundantly restricted in the approach used by conventional systems [43, 44].

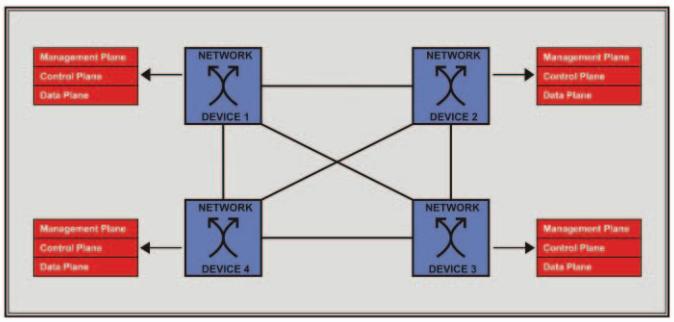



Fig. 4: Network architecture of traditional networks.

## 2. SDN ENABLED CLOUD COMPUTING

Combining SDN and cloud gives birth to new methods and possibilities through which new interfaces automatically become programmable [45]. Although there has been a tremendous rise in the field of cloud computing, those that are playing the vital role of managing cloud networks are facing a significant challenges on which the overall performance is dependent [46-47]. The issues in cloud networks include inflexibility, vendor-specific firewalls, scalability, manageability, etc. Consequently, cloud networks need a cloud infrastructure in which the aforementioned disadvantages are absent. In the future of networking SDN is an emerging model; with its basic principle of separation of the intelligence of the networks with the network, hardware with improved flexibility and scalability hence makes cloud networks programmable which increases the manageability and agility [48]. Out of all the issues in cloud networks, scalability is one of the main issues that need to be considered. SDN is the solution that can be used to upgrade the scalability of the cloud and fulfil the need of the hour. SDN can furnish a new architecture which supports the dynamism in network architecture and then further transforms the usual cloud networks into prosperous provider delivery platforms [49-50]. This is the reason for introducing SDN into the cloud. Making use of SDN in cloud computing can be referred to as SDN- Enabled cloud computing. The SDN based cloud is a new type of cloud in which SDN is introduced and network control is at one single place [51]. From centralized SDN controller, instructions to move the traffic from one place to another is provided to other networking devices [52-53]. In the cloud based on SDN, cloud computing extends from server virtualization and centralization as nicely as storage virtualization and centralization to community virtualization and centralization. Recently, clouds based on SDN have attracted first-rate attention [54].

When the SDN based cloud is compared with the traditional network used in the cloud, it has many advantages in terms of Quality of service (QoS), VM orchestration, etc. For the purpose of allocation, a required bandwidth is needed. This requirement of QoS for all the priority cloud users can be met by the SDN based cloud that uses an Open Vswitch [53]. SDN based networks are not only limited to provide the QoS but also live VM management in which the server can interrupt or exploit the information from one end to another. It is also believed that if SDN is implemented in the cloud it can offer new opportunities for network security to be enhanced since cloud-based SDN provides faster reaction and greater flexibility when the condition is getting revised [55].

### 3.1 Need for Software Defined Cloud Data Centres (SDN-DC)

To conquer the deficiencies of traditional networks, cloud data centres began embracing the SDN idea in DCN [51]. SDN offers a centralized control plane with a perspective of the whole system. With the help of the controller, it can make use of progressive change in the network. The need for dynamic flow of data from sender to receiver in the cloud environment and its data centres is well suited to SDN because of its centralized controller that can easily dynamically alter the flow of data in the network. On the basis of traditional networking used in DCN architecture, SDN-DC takes over from conventional systems administration hardware with SDN-empowered gadgets [52]. When SDN is implemented in DC, each system's administration that is a part of that data centre is fitted for implementing the features of SDN and then can convey all the aforementioned advantages to the cloud data centre [53]. SDN-empowered Cloud Computing (SDN-cloud or SDN-empowered clouds) specified to SDN-DC, as well as inter-cloud organizing that consumes the SDN utilization over various data centres and WAN. In this way, SDN benefits can be connected to internetworking domains [54, 55].

A more extensive term SDN-DC has been proposed by Jararweh et al. [8] and Buyya et al. [7] in which the network infrastructure is also software defined rather than only defining the networking. The approach was to construct a completely mechanized cloud server farm enhancing configurations, that was powerful and self-governing. Buyya et al. [7] expanded the idea of virtualization from a virtualized server to every other component in cloud data centres. The centre innovations to empower SDN-DC to incorporate server virtualization, SDN, arrange virtualization, and virtual middle boxes. With advances, the reconfiguration and adjustment of physical assets in SDN-DC has turned out to be more achievable and easier to be executed. The suggested architecture was executed and assessed in the simulation environment. Jararweh et al. [8] likewise considered a framework concentrating on different parts of SDC including capacity, organizing, data centres, virtualization, and security. It assembled an exploratory setup for SDN-DC with the examined components in the study to explain the viability of the proposed programming defined cloud architecture.

SDN-DC is more realistic as it is suggested for research purposes and has not yet been investigated broadly.

### 3.2 SDN-enabled Cloud Computing Architecture

SDN empowered architecture is separated into three layers. Figure 5 explains the abstract architecture of cloud computing empowered by SDN, based on the literature [46, 53-55]. The SDN controller controls the system related capacities, which are associated with the director of the cloud through APIs that are north-bound. SDN highlights such as arranging topology disclosure, virtual system administration; organizing checks and dynamic system designs are empowered through the controller of SDN. Different types of SDN controller can be practical to expand its adaptability through intercommunicating, with the help of east-west-bound APIs. Cloud resources consist of systems and process administration assets. The cloud administrator convinces hosts to run VMs on a hypervisor, though organizing assets like switches are overseen by the SDN controller by refreshing switches of sending tables through south-bound APIs (e.g. OpenFlow). Note that the SDN controller also manages functions of networking in hosts for examples, virtual switches for VM traffic management and network virtualization.

https://doi.org/10.31436/iiumej.v20i2.1130

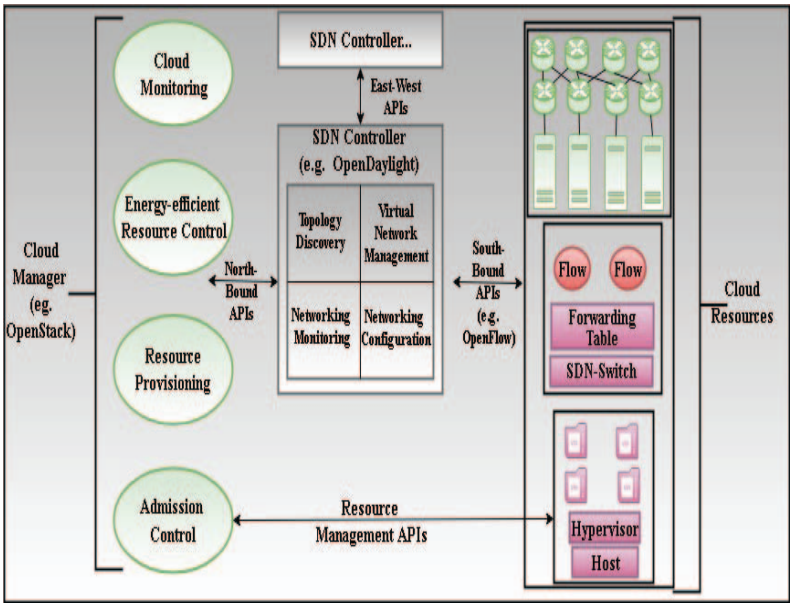



Fig. 5: SDN-enabled cloud architecture.

Recently, Mayoral et al. [56] provided the structure of the cloud with a single SDN controller as well as numerous controllers. The framework arranged the administration of a data centre with the SDN capacity. Keeping in mind the end result to discover the differences between one or more controllers in data centres, the author had assessed the OpenStack [57] architecture and OpenDayLight with respect to SDN. Similarly, the author had also proposed a conclusion to end arrangement of cloud architecture and system assets that oversee the two assets in distributed cloud data centres [58]. Virtual foundations in clouds are powerfully assigned and overseen through the chief. The VM distribution issues were displayed as a mapping figure with switch parameters and arranged joins. Additionally, a set of rules for the calculation of locating the base number of hosts with enough limits was incorporated and chose the shortest way between the hosts. The design of the organization was approved on an exact testbed with OpenVSwitch and OpenStack [57], while the calculation was done in simulation. Fichera et al. [59] exhibited in a detailed testbed setup for the organization of SDN crosswise over the edge, IoT spaces, and cloud in 5G benefits.

The foundation was comprised of isolated orchestrator(s) for the SDN, cloud, and IoT [60]. The SDN orchestrator oversees the cloud and IoT arrangements as well as the transport organized between them. Over individual orchestrators, the administration orchestrator gets direct benefits provisioning for applications with respect to the client's prerequisites. The proposed framework was examined in a testbed which comprised of ONOS (Open Networking Operating System) controller, OpenStack Cloud and Mininet for the edge arrangement. The Control Orchestration Protocol (COP) was suggested by specialists with respect to different organizations to allow overseeing heterogeneous control planes of SDN and Cloud [60]. While OpenFlow means controlling of messages between switches and the SDN controller as the controller southbound interfaces, COP is accountable for organizing among SDN controllers in various system and in addition controllers of cloud working north of controllers. COP brought together transport API for arranging diverse transport in the inter-cloud condition. This convention gives a deliberation of asset topology disclosure, provisioning and way calculation. It was approved through 2 proof-of-idea tests for organizing asset recovery and provisioning [60].

https://doi.org/10.31436/iiumej.v20i2.1130

# 4. SDN BASED IOT CLOUD

The boom of IoT and the quick enhancement of related applied sciences create an extensive connection of "things." Cloud computing is a field for creating big storage statistics and analysis. IoT is attractive on its own, but the enormous amount of data coming from IoT devices, its storage, big data statistics, and analysis must be performed by cloud computing. Real innovation will only come if IoT and cloud computing combine. The aggregation of both cloud computing and IoT will enable new services for monitoring and processing data. For example, sensory information from IoT devices can be uploaded and stored with cloud computing. This data can later be used for statistical analysis and decision making.

In truth, cloud computing and IoT are tightly coupled and are closely related to each other in a similar way that an automobile and gasoline are related. The strength produced with the aid of gas (the IoT) runs the vehicle (Cloud Computing). IoT is no longer solely about connecting random devices nearby and making robots. IoT is an imaginative and prescient tool whereby we are able to talk to gadgets near us in such a way that we can make informed decisions based totally on the statistics accrued by means of these devices.

One of the vital desires of cloud providers is to provide availability to their customers. To supply versatile guide of their clients, cloud vendors work with more than a few fact facilities over distinct geolocations with virtualized asset provisioning. Substantial massive records centres are comprised of heaps of switches interfacing countless servers. Every server can serve a number of utility needs from more than a few consumers by making use of virtualization advances, which have been increased inconceivably in preparation and capacity belongings for the most recent few many years yet not in network resources. As of now, cloud clients can lease virtualized processing assets, i.e., virtualized storage, and virtual machine from various cloud providers that are fixed surely for the client. Be that as it may, the machine asset virtualization innovation is nevertheless a lengthy way from being used in industrial public clouds.

Therefore, it can be concluded that as the data from IoT devices is coming in a huge manner, we need a cloud infrastructure in such a way that it has the capability to handle the rate of growth of the network. SDN has the capability to achieve this. Therefore, if SDN is implemented in the cloud it can increase the capability, efficiency of cloud storage and hence this cloud is called an SDN based IoT cloud.

## 4.1 Issues in SDN-Based IoT Cloud

A wide variety of gadgets enabled with IoT which are used to supply these applications, consisting of actuators, sensors and cellular devices. Communication among devices with one another is done through an integrated community that may additionally be defining the media type i.e. wireless or wired. It constructs the IoT in more necessary theme in its existence in daily practices which offers with the circumstances and the surrounding occasions such as fitness care, commercial agency administration, factories, automation, climate forecasting and many greater areas. Imagine the amount of records generated by means of 10 such government places of work and the quantity of computing power wanted to manage this information. Here, Cloud Computing comes into play. Hence the records gathered via IoT units is stored and processed on the Cloud and primarily based on our learning and processing of that information choices like building renovation, electrical faults, building lifespan is detected. The IoT is generating an extraordinary quantity of data, which in turn places a notable pressure on the Internet

infrastructure. As a result, businesses are working to locate methods to alleviate that stress and resolve the facts problem.

Another problem is Heterogeneity- meaning that IoT objects are changing in their properties and characteristics. As the numbers of devices are increasing day by day different vendors are coming into play to provide the services to various users. Devices that are communicating and getting the services are from different vendors is very hard to manage.

The need to combine SDN and IoT is to compile and resolve many types of issues that will be propagated as the IoT network grows. By making use of SDN, in an IoT network, the IoT controller, which is a high-level module, will interact with the centralized SDN controller for planning and organizing the network and performing the functions related to the IoT.

### 4.2 SDN-Based IoT Cloud Architecture

SDN-empowered architecture is separated into 3 layers. Figure 6 demonstrates the summary shapes of simple SDN-empowered cloud computing structure obtained from the literature. System-related capacities are managed by the SDN controller, which is related to the director of the cloud via northbound APIs. SDN highlights, for example, organize topology disclosure, prepare checking, virtual machine administration and dynamic system designs, are empowered via the SDN Controller. Various SDN Controllers can be sensible to lengthen adaptability through the ability to intercommunicate through the potential of east-west-bound APIs. The cloud property consists of buildings and manner administration assets. Hosts are managed by the network administrator of the cloud for running VMs on a hypervisor, although organizing property like switches is overseen using the SDN controller through clean sending of tables in switches with the resource of southbound APIs. It is very important to understand the fact that the SDN controller manages features of networking in hosts, for example, for community virtualization, VM visitor's administration, and digital switches.

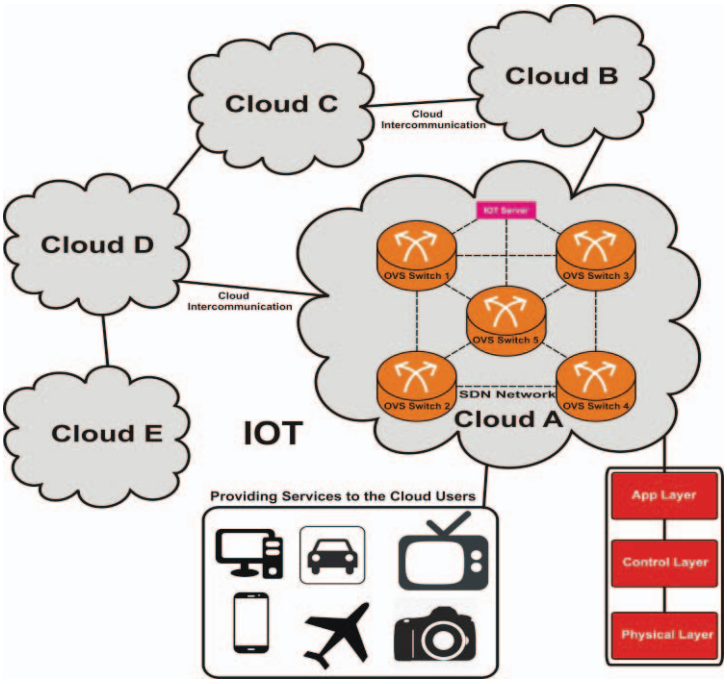



Fig. 6: SDN-based IoT cloud.

https://doi.org/10.31436/iiumej.v20i2.1130

### 4.3 Good Features of SDN for Offering Protection in SDN-Enabled IoT and Cloud Computing

- **Separation of the control plane from the data plane**: SDN decouples the intelligence of the network and management plane which constitutes the networking devices. Hence it permits setting up, besides troubles of large-scale attack and safety experiments. The high configurability of SDN not only allows the network administrator to have clarity among the digital networks but also how to enable experiments on a true environment.
- **Centralized Controller and its vision**: The SDN controller has a full vision and control over the network including traffic monitoring, shaping, switches flow table updating etc. Centralized manipulation of SDN controller allows dynamically resisting threats and attacks as well.
- **Analysis of traffic:** Network traffic can be examined on the basis of evaluation which is a software-based and which is ultimately enables the innovation, as it is possible to enhance the capabilities to change the usage of any technique which is a software-based. Analysis of the network traffic can be carried out in a real-time environment using desktop and getting to know algorithms used, backup databases and any other specific software program tool.
- **Dynamic updating the flow tables:** Flow tables of switches can be dynamically updated through a centralized SDN controller which further helps to prevent much security threats in the network.

## 5. EVALUATION METHOD AND TECHNOLOGY

For quickening improvement and advancement of SDN-empowered cloud computing, toolboxes and devices are required to manufacture a testbed for testing SDN and OpenFlow frameworks in a cloud data centre. The testbed also has the ability to quantify the vitality utilization to assess proposed algorithms. In this section, observational and simulation techniques are clarified as shown in Fig. 7.

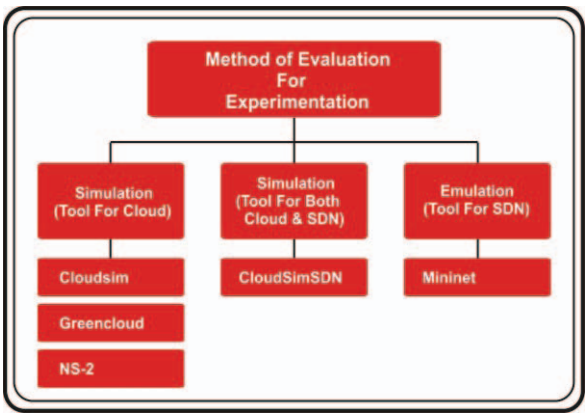



Fig. 7: Evaluation method.

### 5.1 Simulation Platforms for Cloud

Simulation is the first and initial tool for new research and experimental setups. Real Infrastructures are utilized to archetype the performance of the application under certain conditions like patterns of workload, reliability, etc. One of the major examples of

invariable frameworks is Microsoft Azure, Amazon EC2, etc. Therefore, this makes the transcription of results authentic and extremely considerable. On the other hand, it is time-consuming and unvaried to re-configure the parameters across a large-scale. There are some restrictions for the developers as the environment of the cloud is not in the control of them under certain circumstances. Thus, archetype practices in scaling, depending as well as repeating is impossible to execute in the real-world surroundings. An alternative tool that can open up the probability to examine the hypothesis in a manageable environment is a simulation tool. It can also easily recreate the outputs. IT Organizations get remarkable benefits from these kinds of tools.

The simulation platform gives a controlled and reproducible condition for assessment effortlessly of change and setup. For cloud computing, a number of simulation tools have been introduced to evaluate new ways to deal with controlling and overseeing the cloud server farm and different type of situations. CloudSim [61] is a prominent cloud test system executed in Java, giving a discrete occasion-based simulation equipped for mimicking cloud server, VMs, intermediaries, and hosts. Different scenarios can be executed in CloudSim, including the VM position approach, VM relocation arrangements, facilitating strategy and other server farm administration strategy [61]. It likewise bolsters workload reproduction executing in the VMs. With its simple to-utilize discrete occasion-based design, extra components can be added to send and get re-enactment occasions and in addition, stretching out the recent elements to give additional usefulness. As far as there is a concern of network events there is no support available by CloudSim in details [61].

To satisfy the lack or unavailability of system reproduction capacity of CloudSim, NetworkCloudSim [62] is acquainted with re-enacting applications with organizing correspondence assignments. Extra system components are included NetworkCloudSim, including system switches and connections that get arrange occasions and ascertain evaluated organize transmission time. In spite of the fact that NetworkCloudSim [62] incorporates broad system usefulness to mimic server farm system and various applications for message transferring within a data centre, the use of SDN in the experiment execution of CloudSim is clearly not supported [62].

GreenCloud [63] is another NS-2 [64] a based system which is based on a simulation that catches the vitality part of various cloud data centres comprising registration and correspondence of components. With the joining of NS-2 [64], which can catch the system design precisely on the bundle level, GreenCloud [63] can likewise give exact system comes about. The recreation substances incorporate has and switches with control utilization models, for example, DVFS. Workload models are likewise predefined and given in the system to three sorts of occupations, e.g., figure concentrated, information escalated, and adjusted. In NS-2 there is an availability of simulated environment which helps to provide a comprehensive framework with respect to clouds but there is no support for evaluation of SDN based applications.

### 5.2 Emulation Platform for SDN Networks

In order to create the network virtual host, switches and controller links between them to create a network emulated is used called as Mininet [65] emulator. Standard Linux network software is run by Mininet hosts, and Mininet switches support OpenFlow protocol for the purpose of communication and highly flexible custom routing and ultimately creating the networks which are software defined. Mininet not only is a good platform which is useful in providing support to research and developing applications but it is also helpful in making prototypes of networks, debug the networks and then perform testing on them. With the help of Mininet, a complete network including topologies,

server, and everything can be emulated and created in a single laptop. Mininet is the best emulation platform to emulate SDN networks in the absence of real hardware [66].

For SDN support in CloudSim, CloudSimSDN [67] was as of late acquainted with empower mimicking SDN highlights, for example, flexible transfer speed distribution or dynamic system setup. CloudSimSDN [67] is aimed to provide and simulate the highlighted features of SDN that further can be transferred in cloud data centres. Utilizing CloudSimSDN centralized SDN controller can be actualized inside the test system to imitate the conduct of the controller to reduce the unpredictability of copying the OpenFlow controller in the recreation [67]. Despite the fact that it gives just basic outcomes in systems administration reproduction, the disentangled engineering furnishes basic outcomes etc. with the help of SDN features in a very small period of time [67].

Teixeira et al. [68] proposed collaboration of Python-based SDN OpenFlow controller POX with Mininet to support simulation and making use of SDN features in cloud computing. To organize the network topologies Mininet is used to imitate them information activity and then organizing topologies within a data centre in which there is already a presence of OpenFlow SDN controller. By collaborating POX and Mininet, it helps to give handy outcomes and then preparing to-utilize programming. But still, this simulation tool suffers from many cons with respect to help or need in cloud-highlights, for example, characterizing different VM writes or execution of heterogeneous application workloads at the host through simulation.

### 5.3 Evaluation Method for Collaborating Both SDN and Cloud

To overcome the disadvantages occurred in various simulations and emulation tools discussed above. CloudSimSDN [67] has accomplished the task of combining SDN and Cloud together. It is a new simulation tool that provides the framework for SDN networks and Cloud. On top of CloudSim simulator, this new simulation tool is built. CloudSimSDN [67] includes the configuration and dynamic management of the resources which are managed and configure in a cloud data centre through a single point of control that is the SDN controller. The designing and formulation of this framework is done in such a way that it has the ability to perform the evaluation of all the resource management policies and strategies and ultimately which are to be used and appropriate for various SDN-ready cloud data centres [67].

With the help of CloudSimSDN simulation is performed of various cloud data centres, machines, switches in the network, links in between the network, and topologies (virtual) in network. This simulation is performed in order to calculate the approximated values of performance metrics which ultimately guarantees the Quality of service and less energy consumption to satisfy the needs of the environment conservation [67].

### 5.4 Real Platform

OpenStackEmu [57] is providing a platform for the consolidation of a network using emulation with SDN and OpenStack. ONO and ODL (OpenDayLight) [69] are the largest open-source SDN controllers that are making it possible for the SDN to join with OpenStack by means of modules. Modules in the OpenStack suite such as neutron can be designed to utilize as an alternate to the SDN controller rather than utilizing Neutron's own particular capacities.

https://doi.org/10.31436/iiumej.v20i2.1130

# 6. CHALLENGES TO SDN BASED CLOUD

Despite the fact that SDN-based cloud demonstrates numerous great highlights, it faces a few difficulties that must be considered, including execution, accessibility, and versatility security. Although numerous examinations have explored the utilization of SDN innovation in cloud computing, there are as yet a few perspectives that have not been investigated exhaustively [70-71].

## 6.1 Execution

Execution refers to the system hub for the preparing pace considering both latency and throughput [72]. The strategy for SDN was to deal with new packets which bring the network programmability. But, at the same time, it produces execution issues. The authors [73] had explained that nowadays controllers cannot deal with a huge number of flows in connections of 10 Gbps. So the question is how to keep programmability which is requiring for additional examination and to enhance execution.

## 6.2 Accessibility

The duration or the time in which the framework of SDN is in running condition is called accessibility. In the SDN network as there is a dependency on the single centralized controller. Though it brings the programmability, on the other hand, it also brings a challenge with respect to accessibility. One of the advantages that the traditional distributed network architecture is comprised of is that if a switch fails, the accessibility of the system does not affect and still can be looked for [74]. But on the other hand in the SDN network if a single point of control i.e. SDN controller fails the whole network collapses and then the accessibility of the system might be lost completely.

## 6.3 Scalability

It is the ability to increase the network infrastructure so that it can easily incorporate the demand or need for network growth. Then centralized single SDN controller can be converted into a bottleneck of versatility [75]. By introducing peer-with-peer or distributed controller foundation it may share the communication weight to the controller. It may be, a general system view is required to coordinate the correspondences between the controllers utilizing the westward and eastward APIs. Behind controller adaptability, there are some other versatility concerns including the overhead of flow setup and flexibility to failures.

One of the basic issues that need to be addressed is the scalability of the SDN controller [77]. SDN controller accumulates data from the whole system and then handles each and every individual network switch individually; this process makes the controller adaptable and keeps the single point of failure at the controller. In the cloud, this entanglement can become a critical issue because of the size multiple-sided quality of DCN and the normal SLA for the supplier [77]. The SDN controller for the cloud DCN can without much of a stretch turn into a solitary purpose of disappointment as controllers in the DCN can be over-burden because of the gigantic number of network switches. The aforementioned issue can be handled by disseminating controllers or by implementing a proactive approach while communicating between the SDN controller and switch [77].

## 6.4 Virtualization

Previously, NFV is provided by middle boxes, which are embedded with dedicated hardware, such as load balancer, firewall and intrusion detectors. These dedicated and vendor specific hardware devices are meant to perform the single task only and they don’t

https://doi.org/10.31436/iiumej.v20i2.1130

allow the network administrator to reconfigure their codes and hence hardware inflexibility is there which ultimately reduces the cost-effectiveness. Many researchers have already attempted to alter the hardware of middle boxes with a series of SDN switches which are embedded with specific controller programs [5].

### 6.5 Security

As compared to traditional networking SDN brings separation of intelligence of the SDN network constitutes in the control plane and network devices from the data plane. Due to this, there is an increase in the rise of security attacks [78]. Security investigation has demonstrated that the SDN system endures numerous security dangers, including:

1) unapproved get to, e.g., the unapproved controller gets to or unauthenticated application get to.
2) information spillage, e.g., flow control disclosure (side station assault on input cushion) sending strategy revelation (parcel handling timing examination),
3) Alteration in the information, e.g., modification in flow run to alter bundles [78-79].
4) noxious applications, e.g., false run addition controller commandeering.
5) Issues in configuration e.g., absence of TLS (or other verification strategies) reception strategy requirement.
6) Denial of service, e.g., while providing communication between switch-controller there may occur a switch or controller table flooding [80].

For various available open cloud providers, the most important aspect that needs to be addressed is security. While accessing the cloud services user's information is an important factor. Although security in cloud data centres has been explored broadly with multiple methodologies, collaborating SDN into it can make it vulnerable [81]. So, research can be directed towards providing security to the SDN enabled cloud data centres [82-83].

### 6.7 Energy Efficiency

Cloud data centres are consuming a huge amount of electricity which is increasing very rapidly. In 2013, 91 billion kilowatt-hours energy was consumed by US data centres, which is more than enough to power all households in NYC for 2 years. It is even expected that by 2020 it will be increased to approximately 140 billion kilowatt-hours which will cost the US $13 [7] billions annually. Therefore to improve energy efficiency while computing the cloud is becoming a most important research problem in both academics and in industries. There are already many researchers who are working in this area to optimize energy efficiency [7]. With the recent introduction of SDN, the technology of joint optimization is feasible and it can be explored in order to save maximum energy and therefore providing best QoS [7, 84].

### 6.8 Network Function Virtualization

DCN is comprised of many numbers of devices, these devices communicate with each other through multiple switches and devices and virtualization of the network in it is quite difficult because of the complexity of the network. There are gaps which are found while realizing virtualization of the network in clouds, which can be diminished by making of SDN. Although traditional networking technologies as VPN and VLAN not only accepted by industry is making use of them. But by combining SDN with the traditional methods can take even more flexibility, programmability, and security to

clouds. Thus for further research, virtualization of the network can be explored in a cloud data center.

### 6.9 System Function Virtualization

There are gaps in acknowledging virtualization of the system in the cloud, which can be feasible with the coordination of SDN [7]. Though the customary systems administration advancements, like VPN and VLAN has been broadly embraced in industry, co-ordination of SDN with the conventional technologies can bring much better execution and more adaptability to the cloud. Along with SDN, streams among VMs can be controlled more finely all the powerfulness and adaptability by adjusting the evolving activity. Progressively, switches can be continuously arranged in the supervision of the controller, which can direct the whole network. Therefore, virtualization of the network for SLA network guarantee is possible can be investigated in cloud data centres and introducing research in [7, 84].

## 7. FINAL REMARKS AND CONCLUSION

Use of IoT devices is increasing day by day. The foundation behind cloud networks, cloud computing, is making use of large networks to fulfil user demands. The cloud network is growing at a very fast pace but by making use of traditional networks, cloud networking is suffering from inflexibility, less manageability, and vendor specificity. To overcome the situation, SDN can be used in cloud data centre networks. SDN not only achieves programmability, agility, manageability but removes the barrier of vendor-specific by using the principle of separating the control plane and data plane. The job of the network administrator to overlook the network and managing the network can be easily achieved.

In this paper, an overview of traditional networking used in cloud computing with its disadvantages and how SDN can improve its functionality is described. This paper discussed the related work in this collaboration various limitations and advantages of their work is also surveyed. In order to implement cloud and SDN, this paper would be contributing to researchers. Various simulation and emulation platforms for performing the experiments are also discussed. At last research directions towards SDN based cloud are also provided.

## REFERENCES

[1] Senyo, P. K., Effah, J., & Addae, E. (2016). Preliminary insight into cloud computing adoption in a developing country. Journal of Enterprise Information Management, 29(4), 505–524.

[2] Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation computer systems, 25(6), 599-616.

[3] Senyo, P. K., Addae, E., & Boateng, R. (2018). Cloud computing research: A review of research themes, frameworks, methods and future research directions. International Journal of Information Management, 38(1), 128-139.

[4] Bayramusta, M., & Nasir, V. A. (2016). A fad or future of IT? : A comprehensive literature review on the cloud computing research. International Journal of Information Management, 36(4), 635–644.

[5] Son, J., & Buyya, R. (2018). A Taxonomy of Software-Defined Networking (SDN)-Enabled Cloud Computing. ACM Computing Surveys (CSUR), 51(3), 59.

[6] Al-Fares, M., Loukissas, A. and Vahdat, A., (2008), August. A scalable, commodity data center network architecture. In ACM SIGCOMM Computer Communication Review,Vol. 38, No. 4, pp. 63-74.
[7] Buyya, R., Calheiros, R.N., Son, J., Dastjerdi, A.V. and Yoon, Y., (2014), September. Software-defined cloud computing: Architectural elements and open challenges. In Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on (pp. 1-12). IEEE.
[8] Jararweh, Y., Al-Ayyoub, M., Benkhelifa, E., Vouk, M. and Rindos, A.,( 2016). Software defined cloud: Survey, system and evaluation. Future Generation Computer Systems, 58, pp.56-74.
[9] Kreutz, D., Ramos, F.M., Verissimo, P.E., Rothenberg, C.E., Azodolmolky, S. and Uhlig, S., (2015). Software-defined networking: A comprehensive survey. Proceedings of the IEEE, 103(1), pp.14-76.
[10] Badotra, S. and Singh, J., (2017). A Review Paper on Software Defined Networking. International Journal of Advanced Research in Computer Science, 8(3).
[11] McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, G., Peterson, L., Rexford, J., Shenker, S. and Turner, J., (2008). OpenFlow: enabling innovation in campus networks. ACM SIGCOMM Computer Communication Review, 38(2), pp.69-74.
[12] Lantz, B., Heller, B. and McKeown, N., (2010), October. A network in a laptop: rapid prototyping for software-defined networks. In Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks (p. 19). ACM.
[13] Vahdat, A., Clark, D. and Rexford, J., (2015). A purpose-built global network: Google's move to SDN. Queue, 13(8), p.100.
[14] Toosi, A.N., Calheiros, R.N. and Buyya, R., (2014). Interconnected cloud computing environments: Challenges, taxonomy, and survey. ACM Computing Surveys (CSUR), 47(1), p.7.
[15] Mastelic, T., Oleksiak, A., Claussen, H., Brandic, I., Pierson, J.M. and Vasilakos, A.V., (2015). Cloud computing: Survey on energy efficiency. Acm computing surveys (csur), 47(2), p.33.
[16] Jararweh, Y., Al-Ayyoub, M., Benkhelifa, E., Vouk, M. and Rindos, A., (2016). Software defined cloud: Survey, system and evaluation. Future Generation Computer Systems, 58, pp.56-74.
[17] Azodolmolky, S., Wieder, P. and Yahyapour, R., (2013), June. SDN-based cloud computing networking. In Transparent Optical Networks (ICTON), 2013 15th International Conference on (pp. 1-4). IEEE.
[18] Yen, T.C. and Su, C.S., (2014), April. An SDN-based cloud computing architecture and its mathematical model. In Information Science, Electronics and Electrical Engineering (ISEEE), 2014 International Conference on (Vol. 3, pp. 1728-1731). IEEE.
[19] Buyya, R., Broberg, J. and Goscinski, A.M. eds., (2010). Cloud computing: Principles and paradigms (Vol. 87). John Wiley & Sons.
[20] Grover, J. and Sharma, M., (2014), July. Cloud computing and its security issues—A review. In Computing, Communication and Networking Technologies (ICCCNT), 2014 International Conference on (pp. 1-5). IEEE.
[21] Yang, H., & Tate, M. (2012). A descriptive literature review and classification of cloud computing research. Communications of the Association for Information Systems, 31(1), 2.
[22] Kaur, S. and Singh, A., (2012). The concept of cloud computing and issues regarding its privacy and security. International Journal of Engineering Research & Technology (IJERT), 1(3).
[23] Choo, K.K.R., (2010). Cloud computing: challenges and future directions. Trends and Issues in Crime and Criminal justice, (400), p.1.
[24] Zhang, Q., Cheng, L. and Boutaba, R., (2010). Cloud computing: state-of-the-art and research challenges. Journal of internet services and applications, 1(1), pp.7-18.
[25] Kumar, D.K., Rao, D.G.V. and Rao, D.G.S., (2012). Cloud computing: An analysis of its challenges & security issues. International Journal of Computer Science and Network (IJCSN), 1(5).

[26] Dogra, N. and Kaur, H., (2013). Cloud computing security: issues and concerns. International Journal of Emerging Technology and Advanced Engineering, 3(3).
[27] Al-Fares, M., Loukissas, A. and Vahdat, A., (2008), August. A scalable, commodity data center network architecture. In ACM SIGCOMM Computer Communication Review (Vol. 38, No. 4, pp. 63-74). ACM
[28] Clos, C., (1953). A study of non-blocking switching networks. Bell System Technical Journal, 32(2), pp.406-424.
[29] Guo, C., Wu, H., Tan, K., Shi, L., Zhang, Y. and Lu, S., (2008), August. Dcell: a scalable and fault-tolerant network structure for data centers. In ACM SIGCOMM Computer Communication Review (Vol. 38, No. 4, pp. 75-86). ACM.
[30] Guo, C., Lu, G., Li, D., Wu, H., Zhang, X., Shi, Y., Tian, C., Zhang, Y. and Lu, S., (2009). BCube: a high performance, server-centric network architecture for modular data centers. ACM SIGCOMM Computer Communication Review, 39(4), pp.63-74.
[31] Zhang, Z., Deng, Y., Min, G., Xie, J. and Huang, S., (2017). ExCCC-DCN: A highly scalable, cost-effective and energy-efficient data center structure. IEEE Transactions on Parallel and Distributed Systems, 28(4), pp.1046-1060.
[32] Bilal, K., Malik, S.U.R., Khalid, O., Hameed, A., Alvarez, E., Wijaysekara, V., Irfan, R., Shrestha, S., Dwivedy, D., Ali, M. and Khan, U.S., (2014). A taxonomy and survey on green data center networks. Future Generation Computer Systems, 36, pp.189-208
[33] Chen, T., Gao, X. and Chen, G.,( 2016). The features, hardware, and architectures of data center networks: A survey. Journal of Parallel and Distributed Computing, 96, pp.45-74.
[34] Xia, W., Zhao, P., Wen, Y. and Xie, H., (2017). A survey on data center networking (DCN): Infrastructure and operations. IEEE communications surveys & tutorials, 19(1), pp.640-656.
[35] Bari, M.F., Boutaba, R., Esteves, R., Granville, L.Z., Podlesny, M., Rabbani, M.G., Zhang, Q. and Zhani, M.F., (2013). Data center network virtualization: A survey. IEEE Communications Surveys & Tutorials, 15(2), pp.909-928.
[36] He, T., Toosi, A. N., & Buyya, R. (2019). Performance evaluation of live virtual machine migration in SDN-enabled cloud data centers. Journal of Parallel and Distributed Computing, 131, 55-68.
[37] Abts, D., Marty, M.R., Wells, P.M., Klausler, P. and Liu, H., (2010). Energy proportional datacenter networks. ACM SIGARCH Computer Architecture News, 38(3), pp.338-347.
[38] Amarasinghe, H., Jarray, A. and Karmouch, A., (2017), May. Fault-tolerant IaaS management for networked cloud infrastructure with SDN. In Communications (ICC), 2017 IEEE International Conference on (pp. 1-7). IEEE.
[39] Nunes, B.A.A., Mendonca, M., Nguyen, X.N., Obraczka, K. and Turletti, T., (2014). A survey of software-defined networking: Past, present, and future of programmable networks. IEEE Communications Surveys & Tutorials, 16(3), pp.1617-1634.
[40] Alsaeedi, M., Mohamad, M. M., & Al-Roubaiey, A. A. (2019). Toward Adaptive and Scalable OpenFlow-SDN Flow Control: A Survey. IEEE Access, 7, 107346-107379.
[41] Bilal, R., & Khan, B. M. (2019). Software-Defined Networks (SDN): A Survey. In Handbook of Research on Cloud Computing and Big Data Applications in IoT (pp. 516-536). IGI Global.
[42] T Karakus, M., & Durresi, A. (2017). A survey: Control plane scalability issues and approaches in software-defined networking (SDN). Computer Networks, 112, 279-293.
[43] Kim, H. and Feamster, N.,(2013). Improving network management with software defined networking. IEEE Communications Magazine, 51(2), pp.114-119.
[44] Jarraya, Y., Madi, T. and Debbabi, M., (2014). A survey and a layered taxonomy of software-defined networking. IEEE communications surveys & tutorials, 16(4), pp.1955-1980.
[45] Azodolmolky, S., Wieder, P. and Yahyapour, R., (2013). Cloud computing networking: Challenges and opportunities for innovations. IEEE Communications Magazine, 51(7), pp.54-62.
[46] Banikazemi, M., Olshefski, D., Shaikh, A., Tracey, J. and Wang, G., (2013). Meridian: an SDN platform for cloud network services. IEEE Communications Magazine, 51(2), pp.120-127.

[47] Priya, I. D., & Silas, S. (2019). A Survey on Research Challenges and Applications in Empowering the SDN-Based Internet of Things. In Advances in Big Data and Cloud Computing (pp. 457-467). Springer, Singapore.
[48] Yan, Q., Yu, F.R., Gong, Q. and Li, J., (2016). Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges. IEEE Communications Surveys & Tutorials, 18(1), pp.602-622.
[49] Jain, R. and Paul, S.,( 2013). Network virtualization and software defined networking for cloud computing: a survey. IEEE Communications Magazine, 51(11), pp.24-31.
[50] Ma, Y.W., Chen, Y.C. and Chen, J.L., (2017), February. SDN-enabled network virtualization for industry 4.0 based on IoTs and cloud computing. In Advanced Communication Technology (ICACT), 2017 19th International Conference on (pp. 199-202). IEEE.
[51] Li, F., Cao, J., Wang, X., Sun, Y. and Sahni, Y., (2017), June. Enabling Software Defined Networking with QoS Guarantee for Cloud Applications. In 2017 IEEE 10th International Conference on Cloud Computing (CLOUD) (pp. 130-137). IEEE.
[52] Bruschi, R., Davoli, F., Lago, P., Lombardo, A., Lombardo, C., Rametta, C. and Schembra, G., (2017). An SDN/NFV Platform for Personal Cloud Services. IEEE Transactions on Network and Service Management, 14(4), pp.1143-1156.
[53] Papagianni, C., Androulidakis, G. and Papavassiliou, S., (2014), February. Virtual topology mapping in SDN-enabled clouds. In Network Cloud Computing and Applications (NCCA), 2014 IEEE 3rd Symposium on (pp. 62-67). IEEE.
[54] Alhazmi, K., Shami, A. and Refaey, A., (2017). Optimized provisioning of SDN-enabled virtual networks in geo-distributed cloud computing datacenters. Journal of Communications and Networks, 19(4), pp.402-415.
[55] Nguyen, T. G., Phan, T. V., Nguyen, B. T., So-In, C., Baig, Z. A., & Sanguanpong, S. (2019). SeArch: A Collaborative and Intelligent NIDS Architecture for SDN-Based Cloud IoT Networks. IEEE access, 7, 107678-107694.
[56] Buyya, R., Yeo, C.S., Venugopal, S., Broberg, J. and Brandic, I., (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation computer systems, 25(6), pp.599-616.
[57] Son, J., Dastjerdi, A.V., Calheiros, R.N. and Buyya, R., (2017). SLA-aware and energy-efficient dynamic overbooking in SDN-based cloud data centers. IEEE Transactions on Sustainable Computing, 2(2), pp.76-89.
[58] Pajila, P. B., & Julie, E. G. (2019, February). Detection of DDoS Attack Using SDN in IoT: A Survey. In Intelligent Communication Technologies and Virtual Mobile Networks (pp. 438-452). Springer, Cham.
[59] Martínez, R., Mayoral, A., Vilalta, R., Casellas, R., Muñoz, R., Pachnicke, S., Szyrkowiec, T. and Autenrieth, A., (2017). Integrated SDN/NFV orchestration for the dynamic deployment of mobile virtual backhaul networks over a multilayer (packet/optical) aggregation infrastructure. Journal of Optical Communications and Networking, 9(2), pp.A135-A142.
[60] OpenStack Foundation. (2018). Open Source Software for Creating Private and Public Clouds. Retrieved from https://www.openstack.org/
[61] Mayoral, A., Munoz, R., Vilalta, R., Casellas, R., Martínez, R. and López, V., (2017). Need for a transport API in 5G for global orchestration of cloud and networks through a virtualized infrastructure manager and planner. IEEE/OSA Journal of Optical Communications and Networking, 9(1), pp.A55-A62.
[62] Fichera, S., Gharbaoui, M., Castoldi, P., Martini, B. and Manzalini, A., (2017), July. On experimenting 5G: Testbed set-up for SDN orchestration across network cloud and IoT domains. In Network Softwarization (NetSoft), 2017 IEEE Conference on (pp. 1-6). IEEE.
[63] Stallings, W. (2015). Foundations of modern networking: SDN, NFV, QoE, IoT, and Cloud. Addison-Wesley Professional.
[64] Calheiros, R.N., Ranjan, R., Beloglazov, A., De Rose, C.A. and Buyya, R., (2011). CloudSim: a toolkit for modeling and simulation of cloud computing environments and

evaluation of resource provisioning algorithms. Software: Practice and experience, 41(1), pp.23-50.
[65] Badotra, S., & Singh, J. (2017). A Review Paper on Software Defined Networking. International Journal of Advanced Research in Computer Science, 8(3).
[66] Kliazovich, D., Bouvry, P. and Khan, S.U., (2012). GreenCloud: a packet-level simulator of energy-aware cloud computing data centers. The Journal of Supercomputing, 62(3), pp.1263-1283.
[67] Network Simulator NS-2. Retrieved from http://www.isi.edu/nsnam/ns/
[68] Ahn, G. J., Gu, G., Hu, H., & Shin, S. (2019, March). SDN-NFV security 2019 preface. In SDN-NFV 2019-Proceedings of the ACM International Workshop on Security in Software Defined Networks and Network Function Virtualization, co-located with CODASPY 2019 (p. III).
[69] Riley, G. F., & Henderson, T. R. (2010). The ns-3 network simulator. In Modeling and tools for network simulation (pp. 15-34). Springer, Berlin, Heidelberg.
[70] Son, J., Dastjerdi, A.V., Calheiros, R.N., Ji, X., Yoon, Y. and Buyya, R., (2015), May. Cloudsimsdn: Modeling and simulation of software-defined cloud data centers. In Cluster, Cloud and Grid Computing (CCGrid), 2015 15th IEEE/ACM International Symposium on (pp. 475-484). IEEE.
[71] Teixeira, J., Antichi, G., Adami, D., Del Chiaro, A., Giordano, S. and Santos, A., (2013), October. Datacenter in a box: Test your sdn cloud-datacenter controller at home. In Software Defined Networks (EWSDN), 2013 Second European Workshop on (pp. 99-104). IEEE.
[72] Badotra, S. and Singh, J., (2017). Open Daylight as a Controller for Software Defined Networking. International Journal of Advanced Research in Computer Science, 8(5).
[73] Bhushan, K. and Gupta, B.B., (2017). Security challenges in cloud computing: state-of-art. International Journal of Big Data Intelligence, 4(2), pp.81-107.
[74] Bhushan, K., & Gupta, B. B. (2019). Distributed denial of service (DDoS) attack mitigation in software defined network (SDN)-based cloud computing environment. Journal of Ambient Intelligence and Humanized Computing, 10(5), 1985-1997.
[75] Sezer, S., Scott-Hayward, S., Chouhan, P.K., Fraser, B., Lake, D., Finnegan, J., Viljoen, N., Miller, M. and Rao, N., (2013). Are we ready for SDN? Implementation challenges for software-defined networks. IEEE Communications Magazine, 51(7), pp.36-43.
[76] Jarschel, M., Oechsner, S., Schlosser, D., Pries, R., Goll, S. and Tran-Gia, P., (2011), September. Modeling and performance evaluation of an OpenFlow architecture. In Proceedings of the 23rd international teletraffic congress (pp. 1-7). International Teletraffic Congress.
[77] Pandya, B., Parmar, S., Saquib, Z. and Saxena, A., (2017), March. Framework for securing SDN southbound communication. In Innovations in Information, Embedded and Communication Systems (ICIIECS), 2017 International Conference on (pp. 1-5). IEEE.
[78] Yan, Q., Yu, F. R., Gong, Q., & Li, J. (2015). Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges. IEEE Communications Surveys & Tutorials, 18(1), 602-622.
[79] Sezer, S., Scott-Hayward, S., Chouhan, P.K., Fraser, B., Lake, D., Finnegan, J., Viljoen, N., Miller, M. and Rao, N.,(2013). Are we ready for SDN? Implementation challenges for software-defined networks. IEEE Communications Magazine, 51(7), pp.36-43.
[80] Yeganeh, S.H., Tootoonchian, A. and Ganjali, Y., (2013). On scalability of software-defined networking. IEEE Communications Magazine, 51(2), pp.136-141.
[81] Jantila, S. and Chaipah, K., (2016). A security analysis of a hybrid mechanism to defend DDoS attacks in SDN. Procedia Computer Science, 86, pp.437-440.
[82] Dargahi, T., Caponi, A., Ambrosin, M., Bianchi, G. and Conti, M., (2017). A survey on the security of stateful SDN data planes. IEEE Communications Surveys & Tutorials, 19(3), pp.1701-1725.
[83] Wang, B., Zheng, Y., Lou, W. and Hou, Y.T., (2015). DDoS attack protection in the era of cloud computing and software-defined networking. Computer Networks, 81, pp.308-319.

