IIUM ENGINEERING JOURNAL

Volume 22 Number 2

July 2021

IIUM Press

INTERNATIONAL ISLAMIC UNIVERSITY MALAYSIA

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

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

IIUM ENGINEERING JOURNAL

AIMS & SCOPE OF IIUMENGINEERING JOURNAL

The **IIUM Engineering Journal**, published biannually (January and July), 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 at http://journals.iium.edu.my/ejournal. Further correspondence on the status of the paper could be done through the journal website.

INTERNATIONAL ADVISORY COMMITTEE

Published by:

IIUM Press

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

ISSN 1511 - 788X

9 771511 788008

IIUM Engineering Journal
ISSN: 1511-788X E-ISSN: 2289-7860

IIUM ENGINEERING JOURNAL

Volume 22, Issue 2, July 2021

https:/doi.org/10.31436/iiumej.v22i2

Table of Contents

https://doi.org/10.31436/iiumej.v22i2.1563

ACOUSTIC WAVE PROPAGATION IN HIGH SCALE IMPEDANCE MISMATCH MEDIUMS

MD RABIUL AWAL[1*], MUZAMMIL JUSOH[2], MUHAMMAD SYARIFUDDIN YAHYA[1], NURUL ADILAH ABDUL LATIFF[1], SALISA ABDUL RAHMAN[1], AHMAD NAZRI DAGANG[1], HIDAYATUL AINI ZAKARIA[1] AND SHAKIR SAAT[3]

[1]Faculty of Ocean Engineering, Technology and Informatics, Universiti Malaysia Terengganu, 21030 Kuala Nerus, Terengganu, Malaysia
[2]BioEM, School of Computer and Communication Engineering, Universiti Malaysia Perlis (UniMAP), Kampus Pauh Putra, 02600 Arau, Perlis, Malaysia
[3]School of Computing and Informatics, Albukhary International University, Jalan Langgar, Alor Setar Kedah, Malaysia

[*]Corresponding author: rabiulawal1@gmail.com

(Received:29th July 2020; Accepted: 20th February 2021; Published on-line: 4th July 2021)

ABSTRACT: A finite element analysis of acoustic propagation in a multilayered medium is presented in this paper. A circular transmitter (diameter 14 mm, thickness 3 mm) and a rectangular receiver (20×10×0.5 mm^3) are set to detect the variations in the propagation pattern. A complex medium (70×40×60 mm^3) composed of skin, fat, muscle, bone and liquid is designed in a simulated environment. A scale of frequencies (10 kHz to 2 MHz) is applied to trace the impact on the propagation pattern as well. It is found from the analysis that fat and liquid layers affect the acoustic propagation the most (-69 dB), which results in a significant drop in the received sound pressure level at the receiving end. Again, other than skin and fat layers, low frequencies (less than 1 MHz) are more beneficial in terms of sound pressure level. However, higher frequencies contribute to lower displacements at the receiving end, which will cause less power potential as well.

ABSTRAK: Analisis elemen terhingga bagi penyebaran akustik dalam medium berlapis dibentangkan dalam kajian ini. Pemancar bulat (diameter 14 mm, ketebalan 3 mm) dan penerima segi empat tepat (20 × 10 × 0.5 mm^3) diatur bagi mengesan perubahan pola penyebaran. Medium kompleks (70 × 40 × 60 mm^3) yang terdiri daripada kulit, lemak, otot, tulang dan cecair direka dalam persekitaran simulasi. Skala frekuensi (10 kHz hingga 2 MHz) digunakan bagi mengesan corak penyebaran. Dapatan kajian menunjukkan bahawa lapisan lemak dan cecair mempengaruhi penyebaran akustik (-69 dB), yang mengakibatkan penurunan mendadak tahap penerimaan tekanan bunyi di hujung penerima. Selain lapisan kulit dan lemak, frekuensi rendah (kurang dari 1 MHz) adalah lebih berguna dari segi tahap tekanan suara. Walau bagaimanapun, frekuensi lebih tinggi menyebabkan kurang anjakan di hujung penerima, sekaligus mengurangkan potensi daya tenaga.

KEYWORDS: *acoustic energy transfer; acoustic wave; ZnO, multilayered medium; impedance mismatch*

1. INTRODUCTION

Wireless power transfer (WPT) has drawn significant industrial attention in recent times. The initial intention of this scientific module was to provide mobility to handheld devices. Soon, it was realized that over kW range power can be transferred wirelessly by

WPT. To do so, inductive, capacitive, microwave, and acoustic coupling can be used. Among these, inductive WPT is widely practiced and commercially available with high efficiency and power (30 kW) [1-3]. It was recently reported that 5 to 7 meter separation distance between receiver and transmitter was achieved by inductive WPT [4, 5]. Nevertheless, inductive WPT suffers from coupling misalignment and penetration losses [6]. Capacitive WPT is still in the developing stage. Unfortunately, the inverse property of capacitance to distance limits the separation (<1 mm) and efficiency of this WPT [7-10]. Microwave WPT can provide very high efficiency and wide propagation area as well, due to its high energy density penetration [11-13]. Again, optical WPT is capable of transmitting energy over a long range (>km), however, suffers from the requirement for a line of sight [14-16].

Acoustic energy transfer (AET), unlike the aforementioned WPTs, transfers power by propagating energy as sound or vibration waves [17]. The propagated energy is then collected by a receiver which converts the vibration energy to useful electrical energy. To do so, piezoelectric conversion termed as piezoelectricity is practiced so far. Recent interest in wireless powering of implantable devices added to the need to study acoustic propagation in a complex multilayered medium with high scale impedance mismatch, which closely characterizes the human body. Hence, in this paper, we have designed a multilayer medium that is composed of the same type of layers available in the human body. We have considered a wireless pacemaker system and hence, we have limited our study to the human chest only.

The scope of this paper includes a finite element analysis of the acoustic wave propagation in a multilayered medium. The medium is specific for very high impedance mismatched layers. As we are considering AET for implants, it is worthy to investigate the impact of the medium with different physical characteristics. Also, the effect of applying a wide range of frequencies (10 kHz to 2 MHz) is considered in the scope. Nevertheless, dynamic receivers or transmitters are not considered in this study, rather we have considered the fixed locations for the transmitter and receiver.

2. MODELLING IN FINITE ELEMENT ANALYSIS

It is possible to analyze the acoustic propagation under a simulated environment using finite elements. Hence, the system is designed and analyzed with finite elements. As such, a transmitter-receiver coupling is designed for the finite element analysis. The transmitter is a simple circular transceiver with a zinc oxide (ZnO) layer and an aluminum (Al) layer. It is 14 mm in diameter and 3 mm in thickness (Al layer 1 mm and ZnO 2 mm). In contrast, the receiver is a simple ZnO rectangular block with a $20\times10\times0.5$ mm^3 dimension. The reason to keep the transmitter in circular shape is to use the benefit of the directional acoustic beams. Also, a rectangular configuration is more beneficial as receiver [18]. The transmitter and receiver are depicted in Fig. 1.

The medium is designed to resemble human chest layers. It is a $70\times40\times60$ mm^3 block with skin, fat, muscle, bone and liquid (water) layers. Hence, the medium reflects a very high scale impedance mismatch within the layers. An additional liquid block is added to compensate for the medium impedance. The medium details are given in Fig. 2 and Table 1.

https://doi.org/10.31436/iiumej.v22i2.1563

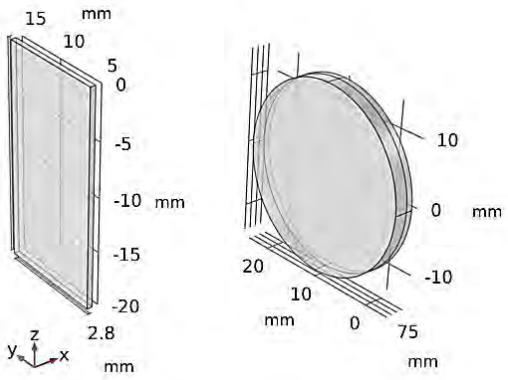


Fig. 1: Transmitter and receiver.

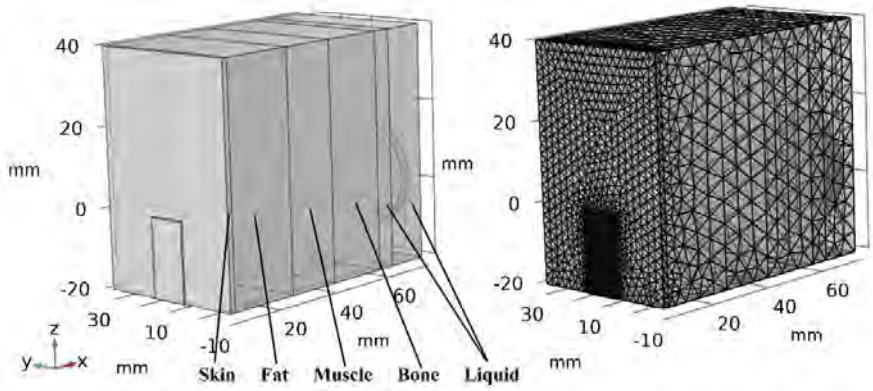


Fig. 2: Propagation medium model and mesh.

Table 1: Human chest as medium

Medium	Dimension [mm^3]	Density[kg/m^3]
Medium	70×40×60	
Skin	1.5×40×60	1109
Fat	20×40×60	911
Muscle	15×40×60	1090
Bone	17×40×60	1908
Liquid 1	5×40×60	997
Liquid 2	11.5×40×60	997

Table 2: SPL in the medium layers

Medium	Max SPL [dB]	Min SPL [dB]
Skin	129	82
Fat	129	74
Muscle	129	70
Bone	130	52
Liquid 1	126	57
Liquid 2	132	68

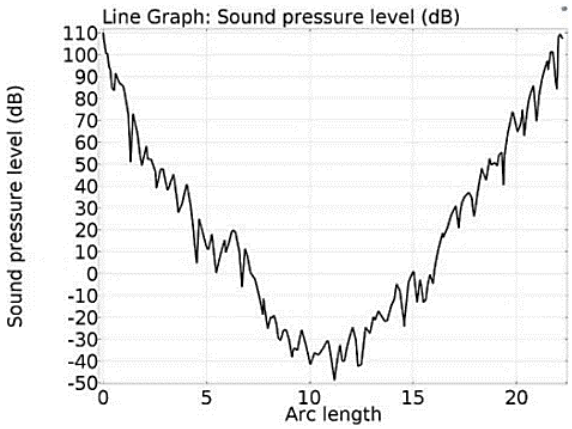


Fig. 3: Sound pressure level received by the PZT receiver.

https://doi.org/10.31436/iiumej.v22i2.1563

3. ACOUSTIC PROFILE AND ANALYSIS

Figures 3, 4, and 5 describe the propagation pattern and received sound pressure level by the rectangular receiver. From the figures, it is clear that fat and liquid layers contribute to decreasing the acoustic pressure field and sound pressure level. However, the irregular spatial acoustic beam formation is observed as well. The sound striking the receiving cantilever does so with a pressure value of approximately 111 dB to as high as 114 dB. A more focused presentation of the received sound pressure level is placed in Fig. 5. The electric potential generated in the receiver terminal has the range of -9.0320e-15 to -5.6501e-13 volt. The total displacement in the receiver is found to be 6.94e-9 mm. Again, the maximum stress on the receiver is 31.4 Pa.

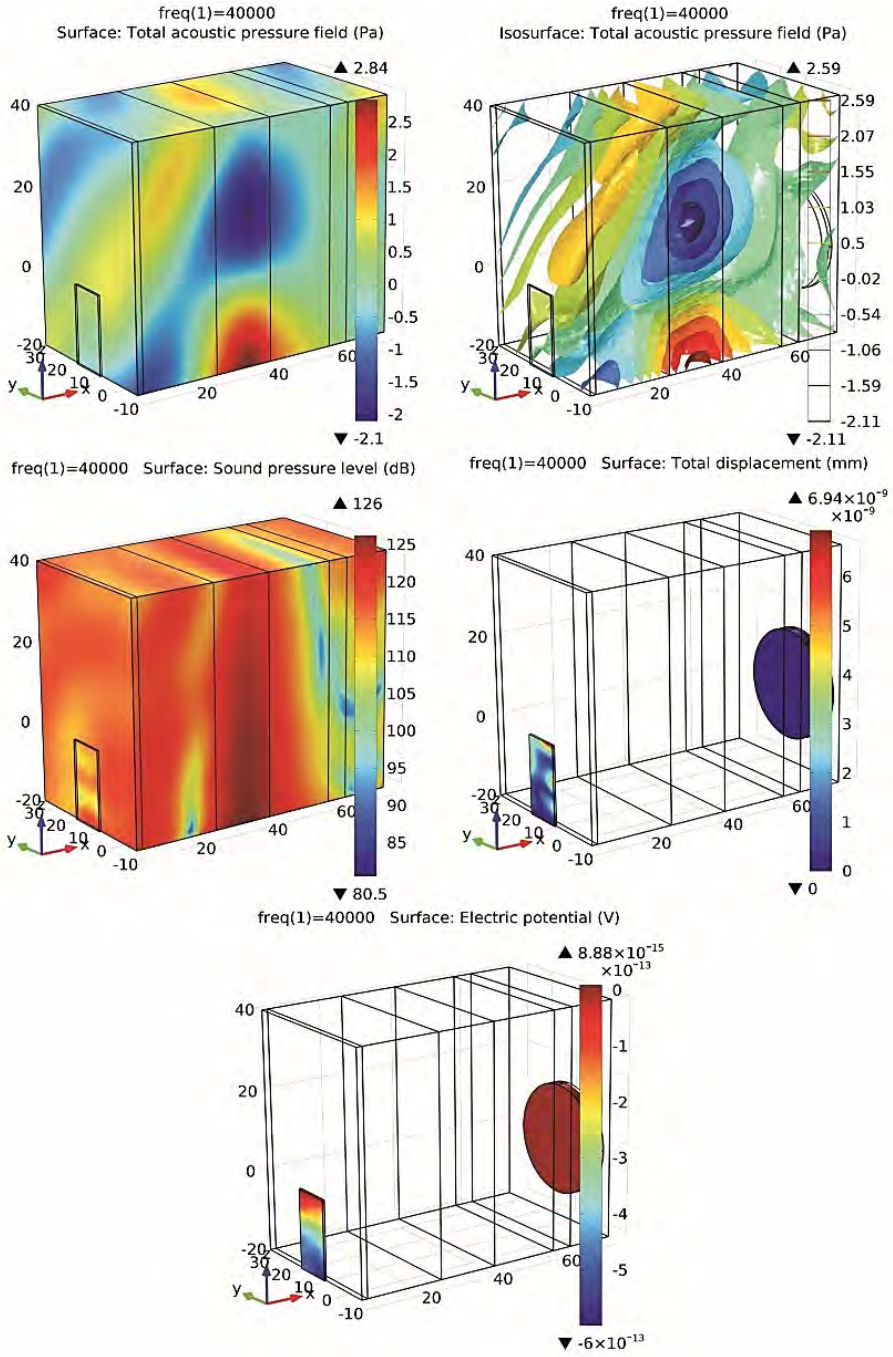


Fig. 4: Acoustic propagation within the layered medium (sample data at 40 kHz).