[84] Negru, C., Mocanu, M., Cristea, V., Sotiriadis, S. and Bessis, N., (2017). Analysis of power consumption in heterogeneous virtual machine environments. Soft Computing, 21(16), pp.4531-4542.

https://doi.org/10.31436/iiumej.v20i2.1142

# DESIGN AND IMPLEMENTATION OF TWO-PHASE INTERLEAVED VOLTAGE SOURCE INVERTER FOR PV APPLICATIONS

**HARIKA SRIDHARAN, SEYEZHAI RAMALINGAM**

*Department of Electrical and Electronics Engineering,*
*SSN College of Engineering, Chennai, India.*

**Corresponding author: sriharkal@gmail.com*

*(Received: 10th March 2018; Accepted: 29th June 2019; Published on-line:2nd December 2019)*

***ABSTRACT:***The design of a voltage source inverter is challenging for PV-grid connected systems due to power quality issues. To improve the power quality,a two-phase interleaved voltage source inverter (IVSI) is proposed in this paper. IVSI phase shifts two voltage source inerters connected in parallel. The inverter topology is interfaced with a 40W PV panel employing a multiple maxima search (MMS) MPPT algorithm. This algorithm results in higher tracking efficiency compared to existing methods. Modelling of the PV with MPPT and the circuit configuration of the interleaved inverter is simulated in MATLAB/SIMULINK. The switches in the IVSI is controlled by employing a unipolar PWM technique. The performance of the IVSI is investigated in terms of weighted total harmonic distortion (WTHD), distortion factor (DF), harmonic spread factor (HSF), and inductor current ripple and compared with classical VSI. From the analysis, it is inferred that the proposed inverter results in reduced total harmonic distortion(THD) and decreased inductor current ripple thereby producing a high-quality output. The hardware model of the two-phase IVSI is developed and interfaced with PV where gating pulses are generated in FPGA. The simulation results are validated experimentally.

***ABSTRAK*:** Rekaan pembalik sumber voltan adalah mencabar pada sistem gabungan grid-PV disebabkan isu kualiti tenaga. Bagi memperbaharui kualiti tenaga, pembalik sumber voltan antara lembar dua-fasa (IVSI) telah dicadangkan dalam kajian ini. Anjakan fasa IVSI pembalik sumber voltan dua-fasa telah dihubungkan secara selari. Topologi pembalik telah diantaramukakan dengan 40W panel PV dengan mengunakan carian maksima berganda (MMS) algoritma MPPT. Hasil algoritma menunjukkan lebih tinggi keberkesanan dibandingkan dengan kaedah sedia ada. PV model dengan MPPT dan konfigurasi litar pembalik antaramuka telah disimulasi menggunakan MATLAB/SIMULINK. Suis dalam IVSI dikawal dengan menggunakan teknik PWM unipolar. Dapatan kajian ke atas IVSI adalah berdasarkan berat pengherotan total harmoni (WTHD), faktor pengherotan (DF), faktor sebaran harmoni (HSF), getaran arus pembalik dan dibandingkan dengan kaedah klasik VSI. Analisis menunjukkan kesimpulan dapatan kajian ke atas pembalik yang dicadangkan telah berkurang jumlah pengherotan harmoni (THD) dan getaran arus pembalik telah berkurang, oleh itu menghasilkan hasil akhir yang berkualiti tinggi. Model perkakasan IVSI dua-fasa telah dibangunkan dan diantaramukakan dengan PV di mana signal pengegetan dihasilkan dalam FPGA. Dapatan hasil simulasi telah disahkan secara eksperimen.

***KEYWORDS:****IVSI; total harmonic distortion(THD);ripple; MPPT*

https://doi.org/10.31436/iiumej.v20i2.1142

# 1. INTRODUCTION

Nowadays, renewable energy has been employed for various applications in order to meet the electricity demand and in particular, solar PV power generation is mostly used due to its eco-friendly nature and abundant availability [1]. The difficulty with this power generation is tracking peak power. To obtain maximum power from PV, various maximum power point algorithms have been investigated in the research work. [2]. The traditional MPPT algorithms are perturb-and-observe (P&O), incremental conductance (INC), fractional open circuit voltage, fractional short circuit current, and current sweep. The P&O and INC suffer from the drawback that they lose the MPP tracking if the solar irradiation varies rapidly. Existing fractional open circuit voltage and fractional short circuit current have the same drawback as their MPP tracking is not real, leaving the computation based on approximation. The sweep waveform of the PV array current is used in the case of the current sweep tracking method, which obtains MPP if the sweep is instantaneous but is difficult to apply in practical situations. Thus the traditional algorithms fail to find global MPP under variable irradiation conditions [3]. Thus, this research work proposes a MMS algorithm that follows the direct search technique (DST) to track a maximum operating point that overcomes the cons of the traditional MPPT algorithms. DST samples the sampling interval by assuming one point to be the center, dividing that intotwo intervals, and proceeding in the same way for the other intervals[4, 5]. In this way, an optimum operating point is obtained. With this MMS algorithm, high tracking efficiency can be achieved even in a partially shaded condition. Tracking efficiency of 97% is achieved with DST, based on the dividing rectangle algorithm. It is also possible to achieve a tracking efficiency of 99 % if this direct search algorithm is employed to get a globalpoint alongwith the traditional algorithm. Thus, the MMS MPPT algorithm is implemented, as itis suited for PV [6, 7].

A voltage source inverter is the most commonly used interface circuit for PV [8] but it suffers from various issues like high THD and high inductor current ripple content. Thus, this research work focuses on interleaving two voltage source inverters, known as the two-phase interleaved voltage source inverter (IVSI) [9]. IVSI is defined as a parallel operation of two voltage source inverters with an interleaving angle (κ) that is chosen based on the equation given by,

$$\kappa = \frac{360}{N} \qquad (1)$$

where, $N$ = number of the inverter connected in parallel.

The paper deals with twoinverters connected in parallel ($N$ = 2). From equation (1), it is found that the interleaving angle is 180 degrees. The performance of the VSI depends upon the control strategy [10]. So, the proposed topology is investigated with different modulation strategies such as fundamental frequency switching, unipolar sine PWM, and space vector modulation. The performances of IVSI with different modulation strategies are compared and unipolar SPWM was found to have low THD and reduced inductor current ripple [11]. Thus, to improve the powerquality of the waveform, IVSI with unipolar SPWM technique is studied in MATLAB/SIMULINK. Finally, a prototype model of the two-phase interleaved voltage source inverter is built and interfaced with a 40W solar panel. The schematic block diagram of the overall work is depicted in Fig.1.

The performance parameters computedfor comparisonare total harmonic distortion ($THD_v$) and inductor current ripple. From the results, it is found that the two-phase IVSI has less THD and reduced inductor current ripple compared to traditional VSI.

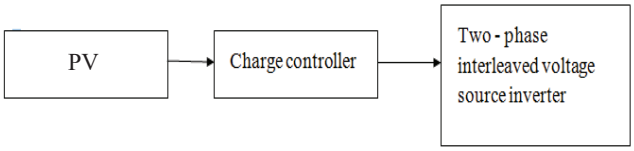



Fig.1: Block diagram of the proposed work.

## 2. TWO-PHASE INTERLEAVED VOLTAGE SOURCE INVERTER

The main feature of the two-stage IVSI is that it eliminates the zero voltage level, which is present in conventional voltage source inverters [12,13]. With the single-level inverter, the voltage levels are $+V_{dc}/3$, 0, $-V_{dc}/3$ but with a two-levelinverter, the voltage levels are $+V_{dc}/2$ and $-V_{dc}/2$. This shows that IVSI suits the PV application [14]. The circuit diagram of the proposed parallel connected voltage source inverter is shown in Fig.2.

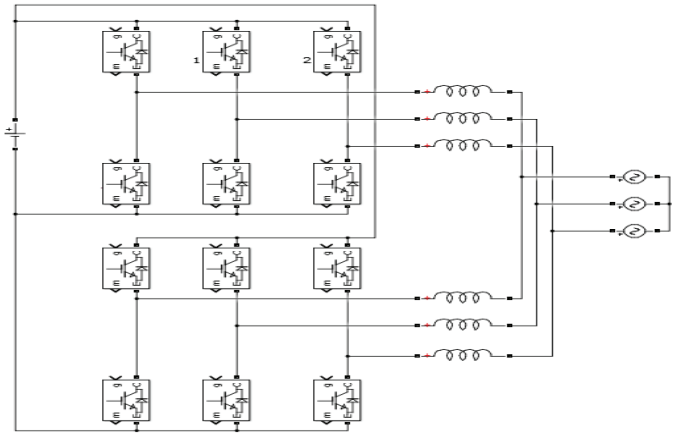


Fig.2: Circuit diagram of the parallel connected voltage source inverter.

Figure 2 shows that the two voltage source inverters are parallel through the filter inductance and connected to the grid. The operation of the IVSI is elaborated into six periods, $T_1$ to $T_6$. The switching states are different for different intervals [15, 16]. During $T_1$, the inductor voltage is zero if a single inverter is employed whereas, with two parallel connected inverters, the voltage across the inductor is $V_{dc/2}$. Similarly, during $T_2$, if a single inverter is employed, the inductor voltage will be $-V_{dc/3}$ whereas for interleaved inverter the voltage level will be - $V_{dc/2}$. Thus, the zero voltage level is eliminated in the interleaving operation of the VSI. The circuit diagram of the interleaved VSI during $T_1$ and $T_2$ is depicted in Fig.3.

## 3. MODULATION STRATEGIES

The proposed two-stage IVSI is analyzed with different modulation strategies such as fundamental frequency switching, unipolar SPWM, and space vector modulation [17]. The results of the different modulation strategies are compared in terms of current THD (total harmonic distortion) and inductor current ripple content [18]. From the comparison, the unipolar SPWM is chosen since it has low $THD_i$ and reduced ripple content. Then, the outputs of the interleaved inverter and conventional voltage source inverter are compared to prove that the proposed inverter topology has low THD, has low filter requirements, and operates at a high switching frequency.

https://doi.org/10.31436/iiumej.v20i2.1142

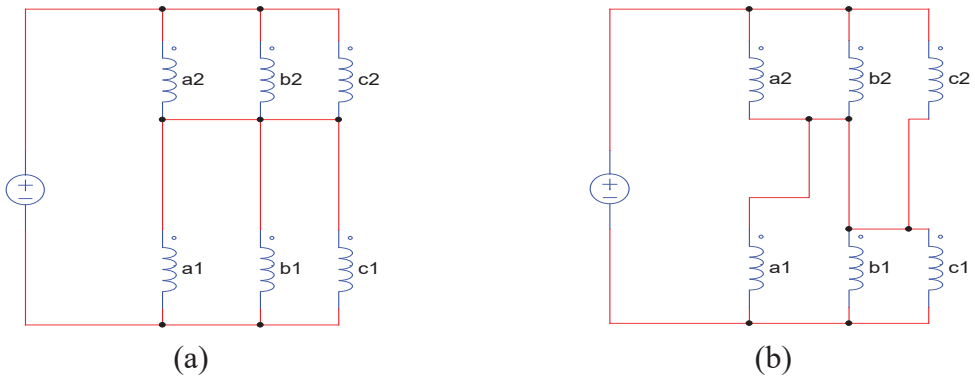



Fig.3: Operation of the proposed inverter (a) during $T_1$ and (b) during $T_2$.

The unipolar PWM requires one sinusoidal reference waveform and one triangular carrier waveform. The switches in the same leg are triggered in a complementary manner with one switch turned on and other switch turned off [19]. Thus, only the gating pattern of the upper switch is considered. The three sinusoidal waveforms with 120-degree phase shift are compared with the triangular carrier waveform for $VSI_1$ as depicted in Fig. 4, whereas for $VSI_2$ , the three reference sinusoidal waveform is compared with the 180 degree phase shifted carrier waveform. The pulse pattern for two-stage interleaved VSI is shown in Fig.5. $S_1$, $S_2$, and $S_3$ denote pulse output for $VSI_1$ and $S_1$',$S_2$',and $S_3$' denote pulse output for $VSI_2$.

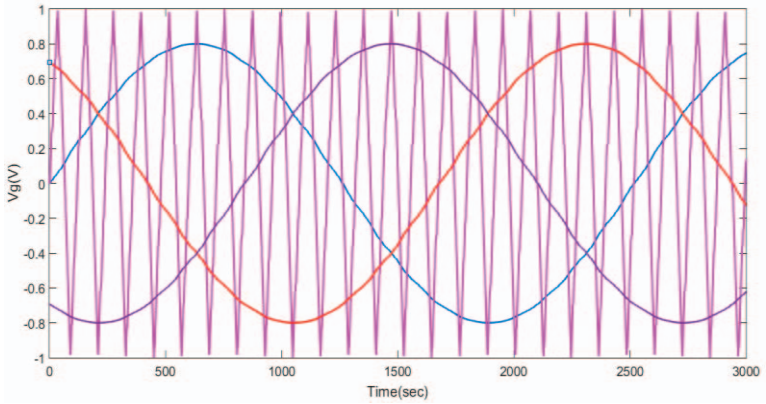



Fig.4: Reference and carrier waveforms for Unipolar PWM.

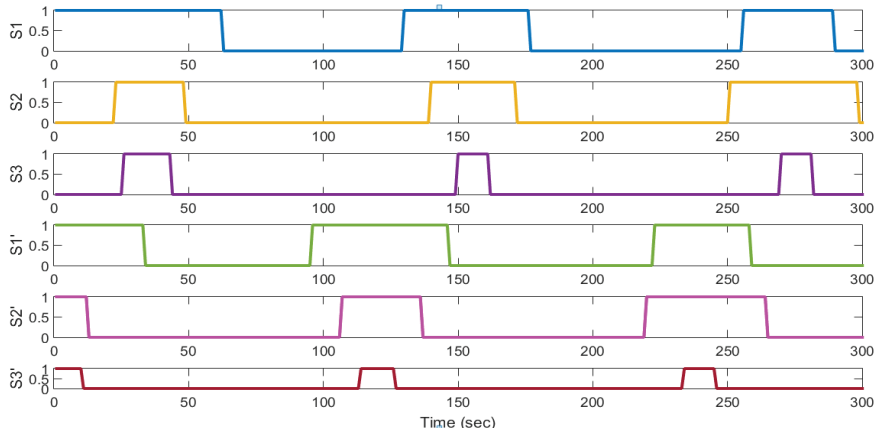



Fig.5: Gating pattern of IVSI.

https://doi.org/10.31436/iiumej.v20i2.1142

## 4. SIMULATION RESULTS

The interleaved voltage source inverter has been simulated in MATLAB/Simulink and the various PWM strategies have been implemented. The simulation model is depicted in Fig.6 and the simulation parameters are shown in Table. 1. The output voltage of the interleaved inverter is shown in Fig.7.

Table 1: Simulation Parameters

| PARAMETER | VALUES |
|---|---|
| DC voltage($V_{dc}$) | 400 V |
| Inductor(L) | 50mH |
| Modulation Index ($m_a$) | 0.8 |
| Switching Frequency | 1050 Hz |
| Interleaving Angle | 180 Degree |

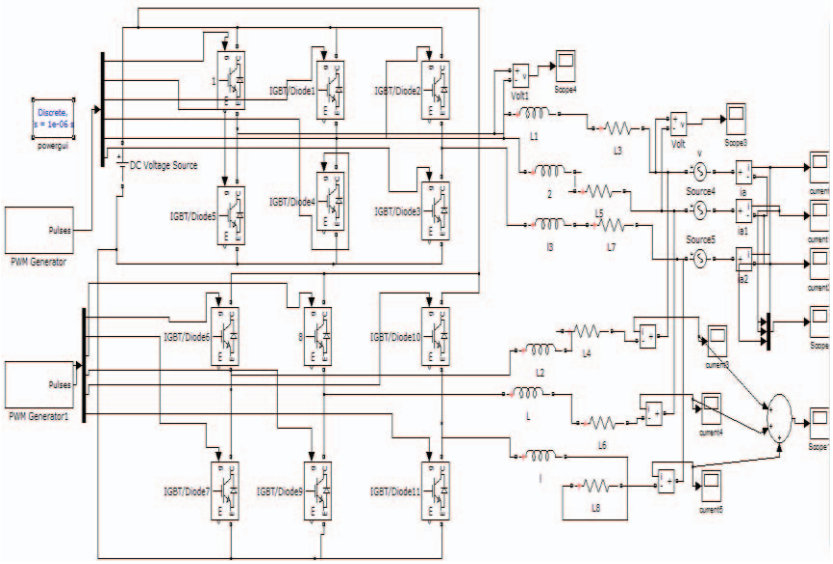



Fig.6: Simulink model of the proposed two-stage IVSI.

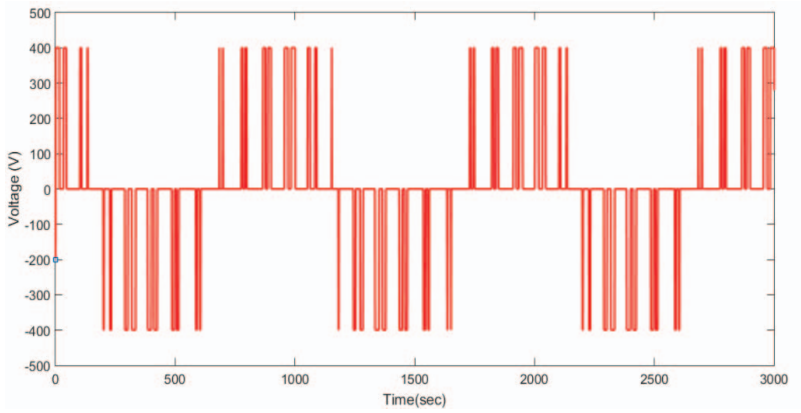



Fig. 7: Output Voltage of the parallel connected VSI.

https://doi.org/10.31436/iiumej.v20i2.1142

Table 2: Comparison of different modulation strategies for the interleaved inverter

| Modulation techniques | $THD_i$% | $I_{Lr}$ |
|---|---|---|
| **Unipolar SPWM** | 0.97 | $L_{ia}$=0.82, $L_{ib}$=0.82, $L_{ic}$ = 0.82 |
| **Fundamental frequency switching** | 1.07 | $L_{ia}$ = 1.46, $L_{ib}$ = 1.38, $L_{ic}$ = 1.20 |
| **Space vector PWM** | 2.92 | $L_{ia}$ = 2.59, $L_{ib}$ = 2.60, $L_{ic}$ = 2.59 |

Table 3: Comparison of interleaved inverter and conventional VSI

| Modulation techniques | $THD_i$% | $I_{Lr}$ |
|---|---|---|
| **Conventional single-stage voltage source inverter** | 2.33 | $L_{ia}$ = 2.33, $L_{ib}$ = 2.33, $L_{ic}$ = 2.32 |
| **Unipolar SPWM** | 0.97 | $L_{ia}$ = 0.82, $L_{ib}$ = 0.82, $L_{ic}$ = 0.82 |

The comparison of the simulation results are depicted in Table 2 and Table3, showing that the two-stage IVSI with unipolar PWM has low THD and reduced ripple content compared to single-stage VSI. Thus, the performance of the proposed topology with unipolar SPWM is analysed for different modulation indexes. The performance parameters are total harmonic distortion(THD), weighted total harmonic distortion(WTHD), distortion factor(DF), and harmonic spread factor(HSF).

### 4.1 Total Harmonic Distortion (THD)

A deviation in the shape of an obtained waveform from its fundamental component is called total harmonic distortion, the expression for THD is

$$\text{THD} = \frac{1}{(V_o)_1}\sqrt{\sum_{3,5,7}^{\infty}(V_o)_n^2} \qquad (2)$$

where$(V_0)_n$ = RMS value of the harmonic component, $(V_0)_1$= RMS value of the fundamental component.

### 4.2 Weighted Total Harmonic Distortion (WTHD)

Evaluating the quality of pulse width modulation (PWM) inverter waveforms is commonly called a weighted total harmonic distortion (WTHD). The frequency and the measured voltage harmonics are inversely proportional to each other.

$$WTHD = \frac{\sqrt{\sum_{n=2}^{\infty}(\frac{V_n}{n})^2}}{V_1} \quad (3)$$

### 4.3 Distortion Factor (DF)

The distortion factor is a measure of the intensity of the deviations of the measured waveform from its fundamental waveform and the mathematical expression of the DF is

$$\text{DF} = \frac{\sqrt{\sum_{n=2,3}^{\infty}(\frac{V_n}{n^2})^2}}{V_1} \quad (4)$$

https://doi.org/10.31436/iiumej.v20i2.1142

where $V_1$ = Fundamental voltage, $V_n$ = total harmonics voltage, $n$ = order of the harmonics.

### 4.4 Harmonic Spread Factor(HSF)

Harmonic Spread Factor defines the noise generation level in the motor and measures the voltage spectrum quality of the inverters.

$$\text{HSF} = \sqrt{\frac{\sum_{j=2}^{N}(H_j - H_o)^2}{N}} \quad (5)$$

where $H_j$= value of $j^{th}$ harmonic, $H_0$ = average value of all $N$ harmonics.

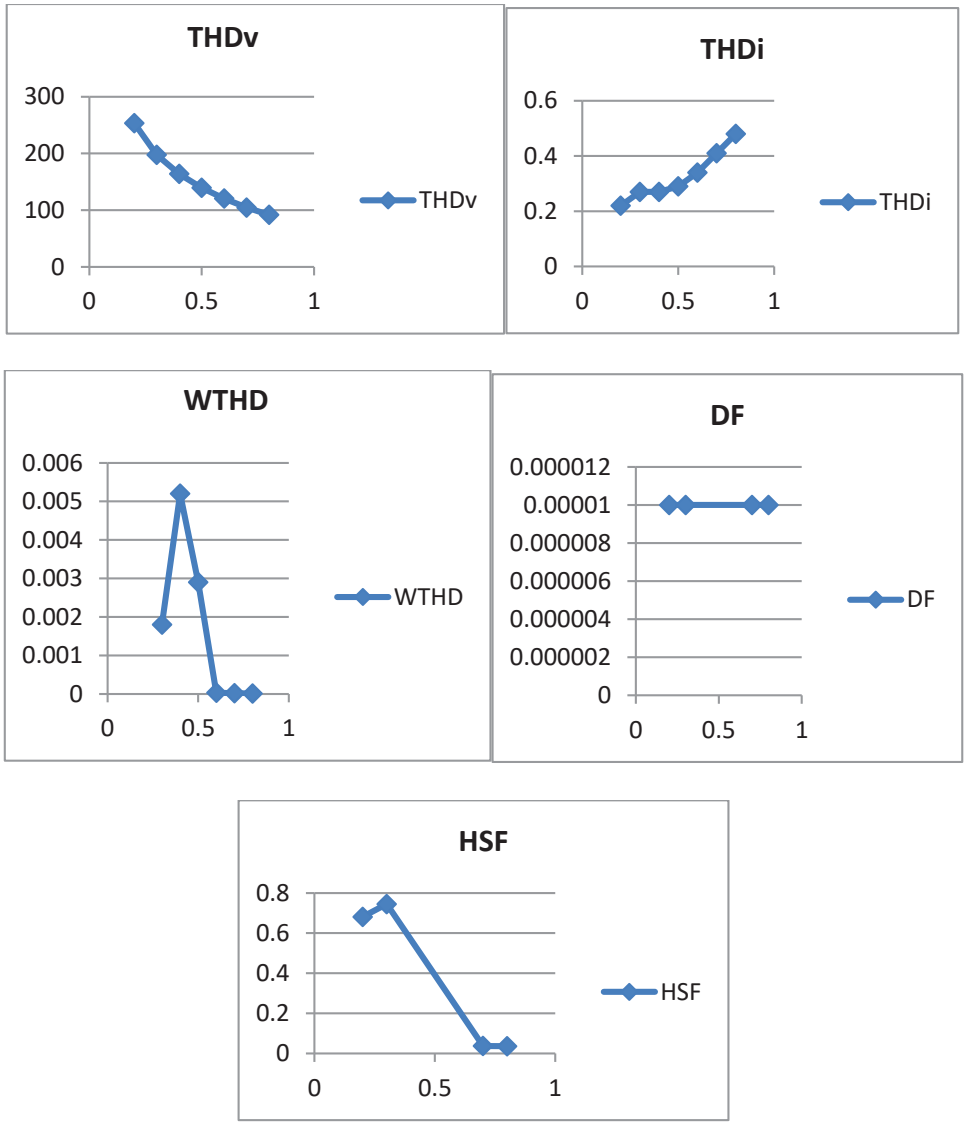



Fig. 8: Performance analysis of the interleaved VSI.

From Fig. 8, it is proven that the total harmonic distortion(THD), weighted total harmonic distortion(WTHD), distortion factor(DF), and harmonic spread factor(HSF) are low at a modulation index of 0.8. Hence, 0.8 is chosen as the modulation index for the proposed topology.

https://doi.org/10.31436/iiumej.v20i2.1142

## 5. INTERFACE OF INTERLEAVED VSI WITH PV

A PV panel is modeled in MATLAB/SIMULINK. The parameters are obtained from the PV panel datasheet and the important electrical parameters need to be considered to get electrical characteristics of the PV panel for the open-circuit voltage and the short-circuit current [20]. Practically, V-I and P-V characteristics are observed for a 40W solar panel as shown in Fig.9.

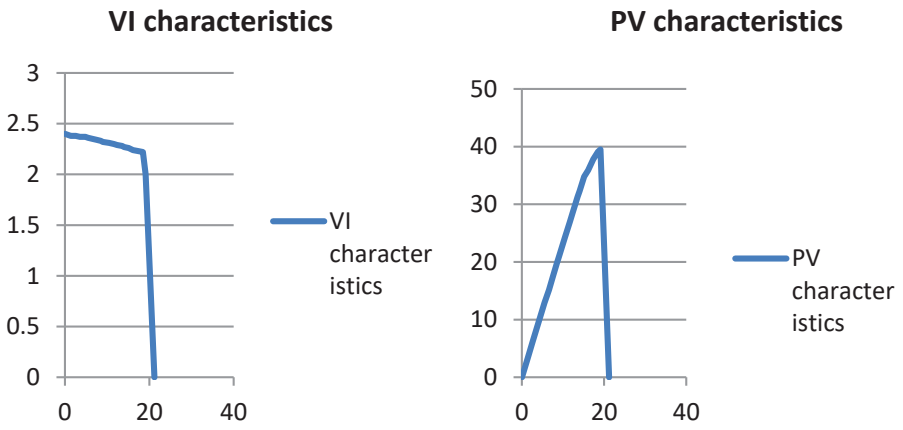



Fig. 9: V-I and P-V characteristics of PV model for 40W panel.

To track the maximum power from the PV, a MMS MPPT algorithm is used. The direct search technique based on the dividing rectangle strategy is clearly explained in Fig.10.

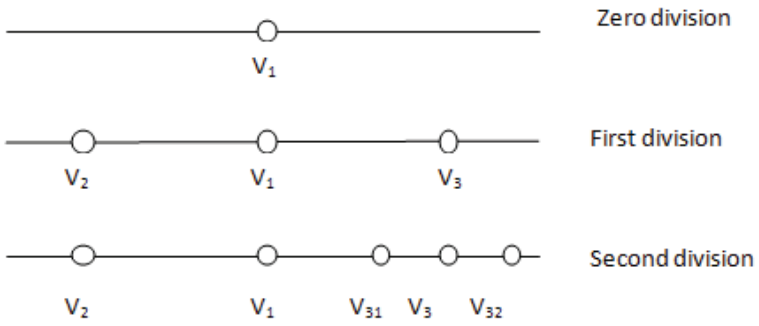



Fig.10: Dividing rule.

Figure 10 shows that $V_1$ is taken as a sample and assumed as its centerpoint. When compared to $V_1$ and $V_2$, $V_3$ has the maximum peak and it is further divided into three samples as $V_{31}$, $V_3$,and $V_{32}$. By doing so, the MMS MPPT algorithm tracks the maximum power, even under the partially shaded condition, and reaches a high tracking efficiency of 97%. The results are compared with the traditional perturb-and-observe method. Both MMS and P&O algorithms are modelled in MATLAB/SIMULINK and results are shown in Fig. 11.

From Fig.11,tracking efficiency is calculated:

With multiple maxima search MPPT,

$$\text{Tracking Efficiency} = \frac{48}{52} = 92\%$$

https://doi.org/10.31436/iiumej.v20i2.1142

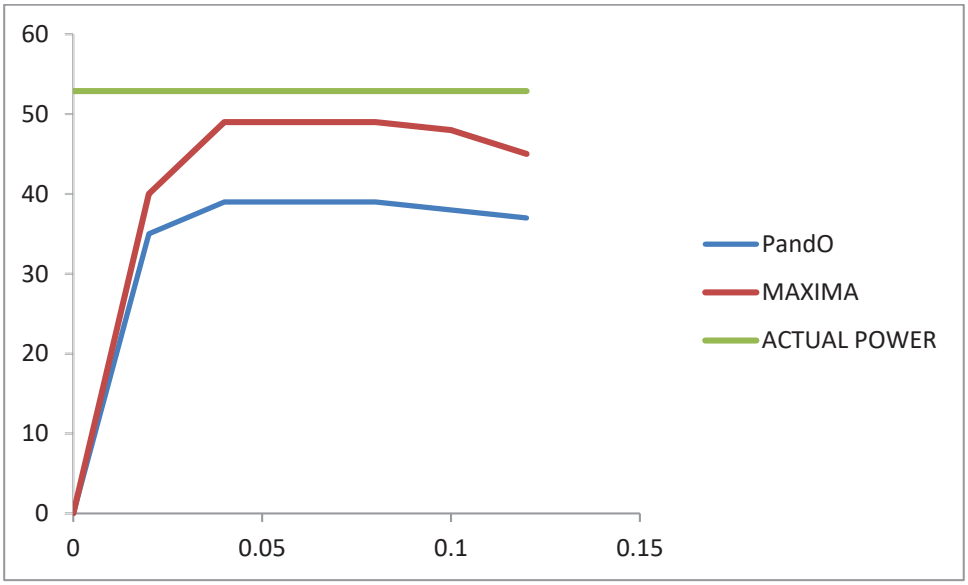



Fig.11: Actual power and tracked power.

With perturb-and-observe,

$$\text{Tracking Efficiency} = \frac{39}{52} = 75\%$$

Thus, it is proven that the multiple maxima search MPPT has high tracking efficiency compared to a traditional algorithm. The overall circuit diagram of the proposed interleaved inverter interfaced with PV panel and output is shown in Fig.12 and Fig.13.

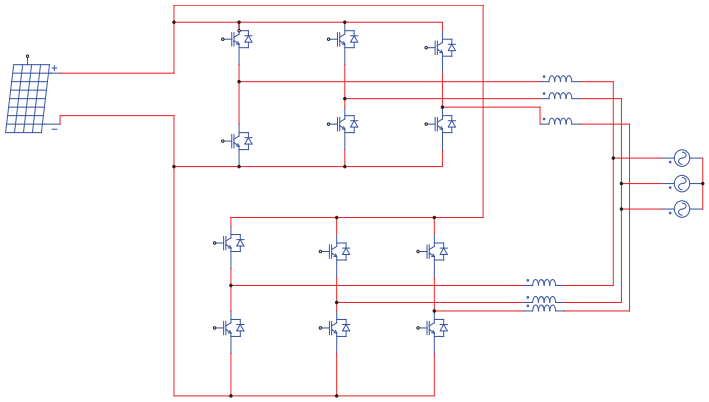


Fig.12: PV interfaced with IVSI.

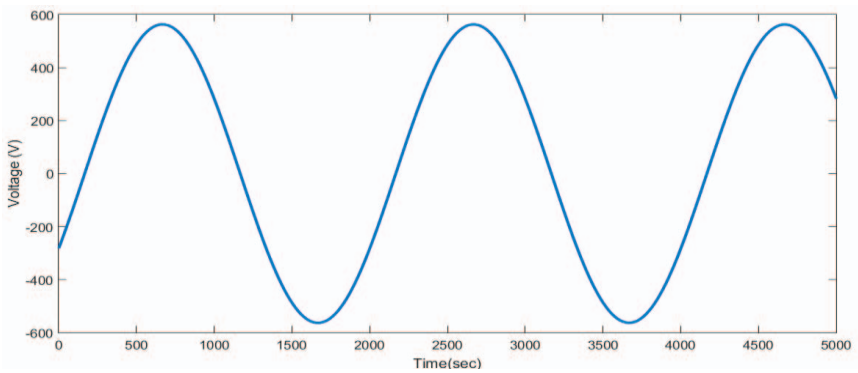



Fig.13: Output Voltage (filtered output) of inverter interfaced with PV.

https://doi.org/10.31436/iiumej.v20i2.1142

## 6. HARDWARE RESULTS

The two-stage IVSI with unipolar SPWM is implemented in hardware, which includes a gating circuit, pulse generation using FPGA-SPARTAN 3E, and design and development of a three phase two-stage interleaved voltage source inverter. The gating for the two-stage interleaved voltage source inverter is generated using Xilinx Spartan 3E FPGA. The generated pulse is given to the switch (MOSFET IRF 840) using an optocoupler IC (TLP 350). The gating circuit is important for providing gate pulses to switches used in the inverter in order to isolate the low power control circuit and the high power inverter circuit. The pulse generated by FPGA is given through the gating circuit. The logic for pulse generation of the proposed inverter is shown in Figs.14 - 16.

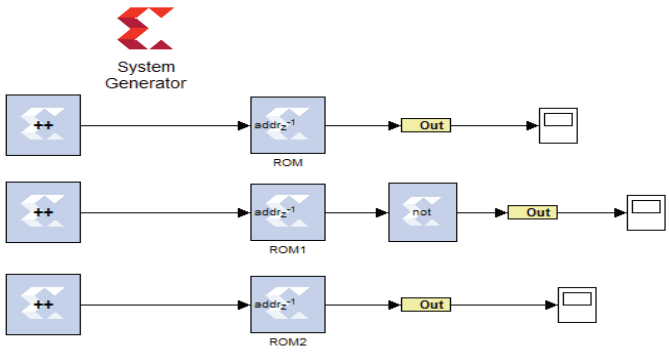



Fig.14: Sine wave generation.

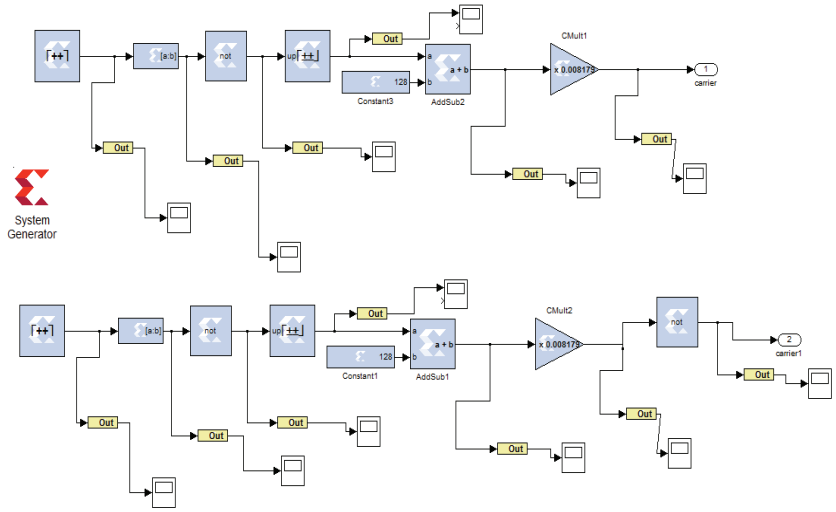
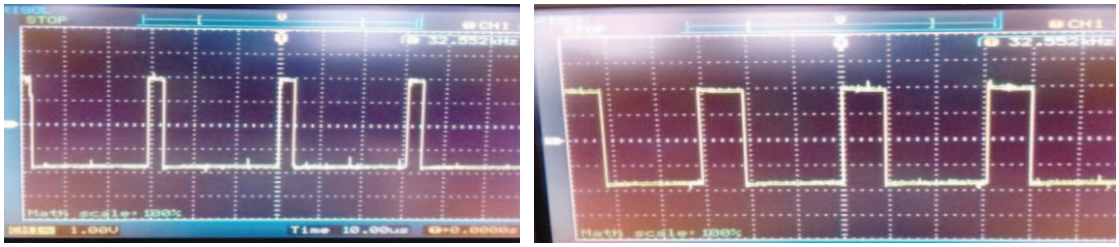



Fig.15: Triangular wave generation.



(a) Switch S1 (b) Switch S2

Fig.16: Gating pulse for interleaved VSI.

Figure 17 shows the prototype model of a two-stage interleaved voltage source inverter with gating circuit. The line-to-line voltage of the IVSI has a THD of 6.576% and its bar graph representation is shown in Fig.18.The phase voltage of the interleaved VSI has a THD of 6.58% and its bar graph representation is shown in Fig.19.

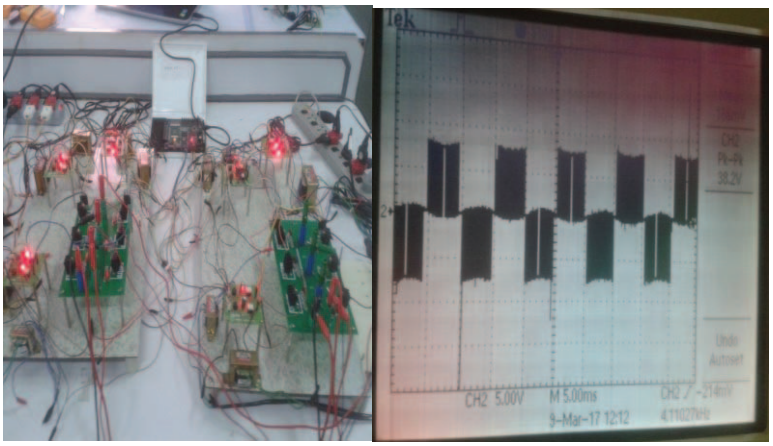


Fig.17: A prototype model of a two-stage IVSI.