https://doi.org/10.31436/iiumej.v22i2.1563

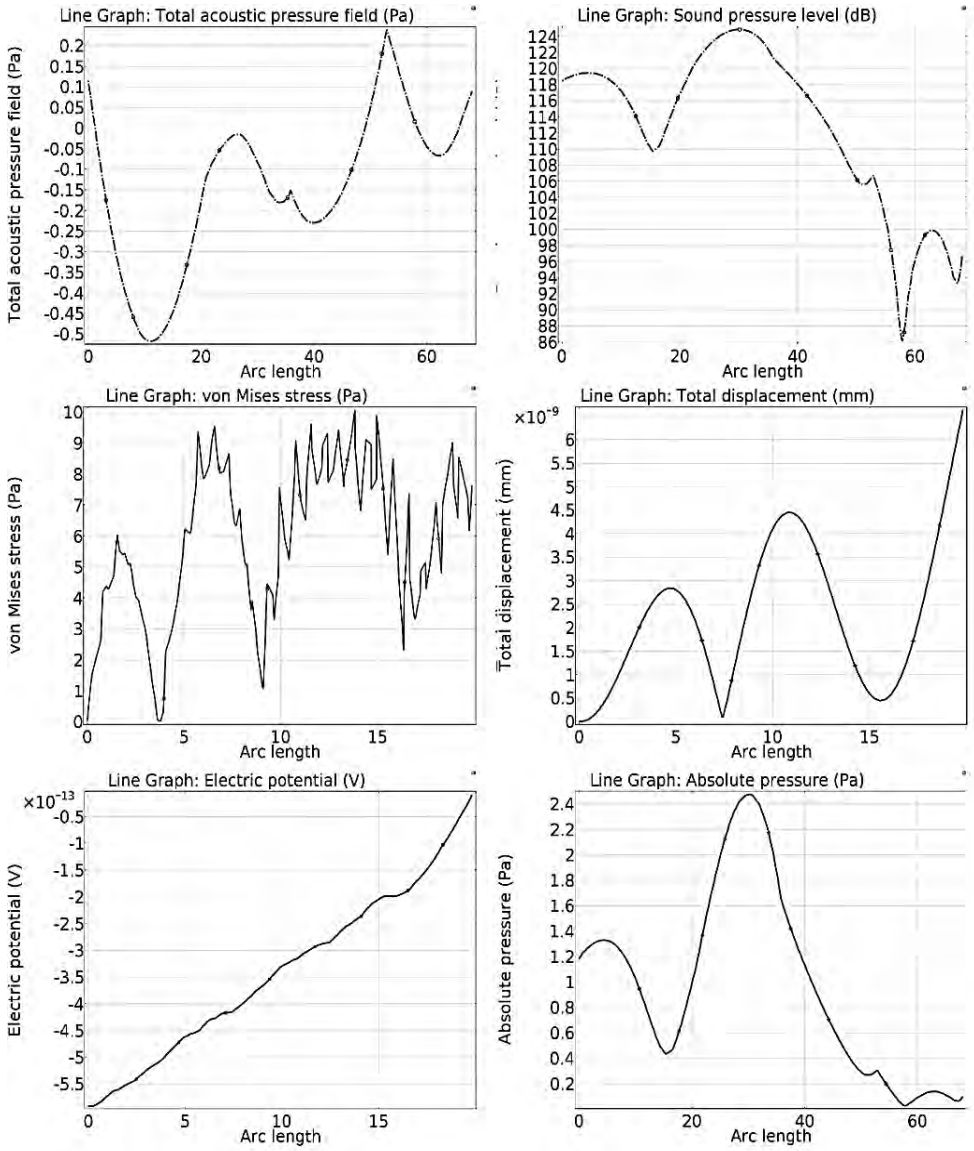


Fig. 5: Descriptions of the wave parameters (sample data at 40 kHz).

Figures 4 and 5 depict wave parameters of the transmitted acoustic waves on the attached wall with the receiver (which has a height of 60 mm) of the considered medium. The impacts of an applied frequency of 40 kHz is presented here. The results suggest that low-density mediums mostly impact the acoustic pressure fields. These result in weaker SPL to the next medium. However, if we increase the applied frequency, the reflections within the medium become more severe. In fact, at 2 MHz, the self-reflection becomes visible and mainly located between the two liquid layers. Figure 3 presents the SPL reflected by the receiver. It is clear from the figure that the reflected SPL can vary from 110 to -50 dB. Also, the reflections mainly occurred in the middle of the receiver even though it is placed at the bottom end of the medium.

Figure 6 presents the SPL against the applied frequencies on the layer with the attached receiver. The applied frequency range is 10 kHz to 2 MHz. The SPL varied from as low as 68 dB to as high as 132 dB. The lowest SPL is found at 1.45 MHz (68 dB) whereas the highest SPL values are available close to 1.21 MHz (132 dB). Hence, it is clear from the figure that the SPL distribution is not benefitted by the frequency increment.

https://doi.org/10.31436/iiumej.v22i2.1563

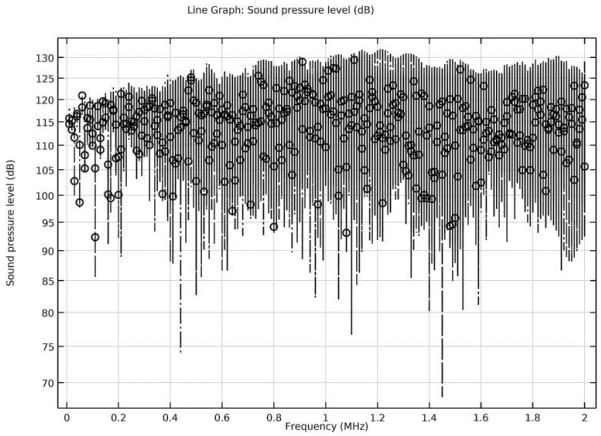


Fig. 6: SPL versus frequency plot.

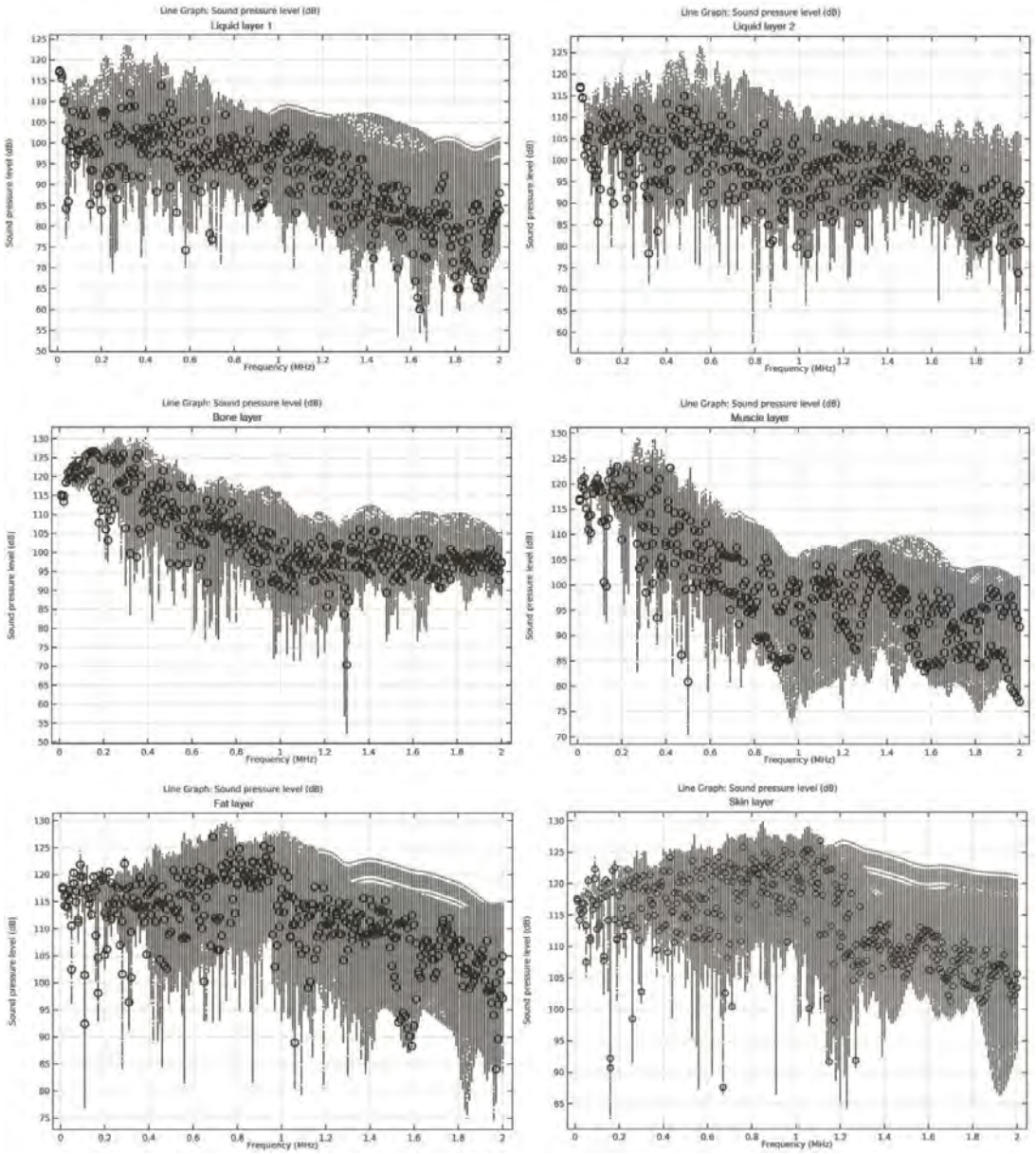


Fig. 7: SPL variations in different layers of the medium (applied frequency range 10 kHz – 2 MHz).

Figure 7 and Table 2 describe the more specified SPL profile according to the layers of the medium. According to the design, there are six layers in total between the transmitter and the receiver. Layer 1 and layer 2 are of liquid material. In layer 1, min observed SPL is approximately 68 dB at 1.45 MHz, whereas, max SPL is 132 dB at 1.21 MHz. In layer 2, min SPL drops down to approximately 57 dB at 0.79 MHz, and max SPL to 126 dB at 0.55 MHz. The SPL drops even more in layer 3 to min approximately 52 dB at 1.3 MHz whereas, max 130 dB is found at 0.27 and 0.35 MHz. Layer 4, 5, and 6 are of muscle, fat, and skin, respectively. For the muscle layer, min SPL received is approximately 70 dB at 0.5 MHz, and max is approximately 129 dB at 0.27 MHz.

Layer 5, which is the fat layer of 20 mm thickness, experiences min SPL of approximately 74 dB at 1.96 MHz, and max approximately 129 dB at 0.73, 0.74, and 0.75 MHz frequencies. Comparatively, in layer 6, the skin layer, min SPL is approximately 82 dB at 1.6 MHz, and max SPL is approximately 129 dB at 0.84, 0.85, 1.05 and 1.06 MHz frequencies.

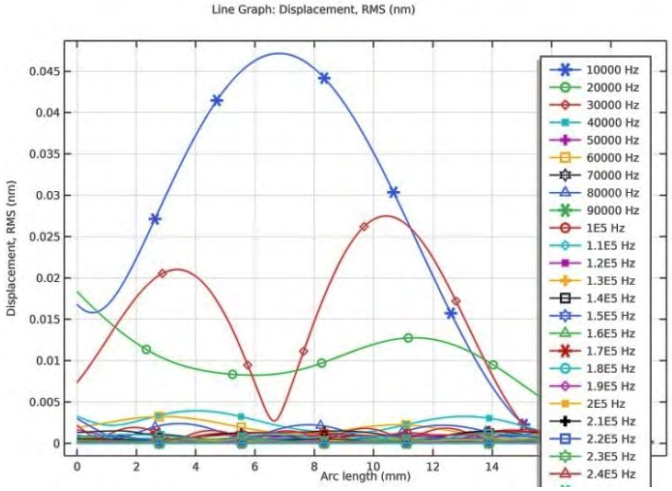


Fig. 8: Displacements occurred in the receiver (top blue, green and red are 10 kHz, 20 kHz and 30 kHz respectively).

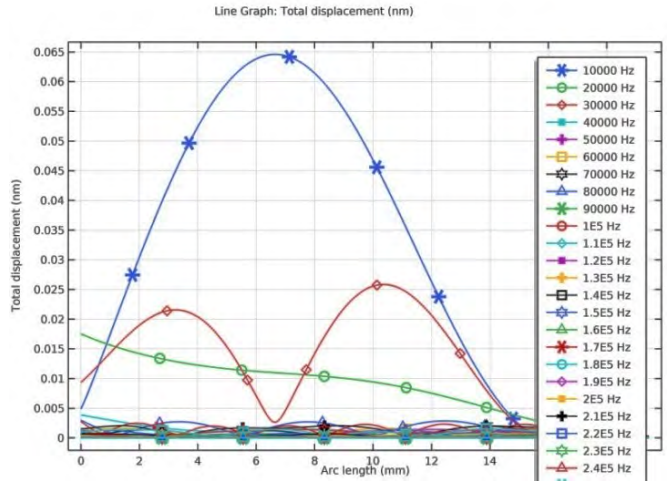


Fig. 9: Total displacements occurred in the receiver (top blue, green and red are 10 kHz, 20 kHz and 30 kHz respectively).

Figures 8 and 9 show the displacements that occurred in the receiver. Root mean square (RMS) of displacements is presented in Fig. 8 while the total displacements are shown in

Fig. 9. Interestingly, increased frequencies have decreased the displacements in the receiver. As the wavelength becomes smaller, it is losing power as well. From the figure we can see, 0.065 nm total displacements are observed (RMS 0.047 nm) for 10 kHz applied frequency. In comparison, 0.04 nm (RMS 0.028 nm) for 30 kHz, and less than 0.01 nm (RMS 0.047 nm) for 20 kHz. The rest of the frequencies will cause less than 0.004 nm displacements.

From the aforementioned results, it is noticeable that increasing frequencies is not going to be beneficial in the context of displacements that will directly impact the AET performance. For acoustic energy transfer, the transmitted energy will be collected by the receiver in the form of acoustic vibrations. Now, vibration will cause certain displacements in the receiver and the relationship is linear. Higher vibration will certainly cause higher displacements. In relation, higher displacements in piezoelectric layers will benefit with a higher electrical output which is the ultimate goal of a wireless power transfer system. Therefore, it is evident that for AET lower frequencies are preferable for implantable devices.

4. CONCLUSION

Finite element analysis of acoustic wave propagation in a very high scale impedance mismatch medium is presented in this paper. The analysis is mainly focused on acoustic energy transfer for implantable wireless devices as the medium contains a wide range of impedances. A circular transmitter and a rectangular receiver are paired to transfer acoustic waves in a multilayered medium that replicates the high scale impedance mismatch. A prescribed vibration is generated in the transmitter which is collected by the receiver to evaluate the impact of the multilayer configuration on the acoustic propagation. Sound pressure level (SPL) and receiver displacements are particularly focused to trace the effect of wide scale applied frequencies. It is found from the results that the SPL can be varied from as low as 68 dB (1.45 MHz) to as high as 132 dB (1.21 MHz). Hence, it is clear from the figure that the SPL distribution is not benefitted by the frequency increment. The electric potential generated in the receiver terminal has the range of -9.0320e-15 to -5.6501e-13 volts. The total displacement in the receiver is found to be 6.94e-9 mm. Again, the maximum stress on the receiver is 31.4 Pa. Therefore, as conclusion, the increased frequency can benefit the SPL while it decreases the displacements in the receiving end. As the displacement magnitude directly impacts the receiver performances, it is more beneficial to use less than MHz range frequencies for acoustic energy transfer applications.