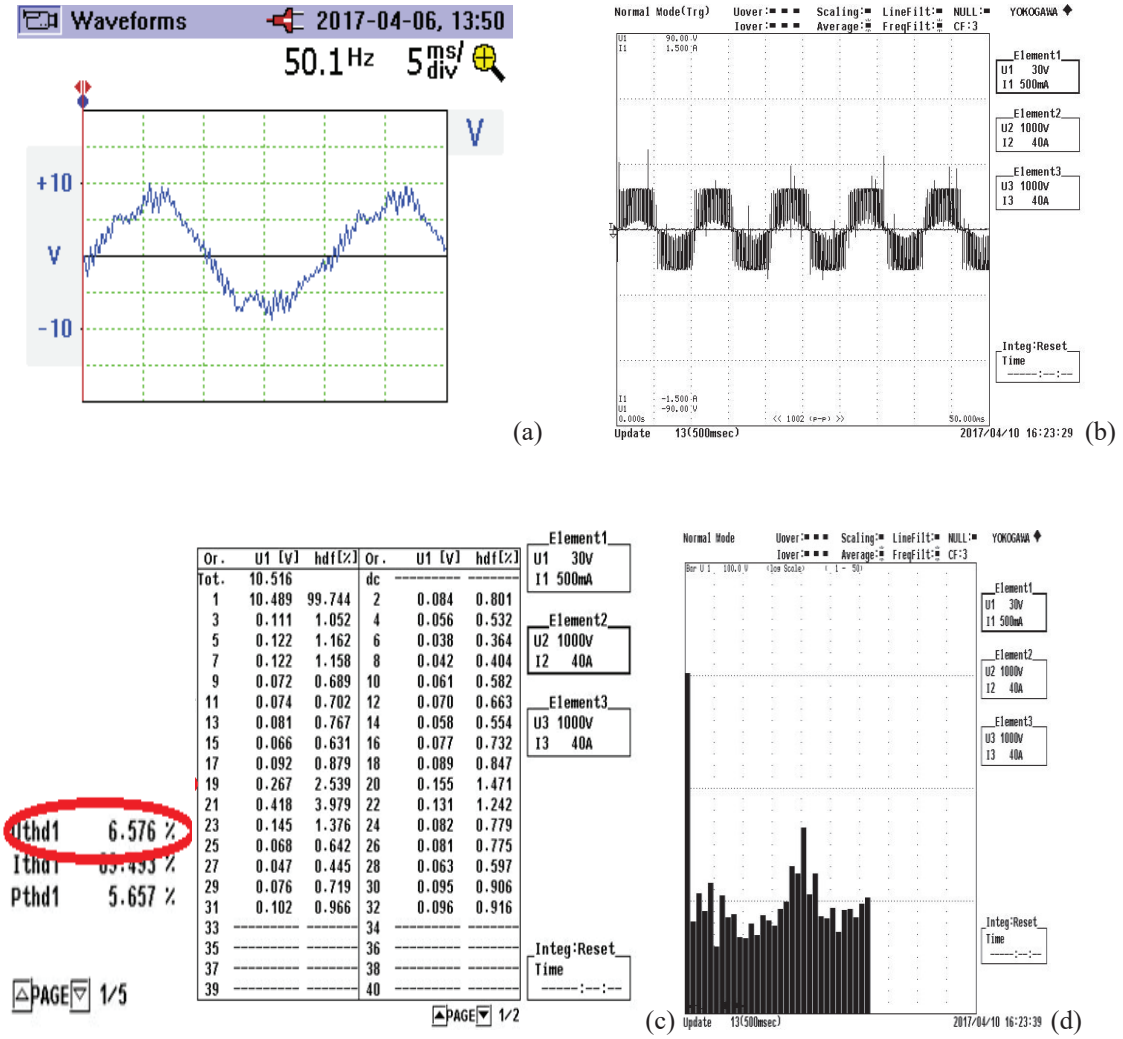



Fig.18: (a) The output voltage of proposed inverter with filter (b) Line-to-Line Voltage of Interleaved Voltage Source Inverter (c) Line-to-Line Voltage THD (d) Bar graph representation of voltage THD.

https://doi.org/10.31436/iiumej.v20i2.1142

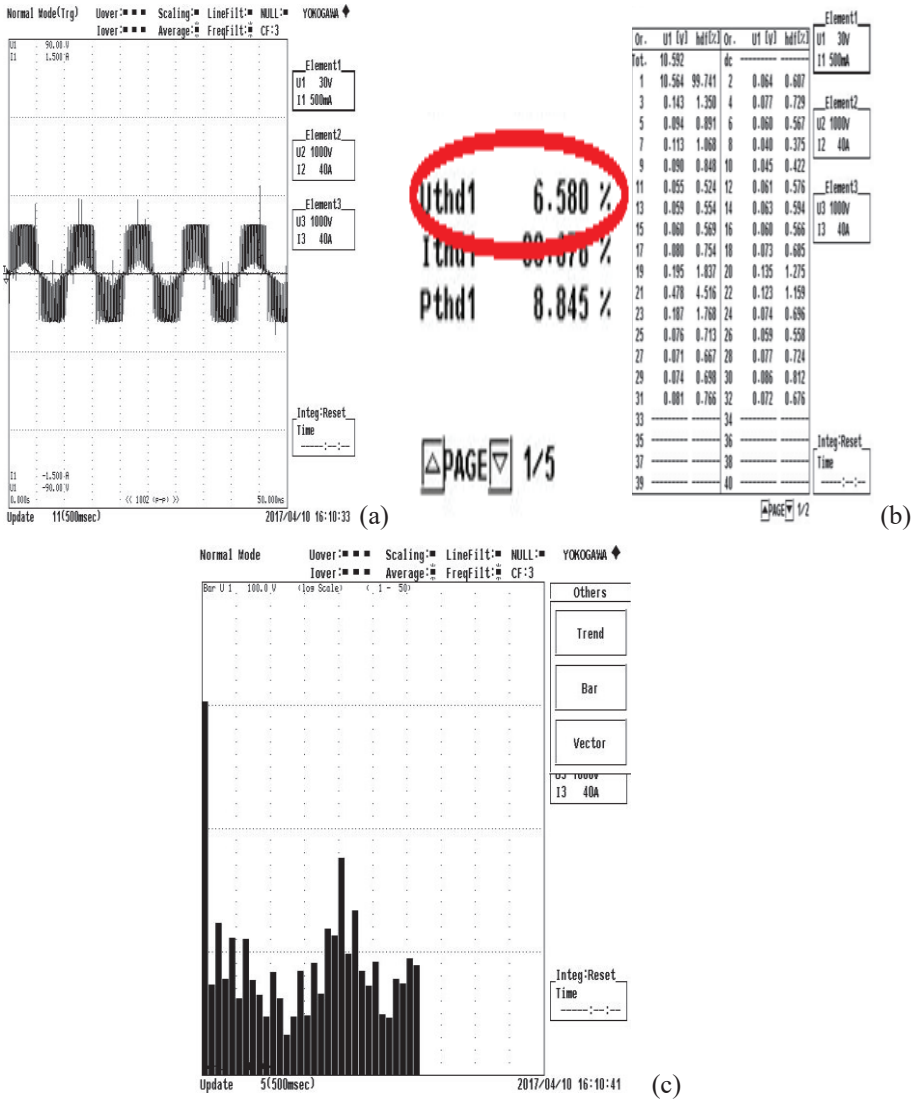



Fig.19: (a) phase Voltage of Interleaved Voltage Source Inverter (b) Phase Voltage THD (c) Bar graph representation of voltage THD.

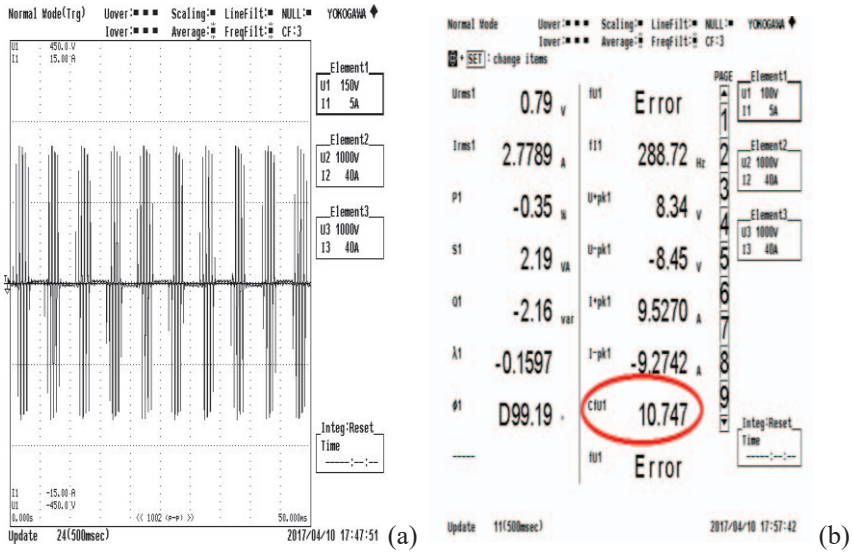



Fig.20: (a) Inductor current (b) Inductor current ripple.

From Fig.20, it is clear that the proposed interleaved VSI has a lower filter inductor current ripple of 10.747%.

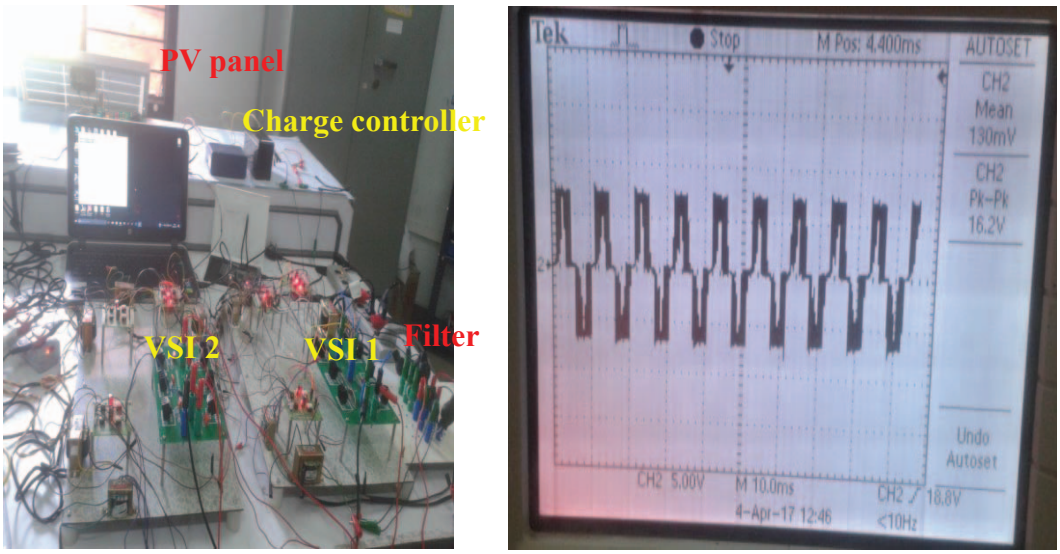



Fig.21: Two-stage VSI is interfaced with PV panel

Figures 18 - 20 prove that the proposed two-phase interleaved voltage source inverter has low THD and reduced inductor current ripple content compared to the conventional voltage source inverter. A prototype model of the interleaved inverter interfaced with PV is shown in Fig.21.The comparison of simulation results and hardware results is listed in Table 4.

Table 4: Simulation andhardware results

| PARAMETERS | SIMULATION RESULTS | HARDWARE RESULTS |
|---|---|---|
| **THD (%)** | 4.1 | 6.58 |
| **Inductor current ripple (%)** | 0.82 | 10.47 |

From Table 4, it is found that the proposed two-phase interleaved VSI has reduced THD (6.58%) and lower inductor current ripple (10.47%). Thus the simulation results are validated through a prototype model.

## 7. CONCLUSION

A two-stage interleaved VSI is studied and analyzed with different modulation strategies. The results of different modulation strategies are compared and it is concluded that the unipolar PWM has lower THD and reduced ripple content in comparison to the other techniques. Then, the two-stage VSI and traditional VSI are analyzed and it is proven that the interleaved technique had better performance. A prototype model of the two-phase interleaved voltage source inverter is built and interfaced with PV through a charge controller. The simulation results are validated experimentally. Hence, the proposed two-phase interleaved voltage source inverter topology is an appropriate choice for a PV grid connected system.

## ACKNOWLEDGMENT

Authors thank the SSN Institution for providing the financial support in carrying out this research work.

## REFERENCES

[1] Kenneth J. Sauer, Thomas Roessler, and Clifford W. Hansen. (2015) Modelling the Irradiance and Temperature Dependence of Photovoltaic Modules in PV syst. IEEE Journal of Photovoltaic, 5(1):152-158.

[2] T. Esram, P.L. Chapman. (2007) Comparison of Photovoltaic Array Maximum Power Point Tracking Techniques. IEEE Transactions on Energy Conversion, 22 (2): 439-449.

[3] L. Piegari, R. Rizzo. (2010) Adaptive perturb and observe algorithm for photovoltaic maximum power point tracking. IET Renewable Power Generation, 4 (4): 317-328.

[4] Weidong Xiao, Fonkwe Fongang Edwin, Giovanni Spagnuolo, JuriJatskevich. (2013) Efficient Approaches for Modelling and Simulating Photovoltaic Power Systems. IEEE Journal of Photovoltaics, 3 (1):500 - 508.

[5] Bader N. Alajmi, Khaled H. Ahmed, Stephen J. Finney, and Barry W. Williams. (2013) A Maximum Power Point Tracking Technique for Partially Shaded Photovoltaic Systems in Microgrids. IEEE Transactions on Industrial Electronics, 60 (4): 1596 - 1606.

[6] Francesco Adamo, Filippo Attivissimo, Attilio Di Nisio, and Maurizio Spadavecchia. (2011) Characterization and Testing of a Tool for Photovoltaic Panel Modelling. IEEE Transactions on Instrumentation and Measurement, 60 (5): 1613 - 1622.

[7] Shibin Qin, Stanton T. Cady, Alejandro D. Dominguez -Garcia, and Robert Carl Nikolai Pilawa-Podgurski. (2015) A Distributed Approach to Maximum Power Point Tracking for Photovoltaic Sub module Differential Power Processing. IEEE Transactions on Power Electronics, 30 (4): 2778-2785.

[8] Feng Yang, Xu Zhao, Cong Wang, ZhiFei Sun. (2013) Research on Parallel Interleaved Inverters with Discontinuous Space-Vector Modulation. Energy and Power Engineering, 5(4): 219-225. https://doi.org/ 10.4236/epe.2013.54B043

[9] Abusara. Mohammad A and Sharkh. Suleiman M. (2013) Design and Control of a Grid-Connected Interleaved Inverter. IEEE Transactions on Power Electronics, 28 (2), 748-764.

[10] G. Shen, D. Xu, L. Cao, and X. Zhu. (2008) An improved control Strategy for grid-connected voltage source inverters with an LCL filter. IEEE Transaction on Power Electronics,23 (4):1899–1906.

[11] S.Harika, Dr.R.Seyezhai. (2016) Analysis of Modulation Strategies for Two-Stage Interleaved Voltage Source Inverter Employed for PV Applications. International Journal of Advanced Research in Basic Engineering Sciences and Technology (IJARBEST), 2 (11):1-12

[12] Jonq-Chin Hwang, Po-Cheng Chen and Chuan-Sheng Liu. (2014) A novel Single-Phase Interleaved Bi-directional Inverter for Grid-Connection Control. IEEE 23rd International Symposium on Industrial Electronics (ISIE):375-379.
https://doi.org/ 10.1109/ISIE.2014.6864642

[13] Christian Nemec, Jörg Roth-Stielow. (2011) Ripple current minimization of an interleaved-switched multi-phase PWM inverter for three-phase machine emulation. 14th European Conference on Power Electronics and Applications (EPE):1-8.

[14] Abusara. M. A, Jamil. M, Sharkh. M.S. (2012) Repetitive Current Control of an Interleaved Grid Connected Inverter. 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG):558-563.
https://doi.org/ 10.1109/PEDG.2012.6254057

[15] Miss. Sangita R Nandurkar, Mrs. Mini Rajeev. (2012) Design and Simulation of three-phase Inverter for grid-connected Photovoltaic systems. Proceedings of Third Biennial National Conference, NCNTE:80-83

[16] Daniel Glose and Ralph Kennel. (2016) Continuous Space Vector Modulation for Symmetrical Six-Phase Drives. IEEE Transactions on Power Electronics, 31 (5):3837 - 3848.

[17] Siva Prasad. J .S and Narayanan. G. (2014) Minimization of Grid Current Distortion in Parallel-Connected Converters Through Carrier Interleaving. IEEE Transactions on Industrial Electronics, 61 (1), 76 - 91.

[18] Brenna. M, Chiumeo.R, and Gandolfi. C. (2011) Harmonic analysis: comparison between different modulation strategies for three-phase inverter connecting Distributed Generation. IEEE International conference on clean Electric power(ICCEP):231-236. https://doi.org/ 10.1109/ICCEP.2011.6036261

[19] Rixin Lai, Lei Wang, Juan Sabate. (2011) A High-Efficiency Two-Phase Interleaved Inverter for Wide Range Output Waveform Generation. IEEE Energy Conversion Congress and Exposition (ECCE):4533-4537. https://doi.org/ 10.1109/ECCE.2012.6342204

[20] Marcelo GradellaVillalva, Jonas Rafael Gazoli, and Ernesto RuppertFilho. (2009) Comprehensive Approach to Modeling and Simulation of Photovoltaic Arrays. IEEE Transactions on Power Electronics, 24 (5).

https://doi.org/10.31436/iiumej.v20i2.1000

# UNWEARABLE MULTI-MODAL GESTURES RECOGNITION SYSTEM FOR INTERACTION WITH MOBILE DEVICES IN UNEXPECTED SITUATIONS

**HANENE ELLEUCH[1]*, ALI WALI[1], ANIS SAMET[2], ADEL M. ALIMI[1]**

[1] *REsearch Groups in Intelligent Machines Lab (REGIM-Lab), University of Sfax, National Engineering School of Sfax (ENIS), BP 1173, Sfax, 3038, Tunisia*

[2]*SiFAST: Software and Computing Services Company, Sfax 3003, Tunisia*

**Corresponding author: Hanene.elleuch@ieee.org*

*(Received: 26th September 2018; Accepted: 14th August 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** In this paper, a novel real-time system to control mobile devices, in unexpected situations like driving, cooking and practicing sports, based on eyes and hand gestures is proposed. The originality of the proposed system is that it uses a real-time video streaming captured by the front-facing camera of the device. To this end, three principal modules are charged to recognize eyes gestures, hand gestures and the fusion of these motions. Four contributions are presented in this paper. First, the proposition of the fuzzy inference system in the purpose of determination of eyes gestures. Second, a new database has been collected that is used in the classification of open and closed hand gesture. Third, two descriptors have been combined to have boosted classifiers that can detect hands gestures based on Adaboost detector. Fourth, the eyes and hand gestures are erged to command the mobile devices based on the decision tree classifier. Different experiments were assessed to show that the proposed system is efficient and competitive with other existing systems by achieving a recall of 76.53%, 98 % and 99% for eyes gesture recognition, detection of fist gesture, detection of palm gesture respectively and a success rate of 88% for eyes and hands gestures correlation.

***ABSTRAK:*** Kajian ini mencadangkan satu sistem masa nyata bagi mengawal peranti mudah alih, dalam keadaan tak terjangka seperti sedang memandu, memasak dan bersukan, berdasarkan gerakan mata dan tangan. Kelainan sistem yang dicadangkan ini adalah ia menggunakan masa nyata video yang diambil daripada peranti kamera hadapan. Oleh itu, tiga modul utama ini telah ditugaskan bagi mengenal pasti isyarat mata, tangan dan gabungan kedua-dua gerakan. Empat sumbangan telah dibentangkan dalam kajian ini. Anggaran pertama bahawa isyarat gerak mata mempengaruhi sistem secara kabur. Kedua, pangkalan data baru telah dikumpulkan bagi pengelasan isyarat tangan terbuka dan tertutup. Ketiga, dua pemerihal data telah digabungkan bagi merangsangkan pengelasan yang dapat mengesan isyarat tangan berdasarkan pengesan Adaboost. Keempat, gerakan mata dan tangan telah digunakan bagi mengarah peranti mudah alih berdasarkan pengelasan carta keputusan. Eksperimen berbeza telah dijalankan bagi membuktikan bahawa sistem yang dicadang adalah berkesan dan berdaya saing dengan sistem sedia ada. Keputusan menunjukkan 76.53%, 98% dan 99% masing-masing telah dikesan pada pengesanan gerak isyarat mata, genggaman tangan dan tapak tangan, dengan kadar 88% berjaya mengesan gerak isyarat mata dan tangan.

***KEYWORDS:*** *Eye gesture recognition, Hand gesture detection, Multimodal interaction, Fuzzy inference system, Human-computer interaction*

https://doi.org/10.31436/iiumej.v20i2.1000

# 1. INTRODUCTION

Nowadays, mobile devices are considered like a ubiquitous computer. No matter where the user is, he can access any information and use it to do anything that he/she wants; capturing photos, paying an invoice, accessing emails, etc. Today, the consumption of mobile devices is changing due to the continuous evolution of the hardware as well as the software. However, the question of how providing the most suitable mode of interaction with these devices has not been really solved. All these years, the mobile human-computer interaction is based on one way: the touchscreen. In fact, the touchscreen is simple, clear and intuitive. However, the limits of this mode are clearly visible when using mobile devices in unexpected situations such as moving, driving or using one hand while the other is busy [1], etc. Besides, the minimization of on-screen targets causes occlusions. To cope with these issues, many solutions are proposed to extend mobile human-computer interaction modalities by introducing natural ways of communications like eyes, hands, and voice. Eyes and hands are the most intuitive way of communication. Eyes are our first channel of communication with the surrounding environment. They play an important role in expressing needs and expressional emotions. Also, hands play an important role in non-verbal communication with the purpose of sending messages correctly. For the ubiquitous interaction with mobile devices, it is clear that eyes and hands have many advantages compared to voice because of the background noise that can affect the quality of communication, on the one hand, and on the other hand, gestures are independent of language and do not require any previous knowledge. Several solutions are presented to extend the ways of interaction with mobile devices. Some of these solutions focus on eye tracking by using EEG [2], EOG [3, 4] and Eye Tribe [5]. Other solutions focused on hand gestures recognition are based on gloves [6, 7]. All these solutions require the addition or modification of the device, which is considered as a barrier to adoption and generalization. In contrast, most mobile devices already contain cameras. For this reason, an image-based solution can be more efficient for eye tracking as well as for the hand gesture recognition at one and the same time.

In this paper, a new mobile human-computer interaction system based on eyes and hands' gestures was proposed for a medical application used by doctors to manipulate patients' information during a surgical intervention. The input of this system is a real-time video captured from the front-facing camera of the device. Two modules are responsible for each modality's gestures. The first module aims to recognize eyes gestures from the captured video. The user's face and his eyes are detected first and then the pupils are determined and tracked. A fuzzy inference system is deployed to recognize the pupils' motions. In this work, the main goal is detecting 8 motions in different directions: right, left, up, down, right-up, left-up, right-down, left-down. In addition to these movements, this module can also detect blinks. The second module aims to recognize hand gestures from the input video. Viola and Jones' algorithm is used to detect 2 gestures: fist and palm. To achieve this goal, a new dataset is collected and used for the training phase. A couple of descriptors are combined to improve the accuracy. The recognized gestures coming from different modules are combined and fused in a separate module to execute actions to monitor mobile devices. This system is evaluated by conducting a study on volunteers in real life conditions, the results obtained are promising.

The remainder of this paper is organized as follows: in section 2, the relevant work was detailed in studying the unimodal and multi-modal interaction systems. In section 3, the proposed system is presented and each module is detailed. In section 4, an assessment is conducted for the purpose of evaluating this system. In this section also, a description of the

experiments of each module separately and then the whole system. Finally, the paper is concluded and an example of application is suggested to prove the practical aspect of the proposed system.

## 2. RELATED WORK

Many research works introduced eyes and hands as alternative modalities to control devices. In this section, the related work was cited that included eyes and hands as the only modality and their combination in the human-computer interaction.

### 2.1. Natural user interface based on unimodal interaction systems

Eye-based interaction is deployed as an important input modality. Gaze estimation is used to determine the point-of-regard on the screen of mobile devices [8,9,10,11 ,12]. Wood and Bulling [8] combines cascade classifiers and a shape-based approach for eyes centers detection, limbus ellipse fitting and limbus back projection for gaze estimation. Pino and Kavasidis [9] used Haar classifier and CAMSHIFT to detect and track eyes respectively. The authors estimated the gaze by determining a correlation between the centroid of eyes rectangle and the user's gaze. [10] proposed a dataset for unconstrained gaze estimation for mobile devices. These works proved their robustness by achieving competitive results, however, they suffered from many limits like the execution time and the need of learning phase and calibration preprocessing. Eyes gestures are also used to interact with mobile devices. [13] proposed a new eyes gestures recognition system for mobile devices. The algorithm is based on image processing algorithms to detect and track the user's eyes from video captured from the device. A face detector based on Viola and Jones algorithm is responsible for detecting the face. Another module tracked the detected face. An eye detector detected the eyes from the face which have been tracked using an eye tracker. The eyes gesture recognition is assured by using template-matching. Other methods are also deployed. Miluzzo et *al.* [11] used contour detection to localize the eyes in their system presented EyePhone that combines blinks detection and gaze's estimation to command mobile devices. In [14], the authors proposed an eye gesture input interaction system for smartphones dedicated for people with ALS. This system recognizes gestures by using template-matching but a calibration step is necessary to use this system for the first time. In [15], the authors proposed a system of eyes gestures recognition. This system recognizes 8 gestures to communicate with mobile devices without any additional peripheries.

Hands gestures are widely used in the human-computer interface. Meng et *al.* [16] presented a new system of hand gesture recognition. The hand detection is assured by using skin color detection algorithm and static approach image to separate the hand from the background. The hand gesture recognition is assured by determining the fingers' number and the angle formed between the fingers. This step is based on hand contour detection and fingertips and palm center extraction. The authors proposed the deployment of cloud computing to solve the problem of the slow processing time. Prasuhn et *al.* [17] proposed a HOG-based hand gesture recognition system deployed on the mobile device and applied it to American Sign Language (ASL). A hand detection area based on the skin color algorithm is first applied. The hand shape features are computed by using the HOG descriptor [18]. The recognition step is assured by finding the best matching between the target image and images from the database. Song et *al.* [19] presented a new system of interaction with mobile devices based on hands gestures. The detection step is based on skin color technique and the recognition phase is assured by the random forest algorithm. The experiments showed promising results. In [20], the author proposed a new system of interaction with mobile devices based on static hands gestures. This system is based on skin color for hand

segmentation and SVM for the recognition phase. The hand detection in all these works is based on skin color detection. This technique was proved as a robust method; however, it causes many false positive detections and occlusions with the face region. The response time for [16] and [17] is relatively high. For this reason, [16] deploys a cloud server for the computation part to have more rapid execution time. [16] and [17] are the only systems developed for mobile devices. However, these systems did not treat the case of the face occlusion and the dynamic hand gesture recognition. [16] did not present the recognition results so the robustness of his method cannot be determined. On the other hand, [17] proved that the recognition rate of their system is only 47%. Besides, this system's algorithm is not totally developed on the device, but it is based on a client-server configuration that made the need of the internet indispensable for the operation process. In this paper, a real-time hand gesture recognition based on the camera was proposed. It is running entirely in the device and it can detect hands gesture without any confusion with user's face and any other body part.

### 2.2. Natural user interface based on multi-modal interaction systems

Many systems are proposed with different combinations of modalities. Gaze and hand gesture were the most popular. Many research works focused on the combination of these modalities in user-computer interaction.

Chatterjee et *al.* [21] used these modalities for the purpose of building an interactive mode with a computer, a series of LEDs and a mobile robot. Each modality is captured by a specific periphery: Eye Tribe Tracker for gaze tracking and Leap motion for recognizing hand gestures. The authors conducted a study to test the efficiency of these modalities applied differently in various scenarios. Hales et *al.* [22] presented a system of object manipulation through gaze and hand gestures. A specific glass assured gaze tracking and a vision-based module assured the recognition of hand gestures. The assessment of this system proved that the combination of these modalities is natural and promising. Pouke et *al.* [23] presented a model of interaction based on gaze and hand gesture for 3D virtual space on tablet devices. The gaze tracking is assured based on corneal reflexion technique and the hand gesture recognition is based on the accelerometer sensor and machine learning algorithms. The author proposed to use the gaze for selecting objects and hand gesture for the manipulation of these objects. Yoo et *al.* [24] evoked the issue of interaction with a large screen cannot be efficient with the traditional tools of interaction like mouse and keyboard cannot be efficient. To this end, the authors presented a new system of human-computer interaction based on gaze and hand gestures.

These works proved the efficacy of multi-modal interaction in comparison with gaze only or hand only gestures. However, in all these works the modalities are treated separately and there is no interaction between them. Besides, these modalities are captured by specific sensors that require adding specific peripheries to the devices.

In this work, eyes and hands gestures modalities are proposed without any additional peripheries. The capture of these modalities is assured by the camera already provided in the mobile device.

## 3. ARCHITECTURE OF THE PROPOSED SYSTEM: MULTIMOB

The proposed system aims to interact with mobile devices through eyes and hands gestures by providing the possibility to do actions like click, swipe etc. through natural modalities instead of touch way. To this end, the input of this system is a real-time video

https://doi.org/10.31436/iiumej.v20i2.1000

captured from the camera of the devices. There are not any additional peripheries. This system is composed of three main modules. The system overview is presented in Fig.1. The first module is responsible for detecting eyes' regions and then capturing eyes positions for each frame and recognizing gestures. The second module is responsible for the detection and recognition of static hands gestures and uses it to the recognition of dynamic hands gestures. Static hand gestures and eyes gestures resulting from each of these two sub-systems are used in the third module and combined in several ways with the purpose of building a new vocabulary to launch final actions that control the device such as click, rotate, wipe etc. In this system, the possibility to use one or more modality is given, the multimodal gestures fusion system is active only when the user chooses to not use eyes gestures and hands gestures only. In the other case, the user can execute specific actions with the corresponding modality.

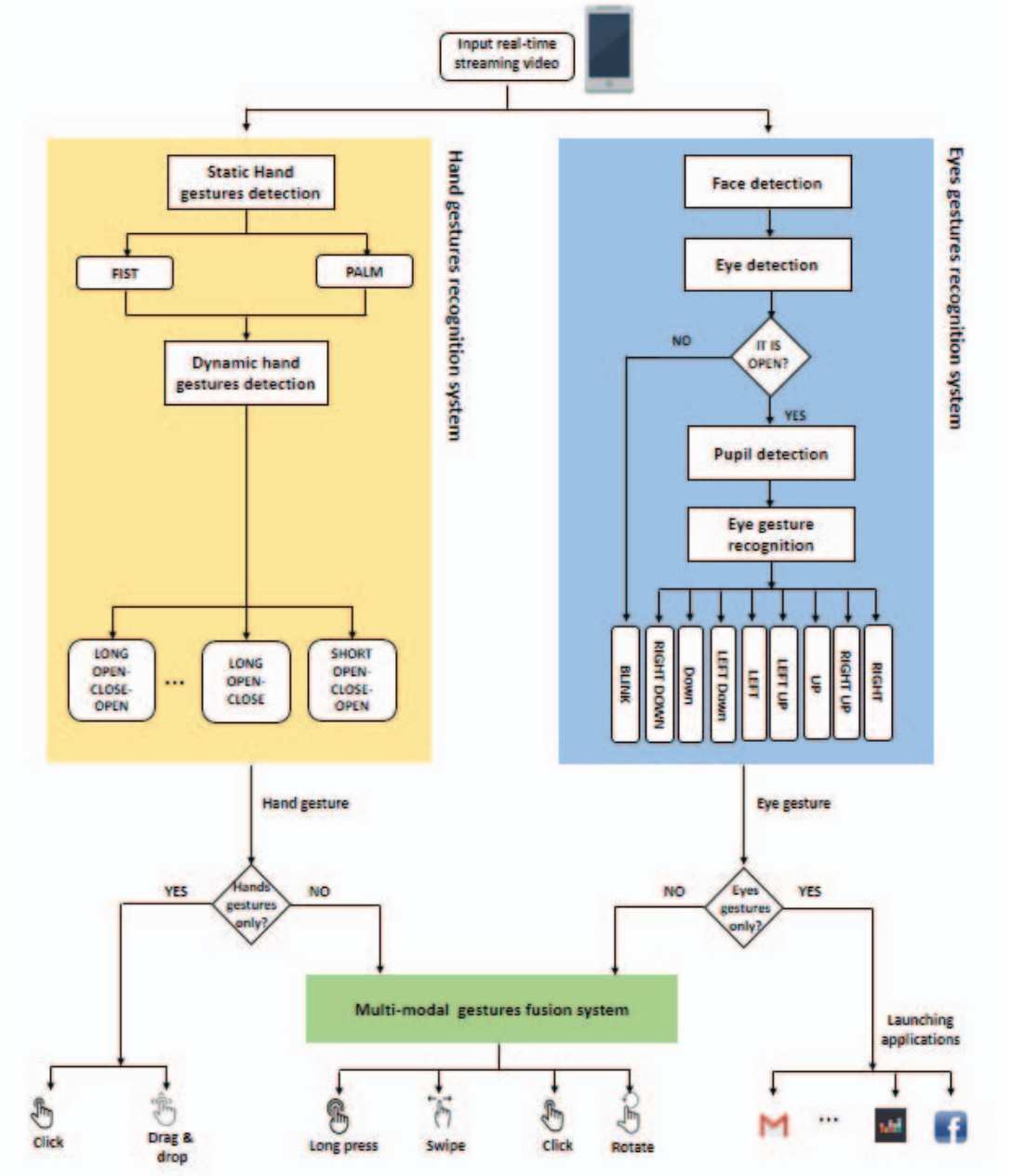



Fig 1 System overview

https://doi.org/10.31436/iiumej.v20i2.1000

### 3.1. Eyes gesture recognition for mobile devices

This subsystem is composed of successive steps.

- Face and eyes detectors based on Viola and Jones algorithm [25] are launched for each given frame.
- The eye detector indicates if the eyes are open or closed. If the eyes are detected as open, the next sub-module is launched.
- The pupil's detection is based on the reduction of the region of interest (ROI) in the eye and the determination of the darkest zone in the eyes [26].
- Once the pupils are detected, the eye gestures recognition sub-module will determine and extract the positions of pupils. These positions in the next step will be provided as inputs for a fuzzy inference system to recognize the corresponding motion.

#### *3.1.1. Face and eyes detection*

In the beginning, this module detects the user's face. This detection is assured by the Viola and Jones algorithm. This classifier, known by its ability to be running in real-time conditions, is provided by the OpenCV library [27] that achieves a true detection rate of 98.5% [28]. The face region is extracted and used as a location for eyes detection. The classification is based on boosted cascade Haar-like features. In fact, the determination of eyes regions includes also a classification about the state of the eyes to determine if they are open or closed.

Eyes blink is an involuntary reflex. This motion is intuitive and very rapid. The duration of the blink is on average between 100 ms and 400 ms [4]. For this reason, to differentiate between the spontaneous and the intentional blinks, a threshold is fixed, fixed at 500 milliseconds, of eye's closing duration. Any blink that lasts for a time longer than this threshold is considered as an intentional blink.

#### *3.1.2. Eyes' gestures recognition*

Once the eyes' region is detected, a reduction of this zone is required. In fact, the result eyes zone contains several parts of the eyes like the eyelid. This part is not needed to detect the pupils; for this reason, the first step in this sub-module is keeping only the small part of the iris and its surrounding. In this ROI, a pupil detector is charged to localize the pupils. In fact, the pupils are characterized by their common features. The pupils are the darkest zone in the eyes. So, to find the pupil's zone, a search of the darkest pixel is applied, once the pupil is localized a tracking task is launched by the template matching [26, 15].

Eye gesture recognition is defined in this work as the motion of pupils in a specific direction. In this module, the target gestures to be detected are the movements of the eyes to the lowing directions: right, right up, up, up left, left, left down, down and down right direction. These gestures have been chosen because they are simple and do not tire the eyes. However, the determination of each gesture cannot be precise. There is an overlap between 2 consecutive gestures for instance between left and left down like Fig.2 shows. To solve this issue, a fuzzy-based approach for the classification of eyes' motions based on pupils tracking was proposed.

https://doi.org/10.31436/iiumej.v20i2.1000

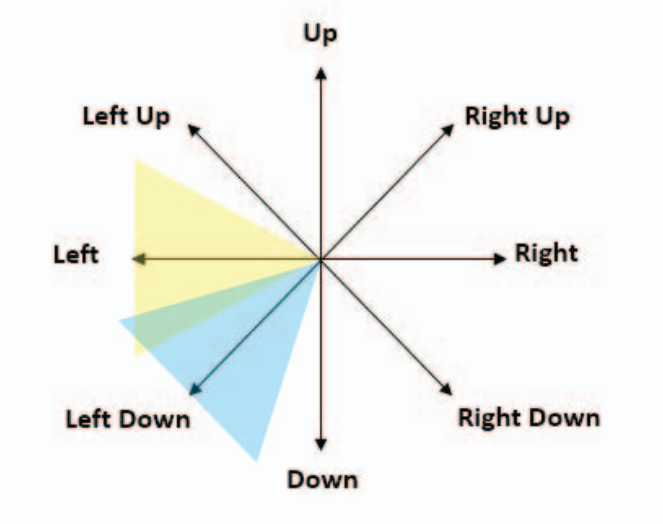



Fig 2 Eyes gestures

The recognition of the eye's gestures used basically the pupils' position extracted from each frame. The variation of these positions in successive frames is exploited in the motion's classification. This variation is computed as shown in the following equations:

$$dx = P2(x) - P1(x) \quad (1)$$

$$dy = P2(y) - P1(y) \quad (2)$$

Where $d_x$ and $d_y$ are the variations between 2 pupils' positions $P_1$ and $P_2$ at the $x$ and $y$-axis respectively computed in pixels. In fact, the eyes are naturally in continuous variation. So, any movement that is lower than a specific threshold $Val_{Threshold}$ cannot be taken in consideration of. The sign of the value of $d_x$ and $d_y$ can be precise in the next step presented by the pseudo-code in Fig. 3

A fuzzy inference system (FIS) based on a set of rules to assure the phase of eyes gestures recognition as presented in Fig.4. These rules are defined in a natural way according to the directions of $d_x$ and $d_y$, for example, if there is a variation on the $x$-axis without any variation on the $y$-axis, the final direction is right.

The FIS used in this approach is a the Mamdani type because there are many outputs so, it is the most intuitive and suitable in this case.

The fuzzification of the inputs and the outputs variables is presented in Fig.5, Fig.6, and Fig.7.

```
if (d_x > Val_Thresold) then
|  d_x ← POSITIVE
else
|  if (d_x < −Val_Thresold) then
|  |  d_x ← NEGATIVE
|  else
|  |  d_x ← ZERO
|  end
end
```

Fig 3 Pseudo-code of defining the sign of $d_x$ and $d_y$

https://doi.org/10.31436/iiumej.v20i2.1000

- Rule 1: if $d_x$ is POSITIVE and $d_y$ is ZERO then Gesture is RIGHT
- Rule 2: if $d_x$ is POSITIVE and $d_y$ is POSITIVE then Gesture is RIGHT UP
- Rule 3: if $d_x$ is ZERO and $d_y$ is POSITIVE then Gesture is UP
- Rule 4: if $d_x$ is NEGATIVE and $d_y$ is POSITIVE then Gesture is LEFT UP
- Rule 5: if $d_x$ is NEGATIVE and $d_y$ is ZERO then Gesture is LEFT
- Rule 6: if $d_x$ is NEGATIVE and $d_y$ is NEGATIVE then Gesture is LEFT DOWN
- Rule 7: if $d_x$ is ZERO and $d_y$ is NEGATIVE then Gesture is DOWN
- Rule 8: if $d_x$ is POSITIVE and $d_y$ is NEGATIVE then Gesture is RIGHT DOWN
- Rule 9: if $d_x$ is ZERO and $d_y$ is ZERO then Gesture is NOGESTURE

Fig 4 Fuzzy rules of eyes gesture recognition

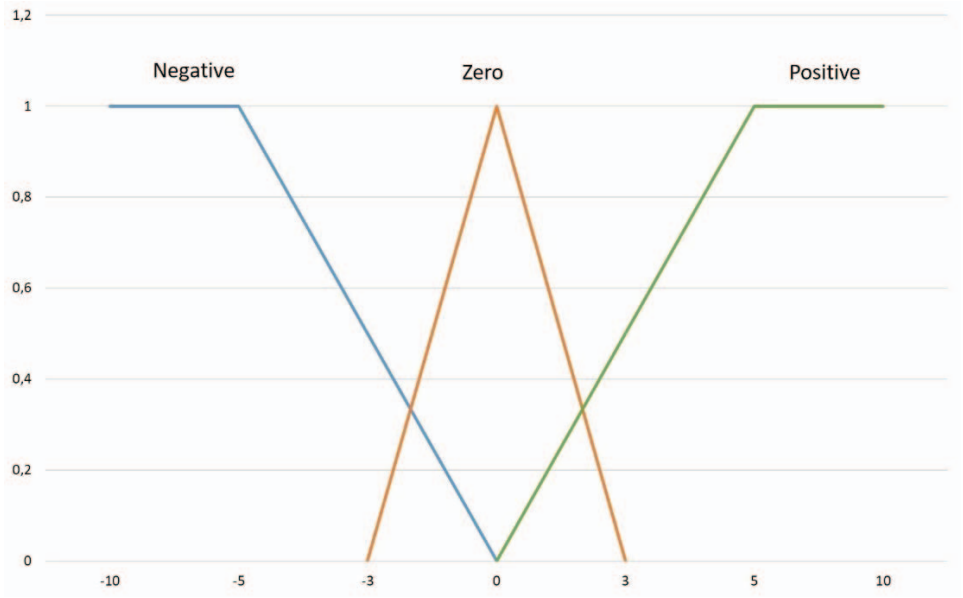



Fig 5 Fuzzification of the input variable $d_x$

### 2.1. Hands gesture recognition for mobile devices

The second system aims to recognize hands gestures. A system of hand gesture recognition is developed. It is based on the 2 types of gestures: the static and the dynamic ones. The definition of the static hand gesture is the stationary posture of a hand formed by fingers and the dynamic hand gesture as the succession of consecutive hand positions.

https://doi.org/10.31436/iiumej.v20i2.1000

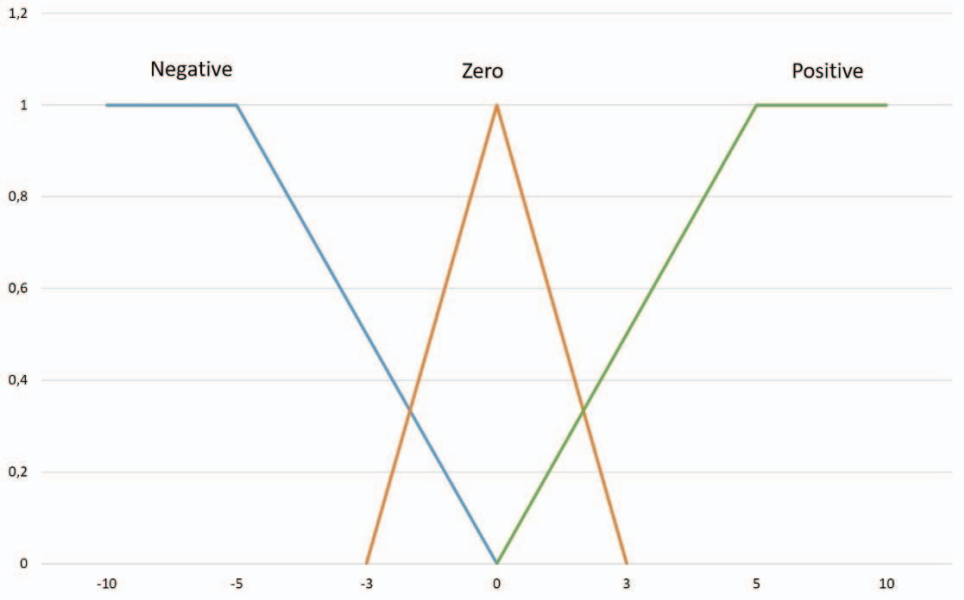



Fig 6 Fuzzification of the input variable $d_y$

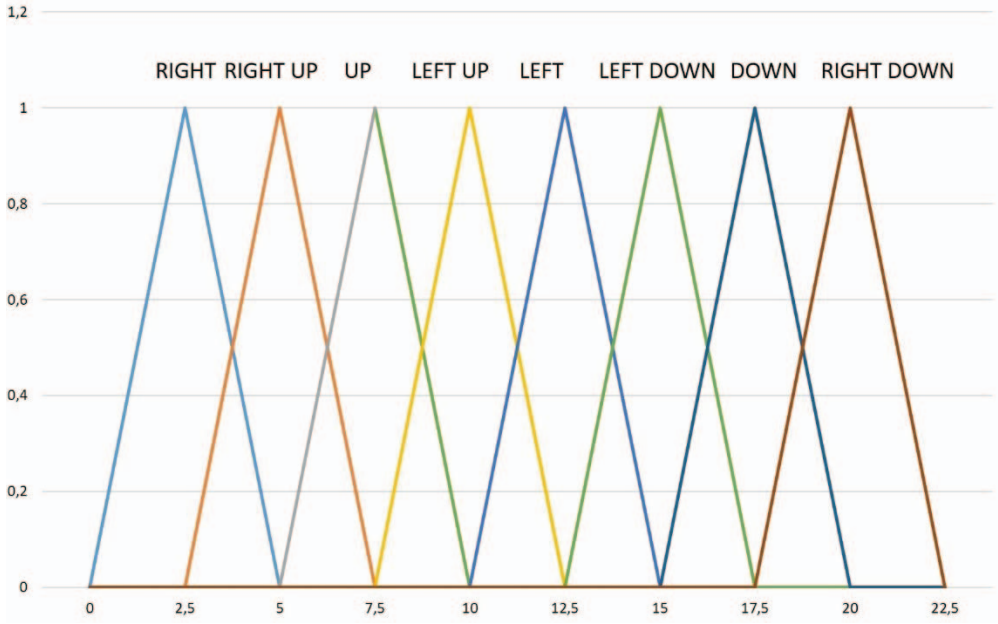



Fig 7 Fuzzification of the output variable Gesture

### *2.1.1. Static hand gesture recognition*

To recognize hand gestures, independent classifiers were created for each gesture based on Viola and Jones algorithm [25]. These detectors are based on images converted in the grayscale. A cascade of strong classifiers is obtained by threshold on scalar features [28] using the Adaboost algorithm. For this system, two hand gestures were chosen: fist (closed hand) and palm (opened hand). To ensure having the highest success recognition rate, 3 classifiers were created for each gesture based on Haar-like, LBP, and HOG descriptors.
*Dataset:* For the training, 7716 and 7611 positive images for the fist and the palm, respectively and 12933 negative images collected from our private dataset [29] and other

https://doi.org/10.31436/iiumej.v20i2.1000

different public databases [30,31] were used. For the collection of this dataset, 21 volunteers have participated. The images collected are in indoor conditions with different backgrounds and hands positions. A low resolution of 72 dpi is used. The parameters of each window are described in the following table:

Table 1 Parameters of hand gestures detectors

| | Windows | Stages |
|---|---|---|
| Fist detector | [32,33] | 5 |
| Palm detector | [42,32] | 5 |

To have the highest success rate, a test of our classifiers is conducted according to three descriptors on 200 images for the closed hand and 200 images for the opened hand. The recall, the precision and the false positive per image (FPPI) were computed. A reminder of the metrics 'equations related to the recall, the precision, and the FPPI is presented in the following equations:

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

$$Precision = \frac{TP}{TP+FP} \tag{4}$$

$$FPPI = \frac{FP}{TP+FN} \tag{5}$$

Where TP, FP, and FN are respectively the value of the true positive, false positive, and false negative of the detection rate. To determine these values, a bounding-box-based representation of the ground truth and the detected hand gestures are created. The PASCAL criterion [32] is applied. It is based on the computation of the bounding bow overlap ($BO_i$) of the detected bounding box ($D_i$)and the ground truth ($G_i$) as in this equation:

$$BO_i = \frac{|D_i \cup G_i|}{|D_i \cap G_i|} \tag{6}$$

A TN detection is obtained if the value of ($BO_i$) exceeds the value of the threshold. An FP detection is determined when a bounding box is detected and the value of ($BO_i$) is less than the threshold. An FN is obtained when there is no bounding box detected although the hand gesture is visible. The test results for the fist and the palm detection are shown in tables 2 and 3, respectively.

To improve our obtained results, the used descriptors are combined with the purpose of increasing the true detection rate and decreasing the false detection.
So, the common region of each bounding box of the two descriptors is computed by using Eq. 6. If the result is more than 0.5, the number of TP is increased and the result is the union of the two bounding boxes.

https://doi.org/10.31436/iiumej.v20i2.1000

Table 2 Recall, precision and FPPI values of the fist detection

| | Recall (%) | Precision (%) | FPPI |
|---|---|---|---|
| HOG | 98 | 62.61 | 0.585 |
| HAAR | 100 | 26.49 | 2.775 |
| LBP | 100 | 38.24 | 1.615 |

Table 3 Recall, precision and FPPI values of the palm detector

| | Recall (%) | Precision (%) | FPPI |
|---|---|---|---|
| HOG | 90 | 50.04 | 0.885 |
| HAAR | 98.5 | 18.97 | 4.2 |
| LBP | 100 | 35.14 | 1.845 |

To have the best result, all possibilities of combinations are tested and the results are presented in tables 4 and 5 for the fist and the palm gesture, respectively.

Table 4 Recall, precision and FPPI of the fist detector for each combination of descriptors

| | Recall (%) | Precision (%) | FPPI |
|---|---|---|---|
| HOG+HAAR | 98.5 | 98.5 | 0.015 |
| HOG+LBP | 97.5 | 97.5 | 0.025 |
| HAAR+LBP | 94 | 94 | 0.09 |

Table 5 Recall, precision and FPPI of the palm detector for each combination of descriptors

| | Recall (%) | Precision (%) | FPPI |
|---|---|---|---|
| HOG+HAAR | 91 | 93.84 | 0.06 |
| HOG+LBP | 99 | 96.11 | 0.04 |
| HAAR+LBP | 79.5 | 85.94 | 0.145 |

https://doi.org/10.31436/iiumej.v20i2.1000


(a) Fist detection results on some examples from our dataset


(b) Fist detection results on some examples from NUS dataset

Fig. 8 Fist detection results

These tables show that the high result is obtained by the combination of HOG and HAAR for the fist detection and HOG and LBP for the palm detection. To check the robustness, these classifiers are tested on a public database NUS [33]. 200 images from this database have been tested for each open and closed hand. The recall, precision and FPPI found are respectively: 95%, 93.59% and 0.065 for the fist and 96%, 97.46% and 0.025 for the palm. These results are promising mainly in comparison with the work of [34] that achieves 94,36% of success rate. Fig. 9 and Fig. 8 presented the detection result for each palm and

https://doi.org/10.31436/iiumej.v20i2.1000

fist gesture for both our database and NUS database. In fact, the detection is successful in several cases as Fig. 8 and Fig. 9 show; there are simple and complex backgrounds and there are pictures captured in outdoor and indoor conditions. This static hand gesture recognition is the same regardless of the color skin, the gender and the size of the hand.

(a) Palm detection results on some examples from our dataset


(b) Palm detection results on some examples from NUS dataset

Fig. 9 Palm detection results

https://doi.org/10.31436/iiumej.v20i2.1000

#### *2.1.2. Dynamic hand gesture recognition*

The second module developed in this system is the dynamic hand gesture recognition. In fact, a dynamic hand gesture is a sequence of static hand gestures.
As the dynamic hand gesture is a natural way of communication, there is a need to differentiate between the intentional one and the gesture used to execute a command. For this reason, vocabulary is built to define each gesture with the commands. The dynamic hand gesture is composed by a sequence of static hand postures that determine the beginning and the end of one gesture.
A tree of vocabulary is presented in Fig.10 to recognize three gestures that each one is responsible for executing a command: a click, a drag, and a drop. All motions begin with the initial gesture which is the opened hand and are finished with the same gesture.

- Click: the user must open his hand, close it and open it again in a duration that does not exceed the minimal time $t_{min}$.
- Drag: the user should open and close his hand for a duration more than the minimal time $t_{min}$ then a tracking module is responsible for following the hand's position.
- Drop: is executed once the user opens his hand after the drag action.

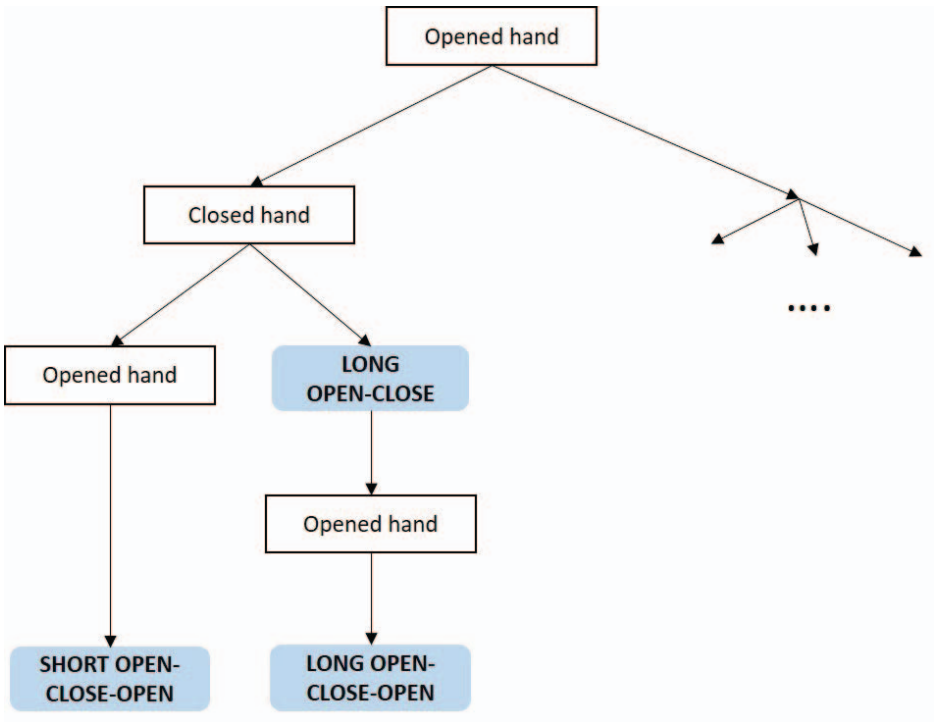



Fig. 10 Dynamic hand gestures vocabulary

### 2.2. Multi-modal fusion based gestures for interaction with mobile devices

The hands and eyes are in related movements. These couple modalities have an important role in the enhancement of the efficiency of controlling mobile devices. The coupled modalities provide for us additional information that can serve in the separation between the intentional and natural motions [35].

https://doi.org/10.31436/iiumej.v20i2.1000

To this end, the fusion of gestures recognized by the previous modules specific for eyes and hands gestures is proposed.

In fact, the fusion domain is divided into three levels: data, feature and decision levels [36].

- Data-level fusion is applied in the phase of receiving information captured from all available inputs. It aims to conserve all this information. However, in this type of fusion, the pre-processing step is absent and so the generated noise cannot be eliminated [36].
- Feature-level fusion combines the features generated from each input modality. Basically, this fusion's category requires that the generated features' vectors have the same representation. This condition cannot be available in many cases particularly when the input modalities have different natures.
- Decision-level fusion is the most used thanks to its capability to operate with different types of modalities. The final decision is taken based on the local decision generated from each modality. In comparison with other fusion levels, decision-level fusion has not any constraint of invariance input or vectors' representation because each module analyses its related modality separately.

The comparison of these fusion-levels leads us to use the decision-level fusion. This is the most convenient in our case. In fact, the eyes and hands gesture recognition modules are captured from the same camera. However, each module analyzed the captured video and generated a decision regarding each modality. The variety of the decisions classified by each module, the ability to fuse them and the respect of the real-time conditions are the important constraints that specify the tool used for the decision fusion step. The decision tree is one of the most known techniques in the supervised classification. It is characterized by its simplicity in analyzing and interpreting thanks to its flexibility in the determination of the input variables and the rules used for the classification task.

The fusion module takes 3 attributes as inputs, one from eyes gestures and two from hand gestures with the purpose of executing actions to command mobile devices. It is noteworthy that the hands gesture recognition module has 2 classes: palm and fist and the eyes gesture recognition module has 9 classes*: right, right up, up, left up, left, left down, down, right down* and *blink*. For the fusion module, a choice is made to consider that *the right*, *right up* and *right down* as *right* direction and *left*, *left up* and *left down* as *left* direction. As final decisions, 6 classes are defined: *long press, rotate*, *click*, *zoom*, *swipe in the right* and *swipe in the left*. Table 6 details how these decisions are organized.

Table 6 Training table of eye and hand gestures fusion module

| Eye gesture | Hand gesture 1 | Hand gesture 2 | Decision |
|---|---|---|---|
| Blink | Close | Close | Long press |
| Right | Close | Close | Rotate |
| Left | Close | Close | Rotate |
| Up | Close | Close | Rotate |
| Down | Close | Close | Rotate |
| Blink | Open | Open | Zoom |
| Right | Open | Open | Swipe in Right |
| Left | Open | Open | Swipe in Right |
| Up | Open | Open | Swipe in Left |
| Down | Open | Open | Swipe in Left |
| Right | Open | Close | Click |
| Left | Open | Close | Click |

| | | | |
|---|---|---|---|
| Up | Open | Close | Click |
| Down | Open | Close | Click |
| Right | Close | Open | Click |
| Left | Close | Open | Click |
| Up | Close | Open | Click |
| Down | Close | Open | Click |

## 3. Experimental results and discussion

In this part, the twofold system is evaluated with the purpose of validating its robustness. To this end, we used an Android-based tablet for the deployment of our system. This device has an android version of 4.2.1 running on NVIDIA Tegra 3 Quad-Core and a 2 Go of RAM. The front-facing camera has a 2 Megapixels resolution.

### *3.1. Eyes gesture recognition for mobile devices*

In this module, these libraries are used: OpenCV [27] image processing and Fuzzy Lite libraries [37] for the classification. 8 participants aged between 25 and 31 years, who are unfamiliar with eyes gestures recognition systems, volunteered to test our system. First, each participant is asked to sit in front of the tablet and save a distance between 25 and 40 cm to keep the whole face captured by the camera. Then, they are asked to close their eyes and move it in the 8 directions. To validate this module, different metrics are computed: the sensitivity, the false positive rate (FPR) defined by these equations:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Sensitivity = \frac{TP}{TP+FN} \quad (2)$$

$$FPR = \frac{FP}{FP+TN} \quad (3)$$

Where the value of TP, FN, FP, TN are computed as in this way:

- There is a TP when a correct eyes gesture is recognized.
- There is a FN when none or a false eyes gesture is recognized.
- There is a FP when an eyes movement is detected but there is no gesture.
- There is a TN when there is no gesture is detected and the user keeps his eyes in a stationary position.

The result of this assessment and a comparison with the most related work are presented in table 7.

Table 7 The result of eyes gesture recognition module

| | Accuracy [%] | Sensitivity [%] | FPR [%] |
|---|---|---|---|
| Our system | 76.53 | 85 | 27 |
| Vaitukaitis et *al.* [13] | 60.0 | 28.3 | 17.6 |

This assessment showed promising results achieved by our system. In comparison with related work, the proposed system attained 76.53% of accuracy while [13] achieved 60%.

https://doi.org/10.31436/iiumej.v20i2.1000

Another test is conducted to investigate the material consumption of this module. The measurement of the physical resources consumption of RAM and CPU is computed. The rapidity of our developed algorithm can be induced in the number of frames per second (FPS). The results are presented in table 8.

Table 8 Computation of RAM and CPU consumption

| | RAM | CPU | FPS |
|---|---|---|---|
| Our system | 90 | 40 | 24 |
| Miluzzo et *al.*[11] | 56.51 | 65.4 | - |

As shown in table 8, our subsystem consumes 90% and 40% of RAM and CPU, respectively. These numbers showed the limit of this module in terms of RAM consumption. It is noteworthy that this module consumes RAM much more than the system proposed by [11], in the contrast to CPU. Besides, this module proved the fact that it can be running in real-time condition by achieving 24 FPS. It is so far advantageous when compared with similar systems [13] and [38] which had lower than 15 FPS.

### 3.2. Hands gesture recognition for mobile devices

For this module, a study is conducted on 10 volunteers that they never use any touch-less application based on eyes and hands gestures. This test is repeated in two scenarios depending on the external environment and the user's state whether he is sitting or walking:

- Scenario 1: the user is in laboratory conditions and takes the tablet when he is sitting.
- Scenario 2: the user is in outdoor conditions and takes the tablet when he is walking.

In this evaluation, each participant is asked to move one hand with the purpose of executing actions corresponding to the vocabulary presented in the subsection 3.2.2. The initial hand position is palm gesture. The results of this experiment are shown in table 9.

Table 9 Experiment result for dynamic hand gesture recognition

| | Scenario 1 | Scenario 2 |
|---|---|---|
| Click | 100% | 90% |
| Drag+Drop | 100% | 100% |

The experiment results prove that the recognition of the dynamic gesture reaches 100% in scenario 1 and 90% in scenario 2 for both motions. In fact, these results depend on the performance to the recognition of the static and dynamic gestures. The error coming from the static gesture causes automatically an error in the dynamic gesture.

https://doi.org/10.31436/iiumej.v20i2.1000

### 3.4. Multi-modal fusion based gestures for interaction with mobile devices

To validate our approach in this module, a study is conducted on 5 volunteers aged between 24 and 30 years. Each participant is asked to operate the tablet to execute the proposed actions. Fig.11 shows that our module can achieve promising results. The success rate attains 100% for Long press and Zoom actions and 80% for others. Indeed, the results of fusion's gestures are due mainly to the result of eyes and hands gestures modules. Note that every error of the eye's gestures recognition or in the hand gestures recognition modules causes a decision error.

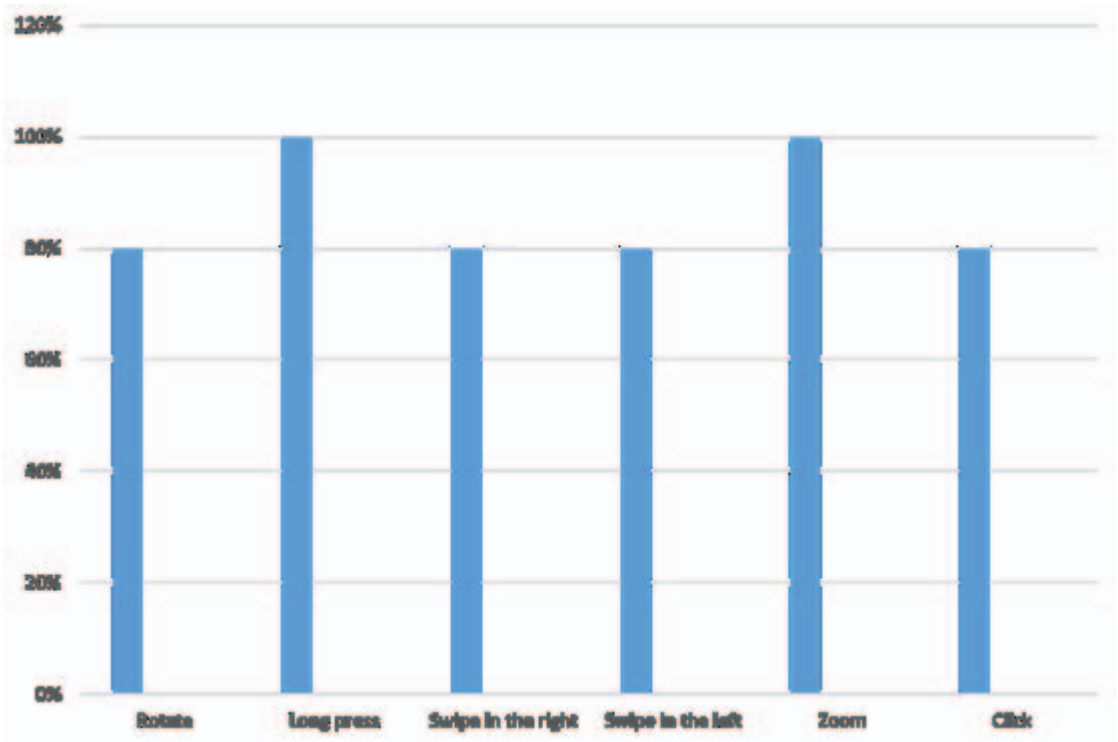



Fig. 11 Experiment results of multi-modal fusion based gestures subsystem

## 4. CONCLUSION

In this paper, a new system for mobile HCI based on eyes and hands gesture recognition is presented. This system uses the eyes' gestures to execute commands based on detection of the eyes and gaze direction. For the hand gesture recognition, two independent classifiers are created and dedicated to fist and palm gesture. A combination of different descriptors is used to have the highest success rate and the minimum false detection. From the static gestures, dynamic gestures are created. They are based on precise vocabulary to execute commands.

To evaluate our system, a study is conducted on 10 participants in two different scenarios for each eye and hands gesture. The experimental assessment showed that this system can reach promising results.

To study the feasibility of the proposed system, an application for medical uses is developed. This application provides the doctor with the possibility to consult and manipulate a patient's medical images and information that he can need during the surgical intervention.

Actually, the surgeon cannot touch anything that is not sterilized during the surgery; also his hands are busy all the time. As a consequence, he is obliged to use touchless ways to interact with his device.

First, the doctor chooses the modality that he/she wants to use to control his/her devices. Then, he/she can swipe the medical images and the patient's information by using their eyes with a blink and an eye movement.

The hand gestures can serve to move a cursor to provide the manipulative actions in the application like clicking on buttons etc.

## ACKNOWLEDGMENT

The research leading to these results has received funding from the Ministry of Higher Education and Scientific Research of Tunisia under the grant agreement number LR11ES48. The research and innovation are performed in the framework of a thesis MOBIDOC financed by the EU under the program PASRI.

## REFERENCES

[1] Nagamatsu T, Yamamoto M, Sato H. (2010). Mobigaze: Development of a gaze interface for handheld mobile devices. In Proceedings of Extended Abstracts on Human Factors in Computing Systems, pp: 3349-3354.

[2] Samadi M R H, Cooke N. (2014). EEG signal processing for eye tracking. In Proceedings 22nd European Signal Processing Conference. pp. 2030-2034

[3] Dhuliawala M, Lee J, Shimizu J, Bulling A, Kunze K, Starner T, Woo W. (2016). Smooth eye movement interaction using EOG glasses. In Proceedings of the 18th ACM International Conference on Multimodal Interaction, 307-311. doi: 10.1145/2993148.2993181.

[4] Bulling A, Ward JA, Gellersen H, Troster G. (2011). Eye movement analysis for activity recognition using electrooculography, IEEE transactions on pattern analysis and machine intelligence, 33(4):741-753

[5] TheEyeTribe [http://www.theeyetribe.com/]

[6] Chen P, Wang P, Wang J, Yao Y. (2017). Design and motion tracking of a strip glove based on machine vision. Neurocomputing. doi: https://doi.org/10.1016/j.neucom.2017.03.098.

[7] Dipietro L, Sabatini A M, Dario P. (2008). A survey of glove-based systems and their applications. IEEE Transactions on Systems, Man, and Cybernetics, Part C (App/lications and Reviews) 38(4): 461-482. doi: 10.1109/TSMCC.2008.923862

[8] Wood E, Bulling A. (2014) Eyetab: Model-based gaze estimation on unmodified tablet computers. In Proceedings of the Symposium on Eye Tracking Research and Applications, pp: 207-210

[9] Pino C, Kavasidis I. (2012). Improving mobile device interaction by eye tracking analysis. In Proceeding of Federated Conference on Computer Science and Information Systems. pp:1199-1202.

[10] Huang Q, Veeraraghavan A, Sabharwal A. (2017). Tabletgaze: dataset and analysis for unconstrained appearance-based gaze estimation in mobile tablets. In Machine Vision and Applications. 28(5): 445-461. doi: 10.1007/s00138-017-0852-4

[11] Miluzzo E, Wang T, Campbell A T. (2010). Eyephone: activating mobile phones with your eyes. In Proceedings of the second workshop on Networking, systems, and applications on mobile handhelds, pp: 15-20.

[12] Iqbal N, Lee H, Lee S Y. (2013). Smart user interface for mobile consumer devices using model-based eye-gaze estimation. In IEEE Transactions on Consumer Electronics, 59(1): 161-166

[13] Vaitukaitis V, Bulling A. (2012). Eye gesture recognition on portable devices. In Proceedings of the 2012 ACM Conference on Ubiquitous Computing, pp: 711-714

[14] Zhang X, Kulkarni, H, Morris M R. (2017). Smartphone-based gaze gesture communication for people with motor disabilities. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems. pp: 2878-2889.

[15] Elleuch, H., Wali, A., Samet, A., Alimi, A.M. (2016). A real-time eye gesture recognition system based on fuzzy inference system for mobile devices monitoring. In Proceedings of International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 172-180.

[16] Meng X, Cheung C M, Ho K L, Lui K S, Lam E Y, Tam V. (2012). Building smart cameras on mobile tablets for hand gesture recognition. In Proceeding of Sixth International Conference on Distributed Smart Cameras, pp: 1-5

[17] Prasuhn L, Oyamada Y, Mochizuki Y, Ishikawa H. (2014). A hog-based hand gesture recognition system on a mobile device. In Proceedings of IEEE International Conference on Image Processing, pp. 3973-3977

[18] Dalal N, Triggs B. (2005). Histograms of oriented gradients for human detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 1:886-893

[19] Song J, Soros G, Pece F, Fanello S R, Izadi S, Keskin C, Hilliges O. (2014). In-air gestures around unmodified mobile devices. In Proceedings of the 27th annual ACM symposium on User interface software and technology, pp: 319-329

[20] Elleuch, H., Wali, A., Samet, A., Alimi, A.M. (2015). A static hand gesture recognition system for real time mobile device monitoring. In Proceedings of the 15th International Conference on Intelligent Systems Design and Applications (ISDA), pp. 195-200.

[21] Chatterjee I, Xiao R , Harrison C. (2015). Gaze+ gesture: Expressive, precise and targeted free-space interactions, Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, 131-138

[22] Hales J, Rozado D, Mardanbegi D. (2013). Interacting with objects in the environment by gaze and hand gestures. In Proceedings of the 3rd international workshop on pervasive eye tracking and mobile eye-based interaction; pp: 1-9

[23] Pouke M, Karhu A, Hickey S, Arhippainen L. (2012). Gaze tracking and non-touch gesture based interaction method for mobile 3d virtual spaces. In: Proceedings of the 24th Australian Computer-Human Interaction Conference, pp. 505-512

[24] Yoo B, Han J J, Choi C, Yi K, Suh S, Park D, Kim C. (2010). 3d user interface combining gaze and hand gestures for large-scale display. In Proceedings of Extended Abstracts on Human Factors in Computing Systems, pp: 3709-3714

[25] Viola P, Jones M J. (2004). Robust real-time face detection. International journal of computer vision. 57(2): 137-154

[26] Elleuch H, Wali A, Alimi A M. (2014). Smart tablet monitoring by a real-time head movement and eye gestures recognition system. In Proceedings of International Conference on Future Internet of Things and Cloud (FiCloud); pp. 393-398.

[27] OpenCV [http://www.opencv.org/]

[28] Gonzalez-Ortega D, Diaz-Pernas F, Martinez-Zarzuela M, Anton-Rodriguez M, Diez-Higuera J, Boto-Giralda D. (2010). Real-time hands, face and facial features detection and tracking: Application to cognitive rehabilitation tests monitoring. Journal of Network and Computer Applications. 33(4): 447-466

[29] REHG [http://www.regim.org/publications/databases/rehg/]

[30] MIT [ http://web.mit.edu/torralba/www/indoor.html]

[31] Memo A, Zanuttigh P. (2018). Head-mounted gesture controlled interface for human-computer interaction. Multimedia Tools and Applications. 77(1): 27-53. doi: 10.1007/s11042-016-4223-3.

[32] Everingham M, Van Gool L, Williams C K, Winn J, Zisserman A. (2010). The pascal visual object classes (voc) challenge. International journal of computer vision. 88(2): 303-338

[33] NUS [https://www.ece.nus.edu.sg/stfpage/elepv/NUS-HandSet/]

[34] Pisharady P K, Vadakkepat P, Loh A P. Attention based detection and recognition of hand postures against complex backgrounds. International Journal of Computer Vision.101(3) :403-419

[35] Elleuch H, Wali A, Alimi M A, Samet A. (2017). Interacting with mobile devices by fusion eye and hand gestures recognition systems based on decision tree approach. In Ninth International Conference on Machine Vision, pp: 103-410.

[36] Dumas B, Lalanne D, Oviatt S. (2009). Multimodal interfaces: A survey of principles, models and frameworks. Human machine interaction. 5440:3-26. doi: https://doi.org/10.1007/978-3-642-00437-7_1

[37] FuzzyLite [http://www.fuzzylite.com/]

[38] Skodras E, Fakotakis N. (2015). Precise localization of eye centers in low resolution color images. Image and Vision Computing. 36: 51-60

https://doi.org/10.31436/iiumej.v20i2.1023

# CULINARY RECOMMENDATION APPLICATION BASED ON USER PREFERENCES USING FUZZY TOPSIS

**FARIED EFFENDY*, BARRY NUQOBA, TAUFIK**

*Information System Study Program, Department of Mathematics,*
*Faculty of Science and Technology, Universitas Airlangga, Surabaya, Indonesia.*

**Corresponding author: faried-e@fst.unair.ac.id*

*(Received: 28th November 2018; Accepted: 29th May 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** Culinary tourism is the experience of finding and enjoying unique and impressive food and drinks in a new region. Each region has a unique food flavor and can be the right choice for culinary tourists who want to try new experiences regarding taste. Differences in tourist preferences in choosing culinary regions sometimes become obstacles in finding a suitable place. Price, distance, facilities, menu variations, and halalness are some of the things that tourists consider in choosing culinary locations. Therefore, it is essential to accommodate tourist preferences to produce appropriate recommendations because the characteristics of tourists in determining their favorite culinary places are very subjective and varied. This study aims to develop a culinary recommendation system based on user preferences using Fuzzy TOPSIS as a recommendation method. Ten criteria can be selected, and their weight value determined by the user. The final result of the recommendation system is an alternative culinary location ranking and map location, along with detailed culinary location information. Result evaluation shows that the system is suitable and useful for users. This study proposes some new criteria that have not been used in other research.

***ABSTRAK:*** Pelancongan masakan merupakan pengalaman mencari dan menikmati makanan dan minuman yang unik dan menarik di sesebuah daerah. Setiap daerah mempunyai rasa unik makanan dan boleh menjadi pilihan tepat untuk pelancong yang ingin mencuba pengalaman baru mengenai rasa. Perbezaan pilihan pelancong dalam memilih tempat masakan kadang-kadang menjadi halangan dalam mencari tempat sesuai. Harga, jarak, kemudahan, variasi menu, dan makanan halal merupakan beberapa faktor yang dilihat pelancong dalam memilih lokasi kulinari. Oleh itu, sokongan cadangan untuk pelancong membuat pilihan adalah penting kerana pilihan makanan adalah sangat subjektif dan berbeza-beza. Kajian ini bertujuan bagi membangunkan sistem cadangan masakan mengikut pilihan pengguna dengan menggunakan kaedah Fuzzy TOPSIS. Sepuluh kriteria boleh dipilih mengikut keutamaan. Matlamat akhir sistem ini adalah sebagai pilihan lain bagi pilihan lokasi makanan dan lokasi peta bersama maklumat terperinci lokasi masakan. Keputusan menunjukkan sistem ini sesuai dan berguna untuk pengguna. Kajian ini mencadangkan beberapa kriteria baru yang tidak digunakan oleh penyelidik lain.

***KEYWORDS:*** *recommender system; fuzzy TOPSIS; culinary ranking; user preferences; GIS*

## 1. INTRODUCTION

Tourism is a complex multidimensional sector that influences the relevance of other areas in tourism activities. Tourism is all activities related to the tourist, and its multidimensional and multidisciplinary nature emerges as a manifestation of the needs of

every person and country as well as interactions between tourists and the local community, fellow tourists, the Government, Regional Government and entrepreneurs [1]

Indonesia is actively developing the tourism industry, especially for Surabaya, as one of the metropolitan cities in Indonesia, also focusing on increasing the visits of both local and foreign tourists. Various efforts have been made, starting from city park beautification, city corner cleaning, growing tourist destinations, and many other things. Surabaya is famous as a city with a variety of culinary destinations that present many colors and choices for tourists who want to enjoy culinary treasures in Surabaya. However, some obstacles are encountered by tourists, including lack of knowledge about culinary locations in Surabaya and different preferences in determining culinary location.

The selection of culinary destinations has many criteria, such as price, location, convenience, type of menu, and others [2]. The choice is also subjective, meaning that each tourist has different preferences in choosing culinary destinations. According to [3], if data attributes of an alternative are not adequately served, contain uncertainties or inconsistencies, the usual Multiple Attribute Decision Making (MADM) method cannot be used. Based on these problems, it is necessary to have a Decision Support System (DSS) with Geographic Information System (GIS), so that the system can provide automatic recommendations on culinary tourism. DSS allows users to make decisions more quickly and carefully by combining hardware, software, and decision processes [4].

There have been many studies discussing Decision Support Systems (DSS) culinary recommendations [2, 5, 7]. Web and mobile-based applications have also been made to facilitate culinary connoisseurs in finding the nearest culinary location. However, most web and mobile apps focus more on distance and rating, so other criteria such as convenience, a variety of menus, prices, and halalness were rarely used as filters or additional rules.

The concept of combining DSS and GIS is now growing. The method in DSS is used to analyze the inputted data, and GIS is used to display location information in the form of mapping. The advantages of DSS in analyzing data and the advantages of GIS in delivering information are combined, making this system better when used [8].

This culinary selection problem requires the consideration of multiple factors and can be categorized as a multi-criteria decision-making problem. Fuzzy TOPSIS is one of the multi-criteria decision-making methods used to find optimal alternatives from several alternatives with certain criteria. Fuzzy is used in linguistic criteria that have uncertainty, while TOPSIS is used to get the best alternative. Fuzzy TOPSIS is easy to use, can take into account all types of criteria (subjective and objective), rational logic, and is easy to understand for practitioners. The calculation process is straightforward, the concept allows the pursuit of the best alternative criteria described in mathematics in a simple way, and essential weights can be entered easily [9, 10]. This study aims to design a mobile application that makes it easier for Surabaya tourists to choose the destination for culinary locations based on their preferences. Criteria to be selected include convenience, menu variation, price, distance, parking area, facilities, service, and halalness.

## 2. RELATED WORK

Several criteria were used to determine tourist destinations according to tourist considerations [11]. The Analytic Hierarchy Process (AHP) method is used as a method for analyzing data to determine the priority of tourist destinations in the Talaud Islands. Promotional patterns that are still traditional for tourism in the Talaud Islands are

considered to be still inefficient. Therefore a web-based system is needed that can become an effective promotional media so that it can increase tourist enthusiasm.

Developing the concept of a combination of GIS with multi-criteria was also used to determine the location of the industry in the Vojvodina area [12]. The use of the idea can provide a location determination solution by analyzing several influential factors. GIS in this study is used as an approach to spatial decision making to determine the location of the industry.

Other research has applied the merge of Fuzzy Multi-Attribute Decision Making (FMADM) and Geographic Information System (GIS) [13]. The incorporation of FMADM and GIS is done to identify and evaluate the application of ecotechnology. The process carried out is to determine the use of ecotechnology methods that are suitable for riverbank walls. Feasibility evaluation refers to specific criteria provided by local area experts. The results of the study show that GIS can provide convenience for users to obtain information regarding geographical problems.

A research was conducted by Taroreh et al. [14] to determine the location of tourism development in the Poso region using feedback from the community (feedback). The Poso region has natural tourism that is attractive to tourists but still has not been developed; the development carried out by the government is based on feedback and input from the community. The results of this study indicate that this system can provide a solution for the government to develop tourism.

Some research [15] used location-based services (LBS) to help tourists find restaurant locations and specific culinary menus. LBS is a service that can send data and information that contains current location information about the user's presence or information that projects several locations of mobile users. LBS is applied as a part of architecture consisting of five components, such as mobile devices, communication networks, positioning components (GPS), provider services, applications, and data providers. Services delivered include map areas, weather conditions, traffic flow conditions, tour guides, shopping information, and so on. LBS requires a precise location to produce useful information.

Afnarius et al. [16] have built a culinary tourism search application in West Sumatra based on the mobile geographic information system. The development of this application aims to help tourists get information about culinary tours in West Sumatra. The app is built using the PHP programming language, Basic4Android, and Google Maps. The database used in the application is PostgreSQL. However, the application still needs to add additional features, such as uploading photos to social networks, accounts for users, and adding new culinary locations.

Geotagging is a procedure for adding geographic metadata information on different media such as adding geographic information on images, audiovisuals, internet sites, SMS messages, QR codes, and RSS in the form of geospatial metadata [17]. Generally, the information added consists of GPS geographic coordinates, location (address), description of data from the place.

The difference between previous research projects lies in the criteria, data analysis, and methods used. This study uses the concept of combining DSS with GIS to create an auto recommendation system, which applies the Fuzzy TOPSIS method. This application is expected to help tourists in Surabaya determine the location of a culinary destination and be able to display the route to the desired culinary tour.

https://doi.org/10.31436/iiumej.v20i2.1023

# 3. METHOD

## 3.1 Data and Information Collection

Survey and literature studies are conducted to find out what criteria are considered by tourists when choosing culinary destinations. At this stage, 50 respondents would select and verify the criteria obtained from literature studies. This verification is used to determine the factors that influence based on user needs. After confirmation, criteria derived will be used as the criteria for selecting culinary destinations in the system. Furthermore, the results of data collection on culinary destinations are used as an alternative in the system. Five culinary experts will assess all of these alternatives.

## 3.2 Fuzzy TOPSIS As Ranking Method

The Recommendation System works by ranking culinary destinations according to user preferences using the Fuzzy TOPSIS method. Following are the ranking steps using the Fuzzy TOPSIS method [18]:

Step 1: Categorizing Positive and Negative Ideal Criteria: at this stage, predetermined selection criteria will be grouped in the categories of positive and negative ideal criteria.

Step 2: Determine fuzzy numbers for weighting criteria and assessment of culinary destinations: at this stage, linguistic variables and TFN (Triangular Fuzzy Number) are associated with establishing the importance weight matrix for each criterion and performance rating for each alternative culinary destination.

Step 3: Determine Importance Weights: at this stage, the importance of each of the criteria used in the system will be determined. The level of importance of the criteria will be denoted by the formula:

$$\widetilde{w}_{ij} = \left(w_{j1}, w_{j2}, w_{j3}\right) \tag{1}$$

Step 4: Construction of Decision Matrix: at this stage, the performance rating of each alternative culinary destination will be determined. The decision matrix that is constructed contains alternative *m*, *n* criteria, and *k* experts. The decision matrix can be denoted by [10]:

$$\tilde{x}_{ij} = \left(l_{ij}, m_{ij}, u_{ij}\right) \tag{2}$$

where $$l_{ij} = min\{l_{ij}^k\}, m_{ij} = \frac{1}{K}\sum_{k=1}^{K} m_{ij}^k, u_{ij} = max\{u_{ij}^k\} \tag{3}$$

Step 5: Normalization of Decision Matrix: the decision matrix obtained in step 4 will be normalized. There is a difference in the normalization technique on the ideal positive criteria (benefit) and negative ideal (cost). Normalized decision matrices can be denoted by:

$$\tilde{R} = \left[\widetilde{r_{ij}}\right]_{m\times n} \tag{4}$$

where

$$\widetilde{r_{ij}} = \left(\frac{l_{ij}}{u_j^+}, \frac{m_{ij}}{u_j^+}, \frac{u_{ij}}{u_j^+}\right) \text{where } u_j^+ = max_i\, u_{ij} \tag{5}$$

for benefit criteria.

$$\widetilde{r_{ij}} = \left(\frac{l_j^-}{u_{ij}}, \frac{l_j^-}{m_{ij}}, \frac{l_j^-}{l_{ij}}\right) \text{where } l_j^- = min_i\, l_{ij} \tag{6}$$

for cost criteria.

Step 6: Weighting Decision Matrix: a normalized decision matrix will be multiplied by the level of importance/weight of the criteria obtained in step 3. The weighted decision matrix can be denoted by:

$$\widetilde{v_{ij}} = \widetilde{r_{ij}} \otimes \widetilde{w_j} \qquad i = 1,2,\ldots,m, j = 1,2,..,n \tag{7}$$

Step 7: Determining Fuzzy Positive Ideal Solution (FPIS) and Fuzzy Negative Ideal Solution (FNIS): determine FPIS and FNIS for each criterion:

$$A^{+} = \{\tilde{v}_1^{+}, \tilde{v}_2^{+}, \ldots, \tilde{v}_n^{+}\} \text{ for FPIS.} \tag{8}$$

$$A^{-} = \{\tilde{v}_1^{-}, \tilde{v}_2^{-}, \ldots, \tilde{v}_n^{-}\} \text{ for FNIS.} \tag{9}$$

where:

$$\tilde{v}_j^{+} = (1,1,1) \text{and } \tilde{v}_j^{-} = (0,0,0) \tag{10}$$

Step 8: Calculating the distance of each alternative to FPIS and FNIS: at this stage, the distance of each alternative that will be ranked against FPIS and FNIS will be calculated using the formula:

$$d_i^{+} = \sum_{j=1}^{n} d\left(\widetilde{v_{ij}} - \widetilde{v_j^{+}}\right) \tag{11}$$

$$d_i^{-} = \sum_{j=1}^{n} d\left(\widetilde{v_{ij}} - \widetilde{v_j^{-}}\right) \tag{12}$$

Step 9: Calculating the Relative Proximity to Ideal Solutions: Calculate the distance of the total criteria of each alternative to FPIS and FNIS so that *x* then calculate the proximity relative to the ideal solution with the formula:

$$S_i = \frac{d_i^{-}}{d_i^{-} + d_i^{+}} \tag{13}$$

Step 10: Ranking: Sort the results of closeness relative to the ideal solution obtained in the previous step from the largest to the smallest value; the higher the value, the better the alternative.

# 4. RESULTS

## 4.1 Data and Information Collection

From the results of data collection and processing, the final results are in the form of criteria that are considered when tourists choose culinary destinations. The full criteria can be seen in Table 1.

All criteria will be used as calculation criteria in the Fuzzy TOPSIS method. Filter criteria are the criteria used to filter out alternative culinary destinations that will be recommended to suit the user's wishes. For the Halal criteria, if a user wants a halal culinary destination, then alternative culinary destinations that will be recommended are alternative culinary destinations with halal menus. The criteria for filters that are not selected will not be a filter for the system; for example, if the halal criteria are not chosen, then all the culinary destinations which have a halal menu and not will be an alternative culinary destination that will be recommended.

https://doi.org/10.31436/iiumej.v20i2.1023

Table 1: Final result criteria

| Code | Criteria | Type |
|---|---|---|
| C1 | Comforts | Linguistic |
| C2 | Menu variation | Linguistic |
| C3 | Area of Place | *Crisp* |
| C4 | Distance | *Crisp* |
| C5 | Parking Area | Linguistic |
| C6 | Facility | Linguistic |
| C7 | Service | Linguistic |
| C8 | Price | *Crisp* |
| C9 | Halal | *Crisp, Filter* |
| C10 | Menu | Linguistic |

Linguistic criteria are non-exact criteria and will be used as criteria given by five experts. Criteria that have a type of crisp value will not be assessed again to facilitate the assessment of alternatives to be carried out by experts. Experts do not need to repeat the assessment of alternative culinary tourist destinations on specific criteria that already have exact numbers.

### 4.2 Fuzzy TOPSIS as Ranking Method

Step 1: Results that have been obtained from the assessment criteria will be grouped into two types of criteria, namely the criteria of benefit and cost. The criteria to be taken are five criteria, namely C1 (Convenience), C2 (Menu Variation), C3 (Place Area), C8 (Price), and C10 (Menu Type). For the calculation example, five criteria will be taken, namely C1 (Convenience), C2 (Menu Variation), C3 (Place Area), C8 (Price), and C10 (Menu Type). Categorizing the assessment criteria can be seen in table 2.