ACKNOWLEDGEMENT

This work is partially supported by Universiti Malaysia Terengganu, under fundamental research grant scheme (Ref: FRGS/1/2019/TK04/UMT/03/1), Ministry of Higher Education, Government of Malaysia.

REFERENCES

[1] Kurs A, Karalis A, Moffatt R, Joannopoulos JD, Fisher P, Soljačić M. (2007) Wireless power transfer via strongly coupled magnetic resonances. Science, 317(5834): 83-86. doi: 10.1126/science.114325

[2] Wang CS, Covic GA, Stielau OH. (2004) Power transfer capability and bifurcation phenomena of loosely coupled inductive power transfer systems. IEEE Transactions on Industrial Electronics, 51(1): 148-157. doi: 10.1109/TIE.2003.822038.

[3] Wang CS, Stielau OH, Covic GA. (2005) Design considerations for a contactless electric vehicle battery charger. IEEE Transactions on Industrial Electronics, 52(5): 1308-1314. doi: 10.1109/TIE.2005.855672.

[4] Waffenschmidt E, Staring T. (2009) Limitation of inductive power transfer for consumer applications. In 2009 13th European Conference on Power Electronics and Applications (pp. 1-10). IEEE.

[5] Park C, Lee S, Cho GH, Rim CT. (2014) Innovative 5-m-off-distance inductive power transfer systems with optimally shaped dipole coils. IEEE Transactions on Power Electronics, 30(2): 817-827. doi: 10.1109/TPEL.2014.2310232.

[6] Choi, BH, Lee ES, Kim JH, Rim CT. (2014) 7m-off-long-distance extremely loosely coupled inductive power transfer systems using dipole coils. In 2014 IEEE Energy Conversion Congress and Exposition (ECCE) (pp. 858-563). IEEE.

[7] Kim A, Ochoa M, Rahimi R, Ziaie B. (2015) New and emerging energy sources for implantable wireless microdevices. IEEE Access, 3: 89-98. doi: 10.1109/ACCESS.2015.2406292.

[8] Hu AP, Liu C, Li HL. (2008) A novel contactless battery charging system for soccer playing robot. In 2008 15th International Conference on Mechatronics and Machine Vision in Practice (pp. 646-650). IEEE.

[9] Kline M, Izyumin I, Boser B, Sanders S. (2011) Capacitive power transfer for contactless charging. In 2011 Twenty-Sixth Annual IEEE Applied Power Electronics Conference and Exposition (APEC) (pp. 1398-1404). IEEE.

[10] Ludois DC, Reed JK, Hanson K. (2012) Capacitive power transfer for rotor field current in synchronous machines. IEEE Transactions on Power Electronics, 27(11): 4638-4645. doi: 10.1109/TPEL.2012.2191160.

[11] Dai J, Ludois DC. (2015) Wireless electric vehicle charging via capacitive power transfer through a conformal bumper. In 2015 IEEE Applied Power Electronics Conference and Exposition (APEC) (pp. 3307-3313). IEEE.

[12] Brown WC. (1984) The history of power transmission by radio waves. IEEE Transactions on Microwave Theory and Techniques, 32(9): 1230-1242. doi: 10.1109/TMTT.1984.1132833.

[13] McSpadden JO, Mankins JC. (2002) Space solar power programs and microwave wireless power transmission technology. IEEE Microwave Magazine, 3(4): 46-57. doi: 10.1109/MMW.2002.1145675.

[14] Karalis A, Joannopoulos JD, Soljačić M. (2008) Efficient wireless non-radiative mid-range energy transfer. Annals of Physics, 323(1):34-48. https://doi.org/10.1016/j.aop.2007.04.017

[15] Raible DE, Dinca D, Nayfeh TH. (2011). Optical frequency optimization of a high intensity laser power beaming system utilizing VMJ photovoltaic cells. In 2011 International Conference on Space Optical Systems and Applications (ICSOS) (pp. 232-238). IEEE.

[16] Cai M, Vahala K. (2000) Highly efficient optical power transfer to whispering-gallery modes by use of a symmetrical dual-coupling configuration. Optics Letters, 25(4): 260-262. https://doi.org/10.1364/OL.25.000260

[17] Awal MR, Jusoh M, Sabapathy T, Kamarudin MR, Rahim RA. (2016) State-of-the-art developments of acoustic energy transfer. International Journal of Antennas and Propagation. https://doi.org/10.1155/2016/3072528

[18] Awal MR, Jusoh M, Ahmad RB, Sabapathy T, Yasin MNM, Mat MH. (2019) Designing cantilever dimension for low power wireless applications. Indonesian Journal Electrical Engineering and Computer Science, 14(2): 758-764. http://doi.org/10.11591/ijeecs.v14.i2.pp758-764

https://doi.org/10.31436/iiumej.v22i2.1566

ENHANCEMENT OF THE CALORIFIC VALUE OF EMPTY FRUIT BUNCH (EFB) BY ADDING MUNICIPAL SOLID WASTE AS SOLID FUEL IN GASIFICATION PROCESS

AMADOU DIOULDE DONGHOL DIALLO, MA'AN FAHMI RASHID ALKHATIB*, MD. ZAHANGIR ALAM AND MAIZIRWAN MEL

Bioenvironmental Engineering Research Centre (BERC),
Department of Biotechnology Engineering, Faculty of Engineering,
International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia

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

(Received: 3rd August 2020; Accepted: 23rd January 2021; Published on-line: 4th July 2021)

ABSTRACT: Empty fruit bunch (EFB), a biomass-based waste, was deemed a potential replacement for fossil fuel. It is renewable and carbon neutral. The efficient management of this potential energy will help to deal with the problem associated with fossil fuels. However, a key parameter for evaluating the quality of raw material (EFB) as a fuel in energy applications is the calorific value (CV). When this CV is low, then its potential utilization as feedstock will be restricted. To tackle this shortcoming, we propose to add municipal solid waste to enhance energetic value. Thus, two major issues will be solved: managing solid residues and contributing an alternative energy source. This study aimed to investigate the possibility of mixing EFB and municipal solid waste (MSW) to make clean energy that is conscious of the environment (climate change) and sustainable development. The selected MSW, comprising of plastics, textiles, foam, and cardboard, were mixed, with EFB at various ratios. Proximate analysis was used to determine moisture content, ash, volatiles, and fixed carbon, whilst elemental analysis, is used to determine CHNS/O for MSW, EFB and their various mixtures. The CV of each element was also measured. The research revealed a significant increase in the calorific value of EFB by mixing it with MSW according to MSW/EFB ratios: 0.25; 0.42; 0.66; 1.00 and 1.50 the corresponding calorific values in (MJ/kg) were 19.77; 21.22; 22.67; 27.04 and 28.47 respectively. While the calorific value of pure EFB was 16.86 MJ/kg, the mixing of EFB with MSW promoted the increase in the CV of EFB to an average of 23.83MJ/kg. Another potential environmental benefit of applying this likely fuel was the low chlorine (0.21 wt. % to 0.95 wt. %) and sulfur concentrations (0.041 wt. % to 0.078 wt.%). This potential fuel could be used as solid refuse fuel (SRF) or refuse-derived fuel (RDF) in a pyrolysis or gasification process with little to no environmental effects.

ABSTRAK: Tandan buah kosong (EFB), sisa berasaskan biojisim, adalah berpotensi sebagai pengganti bahan bakar fosil. Ia boleh diperbaharui dan karbon neutral. Pengurusan berkesan pada potensi tenaga ini dapat membantu mengatasi masalah melibatkan bahan bakar fosil. Namun, kunci parameter bagi menilai kualiti bahan mentah (EFB) sebagai bahan bakar dalam aplikasi tenaga adalah nilai kalori (CV). Apabila CV rendah, potensi menjadi stok suapan adalah terhad. Sebagai penyelesaian, kajian ini mencadangkan sisa pepejal bandaran ditambah bagi meningkatkan nilai tenaga. Oleh itu, dua isu besar dapat diselesaikan: mengurus sisa pepejal dan menambah sumber tenaga alternatif. Kajian ini bertujuan mengkaji potensi campuran tandan buah kosong (EFB) dan sisa pepejal bandaran (MSW) bagi menghasilkan tenaga bersih dari

sudut persekitaran (perubahan iklim) dan pembangunan lestari. Pemilihan MSW, terdiri daripada plastik, tekstil, gabus dan kadbod, dicampurlan dengan pelbagai nisbah EFB. Analisis proksimat telah digunakan bagi mendapatkan kandungan kelembapan, abu, ruapan, dan karbon tetap, manakala analisis asas telah digunakan bagi mendapatkan CHNS/O bersama MSW, EFB dan pelbagai campuran lain. Nilai kalori (CV) setiap elemen turut diukur. Dapatan kajian menunjukkan penambahan ketara dalam nilai kalori EFB dengan campuran bersama MSW berdasarkan nisbah MSW/EFB 0.25; 0.42; 0.66; 1.00 dan 1.50 nilai kalori sepadan (MJ/kg) adalah 19.77; 21.22; 22.67; 27.04 dan 28.47 masing-masing. Manakala nilai kalori EFB tulen adalah 16.86 MJ/kg, campuran EFB dan MSW menunjukkan kenaikan CV dengan EFB pada purata 23.83MJ/kg. Antara potensi semula jadi lain adalah dengan mencampurkan bahan bakar ini dengan kalori rendah (0.21 wt. % kepada 0.95 wt. %) dan kepekatan sulfur (0.041 wt. % kepada 0.078 wt.%). Bahan bakar ini berpotensi sebagai bahan bakar pepejal sampah (SRF) atau bahan bakar yang terhasil dari pepejal sampah (RDF) melalui proses pirolisis atau proses gasifikasi yang sedikit atau tiada kesan langsung terhadap persekitaran.

KEYWORDS: *municipal solid waste; empty fruit bunch; calorific value; energy; refuse derived fuel*

1. INTRODUCTION

Until now, fossil fuels account for almost 84% of global energy demand [1], and are the most reliable sources of energy. Global production of solid waste increases with an increase in population, leading to environmental pollution [2-4]. According to Massarutto [5] the world energy consumption in 2020 amounts to 196 terawatts (TW) of which 76 for electricity and 120 for heat, and that the potential production of energy from waste maybe double the actual figures. The production of solid waste in the world in 2011 amounted to 2 billion tons of waste during the year. And by 2025, it is estimated that there will be 2.2 billion tons of waste per year, after which 9.5 billion tons of waste will be produced per year in 2050 [3,6,7].

Biomass gasification has a high potential for waste treatment compared to other existing techniques, such as soil filling, incineration, etc., because it can accept a wide range of inputs and may produce multiple useful products. Biomass gasification is an intricate process involving the drying of the feedstock followed by pyrolysis, partial combustion of the intermediates, and finally gasification of the resulting products [7]. The calorific value (CV) is the key parameter for assessing the quality of the feedstock (EFB) as fuel in energy applications. However, this calorific value for EFB is low. In order to improve this calorific value, it must be mixed with other raw materials such as MSW.

Municipal solid and biomass wastes are among the most sustainable sources of energy. Vaish et al. [8], reported that the complexity and the increasing quantity of solid waste had made MSW management a challenging task worldwide. Biomass waste is abundant in many countries over the world, like Malaysia, Indonesia, Guinea Conakry, etc. Among all, biomass waste offers significant opportunities for major, renewable, and suitable environmental-friendly energy sources. These residua, instead of being sent to landfill, could be valorized as a source of energy. Another means to manage MSW is by incineration. Heavy metals like mercury (Hg), cadmium (Cd), arsenic (As), chromium (Cr), and lead (Pb), etc., contained in fly ash can cause air pollution by incineration of MSW and soil and water pollution. The emissions of SOx, NOx, COx, and furans can pollute the environment likewise [2,9]. Moreover, greenhouse gases (GHG) emissions should be reduced by the use of suitable technologies. The third-largest source of GHG is MSW at almost 3-4% of the global anthropogenic methane, and 18% of global methane

emission came from total waste sectors [1]. As indicated in the study by Vaish et al. [8] to achieve the goal of sustainable development, the problem of climate change, and other environmental challenges, must be tackled. MSW is usually, managed through disposal at landfills, which experience severe environmental conditions, such as leachate, high salinity, and GHG generation [10]. Sikarwar et al. [7] estimate that the production of electricity from fossils contributes to pollution and the emission of GHGs.

Solid residues are an alternative to provide environmentally friendly and sustainable energy, that is economically profitable when properly managed and processed [6,11,12]. Solid refuse fuel (SRF) and refuse-derived fuel (RDF), can be manufactured in the form of pellets, bricks, etc. It is easy to transport and could contribute to the reduction of pollution problems related to discharge and provides much-needed energy, especially for people suffering a shortage of fuels. Previous studies had promoted SRF and RDF technology, including its characteristics, composition, determination of high heating value, and other parameters [13-17]. In terms of thermal conversion, gasification is one of the many routes to produce clean and environmentally friendly fuel [18,19].

The synthesis gas production by gasification is a process of recovering energy from solid fuels using a high temperature. The quality of RDF depends primarily on the composition of the raw material such as plastic and heavy polymer containers, textiles, foam, etc., which are the basis for the increase in the heating value of fuel oil [13,20-22]. In other words, a higher calorific value is associated, with the content of paper/cardboard, plastic, etc., and their presence in high quantities at (40-80% w/w; weight by weight) can promote the reduction of emissions of CO_2.

This study aims to improve the EFB's calorific value by adding specific amounts of MSW. It could be a promising cleaner alternative solution to polluted fossil fuels. The study focused on the quality of the calorific value and special importance was given to reducing the environmental pollution.

2. MATERIALS AND METHODS

The data collected for this study was based on proximate and ultimate analysis of municipal solid waste and biomass, followed by a calorific value (CV) measurement.

The analysis investigated the impact of moisture content, temperature, steam to biomass ratio, and particle size on gas composition, etc. Adequate heating values make the material promising for applications such as gasification as RDF and SRF technology [12,19].

2.1 Feed Materials

Municipal solid and biomass wastes are abundant in Malaysia; the estimated annual MSW generated is about 13.68 Mt per year, and about 1.17 kg average rate per capita per day of waste, while the amount of EFB waste was estimated at 7.78 Mt per year [17].

Municipal solid waste (plastics, textile, paper/ cardboard, and foam) was collected at the Gombak MSW transportation station landfill and biomass waste (empty fruit bunch) was collected from Sime Darby Research Center at Carey Island, (Selangor). The samples were ground to an average particle size of approximately 0.5 -1 mm, and 0.5-1 g were used as feed material.

2.2 Preparation of Municipal Solid Waste Samples

Municipal solid and biomass waste had been selected as feedstock for the experiment. Some samples (plastic, cardboard/paper, foam, and textile residue) were chosen and characterized.

These components were selected because data and statistics from National Solid Waste Management Malaysia [17] indicated that they are the major constituents of MSW. Characterized samples were dried in the sun to remove moisture. The samples were prepared to a particle size at 0.5-1 mm for foam, plastic, textile, cardboard, and empty fruit bunch.

Following these sizes, each sample was weighed into a certain scale to determine the amount needed for mixing. Then, five selected samples with different ratios were examined.

2.3 Proximate Analysis

Quantitative determination of moisture content, which has an impact on the calorific value; volatile matter, which represents the matter burns in a gaseous state; ash as inorganic waste material; and fixed carbon which amounts to the solid-state is determined using proximate analysis.

2.4 Determination of Moisture Content

Moisture content is considered an important factor that affects the fuel property, seeing that it has an impact on the combustion behavior of the material and its stability. So, the moisture content is determined using ASTM E 871 standards by measuring the weight difference after heating the sample in the oven. It is done by weighing a known mass of the samples in an alumina crucible container and placing them in the oven at a set temperature of 105 ^{o}C for 1 hour. The difference in weight was recorded and calculated as a percentage of the sample weight.

2.5 Determination of Ash Content

The experimental procedure includes preparation of MSW and EFB mixed to a well-defined proportion, then the samples of sizes between 0.5 mm and 1 mm put in an alumina crucible. The muffle furnace Linn High Therm, type: (LM.212.26 DB006031) was initially purged to remove gaseous combustibles in the furnace. The experiment is performed from ambient temperature up to the maximum temperature of 700 ^{o}C at a constant heating rate of 10 ^{o}C/min for 30 minutes (Standard method ASTM D1102-84). The sample is then cooled in air, then in a desiccator, and finally weighed.

2.6 Determination of Volatile Matter

The experimental procedure includes preparation of MSW and EFB mixed to a well-defined proportion, then, the samples of sizes between 0.5 mm and 1 mm were put in an alumina crucible. The muffle furnace Linn High Therm, type: (LM.312.06 DB004031) was initially purged to remove gaseous combustibles in a furnace. The experiment was performed from ambient temperature up to a maximum temperature of 925 ^{o}C at a constant heating rate of 10 ^{o}C/min, for seven minutes (Standard method ASTM E872). Then the sample was cooled in air, then in a desiccator, and weighed. Finally, calculations were performed to determine the percentage of volatile matter in the samples.

The fixed carbon determined through the difference of the sum of the others with the total sample. Equations (1) and (2) are for the determination of fixed carbon and volatile matter, respectively.

$$FC= 1- MC- VM- Ash \quad (1)$$

where: FC, is the fixed carbon in the MSW and EFB that remain in the char during the pyrolysis process after devolatilization. MC stands for moisture content, VM, volatile matter, and Ash is the solid residue of MSW and EFB.

The volatile matter (VM) is determined by the equation:

$$VM= \frac{\text{Loss in weight of sample at } 925\pm20^{\circ}C}{\text{weight of sample taken}} * 100 \quad (2)$$

All analyses were performed in duplicate.

2.7 Ultimate Analysis

The ultimate analysis is used to determine the percentage of the following elements by standard methods, carbon (C), and hydrogen (H), by (ASTM E-777), nitrogen was performed by (ASTM E-778), chlorine, by (ASTM E- 776-87), and sulfur using (ASTM E-775). All these samples were analyzed in Leco Series 628 CHNS. Oxygen was determined by subtracting the sum of all others cited above from the total of samples.

2.8 Heating Value

To perform the calorific value of the MSW and EFB mixture, a Parr 1341 Oxygen Bomb Calorimeter was used for the analysis. It measures the energy released when the sample undergoes complete combustion in the presence of oxygen under a standard condition.

2.9 Chemical Composition of MSW and EFB

The chemical composition of MSW and EFB is shown in Table 1. Table 2 illustrates the ratio of the MSW and EFB mixture.

Table 1: Chemical composition of MSW

No	Components	MSW Percentage (%)	EFB Percentage (%)
1	C	52.96	41.2
2	H	6.58	6.36
3	O	36.78	47.70
4	N	0.65	0.74
5	S	0.028	0.09
6	Cl	0.24	1.010

Table 2: A mixing ratio of MSW and EFB

Samples No	Mixed elements wt.%
1	20 (MSW) +80 (EFB)
2	30 (MSW) +70 (EFB)
3	40 (MSW) +60 (EFB)
4	50 (MSW) +50 (EFB)
5	60 (MSW) +40 (EFB)

3. RESULTS AND DISCUSSION

The analyses of MSW and EFB were carried out individually and for their mixtures in predetermined proportions. Then the calorific value was measured for each sample.

Figure 1 shows the results of the proximate analysis. The minimum value of volatile matter (2.96 wt.%) was found for EFB, while the maximum amount (3.50 wt.%) belonged to MSW. EFB has the highest moisture and ash content (15.4 wt.% and 3.9 wt.% respectively) and the lowest held by MSW (7.5 wt.% and 2.99 wt.% respectively). EFB gasification would incur an additional cost for drying due to high moisture content. EFB had the lowest fixed carbon value compared to MSW (77.74 wt.% and 86 wt.% respectively). The higher fixed carbon, low moisture content, moderate volatile matter, and ash, thus resulting in higher heating value for MSW. A similar result was reported by Afzanizam et al. [23].

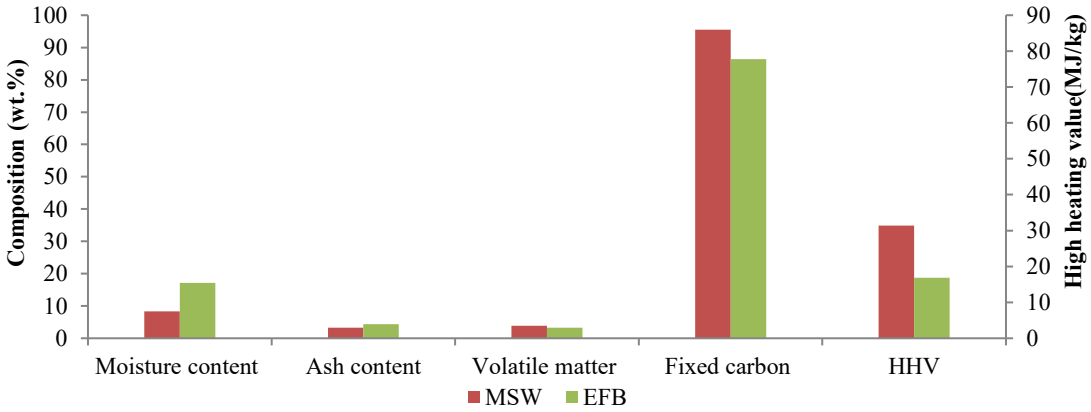


Fig. 1: Proximate analysis of MSW and EFB.

The effect of the mixture (MSW and EFB) ratio on proximate analysis is shown in Fig. 2. It can be seen that the MSW and EFB mixture of 60:40 % has the lowest ash (2.33 wt.%). The highest ash yield (4.22 wt.%), moisture content (12.43 wt.%), volatile matter (3.5 wt.%), and fixed carbon (92.03 wt.%) belong to (50:50; 20:80; 20:80 and 60:40 respectively). Moreover, the moisture content and volatile matter decrease as the quantity of MSW supplied in the mixture increases. Also, ash, fixed carbon, and the calorific value increase proportionally with the increase of MSW in the mixture.

The ultimate analysis of pure MSW and EFB is shown in Fig. 3. The MSW has the highest percentage of carbon and hydrogen (52.96 wt.% and 6.56 wt.% respectively) and the lowest of oxygen and nitrogen (36.79 wt.% and 0.65 wt.% respectively). These two elements, carbon, and hydrogen are significant in the fuel because they increase the calorific value. Similar findings were observed by [23,24]. The high quantity of carbon and hydrogen implies that this raw material could be used as a fuel in thermochemical energy conversion like pyrolysis and gasification for syngas production.

It appears that carbon-hydrogen and calorific value increase with the elevation of the amount of MSW in the mixture. In other words, the calorific value increases proportionally, with MSW due to higher carbon and hydrogen content. In contrast, a small amount of nitrogen promotes the quality of the fuel, because it has no calorific value. The highest value of sulfur found in EFB (0.0908 wt.%) followed by cardboard (0.0886 wt.%) and textile /foam (0.0594 wt.%) and the lowest in plastics (0.00751 wt.%) followed by

MSW (0.0287 wt.%). Sulfur also increases the value of fuel, but a large amount of sulfur leads to a smoky flame and it is harmful to the environment.

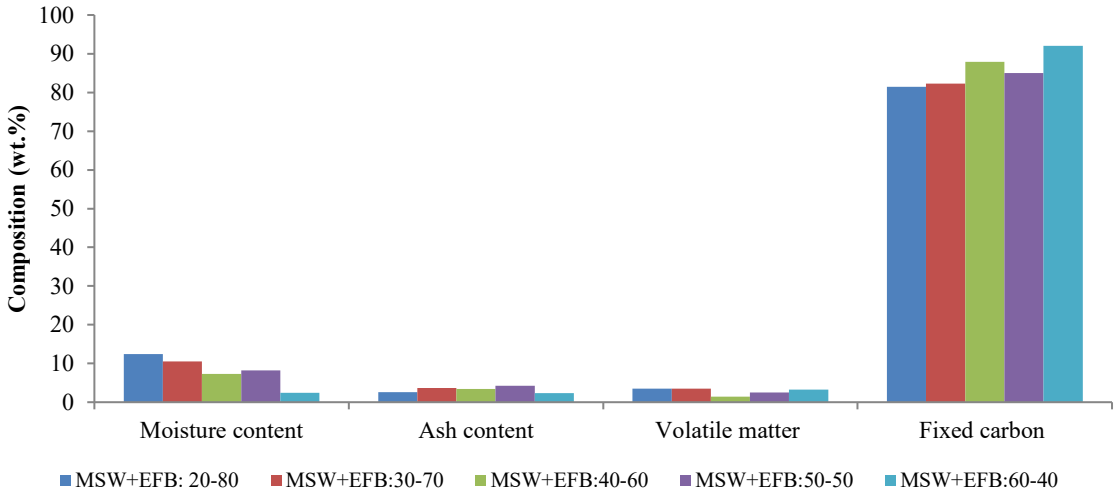


Fig. 2: Proximate analysis of MSW and EFB mixture.

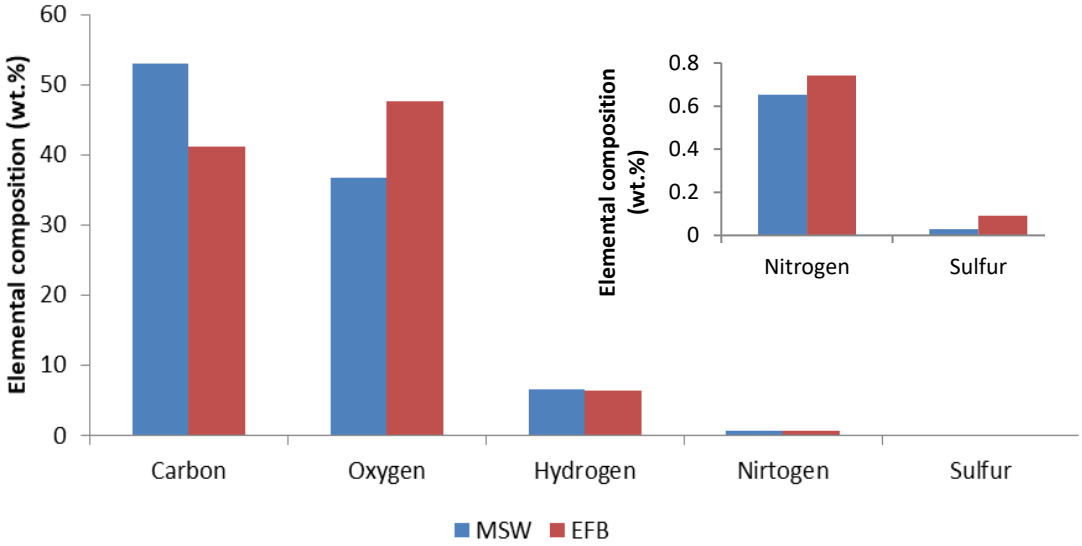


Fig. 3: Ultimate analysis of MSW and EFB (Inset: A zoom in for elemental composition of nitrogen and sulfur for the same materials).

When the proportion of MSW and EFB is (50:50), carbon, and hydrogen got the highest values (52.50 wt.%) and (7.95 wt.%) respectively, while Oxygen and nitrogen have the smallest concentration (34.71 wt.%, and 0.33 wt.% respectively) as seen in Fig. 4.

The comparison of carbon concentrations in Fig. 3 and Fig. 4 demonstrates a gain of 11 wt.% for EFB, which was 41 wt. % for pure EFB compared to 52 wt.% for MSW and EFB mixture of 50:50%. The carbon is one of the sources of the calorific value, the mix of MSW and EFB allows increasing the heating value of EFB. Besides, the calorific value of the mixture gradually increases from 19.77 MJ/kg to 28.47 MJ/kg, as the added amount of MSW increases. It seems likely that these results are due, in fact, to the constituent elements of MSW including carbon, and hydrogen in the plastics portions (72.84 wt.% and 9.46 wt.% respectively). Figure 5 shows the gradual increase in the calorific value of the MSW and EFB mixture.

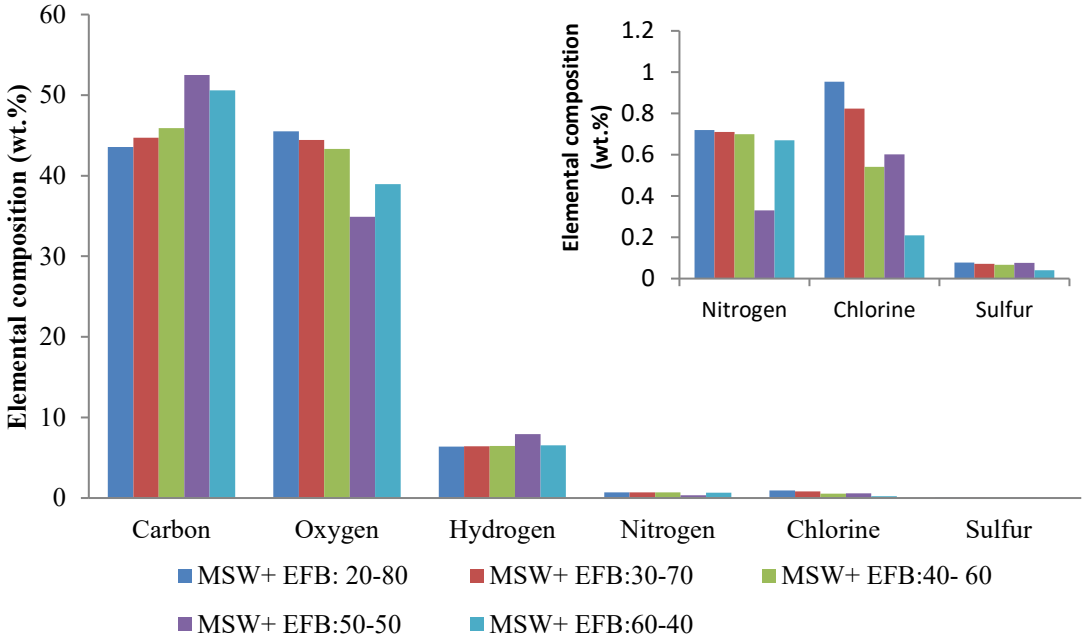


Fig. 4: Ultimate analysis of MSW and EFB mixture (Inset: A zoom in for elemental composition of nitrogen, chlorine and sulfur for the same materials).