Table 2: Grouping of Positive and Negative Ideal Criteria

| Criteria | Type |
|---|---|
| C1 | Ideal Positive Solution |
| C2 | Ideal Positive Solution |
| C3 | Ideal Positive Solution |
| C8 | Ideal Positive Solution |
| C10 | Ideal Positive Solution |

Step 2: Fuzzy Numbers that will be used for weighting criteria, and evaluating each alternative can be seen in Tables 3 and 4.

https://doi.org/10.31436/iiumej.v20i2.1023

Table 3: Importance Rating

| IMPORTANCE RATING | | | |
|---|---|---|---|
| **Linguistic Variables** | **TFNs** | | |
| Very Low (VL) | 0 | 0 | 0.1 |
| Low (L) | 0 | 0.1 | 0.3 |
| Medium Low (ML) | 0.1 | 0.3 | 0.5 |
| Medium (M) | 0.3 | 0.5 | 0.7 |
| Medium High (MH) | 0.5 | 0.7 | 0.9 |
| High (H) | 0.7 | 0.9 | 1 |
| Very High (VH) | 0.9 | 1 | 1 |

Table 4: Assessments Rating

| ASSESSMENTS RATING | | | |
|---|---|---|---|
| **Linguistic Variables** | **TFNs** | | |
| Very Poor (VP) | 0 | 0 | 1 |
| Poor (P) | 0 | 1 | 3 |
| Medium Poor (MP) | 1 | 3 | 5 |
| Fair (F) | 3 | 5 | 7 |
| Medium Good (MG) | 5 | 7 | 9 |
| Good (G) | 7 | 9 | 10 |
| Very Good (VG) | 9 | 10 | 10 |

Step 3: The next step is to choose the importance level of the criteria. The level of importance of the criteria is in the form of linguistic variables that will be converted into fuzzy numbers. For example, this calculation will use the following criteria:
C1: "Very High"
C2: "Very High"
C3: "Very High"
C8: "Medium"
C10: "Medium-High"
So the results are as shown in Table 5.

Table 5: Importance of Weight Matrix

| **Criteria** | **TFN** | | |
|---|---|---|---|
| | ***l*** | ***m*** | ***u*** |
| C1 | 0.9 | 1 | 1 |
| C2 | 0.9 | 1 | 1 |
| C3 | 0.9 | 1 | 1 |
| C8 | 0.3 | 0.5 | 0.7 |
| C10 | 0.5 | 0.7 | 0.9 |

Step 4: In this stage, performance ratings are given by each expert (P) to each alternative to building a decision matrix. The performance rating given by each expert is a linguistic variable. The linguistic variable will then be converted into fuzzy numbers. Experts (P) only provide performance ratings against assessment criteria that are not crisp numbers. The performance

rating for criteria in the form of crisp numbers will be taken directly from the data specifications of culinary destinations. The performance rating given by experts will be aggregated because experts number more than one person. The aggregation calculations for A1 on C1 are as follows: For $k$ = 3 in criterion C1 for alternative A1, for example, $a$ = A1 and $b$ = C1

$$\tilde{x}_{ab} = \left(min(l_{ab}^1, l_{ab}^2, l_{ab}^3), (m_{ab}^1 + m_{ab}^2 + m_{ab}^3)/3\,, max(u_{ab}^1, u_{ab}^2, u_{ab}^3)\right)$$

$$\tilde{x}_{ab} = \left(min(0.9,0.7,0.9), (0.9 + 0.9 + 0.9)/3\,, max(1,1,1)\right)$$

$$\tilde{x}_{ab} = (0.7,0.9,1)$$

For criteria that have crisp values, fuzzy numbers will not be used, but the form of the criteria value will be changed to TFN. For example, criterion C3 (Area of Place) has a value of 8; it will be converted into the form of TFN to (8,8,8). This is done to equalize crisp numbers into calculations using the TFN.

Step 5: At this stage, the normalization of the decision matrix will be carried out. The formula used to normalize depends on the type of criteria that have been grouped in step 1. Calculation of normalization of the decision matrix on positive ideal criteria is as follows:
For i = 5 on C1 belongs to the alternative A1, for example $a$ = A1 and $b$ = C1:

$u_b^+ = max_i(10,10,10,10,9)$ so that $u_b^+ = 10$

$$\tilde{r}_{ab} = \left(\frac{7}{10}, \frac{9}{10}, \frac{10}{10}\right)$$

$$\tilde{r}_{ab} = (0.7,0.9,1)$$

While for the negative ideal criteria, normalization of decision matrices can be calculated as follows:

For $i$ = 5 on C8 belongs to alternative A1, for example $a$ = A1 and $b$ = C8:

$l_b^- = min_i(1.6,1.1,1.8,1.7,2)$ so that $l_b^- = 1.1$

$$\tilde{r}_{ab} = \left(\frac{1.1}{1.6}, \frac{1.1}{1.6}, \frac{1.1}{1.6}\right)$$

$$\tilde{r}_{ab} = (0.69,0.69,0.69)$$

Step 6: Then, the normalized decision matrix will be weighted multiplied by the importance of weights obtained in step 3. Weighting the decision matrix can be calculated as follows:

For criterion C1 in alternative A1, for example, a = A1 and b = C1:

$$\tilde{v}_{ab} = (0.7,0.9,1) \times (0.9,1,1)$$

$$\tilde{v}_{ab} = (0.63,0.9,1)$$

Step 7: FPIS and FNIS have their respective values: $FPIS = (1,1,1)$dan $FNIS = (0,0,0)$

https://doi.org/10.31436/iiumej.v20i2.1023

Step 8: At this stage, each alternative in the weighted decision matrix will be calculated: the distance to the ideal fuzzy positive solution (FPIS) and the negative fuzzy ideal solution (FNIS). Alternative distances to FPIS and FNIS for positive ideal criteria can be calculated as follows:

For C1 and C2 on A1, suppose:

$$a = A1, x = C1, y = C2$$

$$d_a^+ = \sum\left((\mathscr{v}_{ax1}, \mathscr{v}_{ax2}, \mathscr{v}_{ax3}) - (1,1,1)\right) + \sum\left((\mathscr{v}_{ay1}, \mathscr{v}_{ay2}, \mathscr{v}_{ay3}) - (1,1,1)\right)$$

$$d_a^+ = \left|\left((\mathscr{v}_{ax1} - 1) + (\mathscr{v}_{ax2} - 1) + (\mathscr{v}_{ax3} - 1) + (\mathscr{v}_{ay1} - 1) + (\mathscr{v}_{ay2} - 1)\right) + (\mathscr{v}_{ay3} - 1)\right|$$

$$d_a^+ = \left|((0.69 - 1) + (0.9 - 1) + (1 - 1) + (0.63 - 1) + (0.93 - 1) + (1 - 1))\right|$$

$$d_a^+ = 0.91$$

$$d_a^- = \sum\left((\mathscr{v}_{ax1}, \mathscr{v}_{ax2}, \mathscr{v}_{ax3}) - (0,0,0)\right) + \sum\left((\mathscr{v}_{ay1}, \mathscr{v}_{ay2}, \mathscr{v}_{ay3}) - (0,0,0)\right)$$

$$d_a^- = \left|\left((\mathscr{v}_{ax1} - 0) + (\mathscr{v}_{ax2} - 0) + (\mathscr{v}_{ax3} - 0) + (\mathscr{v}_{ay1} - 0) + (\mathscr{v}_{ay2} - 0)\right)\right|$$

$$d_a^- = \left|((0.69 - 0) + (0.9 - 0) + (1 - 0) + (0.69 - 0) + (0.93 - 0) + (1 - 0))\right|$$

$$d_a^- = 5.09$$

So that the distance between FPIS and FNIS can be seen in Tables 6 and 7.

Table 6**:** Alternative Distance to FPIS

| Criteria | A1 | A2 | A3 | A4 | A5 |
|---|---|---|---|---|---|
| C1 | 0.47 | 0.44 | 0.79 | 0.79 | 1.51 |
| C2 | 0.44 | 1.03 | 0.41 | 0.47 | 1.03 |
| C10 | 1.1 | 1.27 | 1.1 | 0.95 | 1.12 |
| C3 | 2 | 1.5 | 2.76 | 2 | 2 |
| C8 | 1.56 | 0.9 | 1.44 | 1.65 | 1.86 |
| Distance | 5.57 | 5.14 | 6.5 | 5.86 | 7.52 |

Table 7: Alternative Distance to FNIS

| Criteria | A1 | A2 | A3 | A4 | A5 |
|---|---|---|---|---|---|
| C1 | 2.53 | 2.56 | 2.21 | 2.21 | 1.49 |
| C2 | 2.56 | 1.97 | 2.59 | 2.53 | 1.97 |
| C10 | 1.9 | 1.73 | 1.9 | 2.05 | 1.88 |
| C3 | 1 | 1.5 | 0.24 | 1 | 1 |
| C8 | 1.44 | 2.1 | 1.56 | 1.35 | 1.14 |
| Distance | 9.43 | 9.86 | 8.5 | 9.14 | 7.48 |

Step 9: The distance of each alternative to the ideal positive fuzzy solution and the ideal negative fuzzy solution will be calculated for the proximity of the ideal solution so that the results can be seen in Table 8. Alternative closeness to the ideal solution can be calculated as follows:

For alternative A1, where a = A1

$$S_i = \frac{d_i^-}{d_i^- + d_i^+}$$

$S_a = \frac{9.39}{9.39+5.61}$ so that $S_a = 0.626$

Table 8: Proximity to ideal solutions

| Distance | A1 | A2 | A3 | A4 | A5 |
|---|---|---|---|---|---|
| di+ | 5.57 | 5.14 | 6.5 | 5.86 | 7.52 |
| di- | 9.43 | 9.86 | 8.5 | 9.14 | 7.48 |
| di- + di+ | 15 | 15 | 15 | 15 | 15 |
| Si | 0.629 | 0.657 | 0.567 | 0.609 | 0.499 |

Step 10: At this stage, an alternative closeness to the ideal solution will be ranked based on the values obtained to determine the best alternative. The higher the value obtained by the alternative, the better. Ranking results can be seen in Table 9.

Table 9: Alternative ranking results

| Ranking | Alternative | Point |
|---|---|---|
| 1 | A1 | 0.657 |
| 2 | A2 | 0.629 |
| 3 | A4 | 0.609 |
| 4 | A3 | 0.567 |
| 5 | A5 | 0.499 |

### 4.3 Implementation

The user of this application can input directly into a text box on what kind of food they want; then they can choose such criteria that they prefer and its importance criteria. There are seven levels of importance criteria that can be selected (Very Low, Low, Medium Low, Medium, Medium High, High, and Very High). The application will display the culinary places that meet the condition and rank it from the highest to lowest; the rank image is displayed from left to right. The app will also view the map and how to reach that place when the picture is clicked. The application image can be seen in Fig. 1.

### 4.4 System Testing

To test the system, black-box testing was used to validate the system in making recommendations by comparing the results of expectations (expert judging) with system output. We used 20 different tests, that cover different criteria, and their weight and calculates the score value with the formula:

$$Score = \frac{Total\ passed\ case}{total\ case} \quad (14)$$

https://doi.org/10.31436/iiumej.v20i2.1023

Table 10 shows number 1 of 20 test cases that have been done; overall result evaluation shows that this system scored 75%, meaning that there were 15 cases that had the same result between system output and expert judgment.

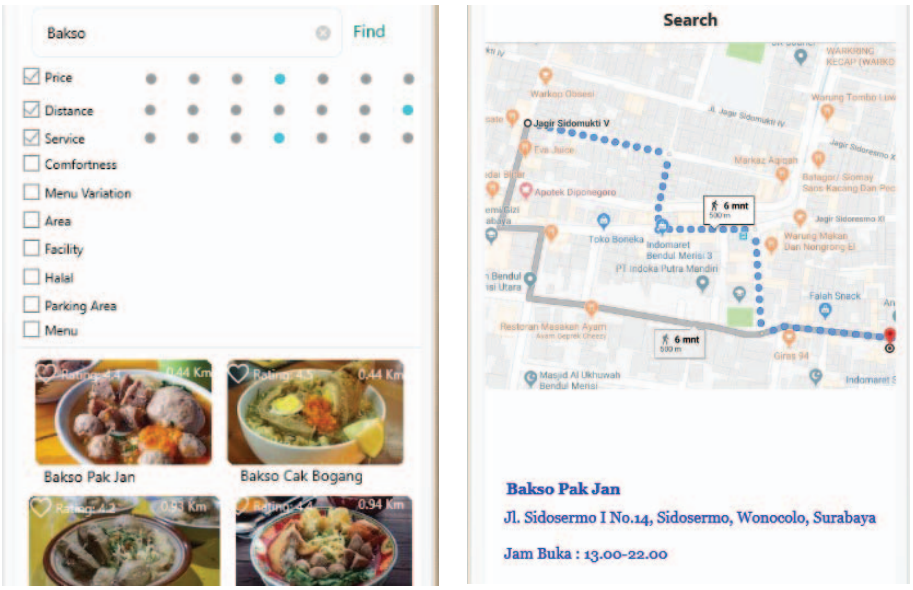



Fig. 1: The Application.

Table 10: Recommendation results

| Test Case | Criteria | Weight | System Output | Experts |
|---|---|---|---|---|
| 1 | Price | Medium | Rank: | Rank: |
| | Distance | Very High | 1.Bakso P.Jan | 1.Bakso P.Jan |
| | Service | Medium | 2.Bakso Cak Bagong | 2.Bakso Cak Bagong |
| | | | 3.Bakso Jenggot | 3.Bakso Jenggot |
| | | | 4.Bakso Bringas | 4.Bakso Bringas |
| | | | 5.Bakso Cak Man | 5.Bakso Cak Man |

## 4.5 System Evaluation

The system evaluation aims to determine the performance accuracy of the Fuzzy TOPSIS method as a decision support system for selecting culinary locations. System evaluation will be carried out by measuring user satisfaction with the system. The users were application users and experts who conduct culinary location selection. The categories used in the evaluation questionnaire are usability, accuracy, functional completeness, performance, and subjective assessment of users of the system as a whole (overall). Each category asked in the questionnaire will calculate the MOS (Mean Opinion Score) value. MOS calculations are used to conclude the ratings given by the user. The application is said to be useful if it has a value of 3 to 4. MOS formula can be seen below:

$$MOS = \frac{\sum_{n=0}^{N} R_n}{N} \quad (15)$$

where R is the individual rating of subject *N* for stimulus or *n* related factor.

This application was evaluated by twenty end-users and five experts; evaluation results show that usability, completeness, accuracy, performance, and overall category each scored 3.5, 3.5, 3.5, 3.65, and 3.6. The MOS results for each category are then calculated with MOS so that the overall system score is obtained. The calculation for the overall system score is as follows:

$$System\ score = \frac{3.5 + 3.5 + 3.5 + 3.65 + 3.6}{5} = 3.55$$

From the results of the system score calculation, it was found that twenty-five system users gave a value of 3.55 from the total value of 4. Therefore, it can be concluded that the system is useful.

## 5. CONCLUSION

The decision support system with the Fuzzy TOPSIS method helps users by identifying the essential criteria in the process of selecting culinary destinations. These criteria were the convenience, a variety of menu, price, distance, parking area, facilities, services, area, halal, and type of menu. In these criteria, there are three types of values, namely linguistic, crisp, and filter. Linguistic use for criteria that are not exact values, crisp values use for criteria in the form of exact values, while filter values use for optional type criteria, which serve as an alternative filter for culinary destinations. The user's level of importance then determines the non-filter criteria. The level of importance of the criteria and assessment of alternative culinary destinations is processed by the Fuzzy TOPSIS method, and the system will display five culinary destinations with the best ranking as recommendations to users as the best culinary destination choice according to their preferences.

Decision support systems that are built have a high impact on solving the problem of choosing culinary destinations by providing recommendations on culinary destinations that suit the needs and preferences of users. This is evidenced by the results of a system testing and evaluation, which shows that 75% of system output has the same result as expected result, and the user considers the features provided in the decision support system with the TOPSIS Fuzzy method worth 3.55 out of 4.

In this study, the app is built with minimal features for end-users so that for further development, it is recommended that this application also has a user-based community where all users can post a review, comments, and also rate so that the performance rating of each alternative can be determined. This feedback can be related to culinary destinations so that this application can be an effective and unbiased platform for people who need culinary destinations recommendation.

## REFERENCES

[1] Haneef SK. (2017) A Model to Explore the Impact of Tourism Infrastructure on Destination Image for Effective Tourism Marketing. (Thesis). UK: University of Salford Manchester.
[2] Maharani S, Hatta HR, Merdiko G. (2014) Decision Support System of Culinary Recommendations Using Ahp And Topsis Methods With Map Visualization. In Bali International Seminar On Science And Technology (BISSTECH) II 2014 (p. C3.2-1-C3.2-6). Bali.
[3] Kusumadewi S. (2006) Fuzzy Multi-Attribute Decision Making (FUZZY MADM). Yogyakarta: Graha Ilmu.

[4] Turban E. (2010) Decision Support and Business Intelligence Systems 9th. Prentice-Hall Press Upper Saddle River, NJ, USA ©2010.
[5] Suteja BR, Guritno S. (2017) A recommendation system for culinary tourists in Jogjakarta based on collaborative filtering. In 2017 Second International Conference on Informatics and Computing (ICIC) (pp. 1–4). Jayapura, Indonesia: IEEE.
[6] Huda AA, Hantono BS, Widyawan. (2015) Sistem Rekomendasi Restoran Dengan Pendekatan Ekstraksi Fitur Pada Menu Makanan. In Seminar Nasional Teknologi Informasi dan Komunikasi 2015 (SENTIKA 2015) (pp. 285-290). Yogyakarta.
[7] Pinandito A, Fanani L, Brata KC. (2015) Alternatives weighting in analytic hierarchy process of mobile culinary recommendation system using fuzzy. ARPN Journal of Engineering and Applied Sciences, 10(19), 8791–8798.
[8] Halim I, Arep H, Kamat SR, Abdullah R, Omar AR, Ismail AR. (2014) Development of a decision support system for analysis and solutions of prolonged standing in the workplace. Safety and Health at Work, 5(2), 97–105. https://doi.org/10.1016/j.shaw.2014.04.002.
[9] Azizi A, Aikhuele DO, Souleman FS. (2015) A Fuzzy TOPSIS Model to Rank Automotive Suppliers. Procedia Manufacturing, 2(February), 159–164. https://doi.org/10.1016/j.promfg.2015.07.028.
[10] Nadaban S, Dzitac S, Dzitac I. (2016) Fuzzy TOPSIS: A General View. Procedia Computer Science, 91(Itqm), 823–831. https://doi.org/10.1016/j.procs.2016.07.088.
[11] Taluay, H. R., Seminar, K. B., & Monintja, D. R. (2015). Development Of Web-Based Tourism Decision Support System In Talaud Island Regency. International Journal of Information Technology and Business Management, 39(1), 36-45.
[12] Rikalovic A, Cosic I, Lazarevic D. (2014) GIS based multi-criteria analysis for industrial site selection. Procedia Engineering, 69: 1054-1063. https://doi.org/10.1016/j.proeng.2014.03.090.
[13] Chou WC, Lin WT, Lin CY. (2007) Application of fuzzy theory and PROMETHEE technique to evaluate suitable ecotechnology method : A case study in Shihmen Reservoir Watershed, Taiwan, 1, 269–280. https://doi.org/10.1016/j.ecoleng.2007.08.004
[14] Taroreh AA, Tanaamah AR, Fibriani C. (2014) Decision Support System Feasibility of Tourism Resort in Poso District used 360 Degree Method. International Journal of Computer Science Issues, 11(2): 201-207.
[15] Brimicombe A, C Li. (2009) Location-Based Services and Geo-Information Engineering. Singapore: Faboulus Printers Pte Ltd.
[16] Afnarius S, Ningsih VM, Frihandana D. (2014) Pembangunan aplikasi wisata kuliner Sumbar berbasis mobile geographic information system. Prosiding Seminar Ilmiah Nasional Komputer Dan Sistem Intelijen, 8(Kommit): 354-360.
[17] Annisa NS, Dwi S. (2012) Perancangan dan Pembangunan Perangkat Fitur Geotagging. Jurnal Teknik POMITS, 1(1): 1-6.
[18] Yayla Y, Ozbek A, Yildiz A. (2012) Fuzzy TOPSIS Method in Supplier Selection and Application in the Garment Industry. Fibres and Textiles in Eastern Europe, 20: 20-23.

https://doi.org/10.31436/iiumej.v20i2.1136

# RADIATION DOSE DELIVERED BY $^{125}$I, $^{103P}$D AND $^{131}$CS AND DOSE ENHANCEMENT BY GOLD NANOPARTICLE (GNP) SOLUTION IN PROSTATE BRACHYTHERAPY: A COMPARATIVE ANALYSIS BY MONTE CARLO SIMULATION

**HAMDA KHAN[1], UMAIR AZIZ[2*], ZAFAR ULLAH KORESHI[2]**

*[1]Department. of Mathematics,*
*National University of Computer and Emerging Sciences, Islamabad, Pakistan.*
*[2]Department of Mechatronics Engineering, Air University, Islamabad, Pakistan.*

*[*]Corresponding author: umair.aziz@mail.au.edu.pk*

*(Received: 10th April 2019; Accepted: 1st August 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** The energy deposition and radiation dose from commonly used radioisotopes, $^{125}$I,$^{103}$Pd, and $^{131}$Cs, used for brachytherapy of cancers is estimated using Monte Carlo (MC) simulations. To enhance the dose, gold nanoparticle (GNP) solutions are injected into the tumor; this results in more effective and shorter therapy duration. It is thus important to estimate the dose enhancement factor (DEF) achievable by a radioisotope. The research presented in this paper thus focuses on a comparative analysis of radioisotopes. To estimate the radiation dose, the Monte Carlo N-particle code MCNP5 was used for a coupled photon-electron simulation of radiation transport from radiation emanating from seeds of radioisotopes implanted in the prostate at positions prescribed to deliver effective doses to the tumor while protecting neighbouring vital organs such as the rectum and urethra. The quantities tallied were the energy deposition (F6 tally) and the pulse heights (*F8 tally) in specified energy bins. The energy deposited in the tumor was used to estimate the absorbed dose to the prostate incorporating the transformations of the radioisotopes during decay. The absorbed dose was subsequently estimated for a GNP-tissue solution with a concentration of 25 mg Au/g of prostate tissue, modelled as a homogenous mixture. From the simulations, it was found that the lifetime absorbed dose is ~96 Gy from 98 seeds, each of 0.31 mCi, of 125I; ~102 Gy, from 115 seeds, each of 1.4 mCi, of 103Pd, and ~90 Gy from $^{131}$Cs seeds replacing 103Pd seeds of the same initial activity. The main advantage of $^{131}$Cs, over $^{125}$I and $^{103}$Pd, is observed in the larger dose rate (~26 cGy/hr) delivered initially i.e. in the first few days which is 1.5 and 5.7 times higher than that for $^{103Pd}$ and $^{125I}$. The absorbed dose for 125I, 103Pd and 131Cs increases to ~245, ~130, ~187 Gy respectively with GNP-tissue solution of 25 mg Au/g tissue. From the analysis, it is found that while the lifetime absorbed dose of all three radioisotopes is of the same order, there are advantages in using $^{131}$Cs; these advantages are further quantified.

***ABSTRAK:*** Pemendapan tenaga dan dos sinaran radiasi daripada radioisotop yang biasa digunakan, $^{125}$I,$^{103}$Pd, dan $^{131}$Cs, digunakan bagi terapibraki kanser dianggar menggunakan simulasi Monte Carlo (MC). Bagi meningkatkan dos, larutan partikel nano emas (GNP) telah disuntik ke dalam tumor; ini lebih memberi kesan dan mengurangkan masa terapi. Oleh itu, adalah penting menganggar faktor dos penggalak (DEF) dapat dicapai dengan radioisotop. Kajian ini mengfokuskan pada analisis perbandingan radioisotop. Bagi menganggarkan dos radiasi, kod Monte Carlo N-partikel MCNP5 telah digunakan pada simulasi pasangan foton-elektron pengangkutan radiasi daripada

pancaran radioaktif benih radioisotop yang ditanam dalam prostat pada posisi yang disebut bagi mencetuskan dos penghantaran yang berkesan pada sel tumor. Dalam masa sama melindungi organ penting seperti rektum dan uretra. Kuantiti diselaras dengan pemendapan tenaga (selaras F6) dan ketinggian denyut (selaras *F8) dalam aras tenaga sebenar. Tenaga yang dienap dalam sel tumor ini telah digunakan bagi menganggarkan dos serapan pada prostat dengan menggabungkan transformasi radioisotop ketika susutan. Dos yang diserap telah kemudiannya dianggarkan bagi larutan tisu-GNP dengan ketumpatan 25 mg Au/g tisu prostat, dimodelkan sebagai campuran homogen. Daripada simulasi, dapatan kajian menunjukkan dos diserap sebanyak ~96 Gy daripada 98 benih, setiap satu daripada 0.31 mCi, 125I; ~102 Gy, dari 115 benih, setiap 1.4 mCi, dari 103Pd, dan ~90 Gy daripada benih $^{131}$Cs menggantikan benih 103Pd pada pemulaan aktiviti yang sama. Keistimewaan utama adalah $^{131}$Cs, ke atas $^{125}$I dan $^{103}$Pd, telah dilihat dalam kadar dos lebih besar (~26 cGy/hr) dikeluarkan pada pemulaannya iaitu dalam beberapa hari pertama iaitu 1.5 dan 5.7 kali lebih tinggi daripada $^{103Pd}$ dan $^{125I}$. Dos yang diserap pada 125I, 103Pd dan 131Cs bertambah kepada ~245, ~130, ~187 Gy masing-masing dengan larutan tisu-GNP sebanyak 25 mg Au/g tisu. Hasil analisis menunjukkan penyerapan seumur hidup dos diserap pada ketiga-ketiga radioisotop dalam aturan yang sama, ini adalah keistimewaan menggunakan $^{131}$Cs; keistimewaan ini akan terus diuji pada masa depan dan diukur kuantitinya.

***KEYWORDS:*** *Monte Carlo; simulation; medical; prostate cancer therapy; gold nanoparticles; dose enhancement*

## 1. INTRODUCTION

Cancer cases globally are expected to grow from 14.1 million in 2012 to 24 million by 2035 [1, 2] with the top three: lung [3], prostate [4-6], and colorectal cases accounting for over 47% of all cancers in men. The treatments for cancer, in order of general preference, are chemotherapy, surgery, and radiation therapy. In the case of recurrent cancers however, brachytherapy is the preferred treatment due to its localized and non-invasive effects.

For brachytherapy, radioisotopes with energy <50 keV, such as $^{125}$I and $^{103}$Pd sources [7-9] with typical implants of 50-80 metallic seeds encasing isotopes, are used as low dose rate (LDR) [10] therapies for the treatment of prostate cancer, uveal melanomas and brain tumours. The dose is estimated by a number of computer codes such as EGSnrc, GEANT[11], PENELOPE [12] and MCNP [13] based on Monte Carlo (MC) methods [14].

Further improvements in the effectiveness of brachytherapy are being considered by the injection of gold nanoparticles through fenestrations of cancer cells. The particle size requirements are in the nanoscale range (~$10^{-9}$ m) which is comparable with the diameter of an atom (~$10^{-10}$ m) [15, 16].

For MC simulations in brachytherapy, it has been demonstrated that DEF is a function of the source energy and concentration of GNP solution [17-20], while the size of GNPs is not so important (above the K-edge energy). Thus, in this work, a homogenous model with considerable savings on the computational effort, required for a full heterogeneous model, is used to extract crucial information on the DEF resulting from the presence of gold in small concentration while MCNP has the capability of modelling very detailed heterogeneous configurations. This paper considers a homogenous model solely for the purpose of carrying out a comparative study for the dose delivered to a prostate tumour by $^{125}$I, $^{103}$Pd and $^{131}$Cs and particularly the enhancement of radiation dose due to the presence of gold.

https://doi.org/10.31436/iiumej.v20i2.1136

The ultimate goal of brachytherapy is to deliver maximum dose to the tumour while minimizing collateral damage to the normal tissue and for maximizing the DEF, there is still no consensus on the optimal size, shape and distribution of GNPs. However, experimental and pre-clinical evidence for mouse tumours, showing a 1-year survival rate of 86% following a dose of 26 Gy with 1.9 nm intravenously administered GNPs *vs* 20% for tumours not laden with GNPs [15], indicate that nanotechnology offers promising improvements in brachytherapy.

The radioisotope cesium-131 ($^{131}$Cs), was introduced to brachytherapy in 2004 [21] and in the first five years it was used in about 3000 prostate implants yielding the required dose in less time compared with $^{125}$I and $^{103}$Pd due to its shorter half-life and high energy, as shown in Table 1.

Table 1: Energy and half-life of $^{125}$I, $^{103}$Pd and $^{131}$Cs

| **Radionuclide** | **E (keV)** | **$T_{1/2}$ (days)** |
|---|---|---|
| $^{103}$Pd | 20.8 | 17 |
| $^{125}$I | 35.49 | 59.4 |
| $^{131}$Cs | 29-30.4 | 9.7 |

Typically,$^{125}$I is used with a radiation dose of 145 Gy or more conforming to the prescribed dose of 145 Gy suggested by the American Association of Physicists in Medicine Task Group 64 [22].

While all three are low-energy sources, so that their dose would not extend to normal tissue in nearby organs [23], one of the main advantages of $^{131}$Cs is that it offers an initial dose rate of ~32 cGy/h at the periphery [24] which is 1.5 and 4 times higher than that from $^{103}$Pd and $^{125}$I respectively. This initial dose rate advantage is a vital radiobiological parameter for rapidly growing tumours [25] as well as for slow-growing tumours such as prostate adenocarcinomas. A set of clinical recommendations for prescribed doses, on the basis of 1200 prostate implants [26] indicated that the high initial dose rate of $^{131}$Cs caused more intense urinary and rectal complication but they also resolved more quickly than that for $^{125}$I and $^{103}$Pd. That study also found that the drop in prostate-specific antigen levels from $^{131}$Cs, with a maximum follow-up of just over three years and median follow-up of 23 months, were equivalent to that from other isotopes. Such promising results from $^{131}$Cs extended its use to gynecological malignancies [27] and to brain radiotherapy [28] from which it was demonstrated that this treatment was well-tolerated and safe for patients.

## 2. MATERIALS AND METHODS

### 2.1 Monte Carlo Simulation

The Monte Carlo code MCNP5 is used to carry out a coupled photon-electron simulation of radiation transport in the range 1 keV-100 MeV to estimate the radiation dose distribution for prostate tumor brachytherapy. The radioisotope sources considered are $^{125}$I, $^{103}$Pd and $^{131}$Cs in the form of 'seeds' modelled as point sources. For $^{125}$I, 98 seeds each of 0.31 mCi and for $^{103}$Pd 115 seeds each of 1.4 mCi were spatially distributed in the tumour tissue, as shown in Fig. 1 [29] for data given for two patients with different size prostates. The placement of needles takes into account a number of factors including the location of vital organs surrounding the prostate, such as the bladder, urethra, and rectum.

For simulating the effect of $^{131}$Cs, the $^{103}$Pd seeds were replaced by $^{131}$Cs seeds of the same initial activity and spatial distribution. The energy deposition track length estimation tally F6 and the pulse height tally *F8 are both used for estimating energy deposition to get reliable estimates in case of a few interactions in a region of interest. The simulation is repeated with gold-tissue solution; material compositions for both simulations are listed in Table 1. The photon and electron data for both tissue and gold are based on ENDF/B-VI (Release 8).

In MCNP the "detailed physics" simulation incorporates coherent (Thomson) scattering and fluorescent photons produced from photoelectric absorption. Electrons produced from photon collisions are transported in a "condensed history" method that accumulates the effects of many individual collisions into single steps sampled probabilistically. The effects of such artefacts for electron transport have been investigated [30-32] with EGSnrc, GEANT and PENELOPE codes and in some cases "large discrepancies" (>3%) have been found between MCNP5 dose distributions and the 'reference codes' concluding that MCNP5 electron transport calculations are not accurate at all energies and in every medium by general clinical standards. It can thus be anticipated that MCNP5 may differ due to its inadequate low-energy treatment of electron transport. Authors of [33] have carried out electron transport comparisons of MCNPX, Penelope and EGSnrc, for electrons of 20-450 keV in water, lead, and tungsten. These comparisons were focused on bremsstrahlung, energy deposition in matter, electron ranges and production of secondary electrons by gamma radiation.

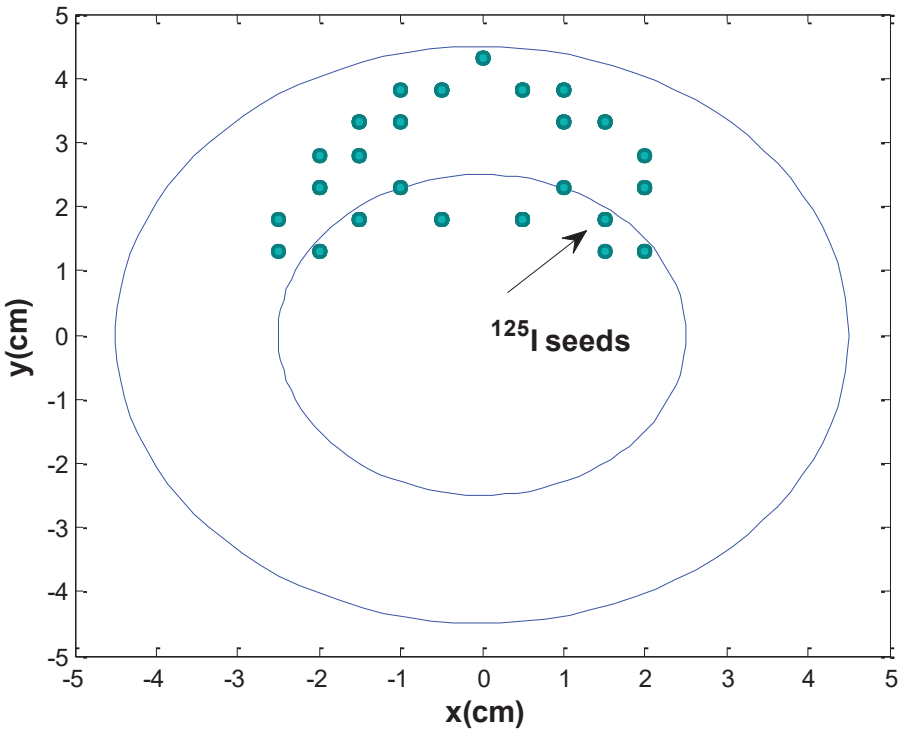



Fig. 1: Placement of 25 needles with 98 seeds of $^{125}$I in the x-y plane.

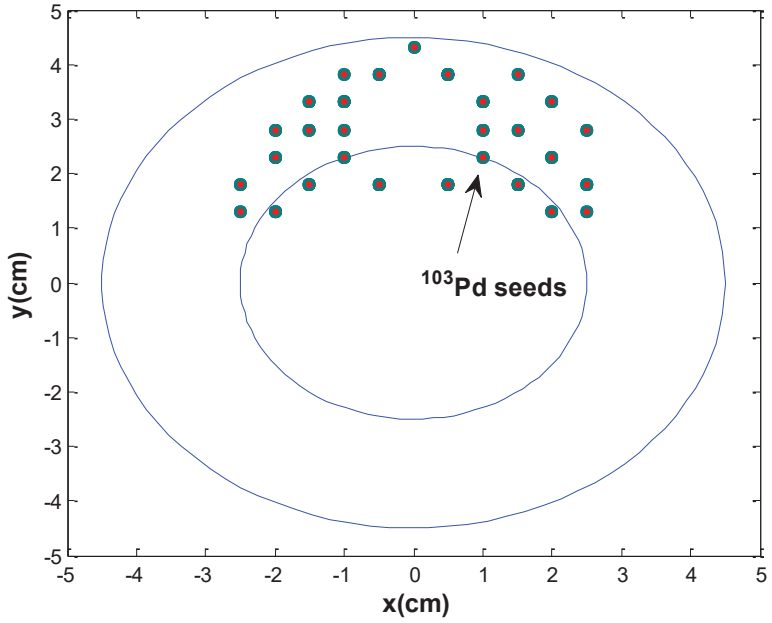



Fig. 2: Placement of 29 needles of 115 seeds of $^{103}$Pd/$^{131}$Cs in the x-y plane.

The F6 tally for dose

$$D = \frac{\rho}{m}\int dE \int dt \int dV \int d\Omega\, \sigma_t(E)H(E)\phi(\bar{r},\widehat{\Omega},E,t)\ \text{MeV g}^{-1}$$

from MCNP is converted from MeV/g to yield the dose rate in Gy/hr by using the activity as shown below.

## 2.2 Activity and Transformations

The activity of a radioisotope is $A(t) = A_o e^{-\lambda t}$ and the number of transformations in an interval $\tau = t_2 - t_1$ is thus

$$\text{Tr}(t_1, t_2) = \frac{A(t_1)}{\lambda}\left(1 - e^{-\lambda\tau}\right) \quad (1)$$

Thus the 'infinite-time' transformations i.e. the number of disintegrations integrated over the time interval $(0, \infty)$ is $A_o/\lambda$.

The absorbed dose is computed as follows:

$$\text{Abs Dose} = 1.6\ X\ 10^{-10}\frac{A_o}{\lambda}\frac{NE}{m}\ \text{Gy} \quad (2)$$

where $A_o$is the initial activity, N is the number of radioisotope seeds, and $E$ is the energy pulse height (*E8) tally (MeV) and $m$ is the mass of the tumor (g).

Thus, the absorbed dose varies directly with the energy deposition and source, which in turn depends on the number of transformations i.e. the product of initial activity and half-life.

https://doi.org/10.31436/iiumej.v20i2.1097

## 3. RESULTS

MCNP simulations were carried out on an Intel(R) Core (TM) i7-2620M CPU @ 2.70GHz with an installed memory of 8.00 GB (3.24 GB usable) and 32-bit Operating system with Windows 7 Professional. The runtime for each simulation of $10^7$ source particles was 35.60, 51.1 and 41 minutes for $^{125}$I, $^{103}$Pd and $^{131}$Cs, respectively

The time-dependent transformations and activities of $^{125}$I, $^{103}$Pd and $^{131}$Cs, for an initial activity of 1 mCi each, are shown in Fig. 3. The basic hypothesis for effectiveness of each isotope arises from its half-life. This factor results in a longer decay time for $^{125}$I while faster decay rates of $^{103}$Pd and $^{131}$Cs translate into lifetime transformations of $2.74 \times 10^{14}$, $7.84 \times 10^{13}$ and $4.47 \times 10^{13}$ for $^{125}$I, $^{103}$Pd and $^{131}$Cs respectively which clearly illustrate the slow build-up, but higher dose from $^{125}$I after ~ six months compared with the much faster effects of $^{103}$Pd and $^{131}$Cs. There is hence a saturation in activities for $^{103}$Pd and $^{131}$Cs with the implications that the effect of these two isotopes is seen much faster than that for $^{125}$I.

Figure 4 shows the absorbed dose for $^{125}$I (initial activity 0.31 mCi, 98 seeds distributed as illustrated in Fig. 1), $^{103}$Pd and $^{131}$Cs (both of initial activity 1.4 mCi 115 seeds distributed as illustrated in Fig. 2) which indicates that $^{131}$Cs is clearly the best in terms of delivering the highest and fastest dose to the tumour, reaching 50 Gy in the first 10 days and 100 Gy when GNPs in a solution of 25 mg/g tissue are injected in the tumour. The absorbed dose on a longer timescale, however, is highest for $^{125}$I with GNP solution but exceeds that from $^{131}$Cs after about three months. For the sake of comparison, when all seeds are of equal initial activity (1.4 mCi) and placed in the configuration of Fig. 2, it is seen in Fig. 5 that the highest absorbed dose for both cases, with and without GNP-solution, is the highest at 117 Gy and ~283 Gy respectively but after ~250 days). Thus for a given requirement, $^{125}$I exceeds the $^{103}$Pd dose in about 200 days; higher doses for stronger tumours can thus not be treated by $^{131}$Cs and $^{103}$Pd unless they are used in conjunction with gold solutions.

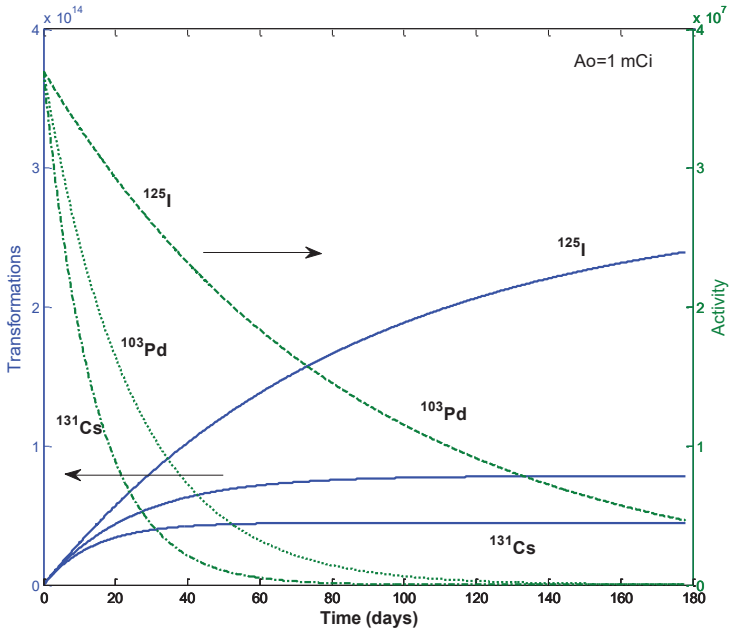



Fig. 3: Transformations and activity of $^{125}$I, $^{103}$Pd and $^{131}$Cs each of initial activity 1 mCi.

https://doi.org/10.31436/iiumej.v20i2.1136

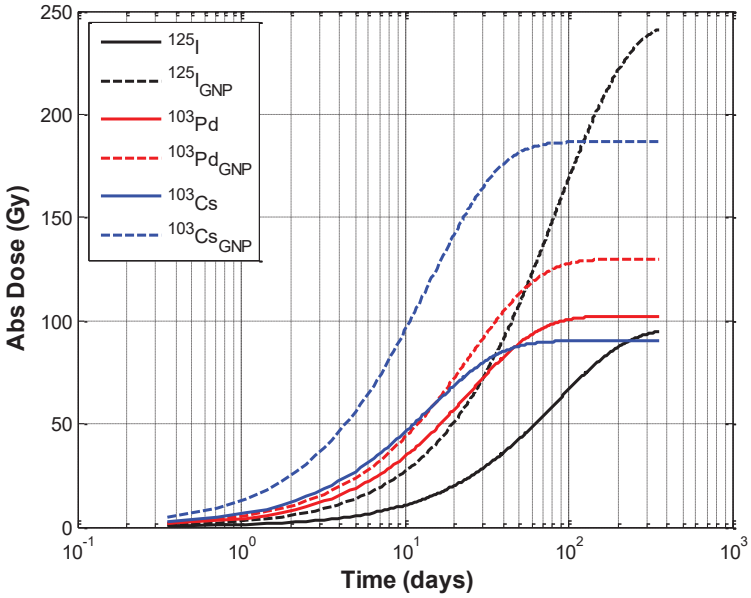



Fig. 4: Absorbed dose (Gy) as a function of time (days) in prostate tissue from 98 seeds each of 0.31 mCi of $^{125}$I, 115 seeds each of 1.4 mCi of $^{103}$Pd and 115 seeds each of 1.4 mCi $^{131}$Cs (replacing $^{103}$Pd seeds).