Hydrogen sulfide reacts with metals to produce the corresponding metallic sulfide. Thus, chlorine is deposited and poses a technical (i.e. corrosion of the wall of the device) and environmental problem during the gasification process. The concentration of chlorine in these samples varied from (0.154 wt.% cardboard to 1.010 wt.% EFB). The average concentration of each component namely cardboard (0.154 wt.%), plastics (0.166 wt.%), textile and foam (0.531 wt.%), municipal solid waste (0.249 wt.%), and empty fruit bunch (1.010 wt.%) was measured.

The finding shows the mixtures of EFB and MSW have high potential to produce RDF or SRF fuel. Based on the European standard EN 15359, the components used in this study show that MSW with 0.249 wt.% of Cl, and 29.74 MJ/kg as low heating value (LHV) is class 1, textiles (0.531 wt.% Cl) class 2, EFB (1.010 wt.% Cl) and 15.09 MJ/kg as low heating value (LHV) is class 3, thus the mixture EFB and MSW (0.21 wt.% Cl), and (0.041 wt.% S) is class1 [16] as shown in Table 3.

The high heating value (HHV) can be converted into the low heating value (LHV) using the following formula:

$$LHVi = HHVi - We.\ (9.Hi + Wi) \quad (3)$$

with LHVi : LHV of i^{th} waste fraction, HHVi: HHV of i^{th} waste fraction, We: standard heat of evaporation of water (2.441 MJ/kg), Hi: hydrogen content of i^{th} waste fraction and Wi : moisture content of i^{th} waste fraction [12]. From the result of Fig. 5, where the calorific value was expressed in HHV and the Table 3 where the calorific value was expressed in LHV, formula (3) was used for the calculation of LHV.

The mixture having a ratio of 60 (MSW) and 40 (EFB) has proven to be the best among others as it has the greatest calorific value (28.47 MJ/kg) and lowest concentrations

of chlorine (0.21 wt.%) and sulfur (0.041 wt.%). Thus, in terms of energy and environment, this mixture can be a promising raw material in the gasification process.

Compared to the European standard, the results of this study show that significant pollutant components such as mercury (Hg), arsenic (As), bromine (Br) are not found in these raw materials. Sulfur (S) and chlorine are lower, hence their compliance with the environmental standard to be used as feedstock in the gasification process. Furthermore, in this study, mercury is not found, being an added advantage for this promising fuel

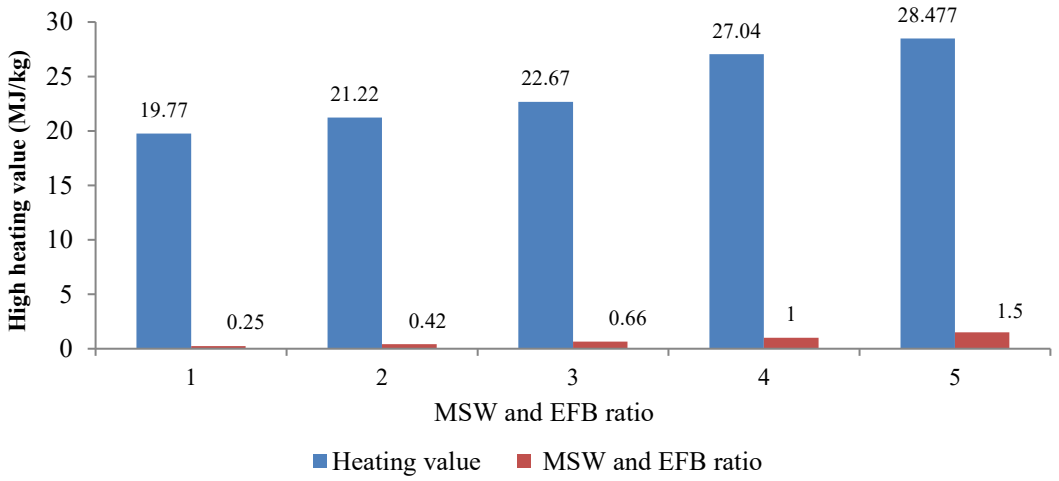


Fig. 5: The high heating value of MSW and EFB mixture.

Table 3: Waste classification criteria as SRF, according to EN 15359 [16]

Parameter	UNIT	CLASSES					CURRENT STUDY		
		1	2	3	4	5	MSW	EFB	60 (MSW) + 40 (EFB)
Lower heating value	MJ/kg	≥25	≥ 20	≥ 15	≥ 10	≥ 3	29.74	15.1	26.98
Chlorine content	%(w/w)	≤0.2	≤ 0.6	≤ 1.0	≤ 1.5	≤ 3	0.249	1.01	0.21
Mercury content	Mg/MJ	≤0.02	≤0.03	≤0.08	≤0.15	≤0.5	0	0	0

4. CONCLUSIONS

The goal of this investigation was to assess the efficacy of mixing municipal solid waste and biomass to improve the calorific value. The materials establish that the concentrations of carbon, hydrogen, nitrogen, and oxygen obtained from the analyses carried out are sufficient to produce a fuel of high calorific value. Moreover, the environmental parameters (Cl and S) are within the prescribed standards. These findings show that the calorific value (CV) of EFB which is 16.86 MJ/kg can be enhanced depending on the quantity of added MSW and can reach up to 28.47MJ/kg. This potential energy can be used as a raw material in a pyrolysis or gasification process with little to no environmental impacts.

REFERENCES

[1] Kumar A, Samadder SR. (2017) A review on technological options of waste to energy for effective management of municipal solid waste. Waste Management, 69: 407-422. https://doi.org/10.1016/j.wasman.2017.08.046

[2] Liu Z. (2019) Gasification of municipal solid wastes: a review on the tar yields. EnergySources, Part A: Recovery, Utilization and Environmental Effects, 41(11): 1296-1304. https://doi.org/10.1080/15567036.2018.1548508

[3] Lombardi L, Carnevale E, Corti A. (2014) A review of technologies and performances of thermal treatment systems for energy recovery from waste. Waste Management, 37: 26-44. https://doi.org/10.1016/j.wasman.2014.11.010

[4] Makarichi L, Jutidamrongphan W, Techato K. (2018) The evolution of waste-to-energy incineration: A review. Renewable and Sustainable Energy Reviews, 91: 812–821. https://doi.org/10.1016/j.rser.2018.04.088

[5] Massarutto A. (2014) Economic aspects of thermal treatment of solid waste in a sustainable WM system. Waste Management. https://doi.org/10.1016/j.wasman.2014.08.024

[6] Beyene HD, Werkneh AA, Ambaye TG. (2018) Current updates on waste to energy (WtE) technologies: A review. Renewable Energy Focus, 24: 1-11. https://doi.org/10.1016/j.ref.2017.11.001

[7] Sikarwar VS, Zhao M, Clough P, Yao J, Zhong X, Memon MZ, Shah N, Anthony EJ, Fennell PS. (2016) An overview of advances in biomass gasification. Energy and Environmental Science, 9(10): 2939-2977. https://doi.org/10.1039/c6ee00935b

[8] Vaish B, Sharma B, Srivastava V, Singh P, Ibrahim MH, Singh RP. (2019) Energy recovery potential and environmental impact of gasification for municipal solid waste. Biofuels, 10(1): 87-100. https://doi.org/10.1080/17597269.2017.1368061

[9] Nobre C, Gonçalves M, Vilarinho C. (2019) A brief assessment on the application of torrefaction and carbonization for refuse derived fuel Upgrading. Lecture Notes in Electrical Engineering, 505: 633-640. https://doi.org/10.1007/978-3-319-91334-6_86

[10] He P, Chen L, Shao L, Zhang H, Lü F. (2019) Municipal solid waste (MSW) landfill: A source of microplastics? Evidence of microplastics in landfill leachate. Water Research, 159: 38-45. https://doi.org/10.1016/j.watres.2019.04.060

[11] Brunner PH, Rechberger H. (2015) Waste to energy – key element for sustainable waste management. Waste Management, 37: 3-12. https://doi.org/10.1016/j.wasman.2014.02.003

[12] Chol ON, Song PH, Chol SY, Hyok RY, Nam KY. (2018) Environmental effects A feasibility study of energy recovery of RDF from municipal solid waste. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects, 40(24): 2914-2922. https://doi.org/10.1080/15567036.2018.1514431

[13] De Gisi S, Chiarelli A, Tagliente L, Notarnicola M. (2018) Energy, environmental and operation aspects of a SRF-fired fluidized bed waste-to-energy plant. Waste Management, 73: 271-286. https://doi.org/10.1016/j.wasman.2017.04.044

[14] Elisabete M, Brás I, Silva ME. (2017) Refuse Derived Fuel from Municipal Solid Waste rejected fractions- Case Study The 15th International a Symposium on District Heating and Cooling. Energy Procedia, 120: 349-356. https://doi.org/10.1016/j.egypro.2017.07.227

[15] Farzad S, Mandegari MA, Görgens JF. (2016) A critical review on biomass gasification, co-gasification, and their environmental assessments. Biofuel Research Journal, 3(4): 483-495. https://doi.org/10.18331/BRJ2016.3.4.3

[16] Garcés D, Díaz E, Sastre H, Ordóñez S, González-lafuente JM. (2015) Evaluation of the potential of different high calorific waste fractions for the preparation of solid recovered fuels. Waste Management, 47(B): 164-173. https://doi.org/10.1016/j.wasman.2015.08.029

[17] Hamzah, N., Tokimatsu, K., & Yoshikawa, K. (2019). Solid fuel from oil palm biomass residues and municipal solid waste by hydrothermal treatment for electrical power generation in Malaysia: A review. Sustainability (Switzerland), 11(4): 1-23. https://doi.org/10.3390/su11041060

[18] Abdullah N, Sulaiman F, Taib RM. (2013) Characterization of banana (*Musa* spp.) plantation wastes as a potential renewable energy source. AIP Conference Proceedings, 1528: 325-330. https://doi.org/10.1063/1.4803618

[19] Aluri S, Syed A, Flick DW, Muzzy JD, Sievers C. (2018) Pyrolysis and gasification studies of model refuse derived fuel (RDF) using thermogravimetric analysis. Fuel Processing Technology, 179: 154-166. https://doi.org/10.1016/j.fuproc.2018.06.010

[20] Nizami AS, Shahzad K, Rehan M, Ouda OKM, Khan MZ, Ismail IMI, Almeelbi T, Basahi J. M, Demirbas A. (2017) Developing waste biorefinery in Makkah: A way forward to convert urban waste into renewable energy. Applied Energy, 186: 189-196. https://doi.org/10.1016/j.apenergy.2016.04.116

[21] Ouda OKM, Raza SA, Nizami AS, Rehan M, Al-waked R, Korres NE. (2016) Waste to energy potential: A case study of Saudi Arabia. Renewable and Sustainable Energy Reviews, 61: 328-340. https://doi.org/10.1016/j.rser.2016.04.005

[22] Komilis D, Kissas K, Symeonidis A. (2014) Effect of organic matter and moisture on the calorific value of solid wastes: An update of the Tanner diagram. Waste Management, 34(2):255. https://doi.org/10.1016/j.wasman.2013.09.023

[23] Afzanizam N, Nazri M, Jaafar M, Tung C, Johan N. (2015) A review of palm oil biomass as a feedstock for syngas fuel technology. Jurnal Teknologi, 5: 13-18.

[24] Onoja E, Chandren S, Abdul Razak FI, Mahat NA, Wahab RA. (2018) Oil alm (*Elaeis guineensis*) biomass in Malaysia: The present and future prospects. Waste and Biomass Valorization, 10: 2099-2117. https://doi.org/10.1007/s12649-018-0258-1

https://doi.org/10.31436/iiumej.v22i2.1608

MICROWAVE IRRADIATION OPTIMIZATION FOR EFFICIENT LIGNIN REMOVAL FROM COCOA SHELL WASTE USING ALKALI

FARAH HAFIDZAH AHMAD[1], NURHAMIZAH IBRAHIM[1], MOHD SABRI MAHMUD[1], WALEED ALI[1], MOHD BIJARIMI MAT PIAH[1], SURIYATI SALLEH[2], AND SAID NURDIN[3]

[1]Department of Chemical Engineering, College & Natural Resources Engineering,
[2]Faculty of Chemical and Process Engineering Technology,
Universiti Malaysia Pahang, Gambang, 26300 Kuantan, Pahang, Malaysia
[3]Chemical Engineering Department, Faculty of Engineering-Syiah Kuala University, Darussalam, Bandar Acheh, 23111 Indonesia

[*]Corresponding author: mohdsabri@ump.edu.my

(Received: 10th September 2020; Accepted: 5th March 2021; Published on-line: 4th July 2021)

ABSTRACT: This paper reports a study to determine the optimum conditions of microwave in assisting alkali treatment for removing lignin from cocoa shell waste (CSW). The CSW was mixed with 5% of NaOH solution at the ratio of 1: 10 of weight to volume of the alkaline before being irradiated in a microwave oven. Various microwave powers (200-400 W), temperature settings (60-80 °C) and irradiation times (10-20 min) were tested on 15 samples set by the Box-Behnken design. The lignin removal was analysed using a 72% sulfuric acid treatment method. A quadratic equation was employed to the response surface and statistical analysis conducted to confirm the adequacy of the model. The plots show that the optimum microwave conditions are 400 W, 76 °C and 19 min, which were capable to remove 86.57% of lignin. Thermogravimetric analysis and micrographs revealed different decomposition temperature of lignin and morphology of extensively-pored surface of treated CSW, respectively.

ABSTRAK: Kajian ini berkaitan penentuan keadaan optimal ketuhar gelombang mikro bagi membantu membuang lignin daripada sisa kulit biji koko (CSW) menggunakan rawatan alkali. CSW dicampurkan dengan larutan NaOH 5 % pada nisbah 1:10 berat kepada isipadu larutan alkali sebelum campuran dipanaskan ke dalam ketuhar gelombang mikro. Pelbagai ujian dibuat pada ketuhar gelombang mikro seperti tenaga (200-400 W), suhu ketuhar (60-80 °C) dan masa pemanasan (10-20 min) ke atas 15 sampel mengikut reka bentuk Box-Behnken. Kadar pembuangan lignin ditentukan dengan menggunakan kaedah rawatan larutan asid sulfurik berkepakatan 72%. Persamaan kuadratik telah digunakan ke atas permukaan respon dan analisis statistik telah dilakukan bagi memastikan kesesuaian model. Plot-plot menunjukkan keadaan optima ketuhar gelombang mikro adalah pada 400 W, 76 °C dan 19 min iaitu berupaya menyingkirkan sebanyak 86.57% lignin. Analisis thermogravimetri dan mikrograf masing-masing menunjukkan perbezaan suhu penguraian lignin dan morfologi permukaan CSW yang dirawat didapati berliang dengan banyaknya.

KEYWORDS: *cocoa shell waste; microwave assisted; alkali treatment; lignin removal; non-wood fibers; response surface methodology*

1. INTRODUCTION

By-product discarded from food industries may harm environment in large quantities [1]. Cocoa (*Theobroma cacao*) shell waste (CSW) is an example of the by-products that is usually used, among others, for dietary, food additives, fuel burning, or fertilizer in order to prevent accumulation [2–6]. An alternative approach to add value is by utilizing CSW as non-wood for bio-composite preparation. However, there is a drawback that needs to be considered since CSW is mainly composed of cellulose, lignin, hemicelluloses, pectin, and wax. Weak binding between fiber and polymer due to moisture absorption [7], low impact strength [8], faced dimension instability for fibers composites [9] and low microbial resistance [10] of CSW application are reported that are mainly caused by the presence of hydroxyl groups from lignin in its fiber. Therefore, reduction of lignin is essentially required.

Pre-treatment using alkali is a vital step in converting the lignocellulosic material into less wood material. Several studies suggested the use of microwave heating to accelerate the pre-treatment process due to some advantages over conventional heating [11,12]. In particular, microwave-assisted alkaline pre-treatment for hydrolysis and fermentation increase cellulose breakdown and the activity of the reactions [13]. Interface of fibers with lower hydroxyl group will bond with matrix better and enhance strengths of the composites [9,14]. Sodium hydroxide (NaOH) is an effective alkaline in breaking down lignin structure [15,16] that leads to an increase in surface area and pore volume in the substrate. Meanwhile, microwave radiation results in rapid heating due to interaction between an irradiated object and electromagnetic field [17,18]. Thus, this high efficiency of treatment eases the implementation [19,20] of breaking down lignocellulosic material bonding.

Several studies reported on the optimization of treatment process parameters for other agricultural wastes. For instance, Sombatpraiwan et al. [17] conducted microwave –assisted on cassava rhizome by varying microwave power (300-900 W), irradiation time (5-15 min), and NaOH concentration (3%–7% w/v). It was found that optimal conditions for lignin removal were 840 W microwave, 9 min, and 3% w/v NaOH. Lai et al. [21] found that 22.38% lignin removal was optimal at 900 watt, 100 ºC, and 80 min from a two-level factorial analysis on the range of microwave power, temperature and heating duration range spanning from 500 to 900 watt, 80 to 100 °C and 40 to 80 min on oil palm trunk.

The aim for this study is to determine optimal microwave–assisted alkali treatment condition on CSW. The pre-treatment parameters are microwave power, temperature setting and irridiation time, which were varied between 200 – 400 W and 60 – 80 ℃ for 10 – 20 min, respectively. This treatment was optimized by using the Box–Behnken design (BBD), response surface methodology (RSM).

2. MATERIALS AND METHODS

2.1 Assisted Alkali Treatment

CSW samples were mixed to 5 % of NaOH at 1: 10 according to the ratio of sample weight to volume of the alkaline. The microwave was set to the ranges of 200 – 400 W, 60 – 70 ℃ and 10 – 15 min for power, temperature and time of irradiation, respectively. All these treatment conditions were designed based on BBD using Design Expert Software version 7.

After the treatment, mixtures were filtered to recover solid residue. The residue was rinsed with distilled water to remove the alkali solution and followed by the addition of 5 %

acetic acid solution in dropwise. Then, the slightly acidic residue was rinsed again with distilled water until neutral pH was achieved. The treated CSW was dried at 105 °C until constant weight was achieved.

2.2 Experimental Design

BBD was used to optimize microwave parameters as the factor in assisting the caustic removal of lignin from CSW via 15 experiments in random order with three replications at the centre point. The design of experiment was of three-factor and two-level factorial Box–Behnken design (BBD) [17]. The three factors were microwave power (X1), temperature (X2), and treatment time (X3) with low, moderate, and high levels factor as represented by −1, 0, and 1 as shown in Table 1. Percentage of lignin removal was the response variable. 3D surface plots were generated through fitting of a quadratic polynomial between the factors and the response.

Table 1 Experimental levels of three independent variables

Factors	Units	Factor Levels		
		-1	0	1
Microwave Power	W	200	300	400
Temperature Setting	°C	60	70	80
Irradiation Time	min	10	15	20

2.3 Lignin Removal Determination

Two grams of untreated / treated CSW samples were mixed with 15 mL of 72% sulphuric acid in a flask. The mixtures were vigorously stirred at ambient temperature for 2 hours. Then, hot distilled water was added to the mixtures to reduce acid concentration up to 4%. The mixture was heated immediately in a water bath at 100 °C. After 4 hours, the CSW was filtered and rinsed with hot distilled water to neutralize pH before it was dried at 105 °C until constant weight was achieved. The amounts of lignin removal for all the samples were determined by using Equation (1).

$$\text{lignin removal } \% = \frac{\text{initial CSW sample} - \text{treated CSW sample}}{\text{initial CSW sample}} \times 100\% \quad (1)$$

2.4 Analyses of Thermogravimetry and Morphology

The degradation of the treated CSW sample was measured by thermogravimetric analysis (TGA) using the Hitachi STA-7200 thermal analyser. CSW samples (5 g) were placed in alumina pan and heated from ambient temperature up to 1000 °C at a heating rate of 10 °C min^{-1} under nitrogen flow blanketing. The sample weights were obtained as a function of temperature. The effect of treatment on surface morphology of CSW were observed via scanning electron microscopy (SEM). The plate with CSW samples was and placed at LEO 1450 VP SEM for the analysis with magnifications of 5000 times.

3. RESULTS AND DISCUSSION

3.1 Effect of Process Parameters on Lignin Removal

Box-Behnken design (BBD) was used to determine the samples and plot the responses surface. Fifteen runs were carried out to obtain the results in Table 2. In particular, experimental run#8 (400 W, 70 °C, 20 min) achieved the highest lignin removal, while run#14 (200 W, 70 °C, 10 min) obtained the lowest one.

https://doi.org/10.31436/iiumej.v22i2.1608

Table 2: Lignin removal of CSW samples using different process parameters

Run	Microwave Power Level	Temperature Level	Irradiation Time Level	Lignin Removal (%) Experimental	Lignin Removal (%) Predicted
1	+1	-1	0	63.07	66.06
2	0	-1	-1	61.79	57.54
3	0	+1	-1	54.47	53.80
4	+1	+1	0	83.01	82.42
5	-1	0	+1	50.19	48.94
6	0	-1	+1	61.43	62.14
7	0	+1	+1	64.17	68.41
8	+1	0	+1	86.34	82.69
9	0	0	0	73.52	69.86
10	-1	+1	0	33.75	30.76
11	-1	-1	0	43.97	44.55
12	+1	0	-1	74.67	75.93
13	0	0	0	67.12	69.86
14	-1	0	-1	32.87	36.52
15	0	0	0	68.94	69.86

The effect of three process variables on lignin residue was generally significant as confirmed by the high R^2 value (R = 0.97). The quadratic equation established the level of lignin removal as a function of different microwave power, temperature setting and irradiation time. All the terms regardless of their significance are included in Equation (2). After considering the significance of the terms, Equation (3) was obtained where the symbols were denoted as A = microwave power, B = temperature setting and C = irradiation time.

$$\text{Lignin Removal (\%)} = 69.86 + 18.29*A + 0.64*B + 4.79*C + 7.54*A*B - 1.41*A*C + 2.51*B*C - 6.68*A^2 - 7.23*B^2 - 2.16*C^2 \quad (2)$$

$$\text{Lignin Removal (\%)} = -239.48403 + 0.098481*A + 7.17353*B + 0.88321*C + 7.53792E\text{-}003*A*B - 2.83000E\text{-}003*A*C + 0.050258*B*C - 6.67972E\text{-}004*A^2 - 0.072318*B^2 - 0.086456*C^2 \quad (3)$$

Fig. 1(a) shows the actual experimental data versus predicted data of lignin residue (%) which was generated from the quadratic equation model. The normal probability plot was also presented in Fig. 1(b). The plot indicates that the internal studentized residual (different between actual and predicted data) follow a normal distribution and form an approximately linear line.

Table 3 summarizes the F-test and p-value that were used to assess the significance of the coefficients of the model for respective parameters and the interaction strength among the parameters. P-values that are less than 0.05 indicate significance of model terms, which are shown by A, C, AB, A^2 and B^2. Oppositely, the values greater than 0.1 indicates their insignificance. The "Lack of Fit F-value" (LOF) of 2.61 implies that the LOF is not significant relative to the pure error, thus good model, which means the terms are adequate to fit the model. Meanwhile, the fitness, R^2 and further-refined fitness, Adj R^2 are 0.970 and

0.916, respectively, which indicate that the estimated model that fits the experimental data is reasonable.

Fig. 1(a) shows the actual experimental data versus predicted data of lignin residue (%) which was generated from the quadratic equation model. The normal probability plot was also presented in Fig. 1(b). The plot indicates that the internal studentized residual (different between actual and predicted data) follow a normal distribution and form an approximately linear line.

Table 3: Statistical analysis

Source of variation	Sum of squares	Degree of freedom	Mean square	F Value	p-value probability
Model	3461.364	9	384.596	17.970	0.0027a
A-microwave power	2676.254	1	2676.254	125.049	< 0.0001a
B-temperature	3.302	1	3.302	0.154	0.7106b
C-time	183.792	1	183.792	8.588	0.0326a
AB	227.281	1	227.281	10.620	0.0225a
AC	8.009	1	8.009	0.374	0.5675b
BC	25.259	1	25.259	1.180	0.3269b
A^2	164.746	1	164.746	7.698	0.0392a
B^2	193.104	1	193.104	9.023	0.0300a
C^2	17.249	1	17.249	0.806	0.4105b
Residual	107.008	5	21.402		
Lack of Fit	85.259	3	28.420	2.613	0.2888b
Pure Error	21.750	2	10.875		
Correction total	3568.373	14			
R^2	0.970				
Adj R^2	0.916				

Value of p-Value < 0.0500 indicate model terms are significant
[a]Significant
[b]Not Significant

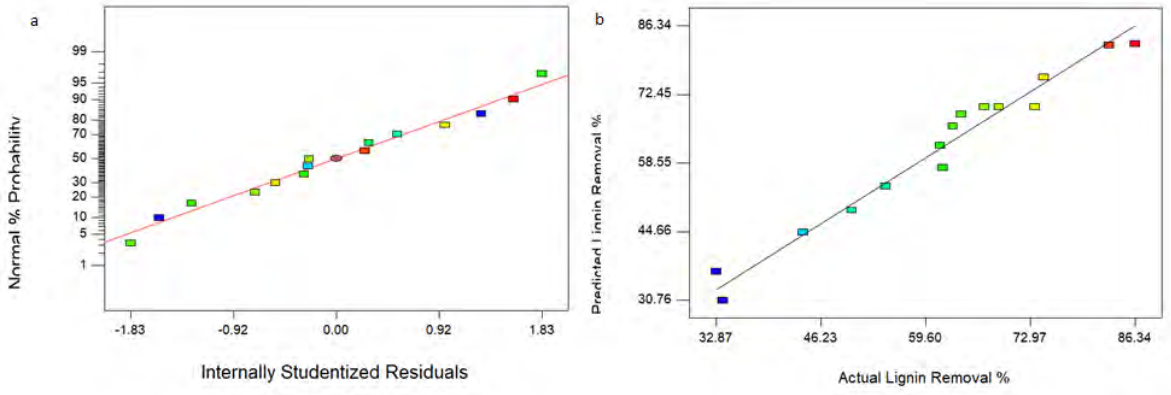


Fig. 1: (a) Correlation of normal probability of internal studentized residuals, and (b) actual conversions and values predicted by model.

3.2 Effect of Individual Parameters: Microwave Power, Temperature Setting and Irradiation Time on Lignin Removal

The effects of three microwave parameters on the alkaline treatment of CSW are shown in Fig. 2. The lignin removal was improved significantly by the increase of microwave power from 200 W to 400 W in Fig. 2(a). In particular, CSW treatment at 80 °C for about 20 minutes increased the lignin removal up to 40% by raising microwave power from 200W to 250W. These results indicate that the microwave power probably help break down the structure of CSW. However, upon increase of temperature and time at 200 W microwave power, the lignin removal only slightly changed as shown in Fig. 2(b) and (c). Upon the increase of temperature from 60 °C to 70 °C, the lignin removal approximately increased up to 14% and reach similar percent removal started from 75°C to 80 °C when CSW exposing 400W of irradiation for about 20 minutes. Lignin removal on CSW at 400 W irradiation and temperature setting about 80°C which varying irradiation time was apparently not effective. Based on the observation in Fig. 2(c), the percentage of lignin removal only improved 7% when the time was increased from 10 to 15 minutes and as of beyond 15 minutes the lignin removal started to decline as the percentage removing was too small. Since the longer irradiation time was less effective for reducing the lignin content, shorter duration would be more economical due to low power consumption [21]. Duan et al. [22] had studied the effects of microwave temperature, microwave time, and hydrochloric acid concentration on characteristics and pyrolysis behaviour of lignin. Their results show that microwave temperature applied the most significant influence on lignin conversion at 600 W. Thus, at high microwave power, temperature is elevated. Response surface area reveals the effect of temperature at different microwave power.

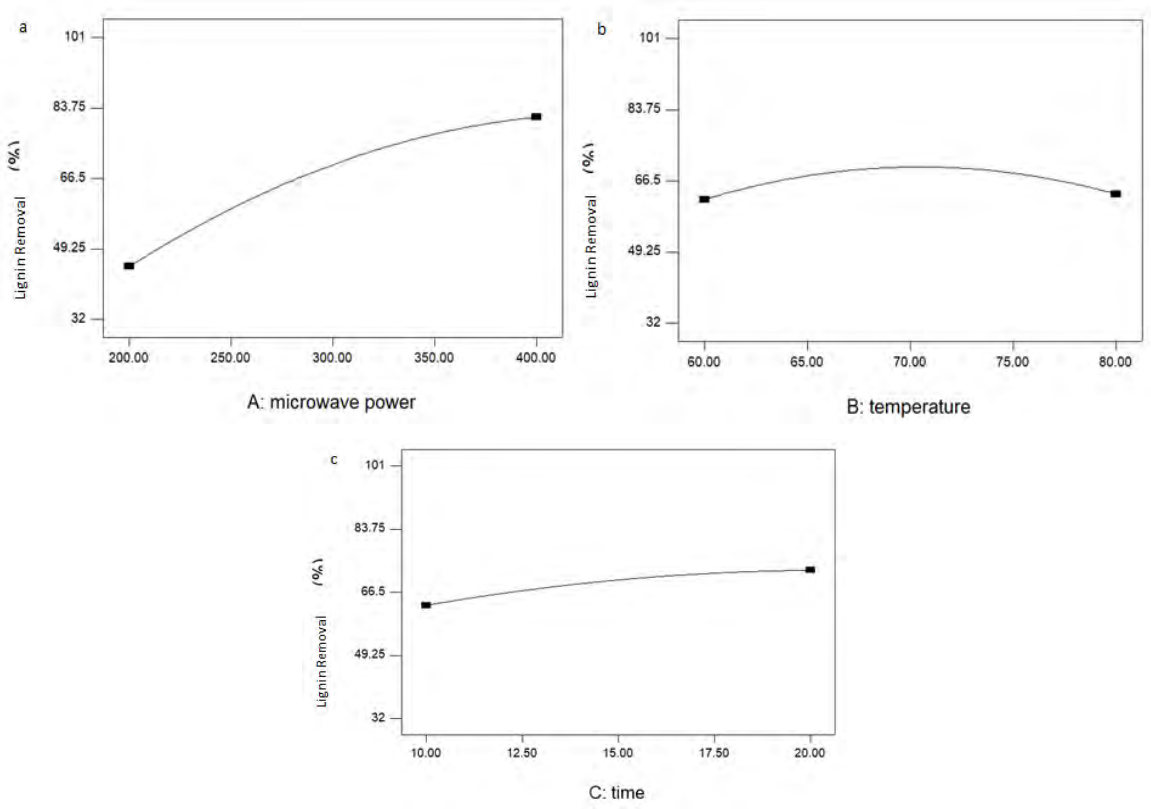


Fig. 2: (a) Effect of various individual parameters: microwave power, (b) temperature setting and (c) irradiation time on the lignin removal from CSW. One parameter was varied while others are kept constant at their centre point.