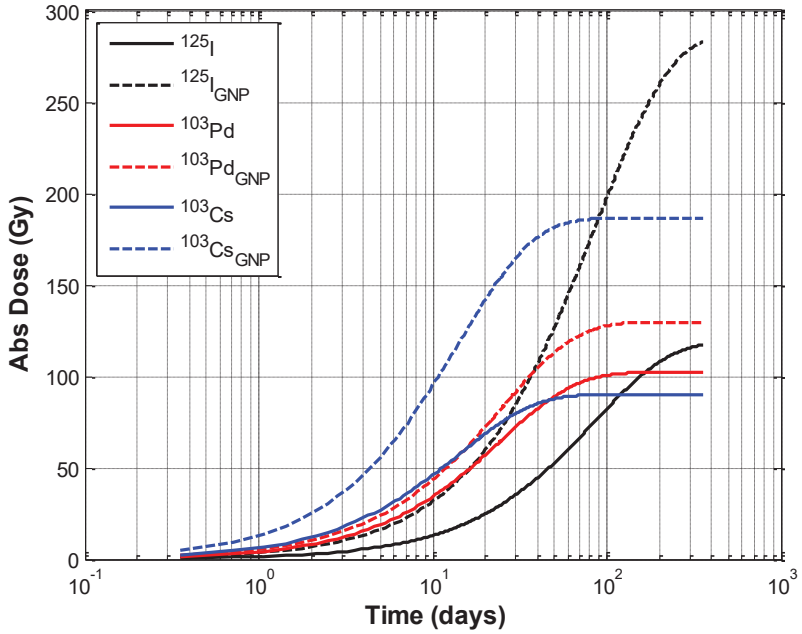



Fig. 5: Absorbed dose (Gy) as a function of time (days) in prostate tissue from $^{125}$I, $^{103}$Pd and $^{131}$Cs (115 seeds each of 1.4 mCi distributed as in Fig. 2)

On a shorter time scale, the differences between all three isotopes is seen from the absorbed dose rates shown in Fig. 6 showing a clear advantage of $^{131}$Cs over the first ~300 hours exceeding 5.7 times and 1.5 times that of $^{125}$I and $^{103}$Pd respectively in the initial period and falling gradually, as shown in Fig. 7. Levelling off is found to be ~300 and ~400 hours respectively for $^{103}$Pd and $^{125}$I.

https://doi.org/10.31436/iiumej.v20i2.1136

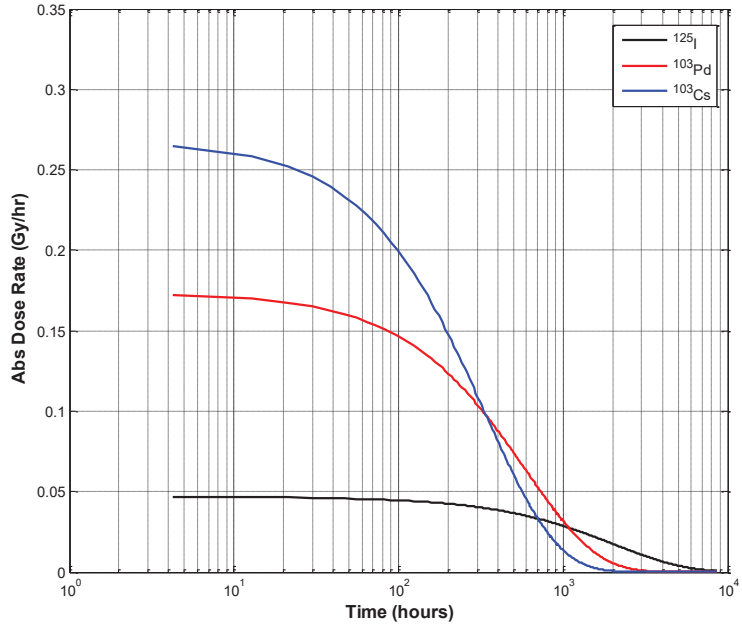



Fig. 6: Absorbed dose rate (Gy/hr) as a function of time (hours) in prostate tissue from 98 seeds each of 0.31 mCi of $^{125}$I, 115 seeds each of 1.4 mCi of $^{103}$Pd and 115 seeds each of 1.4 mCi $^{131}$Cs (replacing $^{103}$Pd seeds).

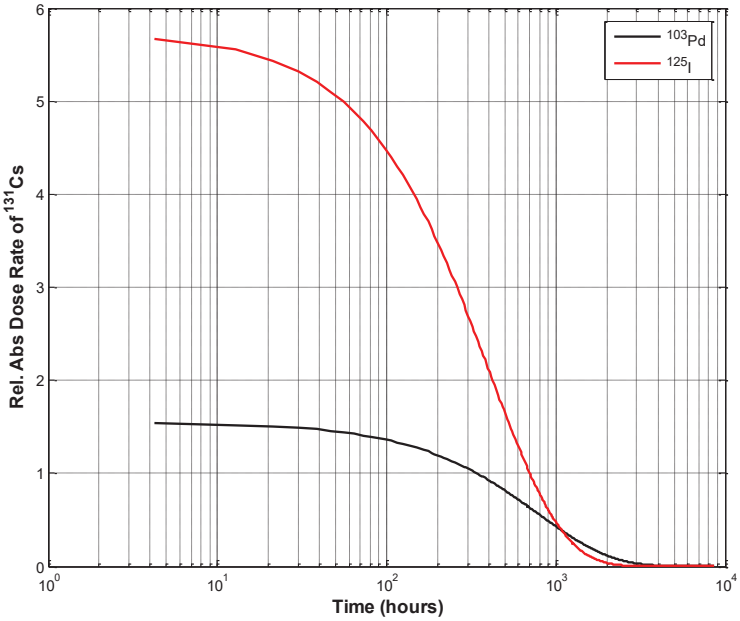



Fig. 7: Relative absorbed dose rate (Gy/hr) of $^{131}$Cs (relative to $^{125}$I and $^{103}$Pd) as a function of time (hours) in prostate tissue from 98 seeds each of 0.31 mCi of $^{125}$I, 115 seeds each of 1.4 mCi of $^{103}$Pd and 115 seeds each of 1.4 mCi $^{131}$Cs (replacing 103Pd seeds).

The lifetime dose delivered by the three sources considered is shown in Table 2. The absorbed dose for $^{125}$I, $^{103}$Pd and $^{131}$Cs increases from 96.30, 102.21 and 90 Gy to ~245, ~130, ~187 Gy with GNP-tissue solution of 25 mg Au/g tissue. These results are in line

with the decay rates of the three isotopes which favour $^{125}$I in terms of energy but result in a slow dose delivery.

The dose enhancement with gold GNP-tissue found in these simulations is the effect of both source energy of each radioisotope and the solution concentration. The important photon interaction for high-Z materials such as gold is the photoelectric effect for which the cross-section varies as $\mu_{PE} \sim \frac{\rho Z^3}{E^3}$ so that low energy and high-Z are desirable for dose enhancement which is localized to the tumor due to the short range of photoelectrons and Auger electrons in the surrounding medium which for electrons of energy 0.1 MeV is ~100 microns in water and ~15 microns in gold so that the effect of gold will require thin layers (of the order of $\ll 100$ nm) to utilize the energy of photoelectrons in water. At this energy, the photoabsorptions were estimated to be ~13.5% of all interactions. Thus, the increases found in this work are a measure of the photon interaction in each radioisotope. As shown in Table 2, the highest dose achievable is ~245 Gy by $^{125}$I, and lesser doses by $^{103}$Pd and $^{131}$Cs for 25 mg Au/g tissue.

When the configuration of Fig. 2 is used for all three sources of 1.4 mCi each, the F6 and F8 tallies for $^{125}$I increase to 1.82733 × $10^{-5}$ (0.0011) MeV/g and 1.24662 × $10^{-3}$ (0.0016) MeV respectively, as depicted in Fig. 5.

Table 2: Lifetime absorbed dose (Gy) to prostate tumour due to $^{125}$I, $^{103}$Pd and $^{131}$Cs. The second result in F6, F8 and Abs. Dose columns refer to Tissue with 25 mg Au/g tissue homogenous mixture

| Source | $T_{1/2}$ (d) | $A_0$ (mCi) | E(kev) $N_{seeds}$ | Tr=$A_0/\lambda$ ($s^{-1}$) | F6 (rel. err.) (MeV/g) | F8 (rel. err.) (MeV) | Abs. Dose* (Gy) |
|---|---|---|---|---|---|---|---|
| $^{125}$I | 59.4 | 0.31 | 35.49 | 8.5 × $10^{13}$ | 1.73548 × $10^{-5}$ (0.0011) | 1.183 × $10^{-3}$ (0.0016) | 96.30 |
| | | | 98 | | 4.30593 × $10^{-5}$ 0.0012 | 3.0087 × $10^{-3}$ 0.0010 | 244.91 |
| $^{103}$Pd | 17 | 1.4 | 20.8 | 1.1 × $10^{14}$ | 1.21460 × $10^{-5}$ (0.0018) | 8.2786 × $10^{-4}$ (0.0015) | 102.21 |
| | | | 115 | | 1.50096 × $10^{-5}$ 0.0019 | 1.0511 × $10^{-3}$ 0.0014 | 129.77 |
| $^{131}$Cs | 9.7 | 1.4 | 30.4 | 6.3 × $10^{13}$ | 1.88145 × $10^{-5}$ (0.0012) | 1.2803 × $10^{-3}$ (0.0015) | 90.17 |
| | | | 115 | | 3.79186 × $10^{-5}$ (0.0013) | 2.64847 × $10^{-3}$ 0.0010 | 186.57 |

*for an organ mass 16.3634 g

## 4. CONCLUSIONS

From the simulations, it was concluded that:

- lifetime absorbed dose is ~96 Gy from 98 seeds, each of 0.31 mCi, of $^{125}$I, ~102 Gy, from 115 seeds, each of 1.4 mCi, of $^{103}$Pd, and ~90 Gy from $^{131}$Cs seeds replacing $^{103}$Pd seeds of the same initial activity,
- there is a saturation in the activities of $^{103}$Pd and $^{131}$Cs with the implications that the effect of these two isotopes is much faster than that of $^{125}$I,
- the main advantage of $^{131}$Cs, is the larger initial dose rate (~26 cGy/hr) for the first few days, which is 1.5 and 5.7 times higher than that for $^{103}$Pd and $^{125}$I,

- with a GNP-tissue mixture, the dose enhancement factors in $^{125}$I, $^{103}$Pd and $^{131}$Cs are 2.5, 1.26 and 2.1 respectively which give another edge to $^{131}$Cs,
- for a specified dose, $^{125}$I exceeds $^{103}$Pd by about 200 days; higher doses for stronger tumours can thus not be treated by $^{131}$Cs and $^{103}$Pd unless they are used in conjunction with gold solutions.

This work has shown results for three candidate radioisotopes in their effectiveness for the treatment of tumours. The basic hypotheses to determine their effectiveness were their radiation energy and half-life. The Monte Carlo simulations carried out in this work have elaborated the effect of radiations from three isotopes and found substantially higher time of treatment, as well as stronger radiation dose with $^{125}$I than for $^{103}$Pd and $^{131}$Cs.

## FUNDING

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

## REFERENCES

[1] Common cancer types, national cancer institute [https://www.cancer.gov/types/common-cancers]
[2] Worldwide cancer data, world cancer research fund. [https://www.wcrf.org/dietandcancer/cancer-trends/worldwide-cancer-data]
[3] Yu E, Lewis C. (2018) Lung cancer brachytherapy. Current Cancer Therapy Reviews, 14(2): 137-148. doi: 10.2174/1573394714666180208145420.
[4] Blanchard P, Pugh TJ, Swanson DA, Mahmood U, Chen HC, Wang X, Graber WJ, Kudchadker RJ, Bruno T, Feeley T, Frank SJ. (2018) Patient-reported health-related quality of life for men treated with low-dose-rate prostate brachytherapy as monotherapy with 125-Iodine, 103-Palladium, or 131-Cesium: Results of a prospective phase ii study. Brachytherapy, 17( 2): 265-276. doi: 10.1016/J.BRACHY.2017.11.007.
[5] Rice SR, Olexa G, Hussain A, Mannuel H, Naslund MJ, Amin P, Kwok Y. (2019) A phase ii study evaluating bone marrow-sparing, image-guided pelvic intensity-modulated radiotherapy (IMRT) with Cesium-131 brachytherapy boost, adjuvant chemotherapy, and long-term hormonal ablation in patients with high risk, nonmetastatic prostate cancer. American Journal of Clinical Oncology, 42(3): 285-29. doi: 10.1097/COC.0000000000000520.
[6] Fahmi MR, Hashikin NA, Yeong CH, Guatelli S, Ng KH, Malaroda A, Rosenfeld AB, Perkins AC. (2019) Evaluation of organ doses following prostate treatment with permanent brachytherapy seeds: A geant4 Monte Carlo simulation study. Journal of Physics: Conference Series, 1248(1): 012049-012049. doi: 10.1088/1742-6596/1248/1/012049.
[7] Radioisotopes in medicine, world nuclear association [https://www.world-nuclear.org/information-library/non-power-nuclear-applications/radioisotopes-research/radioisotopes-in-medicine.aspx]
[8] Park DS. (2012) Current status of brachytherapy for prostate cancer. Korean Journal of Urology, 53(11): 743-749. doi: 10.4111/kju.2012.53.11.743.
[9] Lechtman E, Mashouf S, Chattopadhyay N, Keller BM, Lai P, Cai Z, Reilly RM, Pignol JP. (2013) A Monte Carlo-based model of gold nanoparticle radiosensitization accounting for increased radiobiological effectiveness. Physics in Medicine and Biology, 58(10): 3075-3087. doi: 10.1088/0031-9155/58/10/3075.
[10] Stish BJ, Davis BJ, Mynderse LA, McLaren RH, Deufel CL, Choo R. (2018) Low dose rate prostate brachytherapy. Translational Andrology and Urology, 7(3): 341-356. doi: 10.21037/tau.2017.12.15.
[11] Allison J, Amako K, Apostolakis J, Araujo H, Arce Dubois P, Asai M, Barrand G, Capra R, Chauvie S, Chytracek R, Cirrone GAP, Cooperman G, Cosmo G, Cuttone G, Daquino GG, Donszelmann M, Dressel M, Folger G, Foppiano F, Generowicz J, Grichine V, Guatelli S,

Gumplinger P, Heikkinen A, Hrivnacova I, Howard A, Incerti S, Ivanchenko V, Johnson T, Jones F, Koi T, Kokoulin R, Kossov M, Kurashige H, Lara V, Larsson S, Lei F, Link O, Longo F, Maire M, Mantero A, Mascialino B, McLaren I, Mendez Lorenzo P, Minamimoto K, Murakami K, Nieminen P, Pandola L, Parlati S, Peralta L, Perl J, Pfeiffer A, Pia MG, Ribon A, Rodrigues P, Russo G, Sadilov S, Santin G, Sasaki T, Smith D, Starkov N, Tanaka S, Tcherniaev E, Tome B, Trindade A, Truscott P, Urban L, Verderi M, Walkden A, Wellisch JP, Williams DC, Wright D, Yoshida H. (2006) Geant4 developments and applications. IEEE Transactions on Nuclear Science, 53(1): 270-278. doi: 10.1109/TNS.2006.869826.

[12] Salvat F, José M, Sempau J. (2011) Penelope 2011: a code system for monte carlo simulation of electron and photon transport. Nuclear Energy Agency, doc(2011)5. https://www.oecd-nea.org/science/docs/2011/nsc-doc2011-5.pdf

[13] Briesmeister JF. (2000) Mcnp: a general monte carlo n-particle transport code. version 4c, LA-13709-M. https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-13709-M

[14] Kalos MH, Whitlock PA. (2008) Monte Carlo methods. Wiley-Blackwell, 2008, pp. 203-203.

[15] Jain S, Hirst DG, O'Sullivan JM. (2012) Gold nanoparticles as novel agents for cancer therapy. The British Journal of Radiology, 85(1010): 101-113. doi: 10.1259/bjr/59448833.

[16] Chatterjee DK, Wolfe T, Lee J, Brown AP, Singh PK, Bhattarai SR, Diagaradjane P, Krishnan S. (2013) Convergence of nanotechnology with radiation therapy-insights and implications for clinical translation. Translational Cancer Research, 2(4): 256-268. doi: 10.3978/j.issn.2218-676X.2013.08.10.

[17] Lechtman E, Chattopadhyay N, Cai Z, Mashouf S, Reilly R, Pignol JP. (2011) Implications on clinical scenario of gold nanoparticle radiosensitization in regards to photon energy, nanoparticle size, concentration and location. Physics in Medicine and Biology, 56(15):4631-4647. doi: 10.1088/0031-9155/56/15/001.

[18] Mesbahi A, Jamali F, Garehaghaji N. (2013) Effect of photon beam energy, gold nanoparticle size and concentration on the dose enhancement in radiation therapy," BioImpacts : BI, 3(1): 29-35. doi: 10.5681/bi.2013.002.

[19] Asadi S, Vaez-zadeh M, Masoudi SF, Rahmani F, Knaup C, Meigooni AS. (2015) Gold nanoparticle-based brachytherapy enhancement in choroidal melanoma using a full monte carlo model of the human eye. Journal of Applied Clinical Medical Physics, 16(5): 344-357. doi: 10.1120/jacmp.v16i5.5568.

[20] Banoqitah E, Djouider F. (2016) Dose distribution and dose enhancement by using gadolinium nanoparticles implant in brain tumor in stereotactic brachytherapy. doi: 10.1016/j.radphyschem.2016.06.002.

[21] Kehwar TS. (2009) Use of cesium-131 radioactive seeds in prostate permanent implants. Journal of Medical Physics, 34(4): 191-193. doi: 10.4103/0971-6203.56077.

[22] Yu Y, Anderson LL, Li Z, Mellenberg DE, Nath R, Schell MC, Waterman FM, Wu A, Blasko JC. (1999) Permanent prostate seed implant brachytherapy: Report of the american association of physicists in medicine task group no. 64. Medical Physics, 26(10): 2054-2076. doi: 10.1118/1.598721.

[23] Awan SB, Hussain M, Dini SA, Meigooni AS. (2008) Historical review of interstitial prostate brachytherapy," vol. 5. http://ijrr.com/article-1-345-en.pdf

[24] Yaparpalvi R. *et al.* (2007) Is Cs-131 or I-125 or Pd-103 the “ideal” isotope for prostate boost brachytherapy?– A dosimetric view point. International Journal of Radiation Oncology*Biology*Physics, 69(3): S677-S678. doi: 10.1016/J.IJROBP.2007.07.2038.

[25] Armpilia CI, Dale RG, Coles IP, Jones B, Antipas V. (2003) The determination of radiobiologically optimized half-lives for radionuclides used in permanent brachytherapy implants. International Journal of Radiation Oncology, Biology, Physics, 55(2): 378-385. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/12527051.

[26] Bice WS, Prestidge BR, Kurtzman SM, Beriwal S, Moran BJ, Patel RR, Rivard MJ, Cesium Advisory G. (2008) Recommendations for permanent prostate brachytherapy with 131cs: A

consensus report from the cesium advisory group. Brachytherapy, 7(4): 290-296. doi: 10.1016/j.brachy.2008.05.004.

[27] Wooten CE, Randall M, Edwards J, Aryal P, Luo W, Feddock J. (2014) Implementation and early clinical results utilizing Cs-131 permanent interstitial implants for gynecologic malignancies. Gynecologic Oncology, 133(2): 268-273. doi: 10.1016/j.ygyno.2014.02.015.

[28] Wernicke AG, Smith AW, Taube S, Yondorf MZ, Parashar B, Trichter S, Nedialkova L, Sabbas A, Christos P, Ramakrishna R, Pannullo SC, Stieg PE, Schwartz TH. (2016) Cesium-131 brachytherapy for recurrent brain metastases: Durable salvage treatment for previously irradiated metastatic disease. Journal of Neurosurgery, 126(4): 1212-1219. doi: 10.3171/2016.3.jns152836.

[29] Usgaonker SR. (2004) MCNP modeling of prostate brachytherapy and organ dosimetry. [Online]. Available: https://oaktrust.library.tamu.edu/handle/1969.1/305.

[30] Almansa JF, Guerrero R, Al-Dweri FMO, Anguiano M, Lallena AM. (2006) Dose distribution in water for monoenergetic photon point sources in the energy range of interest in brachytherapy: Monte carlo simulations with penelope and geant4. [Online]. Available: http://fm137.ugr.es/PhotonPointSources/

[31] Archambault JP, Mainegra-Hing E. (2015) Comparison between egsnrc, geant4, mcnp5 and penelope for mono-energetic electron beams. Physics in Medicine and Biology, 60(13): 4951-4962. doi: 10.1088/0031-9155/60/13/4951.

[32] Koivunoro H, Siiskonen T, Kotiluoto P, Auterinen I, Hippeläinen E, Savolainen S. (2012) "Accuracy of the electron transport in MCNP5 and its suitability for ionization chamber response simulations: A comparison with the egsnrc and penelope codes. Medical Physics, 39(3): 1335-1344. doi: 10.1118/1.3685446.

[33] Šídlová V, Trojek T. (2010) Testing Monte Carlo computer codes for simulations of electron transport in matter. Applied Radiation and Isotopes, 68(4-5): 961-964. doi: 10.1016/j.apradiso.2009.12.019.

https://doi.org/10.31436/iiumej.v20i2.1066

# DYNAMIC MECHANICAL AND WATER ABSORPTION PROPERTIES OF MICROCRYSTALLINE CELLULOSE REINFORCED POLYPROPYLENE COMPOSITES: THE EFFECT OF UNCONVENTIONAL IRRADIATION ROUTE

**NOORASIKIN SAMAT*, NURUL HAKIMAH MOHD LAZIM, ZAHURIN HALIM**

*Department of Manufacturing and Materials, Faculty of Engineering,*
*International Islamic University Malaysia,*
*P.O. Box 10, 50728 Kuala Lumpur, Malaysia.*

**Corresponding author: noorasikin@iium.edu.my*

*(Received: 25th January 2019; Accepted: 5th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** The unconventional electron beam (EB) irradiation route in preparing microcrystalline cellulose (MCC) fiber reinforced recycled polypropylene (rPP) composites was studied. In this route, the rPP was first subjected to EB irradiation at various doses (10-50kGy) and was then used as a compatibilizer. Unirradiated and irradiated rPPs were blended at two different ratios (90:10; 50:50) and added with MCC at contents of 5, 20 and 40wt%. Dynamic mechanical analysis (DMA) and water absorption tests were carried out. The DMA spectra exhibited high stiffness and damping behaviour. As the content of MCC increased, the water resistance of composites dropped slightly as compared to the controlled rPP. However, some compositions (50:50/40MCC-10kGy and 50:50/5MCC-50kGy) had shown opposite results. The improvement in the studied properties proved the existence of the compatibility effect that occurred at low irradiation doses, and also depended on the ratio (unirradiated and irradiated rPP) and MCC contents.

***ABSTRAK:*** Kaedah sinaran gelombang elektron secara bukan konvensional dalam penyediaan komposit polipropilena kitar semula (rPP) bersama serat selulosa mikrohablur (MCC) telah dikaji. Dalam kaedah ini, rPP telah didedahkan kepada sinaran gelombang elektron dengan dos yang berbeza (10-50kGy) dan kemudiannya digunakan sebagai penserasi. rPP tidak tersinar dan tersinar telah dicampur dengan nisbah (90:10; 50:50) dan ditambah dengan MCC pada kandungan 5, 20 and 40wt%. Analisis dinamik mekanikal (DMA) dan ujian penyerapan air telah dijalankan. Spektrum DMA menunjukkan sifat kekakuan dan pengenduran yang tinggi. Apabila kandungan MCC bertambah, kerintangan komposit terhadap air berkurang sedikit berbanding rPP terkawal. Bagaimanapun sebahagian komposisi (50:50/40MCC-10kGy and 50:50/5MCC-50kGy) telah menunjukkan keputusan sebaliknya. Penambahbaikan sifat bahan dalam kajian ini membuktikan kewujudan kesan keserasian yang berlaku pada dos sinaran rendah, dan juga bergantung kepada nisbah (rPP tidak tersinar dan tersinar) dan kandungan MCC.

***KEYWORDS:*** *polypropylene; cellulose; compatibilizer; electron beam irradiation; mechanical properties*

## 1. INTRODUCTION

The changing of chemical, structural, and other physical properties of polymers through the application of ionizing radiation (IR) is not a new technique. Of all the

existing irradiation methods, the most common modes of IR used are the gamma irradiation and electron beam (EB) processing. There are a few advantages to the use of EB irradiation. It is fast, simple, can be carried out at room temperature, and the technology is completely environmentally friendly. During the irradiation process, the energy is transferred from the irradiating source to the atoms of irradiated materials. This will result in a few chemical reactions taking place, such as chain scission, branching, and crosslinking [1-2] that subsequently alter the mechanical and thermal properties of the materials [2-7]. Generally, the formation of crosslinking will lead to improvements of properties. Many industries, such as those producing heat shrink tubes, crosslinked wires and cables, packing films, and foams have adopted the EB radiation technique in their product development segment [8-9]. EB processing is also used for sterilization of medical devices, drug delivery systems, preservation of food, and surface curing [10-11]. Due to its wide usage, numerous published works have employed EB techniques in the study of polymer composites.

In the conventional route for polymer composites, the IR is usually applied after the composites have been fabricated [12-14]. However, a new irradiation route for preparing reinforced polymer composites has been reported by Karsli et al. and Gomes et al., where the polypropylene (PP) that was exposed to gamma and EB irradiation respectively, was compounded with short carbon and piassava fibers [15-16]. Tensile tests were performed in both cases to evaluate the function of irradiated PPs as a compatibilizer. This unconventional irradiation route had also been further explored in our previous works [2-5] in recycled polypropylene (rPP) that was reinforced with microcrystalline cellulose (MCC). Here, the possibility of irradiated recycled PP being a compatibilizer was assessed based on the tensile, impact strength and thermal properties. In all of these cases [2-5, 15-16], the properties of composites had been better than the controlled polymer, thereby proving the compatibility effect of the virgin PP or recycled PP through the use of IR.

Apart from using coupling agents to improve the compatibility between natural fiber and polymer [17-19], the use of EB via the unconventional irradiation route would offer an alternative solution in overcoming this drawback. The major advantage of this route is that the modification was carried out on the polymer matrix itself without any chemical modifier or coupling agent. Therefore, the present work aims to further investigate the effects of irradiated recycled PP on dynamic mechanical analysis (DMA) and water absorption properties from the developed composites. The results derived from this analysis is a supplementary research on an earlier study conducted by Lazim et al. [2-3] and Samat et al. [4].

## 2. MATERIALS AND METHOD

### 2.1 Materials

Recycled polypropylene (rPP) from industrial waste was purchased from Top Flow Sdn. Bhd., Shah Alam, Selangor. The reinforcement material was microcrystalline cellulose (MCC) in the form of a fine white powder with a density of 0.6 g/cm$^3$, which was obtained from Sigma-Aldrich Co. (M) Sdn. Bhd. The polymer and the MCC had been dried in an oven at 40-60 °C for at least 7 hours before use.

### 2.2 Preparation of Composites

The rPP pellets were irradiated using a 3MeV electron beam (EB) using the EB model EPS-3000 at room temperature under oxygenated conditions. Irradiation doses were varied

at 5, 10, and 50 kGy, respectively. The unirradiated and irradiated rPP pellets were blended at corresponding ratios of 90:10 wt% and 50:50 wt% according to weight before being added with the various MCC loadings (5, 20, and 40 wt%) using a Brabender twin screw extruder machine at a screw speed of 70-90 rpm. The compounded materials were pelletized and fed into Battenfield HM 600/850 injection moulding to form specimens according to the testing standards (ASTM: D638).

### 2.3 Dynamic Mechanical Analysis (DMA)

Dynamic Mechanical Analysis was conducted over a wide range of temperatures as a means of studying the viscoelasticity behaviour of the materials. DMA was conducted using a Perkin Elmer Pyris Diamond DMTA machine using ASTM D7028. The testing was carried out in three-point bending mode at a frequency of 1 Hz. The temperature range was -50 °C to 150 °C with heating rate at 5 °C/min. The storage modulus (E') and damping factor (tan δ) of the samples were measured as a function of temperature.

### 2.4 Water Absorption

All of the samples had been immersed in distilled water for seven days in a closed container at room temperature. The samples were wiped with tissue and weights were recorded for the initial two 6 hours, which was followed by the next 12 hours and subsequently 24 hours with the whole process repeated until day seven. The percentage of increase in water in relative to the initial weight of each sample was reported as the amount of water absorbed. Six samples of replicated specimens for each composition were used to obtain the average value. The water absorption percentage after being soaked ($M_t$) was calculated according to Eq. (1) (ASTM D570-98):

$$M_t(\%) = \frac{(W_f - W_i)}{W_i} x100 \quad (1)$$

where $W_i$ = initial and $W_f$ = final weight, respectively.

### 2.5 Gel Content Analysis

The result of gel content for neat rPP and rPP/MCC composite samples has been reported in our previous work [2]. The gel content of was determined by the extraction of samples in xylene solvent at 130$^{O}$C for 24 h (ASTM D2765). These samples were dried and weighed and the gel fraction was calculated as:

$$Gel\ fraction\ (\%) = \frac{W}{W_o} x100 \quad (2)$$

where $W_0$ = initial and W = final weight , respectively.

## 3. RESULTS AND DISCUSSION

### 3.1 Dynamic Mechanical Analysis (DMA)

Figure 1 shows the effect of rPP:i-rPP ratio, MCC content and the irradiation dosage to the storage modulus (E') of the rPP reinforced with 40 wt% MCC fibers. It was observed that in the glassy region, the E' of the neat rPP and rPP:i-rPP/MCC composites had remained fairly constant before gradually decreasing across the measured temperature ranges. No rubbery plateau was noticed in these DMA curves. Upon further analysis, it was observed that there had been a greater increase of the E' of the rPP:i-rPP/MCC composites as compared to the neat rPP. This implies that the addition of MCC fibers had enhanced the stiffness by restricting the flow of rPP matrix, which is in line with the results obtained from Young's modulus [3-4]. The storage moduli at specific

https://doi.org/10.31436/iiumej.v20i2.1066

temperatures; -50 °C (glassy region), +27 °C (room temperature) and +100 °C (possible use temperature) were selected as a means of comparing the effect of rPP:i-rPP ratio and irradiation dose. The results are presented in Table 1.

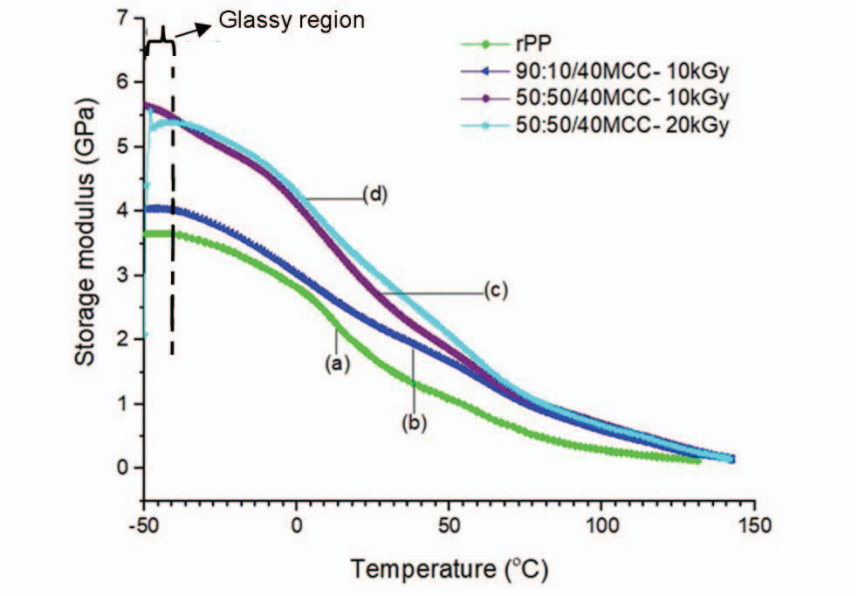



Fig. 1: The storage moduli of (a) neat rPP, (b) 90:10/40MCC-10kGy, (c) 50:50/40MCC-10kGy, and (d) 50:50/40MCC-20kGy.

Table 1: Storage moduli and $T_g$ values of rPP and rPP: i-rPP/MCC composites

| Samples | E' at -50°C (MPa) | E' at +27°C (MPa) | E' at +100°C (MPa) |
|---|---|---|---|
| rPP | 3000 | 1620 | 294 |
| 90:10/40MCC- 10kGy | 4033 | 2184 | 585 |
| 50:50/40MCC-10kGy | 5662 | 2658 | 691 |
| 50:50/40MCC-20kGy | 4310 | 2986 | 658 |

At -50 °C, the increase of the storage moduli of composites with 10 kGy i-rPP at 50:50 ratio was larger than that of 90:10. Moreover, the difference in the storage moduli values of sample 50:50 at 20 kGy is an indication the benefit derived from i-rPP as a second contributor to the enhancement of stiffness. Previously, it was found that at a ratio of 50:50, the increase in the stiffness of composites was dominated by the presence of higher MCC content and that the presence of i-rPP promoted a 'compatibilizing effect' [4]. From the FTIR [2] and morphological analysis [2-4], this effect is attributed through the formation of crosslinking and better filler-matrix interfacial adhesion, respectively. The existence of crosslinking in the 50:50/40MCC-10kGy sample was evident in its FTIR spectra through the absence of the absorption peak assigned to a carbonyl group (C=O) [2]. According to Lazim et al., the disappearance of this absorption peak shows the suppression of the oxidation process associated with the presence of crosslinking structure [2]. In contrast, for the 90:10/40MCC-10kGy sample, the absorption peak of the carbonyl group was pronounced. Therefore, it is likely, a higher i-rPP content had formed harder domains in the rPP through the formation of greater crosslinking, thereby increasing the

E'. The influence of greater crosslinking to the increase of storage modulus of composites was also reported by Banik et al. and Pang et al. [20-21].

It is worth noting that the improvement in the storage modulus had occurred at a low irradiation dose (10 kGy). Earlier on, it was discovered that that higher dosage had led to chain scission; which also reduced the level of crosslinking [2]. From Table 1, when the temperature was increased from -50 to +100 ºC, it was evident that there was a reduction in the values of E', indicating a shift to the rubbery phase. The changing of phase was due to the loosening of chains between the rPP and MCC fibers that led to the increased flexibility of the composite. It was noticed that the E' value of rPP had dropped considerably compared to those of the composite samples.

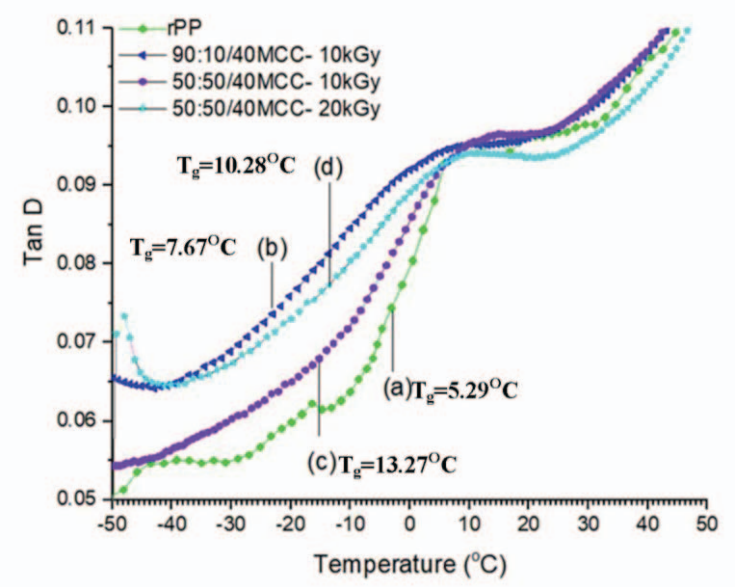



Fig. 2: The tan D of (a) neat rPP, (b) 90:10/40MCC-10kGy, (c) 50:50/40MCC-10kGy, and (d) 50:50/40MCC-20kGy.

Results of the damping behaviour of rPP composites along with their glass transition temperature ($T_g$) values are presented in Figure 2. As seen from the illustration, the neat rPP has the lowest $T_g$ value, which is 5.29 °C. The deviation relative to rPP was evident for the composite samples, where the $T_g$ had shifted to higher values. This finding had also validated the existence of crosslinking in the rPP composites. According to Wongsuban et al., cross-linked structure inhibits the mobility of the polymer chains [22]. Consequently, a higher temperature is required to trigger the change of phases (glassy to rubbery) in the rPP matrix, which was demonstrated by the marked increment in the sample of rPP:i-rPP/40MCC-10kGy ($T_g$=13.27 °C). In fact, this enhancement had almost doubled as compared to rPP ($T_g$=5.29 °C), suggesting that the sample had the most prominent compatibilizing effect. Interestingly, unlike the other natural fiber reinforced polymer system; the intensity of the tan delta peak had declined with the increase of filler content [19, 23-24]. Nevertheless, in our case, the peak curve of tan delta of sample rPP:i-rPP/40MCC-10kGy was observed to be higher than the rPP, which suggests a better damping behaviour. Indeed, the similar sample (50:50/40MCC-10kGy) had also demonstrated substantial improvement of the impact strength as reported in our earlier work [2].

https://doi.org/10.31436/iiumej.v20i2.1066

### 3.2 Water Absorption

Figure 3 depicts the water absorption results for the neat rPP and composite samples. In general, the composites had exhibited a higher water uptake percentage as compared to the neat rPP at all immersion time and composition ranges. This was due to the addition of hydrophilic MCC fibers into the polymer matrices, which had increased the ability of rPP composites to absorb more water. During the initial immersion period (<12hr), the difference in the water absorption percentages for all samples had been negligible. As observed in Figure 3, the water uptake of composites had increased steadily after the immersion period was extended beyond 24 hours. This result was expected due to the increased interaction between the hydrogen molecule from water and the -OH molecule from the MCC fibers. Nevertheless, the overall water uptake of the composites is considered low since the highest percentage shown was less than 5%. Therefore, it is suggested that i-rPP promotes the hydrophobicity of the rPP/MCC composites through the formation of better interfacial adhesion.

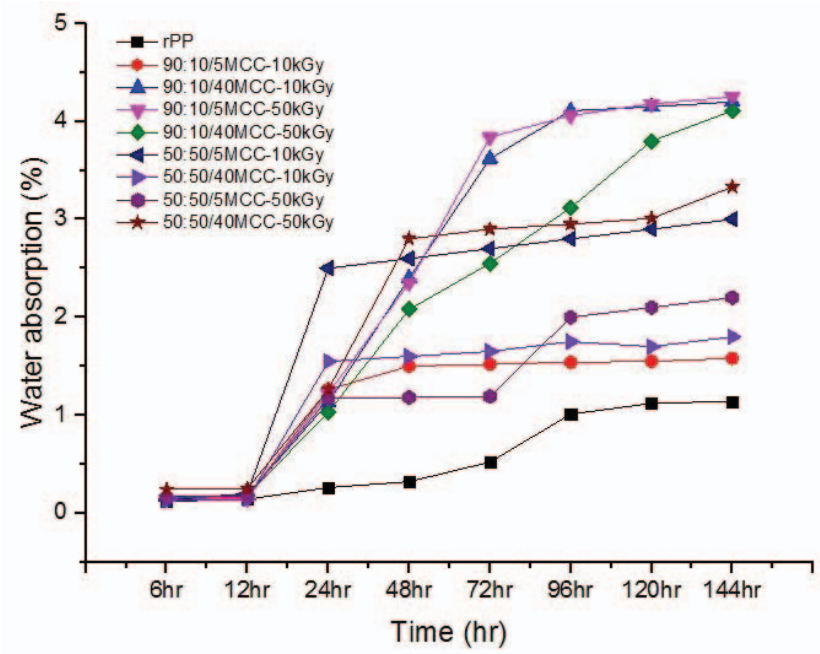



Fig. 3: Water absorption percentage of rPP and rPP/MCC composites over a period of seven days.

As discussed earlier, the varied irradiation doses along with the MCC content caused different crosslinking densities, which subsequently affected the filler-matrix interfacial adhesion. To gain a better perspective on this circumstance, a comparison between the gel content (Table 2) and water uptake results (Fig. 3) was made. Apparently, the composites that possessed higher gel content value had shown a better water resistance than those with lower gel content values. Hence, it is suggested that higher crosslinking densities decrease the diffusion rate of water molecules in the composite, which subsequently leads to lower water uptake. A lower crosslinking density, on the other hand, would only ease the diffusion of water molecules mainly in the amorphous parts [25]. Furthermore, the decrease in the water absorption in this work is similar to the case in natural fiber/polymer composites treated with coupling agent [26-27]. This observation indicates that the introduction of i-rPP was beneficial to improve resistance against water absorption.

Table 2: Water absorption and gel content percentage of rPP/MCC composites

| Ratio | Sample | 12hr | 24hr | 96hr | 144hr | Gel content (%) [2] |
|---|---|---|---|---|---|---|
| 90:10 | 5MCC-10kGy | 0.523 | 1.178 | 1.557 | 1.662 | - |
| | 40MCC-10kGy | 0.127 | 1.184 | 4.083 | 4.271 | 13.16 |
| | 5MCC-50kGy | 0.445 | 1.138 | 3.992 | 4.121 | - |
| | 40MCC-50kGy | 0.118 | 1.021 | 3.036 | 4.057 | 20.18 |
| 50:50 | 5MCC-10kGy | 0.249 | 2.578 | 2.839 | 3.101 | 24.41 |
| | 40MCC-10kGy | 0.353 | 1.531 | 1.675 | 1.911 | 61.43 |
| | 5MCC-50kGy | 0.170 | 1.138 | 2.067 | 2.277 | - |
| | 40MCC-50kGy | 0.379 | 1.413 | 2.996 | 3.349 | 4.65 |

## 4. CONCLUSION

From this investigation, the irregular irradiation route has offered a promising technique on inducing the compatibility behaviour that enhances the recycled PP composites' properties. The enhancement in stiffness (storage modulus) of composites and the damping property had depended on higher levels of MCC content, rPP:i-rPP ratio and crosslinking value. A greater crosslinking density also slows down the absorption of water into the composites. Sample 50:50/40MCC-10kGy exhibits tremendous properties compared to other compositions.

## ACKNOWLEDGEMENT

The authors would like to thank the International Islamic University Malaysia (IIUM) and Ministry of Higher Education (MOHE), Malaysia (FRGS14-163-0404) for the support of this research.

## REFERENCES

[1] Miranda LF, Pereira NC, Faldini SB, Masson TJ, Silva LGA, Silveira LH. (2009) Effect of ionizing radiation on polypropylene composites reinforced with coconut fibers. In Proceedings of the International Nuclear Atlantic Conference:27 September – 2 October 2009; Rio de Janeiro, Brazil; pp 1-9

[2] Lazim NH, Samat N. (2019) The influence of irradiated recycled polypropylene compatibilizer on the impact fracture behavior of recycled polypropylene/microcrystalline cellulose composites. Polymer Composites, 40(S1): E24-E34.

[3] Lazim NH, Samat N. (2017) Effects of irradiated recycled polypropylene compatibilizer on the mechanical properties of microcrystalline cellulose reinforced recycled polypropylene composites. Procedia Engineering, 184:538–543.

[4] Samat N, Lazim NH, Motsidi SNR, Azlina HN. (2017) Performance properties of irradiated recycled polypropylene as a compatibilizer in recycled polypropylene/microcrystalline cellulose composites. Materials Science Forum, 894:62-65.

[5] Samat N, Motsidi SNR, Lazim NH. (2018) Effects of electron beam radiation dose on the compatibilization behaviour in recycled polypropylene/microcrystalline cellulose composites, IOP Conference Series: Mater. Sci. Eng.,290: 012034.

[6] Moldovan A, Patachia S, Vasile C, Darie R, Manaila E, Tierean M. (2013) Natural fibres/polyolefins composites (I) UV and electron beam irradiation. J. of Biobased Materials and Bioenergy, 7(1):58-79.

[7] Rimdusit S, Wongsongyot S, Jittarom S, Suwanmala P, Tiptipakorn S. (2011) Effects of gamma irradiation with and without compatibilizer on the mechanical properties of polypropylene/wood flour composites. J. of Polymer Research, 18(4):801-809.