3.2.1 Response Surface of Lignin Removal

Figure 3 illustrates the plots of interaction between the three microwave parameters on the lignin removal from CSW. These plots exhibit correlation between two variables, with one variable fixed at various levels. The increase of lignin removal had been affected dominantly by the increase of microwave power and temperature as compared to irradiation time. These results conform to the finding made by Vardhini et al. [23] on banana fibre. Among the interaction between respective parameters, the irradiation time was observed to be less effecting in achieving the objective of removing lignin. Furthermore, combination of high microwave power and temperature significantly increased the lignin removal. Overall, the microwave parameters that maximally removed lignin from the CSW under caustic mixture for this study (Fig. 2a, b, and c) were 400 W, 76 °C and 19 min where 86.57% lignin was removed. Further increase of microwave power and temperature may find an optimum point but the study was continued due to the limitation of our microwave oven.

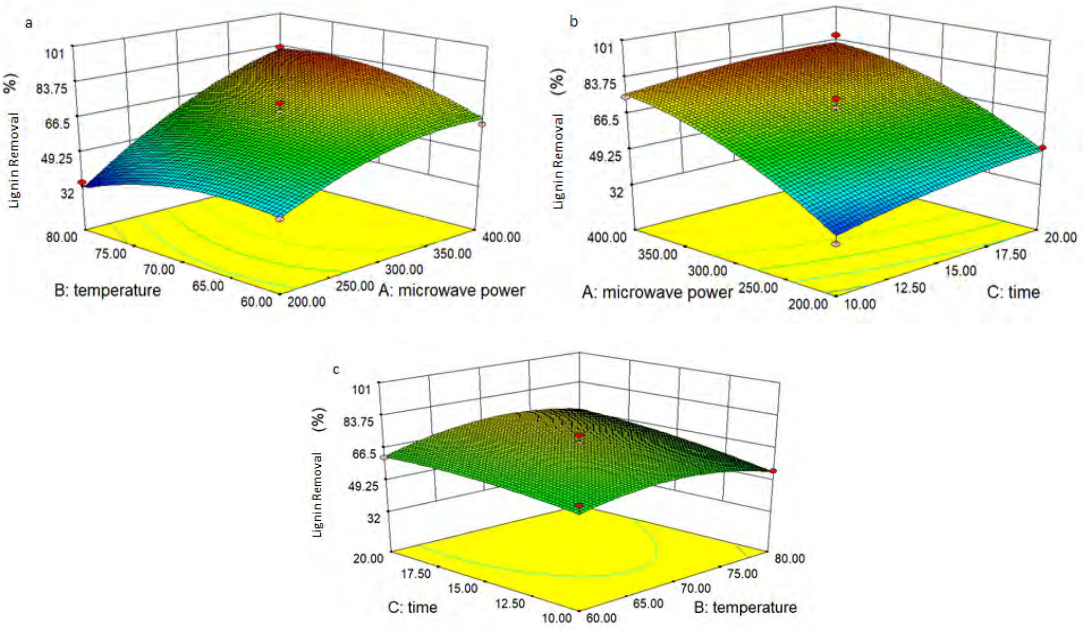


Fig. 3: RSM 3D plots of interaction between (a) microwave power and temperature, (b) microwave power and irradiation time and (c) microwave temperature and irradiation time in removing lignin from CSW in the presence of NaOH.

3.2.2 Thermogravimetric Analysis (TGA)

Thermogravimetric analyses were performed to observe any change in thermal degradation profile of CSW after the treatment. Fig. 4 exhibits the TGA result of raw CSW and treated CSW. The weight of both samples firstly dropped until the temperature 120°C due to removal of unbound and bound water [24] or also known as drying period [25]. When heating temperature reached 230 °C, components which are believed [26] hemicellulose and cellulose started to decompose. Finally, different dropping trend possessed by raw CSW at higher temperature indicating that lignin content in the treated CSW has been decreased [26]. After analysis, it was found that the residual weight of raw CSW 11% lower compare to raw CSW. When the treated fibres were subjected to heat, less chars were generated and resulting in lower amount of residual weight [27].

3.2.3 Scanning Electron Microscopy (SEM)

The changes in morphology of CSW were observed using scanning electron microscopy (SEM). Fig. 5 shows the significant effect of microwave of the alkali treatment of CSW. The surface of treated CSW in Fig. 5(b) seems rougher as compared to un-treated CSW in Fig. 5(a). The result is supported by the study carried out by Pua et al. [28] who found similar morphology result when they treated cocoa pod husk (CPH) using the same alkaline. During their treatment, hemicellulose component including lignin had been removed and left mesoporous structure on the surface. The rough surface may provide more surface area for adhesion.

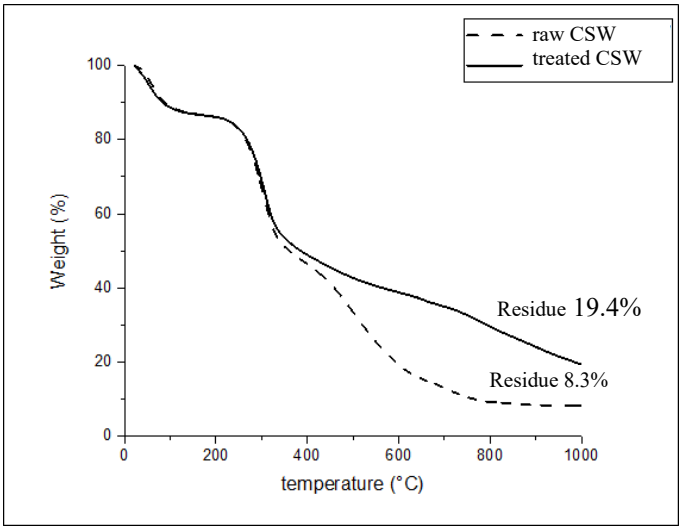


Fig. 4: Thermogravimetric analysis result for both CSW.

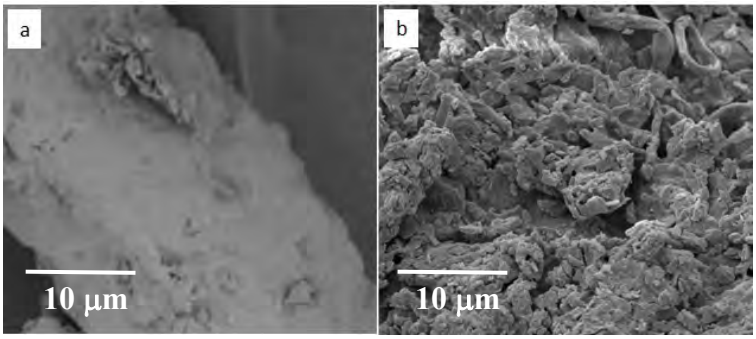


Fig. 5: SEM image with 2000 magnification for
(a) untreated CSW sample and (b) treated CSW sample.

3. CONCLUSION

A microwave-assisted alkali treatment process is an alternative approach in quickly breaking the structure and removing lignin from CSW for preparation of bio composites. This approach is simple, easy and rapid. The results are successfully modelled using Box-Behnken design in the response surface methodology. The maximum removal of lignin up to 86.57% was obtained at the microwave power, temperature and irradiation time of 400 watt, 76 ºC, and 19 min, respectively. Removal of components within lignin decomposing temperature and change of CSW surface were shown by TGA and SEM analyses, respectively.

ACKNOWLEDGEMENT

The authors are thankful for financial assistance by Ministry of Higher Education, Malaysia under research grants number: FRGS/1/2016/TK02/UMP/02/6 (RDU160126), Lembaga Koko Malaysia, Jerantut, Pahang, Malaysia for material supply, Universiti Malaysia Pahang under the grant number PGRS1903125 and Faculty of Chemical and Process Engineering Technology laboratory staff, UMP for technical support and facilities.

REFERENCES

[1] Balentic JP, Ačkar Đ, Jokic S, Jozinovic A, Babic J, Miličevic B, Ubaric D, Pavlovic N. (2018) Cocoa shell: A by-product with great potential for wide application. Molecules, 23 (6): 1-14. doi: 10.3390/molecules23061404.

[2] Handojo L, Triharyogi H, Indarto A. (2019) Cocoa bean shell waste as potential raw material for dietary fiber powder. Int J Recycl Org Waste Agric., 8: 485-491. doi: 10.1007/s40093-019-0271-9

[3] Jokić S, Gagić T, Knez E, Ubarić D, Kerget M. (2018) Separation of active compounds from food by-product (Cocoa Shell) using subcritical water extraction. Molecules, 23(6): 1408. doi: 10.3390/molecules23061408

[4] Fioresi F, Vieillard J, Bargougui R, Bouazizi N, Fotsing PN, Woumfo ED, Brun N, Mofaddel N, Le Derf F. (2017) Chemical modification of the cocoa shell surface using diazonium salts. Journal of Colloid and Interface Science, 494: 92-97. doi: 10.1016/j.jcis.2017.01.069.

[5] Foon CS, May CY, Liang YC, Ngan MA, Basiron Y. (2005) Palm biodiesel: Gearing towards Malaysian biodiesel standards. Palm oil developments (POD) 42: 28-34.

[6] Rojo-Poveda O, Barbosa-Pereira L, Zeppa G, Stévigny C (2020) Cocoa bean shell—a by-product with nutritional properties and biofunctional potential. Nutrients 12(4): 1-29. doi: 10.3390/nu12041123.

[7] Mohammed L, Ansari MNM, Pua G, Jawaid M, Islam MS. (2015) A review on natural fiber reinforced polymer composite and its applications. Int J Polym Sci. 243947. doi: 10.1155/2015/243947

[8] Balaji A, Karthikeyan B, Sundar Raj C (2015) Bagasse fiber – The future biocomposite material: A review. Int. J. ChemTech Res, 7 (1): 223-233.

[9] Sood M, Dwivedi G. (2018) Effect of fiber treatment on flexural properties of natural fiber reinforced composites: A review. Egypt J Pet., 27(4): 775-783. doi: 10.1016/j.ejpe.2017.11.005

[10] Husseinsyah S, Marliza MZ, Selvi E. (2014) Biocomposites from polypropylene and corn cob: Effect maleic anhydride grafted polypropylene. Advances in materials Research, 3(3): 129-137.

[11] Laghari SM. Isa MH, Abdullah A, Laghari AJ, Saleem H. (2014) Microwave individual and combined pre-treatments on lignocellulosic biomasses. IOSR J Eng., 4(2): 14-28. doi: 10.9790/3021-04261427

[12] Nomanbhay SM, Hussain R, Palanisamy K. (2013) Microwave-assisted alkaline pretreatment and microwave assisted enzymatic saccharification of oil palm empty fruit bunch fiber for enhanced fermentable sugar yield. Journal of Sustainable Bioenergy Systems, 3(1): 7-17. doi: 10.4236/jsbs.2013.31002.

[13] Kucharska K, Rybarczyk P, Hołowacz I, Łukajtis R, Glinka M, Kamiński M. (2018) Pretreatment of lignocellulosic materials as substrates for fermentation processes. Molecules 23(11): 1-32. doi: 10.3390/molecules23112937.

[14] Rajesh M, Pitchaimani J, Rajini N. (2016) Free Vibration characteristics of banana/sisal natural fibers reinforced hybrid polymer composite beam. Procedia Engineering, 144: 1055-1059. doi: 10.1016/j.proeng.2016.05.056.

[15] Budiyono, Wicaksono A, Rahmawan A, Hawali Abdul Matin H, Gumilang Kencana Wardani L, Djoko Kusworo T, Sumardiono S. (2017) The effect of pretreatment using sodium hydroxide and acetic acid to biogas production from rice straw waste. MATEC Web Conf. doi: 10.1051/matecconf/201710102011

[16] Modenbach AA, Nokes SE. (2014) Effects of sodium hydroxide pretreatment on structural components of biomass. Transactions of the ASABE, 57(4): 1187-1198. doi: 10.13031/trans.57.10046.

[17] Sombatpraiwan S, Junyusen T, Treeamnak T, Junyusen P. (2019) Optimization of microwave-assisted alkali pretreatment of cassava rhizome for enhanced enzymatic hydrolysis glucose yield. Food and Energy Security (October 2018): e00174. doi: 10.1002/fes3.174.

[18] Agu OS, Tabil LG, Meda V, Dumonceaux T, Mupondwav E. (2019) Pretreatment of Crop Residues by Application of Microwave Heating and Alkaline Solution for Biofuel Processing: A Review. Renew Resour Biorefineries. doi: 10.5772/intechopen.79103

[19] Ethaib S, Omar R, Mazlina M, Radiah A, Syafiie S, Harun MY. (2016) Effect of microwave-assisted acid or alkali pretreatment on sugar release from Dragon fruit foliage. Int. Food Res. J., 23 (Suppl): S149-S154

[20] Safari A, Karimi K, Shafiei M. (2017) Dilute alkali pretreatment of softwood pine: A biorefinery approach. Bioresource Technology, 234: 67-76. doi: 10.1016/j.biortech.2017.03.030.

[21] Lai L, Idris A, Yusof NM. (2014) Lignin extraction from oil palm trunk by microwave-alkali technique. Malaysian Journal of Fundamental and Applied Sciences, 10(2): 59-64.

[22] Duan D, Ruan R, Wang Y, Liu Y, Dai L, Zhao Y, Zhou Y, Wu Q. (2018) Microwave-assisted acid pretreatment of alkali lignin: Effect on characteristics and pyrolysis behavior. Bioresour. Technol., 251: 57-62. doi: 10.1016/j.biortech.2017.12.022

[23] Vardhini KJV, Murugan R, Selvi CT, Surjit R. (2016) Optimisation of alkali treatment of banana fibres on lignin removal. Indian Journal of Fibre and Textile Research, 41: 156-160.

[24] Yang H, Yan R, Chen H, Zheng C, Lee DH, Liang DT. (2006) In-depth investigation of biomass pyrolysis based on three major components: Hemicellulose, cellulose and lignin. Energy & Fuels, 20(1): 388-393. doi: 10.1021/ef0580117.

[25] Sushmitha D, Suranai S. (2019) Microwave-assisted alkali-peroxide treated sawdust for delignification and its characterisation. in: Waste Valoris. Recycl. Singapore, Springer Singapore. pp 527–537.

[26] Waters CL, Janupala RR, Mallinson RG, Lobban LL. (2017) Staged thermal fractionation for segregation of lignin and cellulose pyrolysis products: An experimental study of residence time and temperature effects. Journal of Analytical and Applied Pyrolysis, 126: 380-389. doi: 10.1016/j.jaap.2017.05.008.

[27] Yew BS, Muhamad M, Mohamed SB, Wee FH. (2019) Effect of alkaline treatment on structural characterisation, thermal degradation and water absorption ability of coir fibre polymer composites. Sains Malaysiana, 48(3): 653-659. doi: 10.17576/jsm-2019-4803-19.

[28] Pua FL, Sajab MS, Chia CH, Zakaria S, Rahman IA, Salit MS. (2013) Alkaline-treated cocoa pod husk as adsorbent for removing methylene blue from aqueous solutions. Journal of Environmental Chemical Engineering, 1(3): 460-465. doi: 10.1016/j.jece.2013.06.012.

https://doi.org/10.31436/iiumej.v22i2.1682

ADSORPTION OF METHYLENE BLUE ONTO ACTIVATED CARBON DEVELOPED FROM BAOBAB FRUIT SHELL BY CHEMICAL ACTIVATION: KINETIC EQUILIBRIUM STUDIES

RADHIA NEDJAI*, MA'AN FAHMI RASHID ALKHATIB, MD ZAHANGIR ALAM AND NASSERELDEEN AHMED KABBASHI

Department of Biotechnology Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

Corresponding author: radiaradia19@gmail.com.