[8] Cleland M, Parks L, Cheng, S. (2003) Applications for radiation processing of materials. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 208:66-73.
[9] Tamada, M. (2018) Radiation processing of polymers and its applications. In Radiation Applications, Volume 7. Edited by Kudo H. Singapore, Springer; pp 63-80.
[10] Cai X, Blanchet P. (2015) Electron-beam curing of acrylate/nanoparticle impregnated wood products. BioResources, 10(3): 3852-3864.
[11] Porto KMB, Napolitano CM, Borrely SI. (2018) Gamma radiation effects in packaging for sterilization of health products and their constituents paper and plastic film. Radiation Physics and Chemistry, 142:23-28.
[12] Kim SW, Park K, Lee SH, Kang JS, Kang KH. (2007) Dependence of mechanical and thermal properties of thermoplastic composites on electron beam irradiation. Int. J. Thermophysics, 28:1067-1073.
[13] Moldovan A, Patachia S, Vasile C, Darie R, Manaila E, Tierean, M. (2013) Natural fibres/polyolefins composites (I) UV and electron beam irradiation. J. of Biobased Materials and Bioenergy, 7(1): 58-79.
[14] Yan J, Kim B, Jeong YG. (2015) Thermomechanical and electrical properties of PDMS/MWCNT composite films crosslinked by electron beam irradiation. J. of Materials Science, 50(16):5599-5608.
[15] Karsli GN, Aytac A, Akbulut M, Deniz V, Güven O. (2013) Effects of irradiated polypropylene compatibilizer on the properties of short carbon fiber reinforced polypropylene composites. Radiation Physics and Chemistry, 84:74-78.
[16] Gomes MG, Ferreira MS, Oliveira RR, Silva VA, Teixeira JG, Moura EA. (2013) Influence of electron beam irradiation on pp/piassava fiber composite prepared by melt extrusion process. In Proceedings of the International Nuclear Atlantic Conference: 24-29 Novevember 2013; Recife, PE, Brazil;pp 1-10.
[17] Awanis J, Anis Sofia S, Samat N. (2012) Effect of coupling agent on mechanical properties of composite from microcrystalline cellulose and recycled polypropylene. Advanced Materials Research, 576:390-393.
[18] Samat N, Marini CD, Maritho MA, Sabaruddin FA. (2013) Tensile and impact properties of polypropylene/microcrystalline cellulose treated with different coupling agents. Composite Interfaces, 20(7): 497-506.
[19] Zulkifli N, Samat N, Anuar H, Zainuddin N. (2015) Mechanical properties and failure modes of recycled polypropylene/microcrystalline cellulose composites. Materials & Design, 69:114-123.
[20] Banik I, Bhowmick AK. (2000) Effect of electron beam irradiation on the properties of crosslinked rubbers. Radiation Physics and Chemistry, 58(3):293-298.
[21] Pang Y, Cho D, Han SO, Park WH. (2005) Interfacial shear strength and thermal properties of electron beam-treated henequen fibers reinforced unsaturated polyester composites. Macromolecular Research, 13(5):453-459.
[22] Wongsuban B, Muhammad K, Ghazali Z, Hashim K, Ali Hassan M. (2003) The effect of electron beam irradiation on preparation of sago starch/polyvinyl alcohol foams. Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms, 211(2):244-250.
[23] Essabir H, Achaby ME, Hilali EM, Bouhfid R, Qaissbou A. (2015) Morphological, structural, thermal and tensile properties of high density polyethylene composites reinforced with treated argan nut shell particles. J. of Bionic Engineering, 12:129-141.
[24] Md. Salleh F, Hassan A, Yahya R, Azzahari AD. (2014) Effects of extrusion temperature on the rheological, dynamic mechanical and tensile properties of kenaf fiber/HDPE composites. Composites: Part B, 58; 259-266.
[25] Samat N, Burford R, Whittle A, Hoffman M. (2010) The effects of water and frequency on fatigue crack growth rate in modified and unmodified polyvinyl chloride. Polymer Engineering & Science, 50(2):352-364.
[26] Samat N, Mokhtar WNA, Zulkifli NI. (2014) Properties of eco-friendly composite from recycled polypropylene. Advances in Environmental Biology, 8(8):2760-2766.

[27] Pérez-Fonseca AA, Arrelano M, Rodrigue D, González-Núñez R, Robledo-Ortiz J. (2016) Effect of coupling agent content and water absorption on the mechanical properties of coiragava fibers reinforced polyethylene hybrid composites. Polymer Composites, 37(10): 3015-3024.

# EXPERIMENTAL AND FINITE ELEMENT ANALYSIS OF SOLVENT CAST POLY(LACTIC ACID) THIN FILM BLENDS

**SHARIFAH IMIHEZRI SYED SHAHARUDDIN*, ABDUL RAHMAN MUKHTAR, NUR ATIQAH MOHD AKHIR, NORHASHIMAH SHAFFIAR, MAIZATULNISA OTHMAN**

*Department of Manufacturing and Materials Engineering,*
*Kulliyyah of Engineering, International Islamic University Malaysia,*
*Kuala Lumpur, Malaysia.*

**Corresponding author: shaimihezri@iium.edu.my*

*(Received: 21th March 2019; Accepted: 14th July 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** A combined experimental and finite element analysis (FEA) investigation was performed to study the effect of incorporating poly(propylene carbonate)(PPC) and curcumin on the mechanical properties of poly(lactic acid) (PLA). In addition, the chemical interaction and morphological changes brought upon by each subsequent additive were also observed. The addition of PPC at 30 wt% into PLA causes a decrease in strength and modulus by 51% and 68% respectively whilst inducing higher elongation by 74%. The resultant decrease in strength and modulus of the PLA/PPC blend was recovered by adding a low weight percentage (1 wt%) of curcumin. The images of the fractured surfaces via scanning electron microscope (SEM) revealed the brittle-ductile-brittle progression of PLA due to the addition of PPC and curcumin which corroborates the findings from the tensile test. Fourier-transform infrared spectroscopy (FTIR) revealed that the addition of PPC by 30 wt % resulted in chemical interaction between the carbonyl groups of PLA and PPC as the C=O peak of PLA slightly shifted to a lower wavenumber. The presence of curcumin peaks however was found to be difficult to be identified in the PLA/PPC/curcumin blend. The simulated results for the stress-strain profile using FEA agreed well with the experimental tensile test profile with a relatively low percentage error of less than 6%. Therefore, it was concluded that for these compositions, the developed model can be used for further simulation works to design biomedical devices.

***ABSTRAK***: Gabungan penyelidikan secara eksperimen dan analisis unsur terhad (FEA) telah dijalankan bagi mengkaji kesan campuran poli (propilen karbonat) (PPC) dan kurkumin pada sifat mekanikal poli (asid laktik) (PLA). Tambahan, interaksi kimia dan perubahan morfologi pada setiap penambahan berikutnya turut diperhatikan. Penambahan PPC pada 30 wt% ke dalam PLA menyebabkan pengurangan pada tenaga dan modulus sebanyak 51% dan 68% masing-masing sementara menyebabkan kenaikan pemanjangan sebanyak 74%. Hasil pengurangan pada tenaga dan modulus campuran PLA/PPC diseimbangkan dengan mencampurkan peratus kurkumin kurang berat (1 wt%). Melalui mikroskop imbasan elektron (SEM), didapati imej permukaan retak menunjukkan PLA berturutan rapuh-mulur-rapuh disebabkan penambahan PPC dan kurkumin di mana ianya menyokong dapatan kajian ini melalui ujian kekuatan tegangan. Spektroskopi Inframerah Jelmaan Fourier (FTIR) menunjukkan dengan penambahan PPC sebanyak 30 wt %, interaksi kimia antara kumpulan karbonil PLA dan PPC pada puncak C=O PLA telah berubah sedikit kepada nombor gelombang lebih kecil. Walau bagaimanapun, kehadiran puncak kurkumin adalah sukar dikenal pasti dalam campuran PLA/PPC/kurkumin. Dapatan hasil simulasi pada profil strain-tekanan menggunakan FEA adalah sama dengan ujian kekuatan tegangan dengan peratus ralat yang agak rendah iaitu kurang daripada 6%.

Oleh itu, komposisi model yang dibangunkan ini adalah sesuai bagi meneruskan kerja-kerja simulasi iaitu bagi mereka bentuk alat biomedikal.

***KEYWORDS:*** *PLA; PPC; curcumin; tensile; FEA*

# 1. INTRODUCTION

Poly(lactic acid) (PLA), recently, has been the main thermoplastic polymer candidate for applications that require tailored biodegradability feature such as in the biomedical as well as food and packaging industry [1]. The biomedical application of PLA is due to its ability to fulfil complex requirements in terms of biodegradability, and thermoplastic processability [2, 3]. In comparison to other traditionally studied biodegradable polymers, PLA is able to provide excellent features at low cost [4]. PLA offers good thermal processability, thus, the polymer can be processed via various techniques such as injection moulding, film extrusion, blow moulding and fiber spinning. Despite the polymer's copious advantages, the application of PLA has been restricted to a certain extent due to its inherent brittleness, hydrophobicity, poor heat resistance and poor degradation rate [5,6]. Specifically, one of the main drawbacks of PLA for biomedical application is attributed to its high brittleness [7], since it reduces the flexibility and plasticity [8] of the polymer. The brittleness of PLA is a result of the formation of large spherulites due to its slow crystallization rates [9]. A simple and efficient method to modify the flexibility, strength, processability, and impact toughness of PLA is by blending with various plasticizers such as poly(propylene carbonate) (PPC) [10,11], polyethylene glycol (PEG) [12,7], glucosemonoesters , poly(vinyl acetate) (PVAc) [13], and partial fatty acid esters.

The biodegradable poly(propylene carbonate) (PPC) is synthesized from copolymerization of renewable $CO_2$ and propylene oxide (PO) and is typically amorphous [14]. PPC has a chemical structure akin to that of PLA as shown in Fig. 1. The polymer offers good mechanical property, in particular ductility that could be used to improve the toughness of PLA [11]. In comparison to PLA, PPC has a relatively low glass transition temperature (around 35 °C) due to the lack of a strong polar group in its chain, therefore enabling the PPC chain segment to move easily [15]. Hence, PPC displays a very flexible structure. The polymer's elongation at break is nearly 330%, which is at least five-fold greater in comparison to other polyesters [16].

(a) (b)

Fig. 1: The chemical structure of (a) Polylactic acid, PLA and (b) Poly(propylene carbonate), PPC [11].

Curcumin, an orange-yellow crystalline powder with a chemical structure shown in Fig. 2 is prepared from solvent extraction followed by purification of ground rhizomes of turmeric [17]. Curcumin is commonly used in various applications such as foods and cosmetics. Due to its medicinal properties, there is a high tendency for curcumin to be used in medical applications for antibacterial, antitumor, and antioxidant purposes [18-21]. The bioactivities of curcumin such as anti-cell adhesion and motility, anti-angiogenesis and anti-microbe properties have been elucidated in *in vivo*/*in vitro* studies [22]**.** Various curcumin/polymer blends have been electrospun into nanoscale fiber scaffolds or mats [21,

23]. Several studies of wounded animals treated with PLA/curcumin nanofiber mats have been noticed to undergo improved wound closure rates [24, 25]

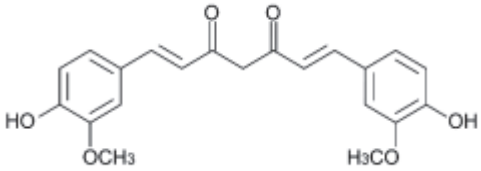



Fig. 2: The chemical structure of curcumin [26].

Thin films often experience very high stresses during service which leads to distortion or deformation [27]. Finite element analysis (FEA) provides insight in identifying critical locations of the device (high stress/strain regions), reduces design iterations/optimization, and reduces the need for bench-top testing; thereby reducing the cycle time during product development [28]. FEA has been used not only to model the mechanical properties [29,30] but also degradation performances [31,32] of various materials. Previously, flat tension specimens of various ceramic matrix composites and monolithic ceramics have been studied via FEA for creep performance [33]. This work was prompted by the scarcity of FEA study on flat thin films subjected to uniaxial tensile tests. The motivation for this study is twofold: first, to study the effect of PPC and curcumin addition on PLA upon its structural, morphological and mechanical properties, and second, to simulate the mechanical effect of PPC and curcumin addition in the elastic and plastic region using FEA (ABAQUS).

## 2. METHODOLOGY

### 2.1 Sample Preparation

PLA (IngeoTM Biopolymer 3052D) was obtained from Nature-works (USA) whilst PPC (QPAC 40) was sourced from Empower Materials (Malaysia). Analytical grade chloroform (R&M Chemicals, United Kingdom) was used as solvent. The curcumin (Live-well$^{TM}$) used in this study was manufactured by Pahang Pharmacy Sdn. Bhd. (Malaysia). Both PLA and PPC pellets were dried at 70 °C and 40 °C respectively for 2 hrs in an oven prior to being dissolved in chloroform. Subsequently, PLA, PPC and curcumin were prepared at different weight % ratios (PLA:PPC -100:0, PLA:PPC - 70:30 and PLA:PPC:curcumin; 69:30:1). Both PLA and PPC were dissolved in chloroform at the ratio of 1:10. Upon complete dissolution, curcumin was then added and stirred for 1 hour. Each blend was then poured into a Teflon covered container and was left to dry for 24 hours at ambient temperature.

### 2.2 FTIR Analysis

Identification of the functional groups of the PLA and PLA/PPC blends were obtained by using Tensor 27 FTIR Spectrometer (Bruker, USA). The spectra were collected within 4000 – 600 $cm^{-1}$ region with a resolution of 4 $cm^{-1}$ and 16 scans.

### 2.3 Morphological Observation

The surface of the fractured tensile test specimens were sputtered with palladium using a Quorum SC7620 Sputter Coater in order to make the surface conductive prior to SEM analysis. The resulting surfaces were then observed via InTouch Scope JSM-IT100 (JEOL).

### 2.4 Tensile Test

The FEA simulations for the studied polymers were based on the stress-strain response of the materials conducted according to ASTM D882. Thin films were cut according to the following dimensions of 50 mm × 10 mm × 0.80 mm with an extra 10 mm at both ends for clamping purposes as shown in Fig. 3. The grey tabs in this figure represent the glued cardboard strips of the same size, by which the sample is clamped between machine grips. Tensile tests were performed using a Shimadzu Universal Tensile Machine (AGS-10K NXD) at a crosshead speed of 20 min/mm with a load cell of 5 kN. The average and standard deviation of the mechanical properties were calculated from the load-elongation curve based on 5 samples.

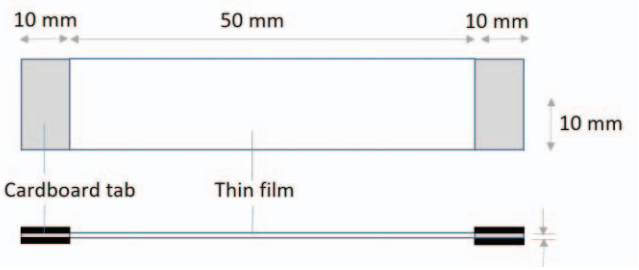



Fig. 3: Preparation of test specimen for tensile test.

## 2.5 Simulation Model

The FE model of the thin films were developed using Abaqus software. The 3D model was discretized into a structured element mesh using 3D linear hexahedral elements (C3D8R) which resulted in total mesh and node values of 1562 and 700 respectively for all samples. A schematic representation of the tensile specimen as well as the boundary condition inflicted on the sample is shown in Fig. 4.

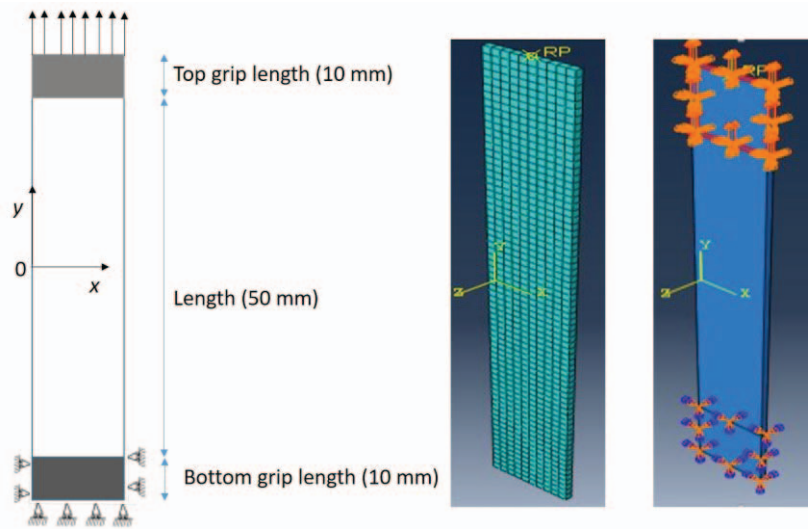



Fig. 4: (a) Schematic representation of the specimen geometry (b) Finite element mesh using 3D linear hexahedral elements (C3D8R) (c) boundary conditions for the simulation of the tensile test.

In the actual tensile test, each specimen was clamped and displaced using hydraulic rigid grips. The additional length of 10 mm at each end of the sample functions as a grip length in which the top grip length will be displaced upward by the load cell whilst the bottom grip

length would be stationary. In order to replicate the experimental tensile test, the boundary and loading conditions were translated as follows [34]; the top grip length (10 mm) and sample test length (50 mm) were kept as fixed in all directions but unconstrained in the direction of the load applied. The top grip was loaded with an upward load of 5 kN. Meanwhile, the bottom grip length had no degree of freedom. The average thickness of each sample was 0.8 mm. The constitutive material data for the elastic model in the simulation were defined by the experimental elastic modulus whilst Poisson's ratio was obtained from the literature [35] as shown in Table 1. The plasticity model was included in the FEA by defining the region of yield stress and corresponding plastic strain that were obtained from the experiment. There were no purposefully designed geometrical flaws such as notches or necking in the simulation model that would trigger localization of the specimen at any desired region.

Table 1: Material properties used for the FEA simulation element

| Composition (wt%) | Young's modulus (MPa) | Yield strength (MPa) | Poisson's ratio |
|---|---|---|---|
| Neat PLA | 936 | 8.4 | 0.36 |
| PLA / PLC (70:30 wt%) | 156 | 3.5 | 0.36 |
| PLA / PPC / curcumin (69:30/1 wt%) | 771 | 11.7 | 0.36 |

# 3. RESULTS AND DISCUSSION

## 3.1 FTIR

Figure 5 shows the FTIR spectra of neat PLA, PPC, and curcumin as well as their blends. The FTIR detects the chemical interaction and compatibility between the molecular chains of different polymer blends by the shift and broadening of the absorption band [36]. The neat PLA spectra shows characteristic stretching frequencies for C=O, $-CH_3$ asymmetric, $-CH_3$ symmetric, and C–O, at 1752 $cm^{-1}$, 2995 $cm^{-1}$, 2947 $cm^{-1}$ and 1085 $cm^{-1}$, respectively. The PPC spectra shows the typical characteristic peaks at 2920 $cm^{-1}$, 1746 $cm^{-1}$, and 1235 $cm^{-1}$ attributed to C-H bond stretching vibration, C=O bond stretching vibration and C-O stretching respectively. In the 3700 – 3100 $cm^{-1}$ region, the absorption of neat PPC and PLA is not identified due to the rarity of terminal O-H groups.

The PLA/PPC spectra in the high wavenumber region, (3100-2800 $cm^{-1}$), relates to the -CH- stretching vibration, which shows similar absorption peaks as pristine PLA. Ma et al. [38] observed that the stretching vibration may shift with higher PPC loading dueto intermolecular interaction between C-H and O-C- or between C-H and O=C-. PLA is made from lactic acid and forms rich carbon chains that serve as the main functional group (C=O groups). The functional group is detectable at the frequency region of 1740-1750 $cm^{-1}$ (denoted as 'A') [37]. The addition of 30 wt % PPC causes the C=O peak of PLA at 1752 $cm^{-1}$to slightly shift to a lower wavenumber $cm^{-1}$ (by 3 $cm^{-1}$). Therefore, implying that the carbonyl groups may have taken part in the interaction between PLA and PPC. Such interaction could be between C-H…O=C-, C=O…O=C or C=O…O-C dipole-dipole interaction [38]. A more significant wavenumber shift of up to 11 $cm^{-1}$ has been observed with higher PPC loading [38]. It was noted that there is no significant peak shift in the C-H bending ('B' region) as well as the C-O stretching vibration ('C' region). It was deduced that despite the similar chemical nature of PLA and PPC, the addition of 30 wt% of PPC only resulted in chemical interaction between the carbonyl groups of these polymers.

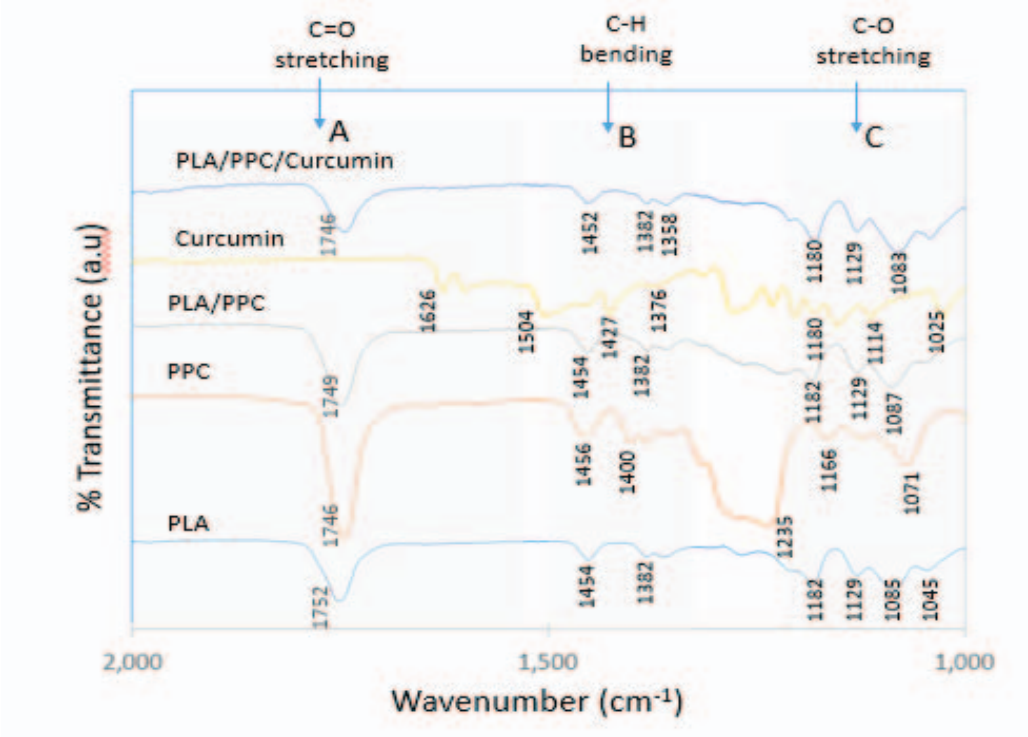



Fig. 5: Fourier transform infrared (FTIR) spectra of PLA, PPC, PLA/PPC, and PLA/PPC/curcumin blends.

Curcumin can be distinguished by its phenolic O-H stretching band at 3508 $cm^{-1}$ [26, 39]. Other significant curcumin bands are the aromatic C-C band (1602 $cm^{-1}$), olefinic C-H (1428 $cm^{-1}$), asymmetric C-O-C (1026 $cm^{-1}$) and phenolic C-O (1276 $cm^{-1}$) [40, 39]. These significant peaks are found to be comparable and appear at 3507 $cm^{-1}$, 1601 $cm^{-1}$, 1428 $cm^{-1}$, 1026 $cm^{-1}$ and 1274 $cm^{-1}$ respectively. A sharp common peak at 1180 $cm^{-1}$, assigned as phenolic C-O is observed in both curcumin and PLA/PPC/curcumin, hence confirming the presence of curcumin in the blend. Apart from the C-O phenolic peak, the lack of presence of other curcumin peaks observed in the PLA/PPC/curcumin trace suggests that these curcumin peaks were either eclipsed by the more generous peaks of PLA at that wavenumber range due to the higher weight % of PLA [41]. The FTIR samples for PLA/PPC/curcumin were performed twice to confirm such a scenario. Previously, Chen et al. [42] reported that curcumin loading below 5% would not give any chemical reaction with PLA.

### 3.2 Tensile Test

The tensile test was conducted to investigate the mechanical effect of adding 30 wt% of PPC and 1% of curcumin at the expense of PLA. Figure 6 illustrates that the neat PLA curve behaves linearly in the low strain, then starts to plastically deform in the region of 3% strain. The strain at break of neat PLA in this study is 11.7% and is higher than the PLA thin films studied by Yao et al. [10]. Aside from the grade of PLA used, the difference in sample thickness, as well as testing speed [13], it is postulated that the strain value in this study may have also been affected by the sample preparation technique since the thin films in their study [10] were prepared by melt compounding.

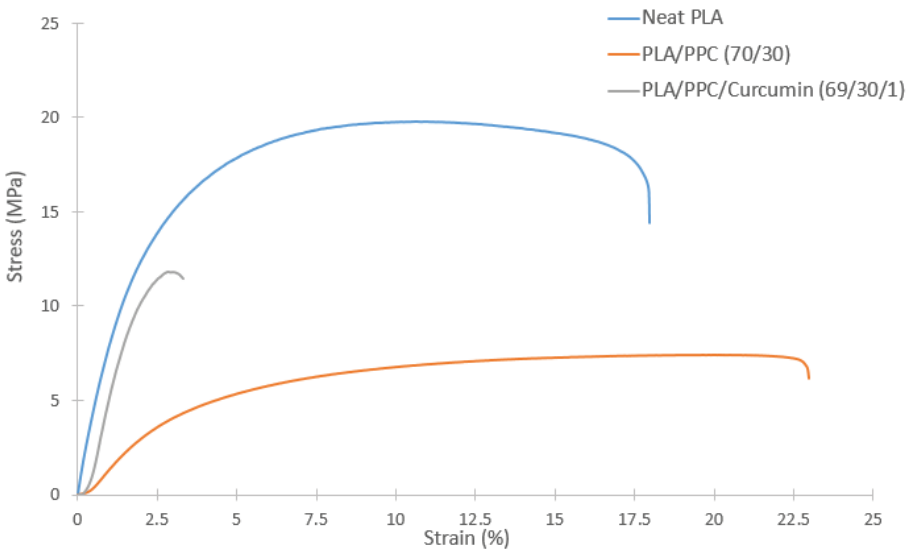



Fig. 6: Stress-strain curves for neat PLA, PLA/PPC, and PLA/PPC/curcumin.

Studies on PLA/PPC blends by Zao et al. [43] and Yu et al. [44] suggest that the addition of 30 wt% of PPC creates the optimum blend after considering the resulting elongation and flexural and impact strength properties. In this study, the addition of 30 wt % of PPC into PLA causes a decrease of strength and modulus by 51% and 68% respectively (Table 2). The trend of decreasing tensile strength and Young's modulus upon blending of PPC into PLA is well documented. Studies by Ma et al. [38] and Zou et al. [43] show a decreasing trend in strength as well as of modulus with further loading of PPC. It is also seen that PPC contributes to the elastic component. As a result, a higher strain at break by 74% was achieved in the PLA/PPC blend in this study. Such an increase in strain value is manifested from the weak molecular chain interaction and the presence of plenty weak flexible, polar C–O–C bonds in the backbone structure of PPC [15, 43]. Sun et al. [45] suggested that the transformation of PLA from brittle to ductile nature indicates a good compatibility between PLA and PPC as evidenced by the chemical interaction found in FTIR.

Table 2: Mechanical properties of neat PLA, PLA/PPC and PLA/PPC/curcumin

| Properties | Neat PLA | PLA / PPC | PLA/ PPC/Curcumin |
|---|---|---|---|
| Modulus (MPa) | 481.3 ± 145.1 | 155.4 ± 26.9 | 726.7 ± 62.4 |
| Strength (MPa) | 15.3 ± 2.8 | 7.4 ± 0.3 | 11.0 ± 1.1 |
| Strain at break (%) | 11.7 ± 3.3 | 20.4 ± 1.4 | 2.4 ± 0.7 |

The addition of 1 wt % of curcumin significantly improves the strength and modulus of PLA/PPC, hence suggesting that the curcumin particles act as reinforcement filler. The presence of curcumin creates resistance that reduces the mobility of internal PLA/PPC structures and resulted in an enhanced stiffness property. However, the inclusion of filler also increases the polymer blends' brittleness characteristics [46]. As seen here, when the PLA/PPC/curcumin was subjected to tensile-mode deformation, the samples failed at very low tensile strain. The composition of curcumin was restricted to only 1 wt% in this study since previous study by the author [47] has shown a decreasing trend in strength with further curcumin additions. An obvious decrease in the tensile strength at higher filler content is

likely to occur as the filler-matrix interaction is replaced by the contact between particles as well as filler agglomeration [46, 48].

### 3.3 Surface Morphology

The photographs of the dried polymer blend films are shown in Fig. 7. The addition of PPC into PLA increases opacity whilst the addition of curcumin imparts yellowish hue due to the bioactive yellow pigment from curcumin with the following structure; 1,7-bis(4-hydroxy-3-methoxyphenyl)-1,6-heptadiene-3,5-dione.

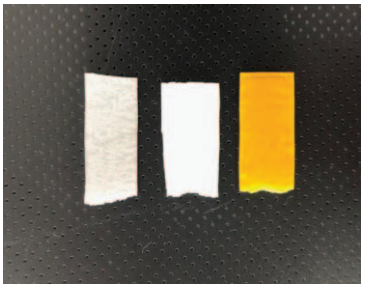


Fig. 7: From left to right, fractured tensile test specimens of PLA, PLA/PPC, and PLA/PPC/curcumin.

The morphological structure of the PLA blend was examined via SEM, and the results are shown in Fig. 8. The SEM image of neat PLA exhibited the fractured surface that occurs at different layers and at different parts of the sample. Similar surface topography was observed from other studies [13, 49] that confirms the brittle nature of PLA. The blending of PPC into PLA creates a surface fracture that is much smoother with no visible aggregation. Such effect signals good dispersion and homogeneity in the blended films. The transition of the PLA and PLA/PPC samples from rough to smooth surface are in agreement with the observations made from the mechanical properties, in which the neat PLA progresses from a brittle to a ductile nature with the decrease in stiffness and an increase in strain values. In comparison, the addition of curcumin induces a much rougher fractured surface with a phase-separated morphology. It was observed that the dispersed curcumin particles created almost spherical and empty voids at the fractured surface. The change in the fractured surface of PLA/PPC blend from smooth to rough correlates to the effect of filler particle addition that changes the PLA/PPC from being ductile to a more brittle nature [46].

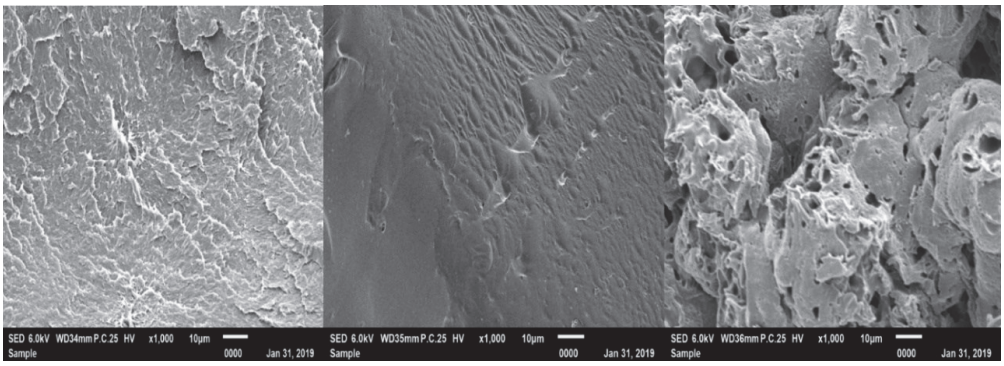



Fig. 8: SEM micrographs at 1000x magnification of the surfaces of the fractured (a) neat PLA (b) PLA/PPC (c) PLA/PPC/curcumin films subjected to uniaxial tensile stress.

### 3.4 Finite Element Analysis of the Tensile Test

The results from FE simulations are compared to the experimental data in Fig. 9 - Fig. 11. The graphs obtained show that the simulated stress-strain behaviour (overall shape) in the elastic and plastic region are in excellent agreement with the experimental data for all compositions studied. The damage criterion model was not implemented in this FEA model, hence the simulation and experimental work are comparable up to the ultimate tensile strength (maximum stress) only. The convergence between the simulation and experimental results near the maximum stress value indicates the reliability of the model since the maximum stress value is generally used as a criteria in design. The accuracy of the experimental and simulated model was obtained by comparing the integrated area beneath the stress-strain curves. Simulation performed using Abaqus FEA models described accurate results with a relatively low percentage error of 1% for both PLA as well as PLA/PPC and 6% for PLA/PPC/curcumin.

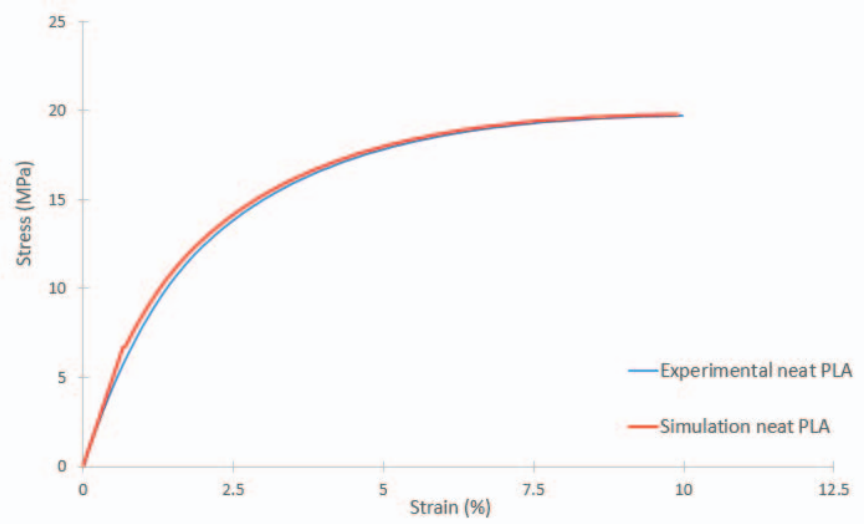



Fig. 9: Stress-strain curves obtained experimentally and via FEA for neat PLA.

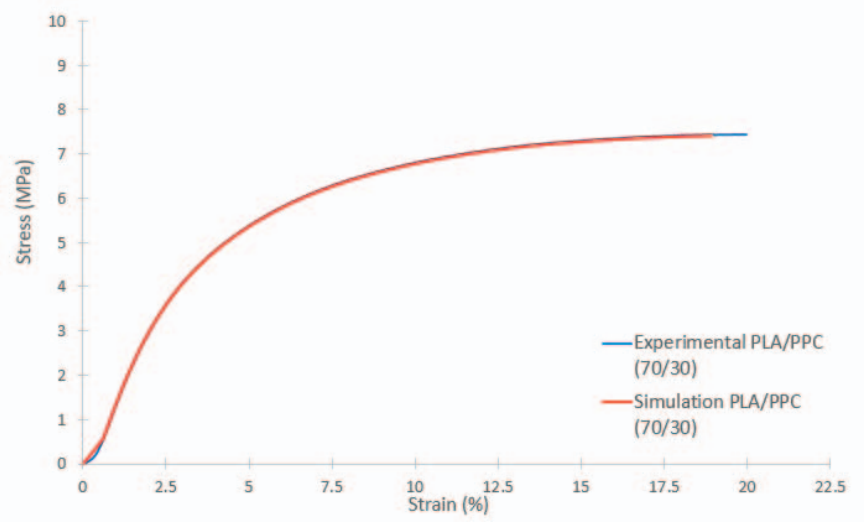



Fig. 10: Stress-strain curves obtained experimentally and via FEA for PLA/PPC.

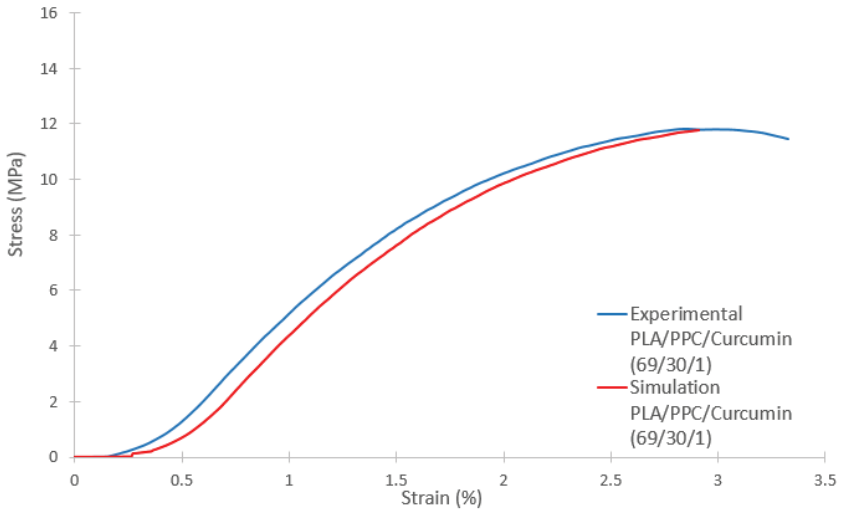
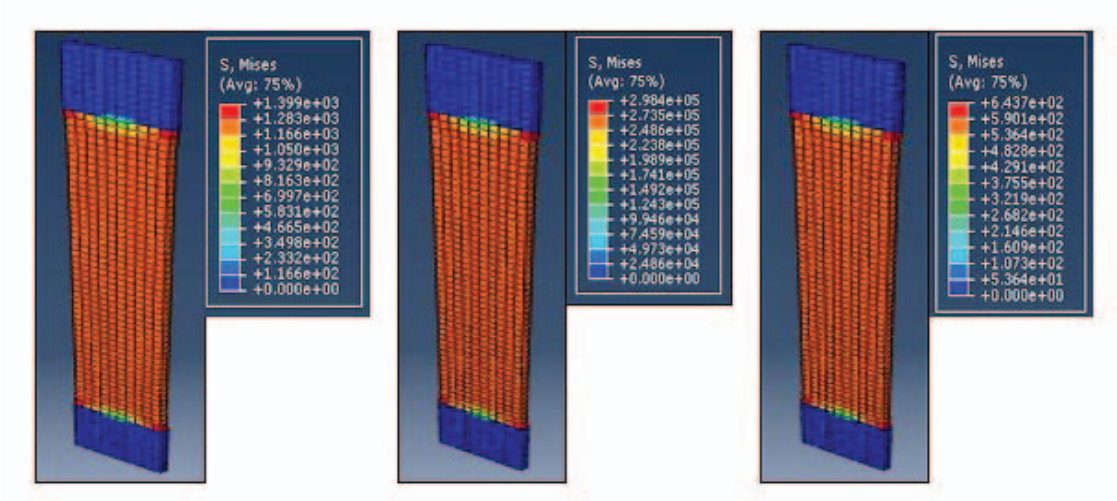



Fig. 11: Stress-strain curves obtained experimentally and via FEA for PLA/PPC/curcumin.




(a) (b) (c)

Fig. 12: Contour plots of computed Von Mises stress along the specimen length for (a) neat PLA, (b) PLA/PPC blend and (c) PLA/PPC/curcumin blend.

Figure 12 shows the distribution of Von Mises stress calculated by the finite element model based on Young's modulus, Poisson ratio, boundary conditions as well as applied load. Similar topography of low to high Von Mises Stress was obtained for all compositions studied. It can be seen from Fig. 13 that the stress distribution is divided into three regions. The first region undergoes no stress since it is the area under grip length, and hence was not deformed. The second region, is noted as the transition phase in which the stress starts to increase. It is also noted that the FE model highlighted the highest Von Mises stress in this transition region between grip and gage length which may indicate the location that possibly initializes the fracture [33]. The model postulated that fracture may occur at either end of the grip length since there were no purposefully designed geometrical flaws such as notches or necking in the simulation model that would trigger localization of the specimen at any desired region. The FE model shows that the addition of PPC into PLA increases the Von Mises stress scale from 1400 MPa to 300 GPa. Whilst the addition of curcumin particles into PLA/PPC blends resulted with reduced Von Mises stress distribution (640 MPa). The values for Young's modulus used for the FE analysis were 936 MPa, 156 MPa and 771 MPa

for PLA, PLA/PPC and PLA/PPC/curcumin respectively. The Young's modulus property influences the stress distribution and hence the elastic deformation of the model [50]. Thus, as the Young's modulus of the sample increases or decreases, correspondingly the upward tensile force respectively causes less or more stress due to the material's resistibility to deformation.

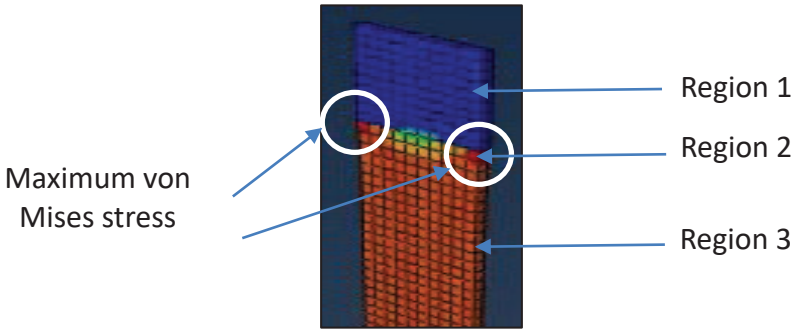



Fig. 13: Identified regions of Von Mises Stress.

## 4. CONCLUSION

The FTIR analysis illustrated that the chemical interaction, in particular between the carbonyl group (C=O), exists when PPC was added at 30 wt% into PLA. Despite the resultant decrease in strength and modulus, the strain at break increased by 74%. The drop in strength and modulus of PLA/PPC blend was recovered by adding a low weight percentage of curcumin. However, the PLA/PPC/curcumin blend failed at very low tensile strain which shows the increase in the brittle nature of the blend. In light of this observation, it was further confirmed by SEM images that the addition of PPC and curcumin respectively increases its ductility and brittleness. The validation and the excellent agreement between the simulation and experimental results are important indices of the simulated model developed using Abaqus. Hence, this model can be used to predict the performance of complex design components based on PLA/PPC/curcumin blends for biomedical applications.

## ACKNOWLEDGEMENT

This work was fully supported by the International Islamic University Malaysia (IIUM), (RIGS17-052-0627). The authors would like to express their deepest gratitude to facilities and technical assistance provided by the International Institute for Halal Research and Training (INHART), IIUM.

## REFERENCES

[1] SuPing L, Darrel U. (2009) Degradability of Polymers for Implantable Biomedical Devices. Int J Mol Sci., 9 (10): 4033-4065.
[2] Kaojin W, Satu S, Julian XXZ. (2017) A Mini Review: Shape Memory Polymers for Biomedical Applications. Front Chem Sci Eng., 11(2): 143-153.
[3] Rajendra PP, Sunil UT, Suresh US, Jalinder TT, Abraham JD. (2014) Biomedical Applications of Poly(Lactic Acid). Recent Pat Regen Med., 1(4): 240-251.
[4] Kotiba H, Mosab K, Yang HW, Fawaz D, Young GK. (2015) Properties and Medical Applications of Polylactic Acid: A Review. Polym Lett., 9(5):435–455.
[5] Broz M, David LV, Newell RW. (2003) Structure and Mechanical Properties of Poly(D,L-lactic acid)/Poly(ε-caprolactone) Blends. Biomat., 24:4181-4190.

[6] Erin H, Margaret F. (2014) Increasing Surface Hydrophilicity in Poly(Lactic Acid) Electrospun Fibers by Addition of PLA-β-PEG Co-Polymers. J Eng Fiber Fabr., 9(2):153-164.
[7] Feng-Jiao L, Shui-Dong Z, Ji-Zhao L, Jun-Zhe W. (2015) Effect of Polyethylene Glycol on the Crystallization and Impact Properties of Polylactide-based Blends. Polym. Adv. Technol., 26(5):465-475.
[8] Yanping L, Hanghang W, Zhen W, Qian L, Nan T. (2018) Simultaneous Enhancement of Strength and Toughness of PLA Induced by Miscibility Variation with PVA. Polymers, 10 (10):1178-1191
[9] Ahmed Mohamed E. (2014) The Effect of Additives Interaction on the Miscibility and Crystal Structure of Two Immiscible Biodegradable Polymers. Polimeros, 24(1):9-16.
[10] Yao M, Deng H, Mai F, Wang K, Zhang Q, Chen F, Fu Q. (2011) Modification of Poly(lactic acid)/Poly(Propylene Carbonate) Blends Through Melt Compounding with Maleic Anhydride. Polym Lett., 5(11):937-949.
[11] Sheng-Xue Q, Cui-Xiang Y, Xue-Yang C, Hai-Ping Z, Lifen Z. (2018) Fully Biodegradable Poly(lactic acid)/Poly(propylene carbonate) Shape Memory Materials with Low Recovery Temperature Based on In Situ Compatibilization by Dicumyl Peroxide. Chinese J Polym Sci., 36:783-790.
[12] Mihir S, Ananda Kumar R, Vipul D, Richard AG, Stephen PM. (1997). Biodegradable Polymer Blends of Poly(lactic acid) and Poly(ethylene glycol). J Appl Polym Sci., 66:1495-1505.
[13] Wannapa C, Varaporn T, Jean-Francois P, Pamela P. (2012) Effect of Poly(Vinyl Acetate) on Mechanical Properties and Characteristics of Poly(Lactic Acid)/Natural Rubber Blends. J Polym Environ., 21:450-460.
[14] Baherah B, Yanwei M, Fariba D (2015). Antimicrobial Packaging for Biomedical Applications from a Biodegradable Polymer in Asia Pacific Confederation of Chemical Engineering Congress 2015; Melbourne, Australia.
[15] Youhua T, Xianhong W, Xiaojiang Z, Ji L, Fosong W. (2006) Crosslinkable Poly(propylene carbonate): High-Yield Synthesis and Performance Improvement. J Polym Sci., 44:5329-5336.
[16] Iman M, Ali F, Hesham B, Sean Ryan D, Ali Negahi S, Fariba D. (2016) Biomedical Applications of Biodegradable Polyesters. Polymers, 8(1):20-52.
[17] Dalbir S, Sharma S, Davinder Pal Singh O, Idress Ahmed W. (2010) Effect of Extraction Parameters on Curcumin Yield from Turmeric. J Food Sci. Technol., 47(3):300-304.
[18] David E R, Biotechnology 2, Global Prospects, Florida: CRC Press, 2011.
[19] Bharat A, Bokyung S. (2009) Pharmacological Basis for the Role of Curcumin in Chronic Diseases: An Age-old Spice with Modern Targets. Trends Pharmacol Sci., 30(2):85-94.
[20] Sahdeo P, Amit Tyagi K, Bharat A. (2014) Recent Developments in Delivery, Bioavailability, Absorption and Metabolism of Curcumin: the Golden Pigment from Golden Spice. Cancer Res. Treat., 46(1): 2-18.
[21] Nithya R, Tamil Selvan N, Sheeja R. (2015) Preparation and Characterization of Electrospun Curcumin Loaded Poly(2-hydroxyethyl methacrylate) Nanofiber-a Biomaterial for Multidrug Resistant Organisms. J Biomed Mater Res., 103(1):16-24.
[22] Xia F, Chun Z, Dong-Bu L, Jun Y, Hua-Ping L. (2013) The Clinical Applications of Curcumin: Current State and the Future. Curr Pharm Des., 19:11-13.
[23] Xiao-Zhu S, Gareth R W, Xiao-xiao H, Li-Min Z. (2013) Electrospun Curcumin-loaded Fibers with Potential Biomedical Applications. Carbo Polym., 94:147-153.
[24] Thuy Thi T N, Chiranjit G, Seong Gu H, Tran Dai L, Jun Seo P. (2013) Characteristics of Curcumin –loaded Poly(lactic) Nanofibers for Wound Healing. J Mater Sci, 48:7125-7133
[25] Bhaarathi D, Saraswathy N, Ramasamy M, Ponnusamy S, Palanisamy V, Sukumar V, Venugopal R. (2013) Electrospinning of Curcumin Loaded Chitosan/Poly (lactic acid) Nanofilm and Evaluation of its Medicinal Characteristics. Front Mat Sci., 7(4):350-361.
[26] Heni R, Yulia Lie Y, Annisa R, Noboyuki M. (2016) Curcumin-loaded PLA Nanoparticles: Formulation and Physical Evaluation. Sci Pharma., 84(1):191-202.

[27] Li-mei J, Yi-chun Z, Yongli H. (2010) Elastic-plastic Properties of Thin Film on Elastic-plastic Substrates Characterized by Nanoindentation Test. T Nonferr Metal Soc., 20(12):2345-2349.
[28] Senthil E, Ashley Kelley J, Jorgen B, Virginia L G (2011). Material Modeling of Polylactide, in Simulia Customer Conference, Barcelona, Spain, 2011.
[29] Harpen L, Ahmed I, Felfel RM, Qian C. (2012) Finite Element Modelling of the Flexural Performance of Resorbable Phosphate Glass Fibre Reinforced PLA Composite Bone Plates. J Mech Behav Biomed Mater., 15:13-23.
[30] Anna B, Susana PF, Jose-Ramon S, Wang W, Peter EM. (2015) Computational Bench Testing to Evaluate the Short-Term Mechanical Performance of a Polymeric Stent. Cardio Eng Tech., 6(4):519-532.
[31] Andrea Costa V, Antonio T. M, Rui Miranda G, Volnei T. (2011) Material Model Proposal for Biodegradable Materials. Procedia Eng., 10:1597-1602.
[32] Joao Silve S, James E M, Kumbakonan RR. (2008) Constitutive Framework for Biodegradable Polymers with Applications to Biodegradable Stents. Asaio J., 54(3): 295-301.
[33] Link RE, Garrell MG, Albert S, Edgar LC, Ronald S. (2003) Finite-Element Analysis of Stress Concentration in ASTM D 638 Tension Specimens. J Test Eval., 31(1):1-6.
[34] Mayank N, Anurag D, Misra R K, Harlal Singh M. (2014) Tensile Test Simulation of CFRP Test Specimen Using Finite Elements. Procedia Mat Sci., 5:267-273.
[35] Jonathan T, Jose C, Justin K, Ali PG. (2015). Mechanical Property Optimization of FDM PLA in Shear with Multiple Objectives. JOM., 67(5):1183-1193.
[36] Shuwen P, Xiuyan W, Lisong D. (2005) Special Interaction Between Poly (propylene carbonate) and Corn Starch. Polymer Composite, 26:37-41.
[37] Kurniawan Y, Yohanes Aris P, Setyo P, Bruce A W, Hadi Karia P, Titi Candra S. (2015) Infrared and Raman Studies on Polylactide Acid and Polyethylene Glycol-400 Blend. in AIP Conference Proceedings of the Third International Conference on Advanced Materials Science and Technology (ICAMST 2015): 6-7 Oct 2015; Semarang. Edited by Sutikno, Khairurrijal, Heru Susanto, Risa Suryana, Kuwat Triyana and Markusdiantoro, AIP Publishing; pp 0201011-0201016.
[38] Xiaofei M, Jiugao Y, Ning W. (2006) Compatibility Characterization of Poly(lactic acid)/Poly(propylene carbonate) Blends. J Polym Sci Pt B., 44: 94-101.
[39] Rosana Aparecida da S, Mateus Ferreira de S, Daniela O, Evandro B. (2016) Preparation of Curcumin-loaded Nanoparticles and Determination of the Antioxidant Potential of Curcumin after Encapsulation. Polimeros, 26(3):207-214.
[40] Harshal Ashok P, Karde M, Mundle N, Jadhav P. (2014) Phytochemical Evaluation and Curcumin Content Determination of Turmeric Rhizomes Collected From Bhandara District of Maharashtra (India). Med. Chem., 4(8):588-591.
[41] Duong Quang L, Trang M, Thi Thu T N, Thi Ngoan N, Thi Cham B, Nguyen Hai B, Thi Bich H P, Tran Dai L, Jun Seo P. (2012) A Novel Nanofiber Cur-loaded Polylactic Acid Constructed by Electrospinning. Adv Nat Sci-Nanosci., 3:1-5.
[42] Yan C, Jie L, Yanna F, Hongbo W. (2010) Preparation and Characterization of Electrospinning PLA/Curcumin Composite Membranes. Fib Polym., 11(8):1128–1131.
[43] Wei Z, Rongyuan C, Guizhen Z, Haichen Z. (2016) Mechanical, Thermal and Rheological Properties and Morphology of Poly (lactic Acid)/Poly (propylene Carbonate) Blends Prepared by Vane Extruder. Polym Adv Technol,. 27(11): 1430-1437.
[44] Cuixiang Y, Ruirui Z, Haiyue W. (2018) The Mechanical Property of PLA/PPC Blends. Int J Trend Res Dev., 5(2):206-208.
[45] Qirui S, Tizazu M, Manjusri M, Amar M. (2016) Novel Biodegradable Cast Film from Carbon Dioxide Based Copolymer and Poly(Lactic Acid). J Polyms Environ., 24(1):23-36.
[46] Nurdina Abdul K, Mariatti J, Samayamutthirian P. (2009) Effect of Single-Mineral Filler and Hybrid-Mineral Filler Additives on the Properties of Polypropylene Composites. J Vinyl Addit Techn., 15(1):20-28.

[47] Sharifah Imihezri S S, Qairol A A B, Noor Azlina H, Nor Khairusshima MK. (2017) Thermal, Structural and Mechanical Properties of Melt Drawn Curloaded Poly(lactic acid) Fibers. Procedia Eng., 184:544 - 551.
[48] Zhao W, Xiangling L, Min Z, Wei Y, Mingbo Y. (2017). Synthesis of an Efficient Processing Modifier Silica-g-poly (lactic acid)/Poly (propylene carbonate) and its Behavior for Poly (lactic acid)/poly (propylene carbonate) Blends. Ind Eng Chem Res., 56 (49):14704-14715.
[49] Baki H. (2014). The Properties of PLA/Oxidized Soybean Oil Polymer Blends. J Polym Environ., 22(2):200-208.
[50] Yongtao W, Liming K, Sizhuo S, Yuxing D, Zhong G, Zhirui Y, Qigang Z. (2013) Finite Element Analysis on Von Mises Stress Distributions of Si DSP. Mater Sci Semicond Process., 16(1):165-170.

https://doi.org/10.31436/iiumej.v20i2.912

# EFFECT OF BONDED COMPOSITE PATCH ON THE STRESS INTENSITY FACTOR FOR A CENTER-CRACKED PLATE

**ABDUL AABID[1], MEFTAH HRAIRI[1*], JAFFAR SYED MOHAMED ALI[1], AHMED ABUZAID[2]**

*[1]Department of Mechanical Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia.*
*[2]Department of Aeronautical Engineering, Sudan University of Science and Technology, Khartoum, Sudan.*

*[*]Corresponding author: meftah@iium.edu.my*

*(Received: 10th March 2018; Accepted: 13th May 2019; Published on-line:2nd December 2019)*

***ABSTRACT:*** Crack propagation until fracture is an important criterion to predict a structure's service life. In order to increase the latter, the cracked component needs to be repaired or replaced. In the present study, a finite element analysis has been carried out to investigate the effects of adhesive thickness, patch thickness and crack length on the passive repair performance of a center-cracked rectangular aluminum plate under mode-I loading condition using finite element ANSYS package. A comprehensive parametric study shows that the stress intensity factor is influenced by the patch thickness, patch size, adhesive material, and adhesive thickness.

***ABSTRAK:*** Penyebaran retak sehingga patah adalah kriteria penting bagi menjangka hayat struktur. Bagi memanjangkan jangka hayat struktur, komponen keretakan perlu dibaik pulih atau diganti. Kajian ini telah menjalankan analisis elemen tak terhingga bagi mengetahui kesan ketebalan pelekat, ketebalan tampalan dan panjang retak pada bahagian keretakan tengah plat petak aluminium yang dibaiki secara pasif, menggunakan pakej ANSYS di bawah beban mod-I. Kajian parametrik yang menyeluruh menunjukkan faktor tekanan intensif dipengaruhi oleh ketebalan tampalan, saiz tampalan, bahan pelekat dan ketebalan pelekat.

***KEYWORDS:*** *stress intensity factor; center crack; composite patch; adhesive; finite element*

## 1. INTRODUCTION

A structure can be damaged by low energy effects and initiated cracks in its inner layer which may propagate and cause failure of the structure. Damages such as delamination, notches, and cracks are inevitable in various fields of engineering, especially in the aerospace field, and these damages are mostly due to fatigue, corrosion, and accidents. For the repair of such damaged structures, composite patches have been widely used and proven effective. Numerous studies have been reported in this type of passive structural repair in the past four decades.

Much of the earlier numerical work, reported in the literature, was based on FRANC2D/L finite element (FE) code. A circular notch in an aluminum 2024-T3 plate repaired with a graphite/epoxy composite patch was investigated using that code [1] where the developed geometry of the plate and patch were not carried out using a 3D element formulation that can be close to experimental work. Using the same FE code,

another study investigated the effect of adhesive layers to repair the cracked plate with bonded composite patch in which distinct types of adhesives were used and the best one was selected to reduce the stress intensity factor (SIF) and strain energy release rate (SERR) [2]. Later, an algorithm was added to that FE code to simulate the temperature effect to determine reduction of SIF [3]. Using the same code, an investigation of composite patch use to repair a cracked plate was carried out. The effect of the thickness of the patch and adhesive were considered and the patch was designed with circular shape and implemented on an edge-cracked rectangular plate [4].

In recent years, simulation of cracked plate with composite patches has been done using ABAQUS and ANSYS software for 2D and 3D analysis. Mhamdia et al. [5] repaired the crack with bonded composite patch under thermo-mechanical load to reduce SIF. The effect of thermal residual stresses resulting from adhesive curing on the performance of bonded composite repair in aircraft structures was analysed with ABAQUS software [5]. Another study reported experimental results of applying thermal and mechanical load to check the effectiveness of a composite patch over the cracked structure. Patch material, patch size, patch shape, and adhesive material were considered effective in SIF reduction. A center crack emanating from a circular notch was employed to study the reduction in SIF [6]. A composite patch bonded to a cracked plate made of aluminum 2024-T3 under mixed-mode loading was investigated [7]. In this study, 20 noded solid186 element types were used to model the plate, patch, and adhesive bond and the geometry was developed according to the angular changes from 0% to 100% to move mode I to mixed mode. Aabid et al. [8-10] numerically analysed the results of stress concentration factor (SCF) for a center-holed rectangular aluminum plate with and without composite patch repair and SIF when a crack is emanating from the hole. It can be found from the above literature survey that much of the work has been done on edge-cracked rectangular plates. Therefore, the aim of this research is to simulate the results for center-cracked rectangular plates.

In the present work, numerical simulation under linear-elastic fracture mechanics (LEFM) was carried out to establish the effect of composite patch and adhesive characteristics on a center-cracked rectangular plate subjected to a uniform tension load leading to a Mode I type of failure.

## 2. PROBLEM DEFINITION

### 2.1 Specimen

The center-cracked rectangular plate under tensile load with integrated composite material patch is considered for this study, as depicted in Fig.1. The material of the cracked rectangular plate is considered to be aluminum 2024-T3 and the adhesive material is considered to be araldite 2014. In the present study, a crack length of $2a = 30$ mm was used with an integrated composite patch on the crack length under an applied tensile load of 1 MPa. The dimension and mechanical properties of the center-cracked rectangular plate, adhesive layer, and composite material are as shown in Table 1 and Table 2. The composite patch is applied to the damaged area to cover the crack length. The dimensions of the patch are such that it completely covers the damaged area. In the present study boron/epoxy is selected as the composite patch material based on its high strength-to-weight ratio.

https://doi.org/10.31436/iiumej.v20i2.912

## 2.2 Patch Shape

A patch with a skewed shape had been shown to be the most optimum patch design. However, one should be careful when designing such patch shapes to ensure the stress level in the plate remains within the design limits [11]. The rectangular shape of the patch was found to be the second best choice. Considering mechanical performances and manufacturing aspects, this rectangular shape represents a better compromise. Therefore, the patch size was chosen as the minimum patch size that completely covers the crack length and improves the mechanical behavior of the cracked plate under the mode I loading [2].

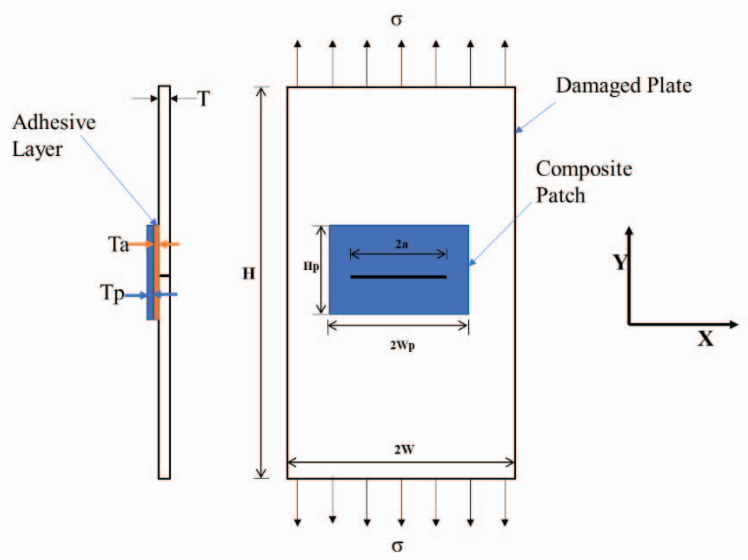



Fig. 1: Rectangular plate with a center crack and integrated composite patch.

Table 1: Dimensions of the host plate, boron/epoxy and adhesive

| Dimensions | Aluminum plate (mm) | Boron/Epoxy (mm) | Adhesive (mm) |
|---|---|---|---|
| Height | H=200 | Hp=40 | Ha=40 |
| Width | 2W=80 | 2Wp=60 | 2Wa=60 |
| Thickness | T=1 | Tp=0.5 | Ta=0.03 |

Table 2: Materials properties of the host plate, boron/epoxy and adhesive

| Parameter | Aluminum plate | Boron/Epoxy | Adhesive (Araldite 2015) |
|---|---|---|---|
| Density | 2715 kg/m$^3$ | | 1160 kg/m$^3$ |
| Poisson's Ratio $\upsilon_{12}$ | 0.33 | 0.3 | 0.345 |
| Poisson's Ratio $\upsilon_{13}$ | | 0.28 | |
| Poisson's Ratio $\upsilon_{23}$ | | 0.28 | |
| Young's Modulus ($E_1$) | 68.95 GPa | 200 GPa | 5.1 GPa |
| Young's Modulus ($E_2$) | | 19.6 GPa | |
| Young's Modulus ($E_3$) | | 19.6 GPa | |
| Shear Modulus ($G_{12}$) | | 7.5 GPa | |
| Shear Modulus ($G_{13}$) | | 5.5 GPa | |
| Shear Modulus ($G_{23}$) | | 5.5 GPa | |

https://doi.org/10.31436/iiumej.v20i2.912

# 3. FINITE ELEMENT MODELLING

The finite element modelling (FEM) is performed using ANSYS mechanical APDL software for simulation purposes [12]. The gradient of the stress and strain fields around the crack front is very high, and to describe such behavior, the displacement extrapolation method uses the nodal displacements around the crack tip. To obtain a good representation of the crack tip displacement, singular elements are used as shown in Fig. 2(a). The nodes used for the crack tip displacements required for the equation are illustrated in Fig. 2(b).

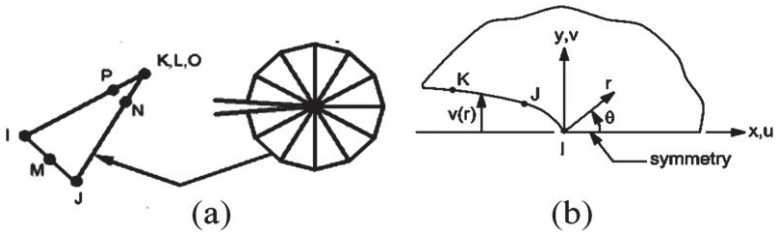



Fig. 2: (a) The singular element near the crack tip and (b) nodes used for the approximate crack tip displacements.

In this work, the SIF is calculated using ANSYS finite element analysis (FEA) software with reasonable accuracy. The displacement near the crack tip is obtained from finite element (FE) analysis, and then the SIF is calculated according to equation (1).

$$K_I = \sqrt{2\pi}\frac{2\mu A}{1+k} \tag{1}$$

## 3.1 Mesh Sensitivity Analysis

To study the mesh effect on the computational results, three mesh sizes were considered as illustrated in Table 3. The processor, which used to run the different cases of the simulation, was an Intel(R) Core i7-3770 CPU @ 3.40 GHz, 16.0 GB RAM, and a Windows-10 64-bit operating system. The parameters selected to run the simulations are: crack length of 10 mm, width of the adhesive bond and the composite patch of 30 mm, height of the adhesive bond and the composite patch of 20 mm, thickness of the adhesive bond of 0.03, mm and thickness of the composite patch of 0.5 mm. The grid structured mesh was used to mesh the adhesive bond and composite patch by applying the size of elements based on mesh type. For the damaged plate, the elements were divided by picking each line, and the generated mesh is of an unstructured type.

As shown in Table 3, it was found that the refinement of the mesh, from intermediate to fine, slightly improved the solution accuracy of the SIF value, leading to a maximum relative difference of 4 %. This comparison indicates that the intermediate mesh provides sufficient resolution and accuracy with around half the computational time, and hence, this mesh size was used for the rest of the simulations.

Table 3: Mesh independence test

| Mesh Type | No. of Elements | CPU Runtime (seconds) | SIF (MPa) |
|---|---|---|---|
| Coarse | 6,832 | 300 | 0.0691 |
| Intermediate | 33,180 | 660 | 0.0667 |
| Fine | 68,920 | 1322 | 0.0640 |

https://doi.org/10.31436/iiumej.v20i2.912

### 3.1 Modelling of the Integrated Structure

In this case, a boron/epoxy composite patch was used, and a SOLID186 element type was used to model the damaged plate, adhesive layer, and composite patch. SOLID186 contains a 20-noded higher order element with 5 degrees of freedom. A typical FE model for the center-cracked plate integrated with a composite patch is shown in Fig. 3. Due to symmetry, only a quarter of the plate was modeled. A total of 10 singular elements were employed to model the crack front. The adopted FE mesh consisted of 7910 coupled-field elements used to model the composite patch, while 21,580 and 3690 high-order reduced integration solid elements were used to model the damaged plate and adhesive bond, respectively.

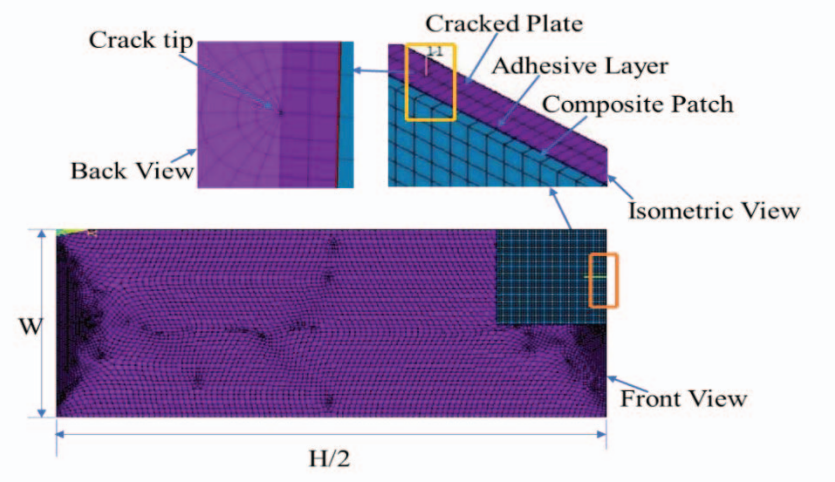



Fig. 3: Finite element mesh and the location of the singular element.

### 3.2 Validation of Finite Element Model

Most of the previous work has been done for the edge cracked plate as it can be easily verified through experimental work. Therefore, an edge-cracked plate was selected to validate the present finite element model. The meshes that are used to model the center and edge cracked plates are almost identical except for their boundary conditions. To verify the simulation results of passive repair, the case study reported in [13] was simulated using the FE model developed in the present work. The edge-cracked rectangular plate integrated composite patch is shown in Fig. 4.

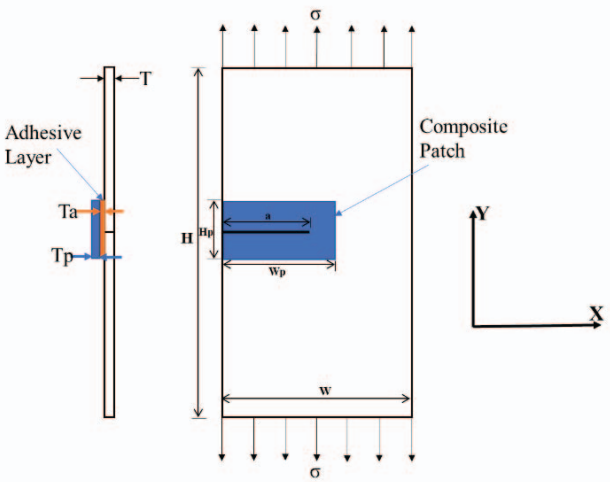



Fig. 4: Rectangular plate through an edge-crack integrated composite patch.

https://doi.org/10.31436/iiumej.v20i2.912

The dimensions and mechanical properties of the cracked plate, adhesive layer and composite material are mentioned in Tables 4 and 5. Applied tensile stress is 58.33 MPa, and the crack length is 30 mm. The results obtained from the current numerical simulation were compared with that reported in [13] and there was a very good agreement as shown in Table 5.

Table 4: Dimensions of the host plate, boron/epoxy patch and adhesive (FM-73)

| Dimensions | Host plate (mm) | Boron/Epoxy (mm) | Adhesive (mm) |
|---|---|---|---|
| Height | H=254 | Hp=75 | Ha=75 |
| Width | W=254 | 2Wp=130 | 2Wa=130 |
| Thickness | T=4.76 | Tp=2.28 | Ta=0.15 |

Table 5: Material properties of the host plate, boron/epoxy patch and adhesive (FM-73)

| Parameter | Aluminum plate | Boron/Epoxy | Adhesive (FM-73) |
|---|---|---|---|
| Plate Strength | 350 MPa | | |
| Poisson's Ratio υ12 | 0.33 | 0.3 | 0.32 |
| Poisson's Ratio υ13 | | 0.28 | |
| Poisson's Ratio υ23 | | 0.28 | |
| Young's Modulus (E1) | 72 GPa | 200 GPa | 2.55 GPa |
| Young's Modulus (E2) | | 19.6 GPa | |
| Young's Modulus (E3) | | 19.6 GPa | |
| Shear Modulus (G12) | | 7.2 GPa | |
| Shear Modulus (G13) | | 5.5 GPa | |
| Shear Modulus (G23) | | 5.5 GPa | |

Prior to repair validation, the FE model of the unrepaired plate (cracked plate without a patch) was also validated using a fracture mechanics analytical solution [14]. Comparing the results gained by Tada's analytical solution [14] according to equation (2) and the present finite element results, there is a good agreement, as shown in Table 6.

$$K_I = \sigma\sqrt{\pi a}\sqrt{\frac{2b}{\pi a}tan\frac{\pi a}{2b}}\frac{0.752+2.02\left(\frac{a}{b}\right)+0.37\left(1-sin\frac{\pi a}{2b}\right)^3}{cos\frac{\pi a}{2b}} \quad (2)$$

Where $a$ is the crack length, and $b$ is the width of the damaged plate.

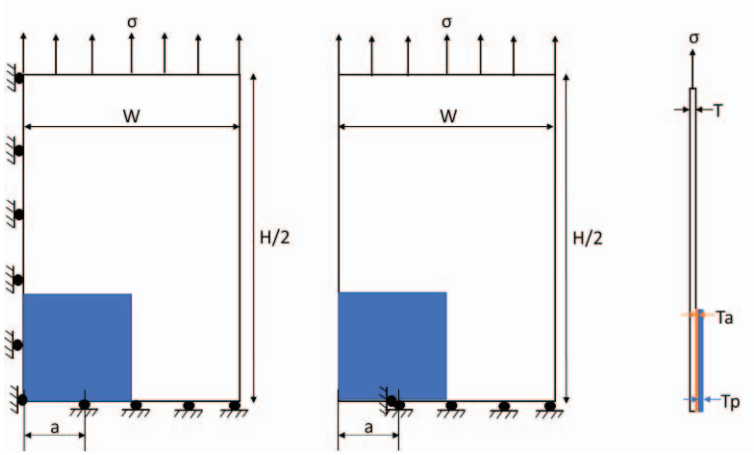
Table 6: Validation of numerical simulation results

| Condition | Ref. [9] $MPa\sqrt{m}$ | Ref. [10] $MPa\sqrt{m}$ | Present Results $MPa\sqrt{m}$ | Relative Error |
|---|---|---|---|---|
| Without Repair | - | 20.049 | 20.432 | **1.874%** |
| Passive Repair | 9.789 | - | 9.623 | **1.695%** |

### 3.3 Boundary Conditions for Center and Edge Cracked Plate

Figure 5 shows the applied boundary conditions for the edge and center cracked plate. Due to symmetry, the model used for the center-cracked plate is a quarter of the plate and for the edge-cracked plate is half the plate. The dimensions of the adhesive bond and composite patch are illustrated in Table 1 for the center cracked plate, and Table 4 for edge cracked plate.

https://doi.org/10.31436/iiumej.v20i2.912




(a) (b) (c)

Fig. 5: Applied Boundary conditions (a) center-cracked (b) edge-cracked (a) side view of plate.

## 4. RESULTS AND DISCUSSION

In this section, finite element analysis was carried out to study the effect of a composite patch placed on the damaged area of a rectangular plate under uniform uniaxial tension load for the problem shown in Fig. 1.

### 4.1 Comparison between Unpatched and Patched Crack

The variation of mode I stress intensity factor is plotted in Fig. 6 for five different values of crack length in the unpatched and repaired cracked plates. It can be seen that SIF strongly decreased when employing the patch repair and more so for bigger crack lengths since the patch carried the load as the crack grew. This is because the stress singularity around the crack tip is reduced due to the composite patch induced shear force on the crack area. Indeed, the SIF reduction reached a maximum value of 68% compared to a 60% reduction for the glass/epoxy patch as obtained in [6]. Therefore, boron/epoxy patches are more effective for reduction of SIF. Furthermore, Fig. 6 shows that for the repaired plate and as the crack length increases, the SIF exhibits an asymptotic behavior.

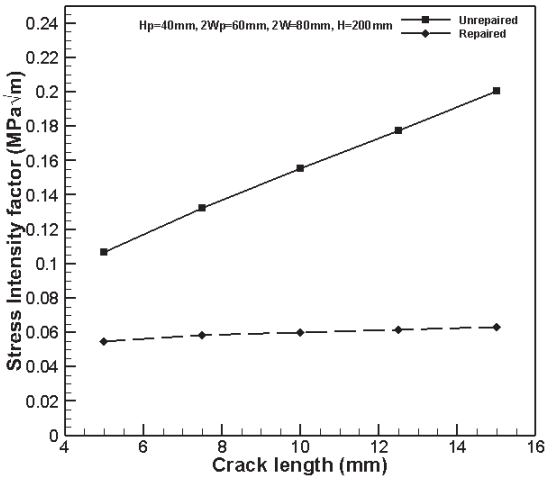



Fig. 6: Crack length variation with and without patch.

### 4.2 Effect of Patch Thickness

Patch thickness represents one of the best means for improving patch repair performance. Thus, many researchers investigated the patch thickness effect on the repair performance in damaged structures [4,5,15]. Mhamdia et al. [16] achieved almost 50% decrease in the crack tip SIF by increasing the patch thickness to repair a cracked plate under mechanical loading.

Figure 7 displays this effect on the normalized SIF variation. It is clearly illustrated that the normalized SIF at the crack tip is reduced proportionally as the patch thickness increases. Indeed, the size of the patch is important to transfer shear load at the damaged area. An increase of 45-50% of the patch thickness reduces the normalized SIF by the same order. Hence to increase the repair performance, the size of the patch should be thicker. Moreover, for repairing cracks, it is preferable to use multi-layered composite patches to improve the distribution of stresses.

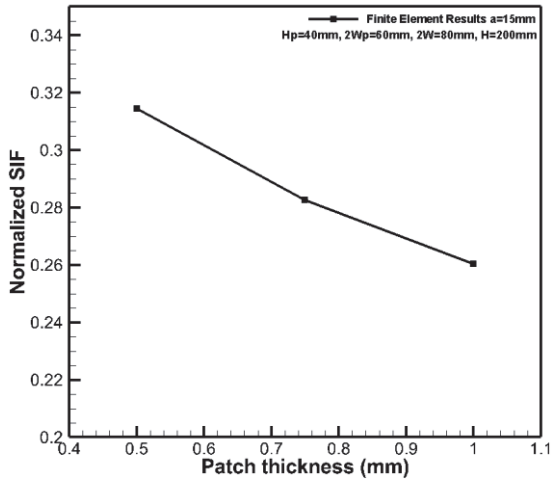



Fig. 7: Effect of patch thickness for passive repair.

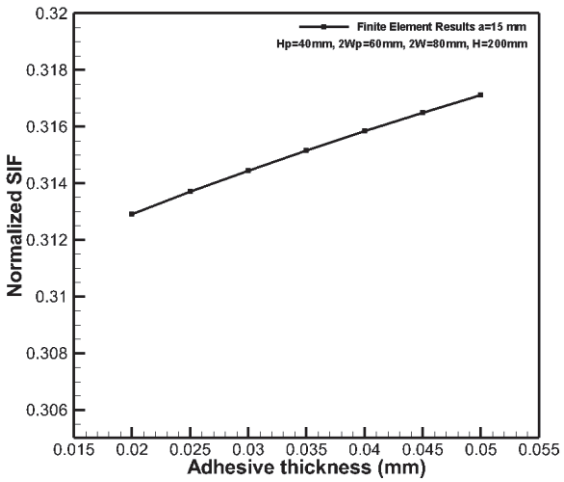



Fig. 8: Effect of adhesive layer thickness for passive repair.

### 4.3 Effect of Adhesive Thickness

Figure 8 shows the variation of the normalized SIF as a function of the adhesive thickness. It can be seen that a very small increment in the adhesive thickness increases the normalized SIF. Therefore, when repairing cracked structures, it is desirable to use

smaller adhesive thickness. This finding correlates well with the results highlighting this effect in [15]. Thus, the adhesive thickness should be chosen carefully as on the one hand, a bigger value of thickness will reinforce adhesion but dissipate the transfer of the loads towards the patch, which decreases the favorable effects of the patch. On the other hand, smaller thickness will transfer the load towards the patch but increases the risk of adhesive failure.

### 4.4 Effect of Adhesive Materials

The mechanical stresses between the composite patch and the host cracked plate increase significantly with the increase of the shear modulus of the adhesive. However, it is not recommended to increase the adhesive shear modulus indefinitely as this will reduce the adhesive strength, which can increase the risk of the adhesion failure. Thus, the adhesive choice should compromise between allowing stress transmission to the patch while preventing the failure of the adhesion.

The variation of the SIF as a function of four different adhesive materials, shown in Fig. 9, confirms the above statement. The mechanical properties of adhesive materials are illustrated in Table 7. It can be seen in Fig. 9, that when a crack length increases, the normalized SIF will decrease. This means that when the crack length increases, there is a chance that the material will fail drastically if not repaired. Therefore, to control this, the composite patch has been used and the adhesive material properties create effective shear loads that depend on shear modulus, as was highlighted in Ref. [5]. The araldite 2014 outperforms the other adhesives as it better transfers the load towards the patch, as highlighted by the higher reduction in normalized SIF.

Table 7: Mechanical properties of adhesive materials

| Material Type | E (GPa) | υ | G (GPa) | ρ kg/m$^3$ |
|---|---|---|---|---|
| Araldite 2014 | 5.1 | 0.345 | 1.89 | 1160 |
| EPON422J | 3.49 | 0.29 | 1.10 | |
| FM-47 | 2.106 | 0.3 | 0.81 | |
| FM-73 | 1.83 | 0.33 | 0.688 | 875 |

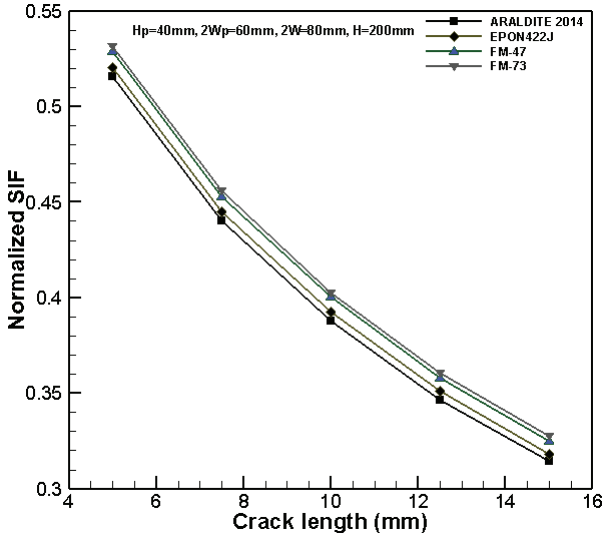



Fig. 9: Effect of adhesive materials for passive repair.

### 4.5 Effect of Crack Length

Figure 10 displays the normalized SIF variation as a function of crack length. It can be seen that the reduction of normalized SIF is higher for larger crack lengths. Indeed, the normalized stress intensity factor is maximum for smaller cracks, and then it decreases and stabilizes towards an asymptotic value for larger cracks. This behaviour leads us to state that the patch stress absorption is less significant for short cracks.

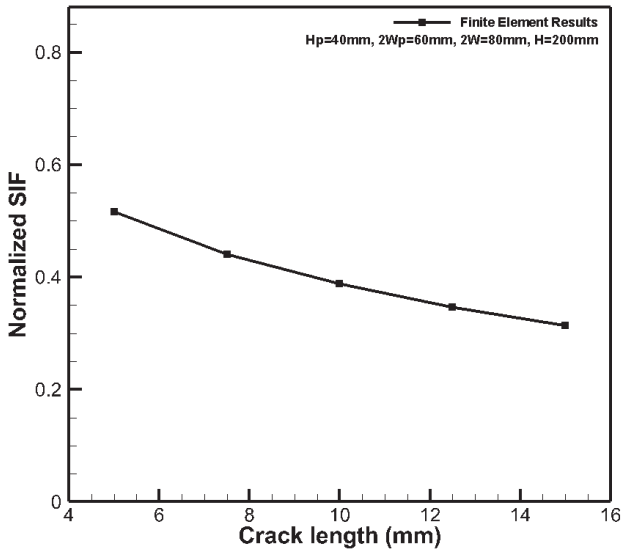



Fig. 10: Effect of crack length for passive repair.

## 5. CONCLUSION

This work describes the numerical simulation of the composite patch for a damaged rectangular plate under structural load by a finite element analysis ANSYS package. A substantial reduction in SIF is observed with an integrated composite patch. Based on present simulations, the following conclusions are drawn;

- About 68% SIF reduction is observed for single side patch repairs for boron/epoxy patches and a crack length of 15 mm.
- Increases of patch thickness reduce the SIF at the crack front.
- Increase in adhesive thickness results in a decrease of SIF reduction at the crack front.
- It has been proven that araldite 2014 adhesive material has a good mechanical property to reduce maximum SIF at the crack front.

The effect of the composite patch on the reduction of SIF on a center-cracked rectangular plate subjected to uniaxial load is simulated, and the results demonstrated that the SIF decreases linearly with the effect of the composite patch.

## ACKNOWLEDGEMENT

The authors would like to thank the Ministry of Higher Education, Malaysia for supporting this work through Research Grant FRGS 16-066-0565.

# REFERENCES

[1] Madani, K, Touzain S, Feaugas X, Benguediab M, and Ratwani M. (2009) Stress Distribution in a 2024-T3 Aluminum Plate with a Circular Notch, Repaired by a Graphite/Epoxy Composite Patch. Int. J. Adhes. Adhes., 29: 225–33.

[2] Ricci F, Franco F, Monrefusco N. (2011) Bonded Composite Patch Repairs on Cracked Aluminum Plates: Theory, Modeling and Experiments. In Advances in Composite Materials - Ecodesign and Analysis, 445-464.

[3] Ergun E, Suleyman T, Muzaffer T. (2012) Stress Intensity Factor Estimation of Repaired Aluminum Plate with Bonded Composite Patch by Combined Genetic Algorithms and FEM under Temperature Effects. Indian J. Eng. Mater. Sci., 19 (1): 17-23.

[4] Ouinas D, Belkacem A, Belabbes B. (2013) The Optimization Thickness of Single/Double Composite Patch on the Stress Intensity Factor Reduction. J. Reinf. Plast. Compos., 32 (9): 654-463.

[5] Mhamdia R, Bachir Bouadjra B, Serier B, Ouddad W, Feaugas X, Touzain S. (2011) Stress Intensity Factor for Repaired Crack with Bonded Composite Patch under Thermo-Mechanical Loading. J. Reinf. Plast. Compos., 30 (5): 416-424.

[6] Vishnuvardhan, Nadimpalli L, Himanshu P, Akhilendra S. (2016) Composite Patch Repair Modelling by FEM. Proceedings of First Structural Integrity Conference and Exhibition (SICE-2016): 4-6 July 2016.

[7] Maleki H N, Chakherlou TN. (2017) Investigation of the Effect of Bonded Composite Patch on the Mixed-Mode Fracture Strength and Stress Intensity Factors for an Edge Crack in Aluminum Alloy 2024-T3 Plates. J. Reinf. Plast. Compos., 36 (15): 1074-1091.

[8] Aabid A, Meftah H, Jaffar Syed Mohammed A, Ahmed A. (2018) Stress Concentration Analysis of a Composite Patch on a Hole in an Isotropic Plate. Int. J. Mech. Prod. Eng. Res. Dev., 6 (Special Issue): 249-255.

[9] Aabid A, Meftah H, Jaffar Syed Mohammed A, Ahmed A. (2018) Numerical Analysis of Cracks Emanating From Hole in Plate Repaired by Composite Patch. Int. J. Mech. Prod. Eng. Res. Dev., 6 (Special Issue): 238-243.

[10] Aabid A, Meftah H, Shaik Dawood MSI. (2019) Modeling Different Repair Configurations of an Aluminum Plate with a Hole. Int. J. Recent Technol. Eng., 7 (6S): 235-240.

[11] Mahadesh Kumar A, Hakeem SA. (2000) Optimum Design of Symmetric Composite Patch Repair to Centre Cracked Metallic Sheet. Compos. Struct., 49 (3): 285-292.

[12] ANSYS Inc. (2017) ANSYS FLUENT 18.0: Theory Guidance. Canonsburg PA.

[13] Ouddad W, Bouadjra BB, Belhouari M, Ziadi A. (2015) Stress Intensity Factor for Repaired Crack with Bonded Composite Patch Repair under Mechanical Loading in Aircraft Structures. In 12th Congress in Mechanical, 416-424. Casablanca (Maroc).

[14] Tada H, Paris PC, Irwin GR. (2000) The Stress Analysis of Cracks Handbook, Third Edition.

[15] Albedah A, Bachir Bouiadjra B, Mhamdia R, Benyahia F, Es-saheb M. (2011) Comparison between Double and Single Sided Bonded Composite Repair with Circular Shape. Mater. Des., 32 (2): 996-1000.

[16] Mhamdia R, Serier B, Albedah A, Bachir Bouiadjra B, Kaddouri K. (2017) Numerical Analysis of the Influences of Thermal Stresses on the Efficiency of Bonded Composite Repair of Cracked Metallic Panels. J. Compos. Mater., 51 (26): 3701-3709.