(Received: 1st November 2020; Accepted: 25th February 2021; Published on-line: 4th July 2021)

ABSTRACT: This article provides results of the usability of baobab fruit shell to produce activated carbons by chemical activation using $ZnCl_2$, H_3PO_4, and KOH. This study indicated that activated carbon produced from baobab fruit shell fruit can be used as a promising adsorbent for the removal of methylene blue from aqueous solutions. Significant changes on the material surface following the activation process were observed through SEM and FTIR analyses. Scanning electron micrographs of BFS-ACs showed that porous structures were formed during activation, while the FTIR results indicated that the carbons have abundant functional groups on the surface. KOH activation led an activated carbon with a high methylene blue adsorption of 95.54 % and maximum adsorption capacity of 113.63 mg/g, which is directly related to the specific surface area of activated carbons. The adsorption isotherm data were fitted to Langmuir and Freundlich adsorption models. The Langmuir isotherm model showed better fit to the equilibrium data than the Freundlich model. The adsorption process was well described by the pseudo-second-order kinetics. The BFS-ACs is an effective and low-cost adsorbent for the removal of MB from an aqueous solution.

ABSTRAK: Kajian ini memberi input tentang kebolehgunaan kulit buah baobab bagi menghasilkan karbon teraktifan melalui aktiviti kimia menggunakan $ZnCl_2$, H_3PO_4, dan KOH. Karbon aktif daripada kulit buah Baobab ini berpotensi sebagai penyerap bagi menyingkir larutan akueus metilin biru. Perubahan ketara pada permukaan bahan diikuti dengan proses pengaktifan dipantau melalui analisis SEM dan FTIR. Imbasan elektron mikrograf BFS-AC menunjukkan struktur porus terhasil semasa proses pengaktifan. Sementara dapatan FTIR menunjukkan karbon mempunyai banyak kumpulan berfungsi pada permukaan. Pengaktifan KOH menghasilkan karbon aktif menggunakan larutan biru metilin yang tinggi sebanyak 95.54 % dan kapasiti maksimum penyerapan 113.63 mg/g, iaitu berkadar langsung dengan tumpuan kawasan permukaan karbon aktif berkaitan. Data isoterma penyerapan dibina pada model penyerapan Langmuir dan Freundlich. Model isoterma Langmuir lebih padan pada data keseimbangan berbanding model Freundlich. Proses penyerapan menunjukkan lebih kinetik order-kedua-pseudo. BFS-AC sangat efektif dan penyerap murah bagi membuang MB daripada larutan akues.

KEYWORDS: *activated carbon; Baobab fruit shells; methylene blue adsorption; adsorption isotherms; adsorption kinetics*

1. INTRODUCTION

The wide use of dyes in various industries such as textile, paper, rubber, printing, plastic, leather, cosmetics and pharmaceuticals, and food has polluted the environment, especially the aquatic eco-system [1,2]. Dye production dramatically increased in the last years due to the demand in these industries. The presence of dyes in industrial effluents, even at very low concentrations, can cause many significant problems, such as reducing light penetration, which disturbs photosynthesis activity in aquatic life. Moreover, these dyes affect human lives due to the bioaccumulation process which is transferred through the food chain [1,3]. This results from the chemical structure of dye molecules that contain double bonds of olefinic, carbon-nitrogen or nitrogen-nitrogen, and heterocyclic or aromatic rings, which make these dyes hard to break down [4]. Thus, dyes are considered mutagenic, toxic, allergenic, carcinogenic, non-degradable substances [5]. There are more than 700,000 dyestuffs produced per year, which are available with more than 100,000 different chemical structures around the world [3,6]. Methylene blue (MB) is considered the most common azo dyes used for coloring in industry [7]. Because of its harmful effect on water, it is crucial to eliminate it from waste streams before discharge to public water sources.

In order to minimize this dye's impacts on the environment, many technologies have been developed to treat wastewater such as flocculation, coagulation, ozonation, chlorination, anodic oxidation, and electrooxidation. However, these techniques are costly and frequently become secondary sources of pollution [8]. Many studies revealed that adsorption by solid adsorbents presents the advantages of high efficiency, simple design, and low initial costs [7]. The activated carbon is one of the most effective adsorbents for the removal of water pollutants. Using agricultural by-products for the production of activated carbon provides an environmentally credible alternative to coal, pitch, petroleum, and coke. It minimizes the production cost of activated carbon since it is inexpensive and abundant. Moreover, it is an efficient solution to the discarding of huge amounts of waste.

Various agricultural by-products have been used as a precursor in producing activated carbon for the removal of methylene blue from the wastewater such as pine cone [1], bamboo [9], karanj (*Pongamia pinnata*) fruit hulls [10], prawn shells [11], cashew nut shells [7], tea fruit peel residues [12], tea (*Camellia sinensis L.)* seed shells [13], dabai (*Canarium odontophyllum*) nutshells [14], date pits [15], cola nutshells [16], bitter kola (*Garcinia kola*) nutshells [17], baobab fruit shells (*Adansonia digitata L.*) [18], palm kernel shell [19], lapsi seed stone [20].

Adsorption capacity is one of the most important characteristics of activated carbon. It is affected by the inherent nature of the precursor and the activation process which is determined by different conditions such as the activating agents, temperature, and activating time [20]. Activating agents have a great influence on the surface and pore formation of produced activated carbon. Various chemical agents proposed for chemical activation such as potassium hydroxide [17,21,22], zinc chloride [17,19], phosphoric acid [17,22,23], potassium carbonate [24,25], calcium chloride [20], and sodium hydroxide [26,27]. Each of these reagents creates different influences on pore development and has its distinctive safety concerns that influence the application of the produced activated carbon. For example, the H_3PO_4 and $ZnCl_2$ have boiling points of 213 °C and 732 °C, respectively, could decompose at the activation temperature and cause toxic fumes. The KOH, on the other hand, has a boiling point of 1327 °C, which is higher than the typical activation temperatures [28].

Baobab fruit shell is not only an organic waste component that has no economic benefits but is also creates problems with waste treatment and disposal. Baobab fruit shells are an abundant new precursor for the production of activated carbon as they contain high volatile

matter (86.73 %) and low ash (5.57 %) [29]. Therefore, the main focus of this study is to utilize Baobab Fruit Shell (BFS) as a renewable and low-cost precursor for producing a high surface area activated carbon (BFS-ACs) by chemical activation using H_3PO_4, KOH, and $ZnCl_2$. Activated carbon could appropriately be prepared from the BFS which would lead to its valuable application. To the best of our knowledge, there has been no study reported on the influence of activating agents on the adsorption capacity of activated carbons made from BFS. Hence, the main objective of this research was to produce high surface area microporous activated carbon from baobab fruit shells using different chemical activation reagents and to study their effects on the adsorption of methylene blue (MB). The adsorption properties of the prepared BFS-ACs was studied with Methylene Blue in aqueous solution under equilibrium and kinetic conditions to obtain the associated adsorption parameters.

2. MATERIALS AND METHODS

2.1 Materials

BFS were used as a precursor for the preparation of activated carbon and were acquired from West Sudan. The samples BFS were washed several times with tap water and finally with distilled water to remove any impurities. The shells were dried in the oven at 105 °C for dehydration until a constant weight was obtained. Then, they were ground and screened to obtain particles of the size of 1 mm.

Phosphoric acid (H_3PO_4), potassium hydroxide (KOH), and zinc chloride ($ZnCl_2$) were utilized as chemicals reagents for the activation process, which were acquired from Macron Fine Chemicals (USA), Sigma-Aldrich (Malaysia), and R&M Chemicals (Malaysia), respectively. Nitrogen gas (99.95 %) was purchased from Fuelink Marketing Sdn. Bhd. (Selangor, Malaysia), that was used for the inert atmosphere during carbonization. Iodine was acquired from Univar (UK). Potassium iodide was purchased from Ducheda Biochemie (Malaysia). All other chemicals of analytical grade were utilized throughout the experiments.

2.2 Preparation of Adsorbate

Methylene blue (C.I.52015) StainPur was obtained from Systerm. Methylene blue solution (1000 mg/L) was prepared through dissolving 1.127 g methylene blue with distilled water in the volumetric flask of 1 L [30]. The prepared solution was diluted with distilled water to obtain the required concentrations for each experiment.

2.2 Activated Carbon Preparation

For the activation process, 10 g of dried BFS was mixed with different chemical reagents ($ZnCl_2$, KOH, or H_3PO_4) at an impregnation ratio of 1:1. Then, 80 mL of distilled water was added and agitated utilizing a rotary shaker for 1 hour at 50 °C. The samples were then oven-dried at 100 °C for 16 hours to get a thick paste. Later, the samples were inserted into a Carbolite horizontal tube furnace at 10 °C/min at 500 °C for 60 min under continuous flowing nitrogen gas (99.95 %). The carbonization step was conducted in a Lenton horizontal tubular quartz reactor (England, UK) consisting of a Carbolite horizontal tube electrical furnace (Model CTF, UK) with a maximum operating temperature of 1200 °C, a pressure regulating valve, and a controller (Fig. 1). The resulting products were finally moved to the desiccator to cool down. The weight of produced activated carbons (BFS-ACs) was measured [31].

https://doi.org/10.31436/iiumej.v22i2.1682

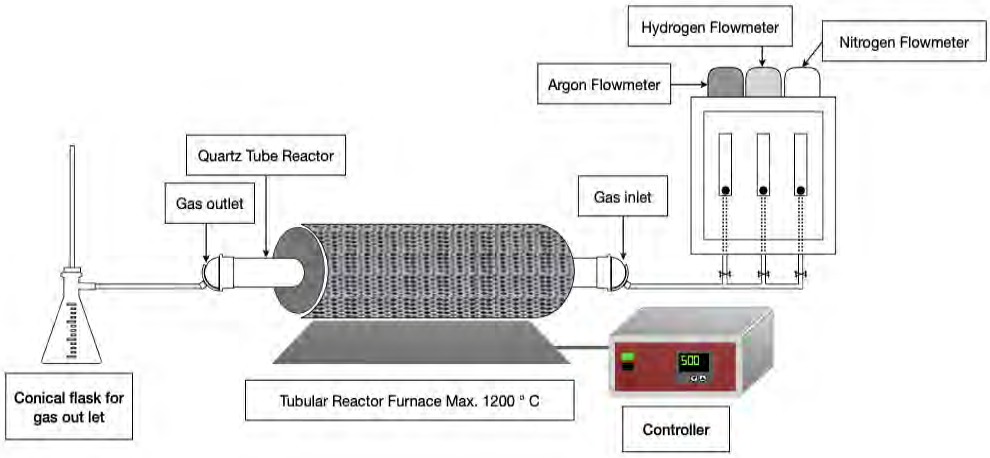


Fig. 1: Schematic diagram of the experimental setup for BFS-ACs production.

2.4 Activated Carbon Purification

BFS-ACs were washed utilizing 50 mL of 0.5 M HCl to remove any residual of the dehydrating agents and any ash in the samples. The BFS-ACs were frequently washed with warm distilled water until a constant pH was achieved and were dried at 110 °C for 24 hours using a vacuum oven [31].

2.5 Characterization of the Carbonized Carbon

The following characteristics of the carbonized BFS carbon were determined:

2.5.1 Proximate Analysis

The ash content was determined by ASTM D 2866-83 [32], and the moisture content of BFS-ACs was measured based on a procedure by ASTM D 2867-91 [33].

2.5.2 Determination of the Yield

Activated carbon yields were calculated based on the equation:

$$Yield(\%) = \frac{W_f}{W_i} \times 100 \quad (1)$$

where W_f is the dry mass of final activated carbon in grams, which is taken after the activated carbon is produced. While W_i is the dry mass of BFS in grams, which is taken at the beginning of the experiment [34].

2.5.3 Determination of the Bulk Density

The density of the prepared product was determined by wrapping five grams of activated carbon in a plastic bag. Then, the sample was placed in a 500 mL graduated cylinder that contained water. The difference in the level of water was measured. The density was calculated using the equation [35]:

$$Bulkdensity(g/cm^3) = \frac{ActivatedCarbonweight(g)}{Volume(cm^3)} \quad (2)$$

2.5.4 Determination of Iodine Number

Iodine number was carried out following the procedure by Vunain et al. [36]. Potassium iodide (4.10 g) was precisely weighed and mixed with 2.70 g of iodine crystals to prepare one liter of iodine solution. Standardization of the stock solution was done using a standard

solution of sodium thiosulfate ($Na_2S_2O_3$). Five hundred milligrams of the produced activated carbon were transferred to Erlenmeyer flasks of 250 mL and was mixed with 100 mL of iodine solution. The mixture was left stirring using a rotary shaker at a rapid speed, and then the mixture was filtered using filter paper. Only 20 mL of the filtrate was transferred into a clean Erlenmeyer flask, where it was titrated with standardized 0.1 N sodium thiosulfate ($Na_2S_2O_3$) until the yellow color disappeared. An amount of starch solution (around 1 mL) was added, and the titration continued pending the disappearance of the blue indicator color. The amount of iodine adsorbed per milligram, I (mg/g) was plotted against the concentration of iodine in the filtrate (C). The residual iodine concentration (C) was calculated and the whole procedure of the experiment was repeated using a different mass of activated carbon if the value of (C) was not within the range of 0.008-0.04 mol/L. The I and C values were calculated by equations (2) and (3), respectively [37].

$$I\left(\frac{mg}{g}\right) = (N_1 \times 126.93 \times V_1) - \left(\left(\frac{(V_1 + V_{HCl})}{V_f}\right) \times (N_2 \times 126.93 \times V_2)\right) \Big/ W_C \quad (3)$$

$$C = N_2 \times V_2 \quad (4)$$

where N_1 is the normality of the iodine solution, N_2 is the normality of sodium thiosulfate solution, V_1 is the additional volume of iodine solution, V_2 is the volume of sodium thiosulfate solution consumed, V_{HCl} is the additional volume of 5 % HCl, V_f is the filtrate volume used in titration and W_C is the activated carbon weight.

2.5.5 Fourier Transform Infrared (FTIR)

Fourier transform infrared (FTIR) spectroscopic was utilized to analyze the surface chemistry of prepared activated carbons (BFS-ACs) and the raw material (BFS) applying a Perkin Elmer Frontier model FTIR-spectrometer. Before starting the analysis, the dry samples were combined with KBr powder. Then, the mixtures were compressed into capsules that were utilized for analysis. The % of absorbance of samples was recorded over 400 to 4000 cm^{-1}.

2.5.6 Screening Electron Microscopy (SEM)

The morphological structure of the BFS and prepared BFS-ACs was observed utilizing scanning electron microscopy (SEM). The samples were prepared on a carbon tape and images were taken using JEOL-IT100 SEM instrument (JEOL, Japan). Pore size measurements were carried out using the measurement tool equipped with the JEOL IT100 SEM (InTouchScope™ version 1.060).

2.6 Batch Equilibrium Studies

In experiments of batch equilibrium adsorption, 0.3 g of different BFS-AC samples were weighed and taken in several conical flasks. Methylene blue solution (100 mL) at different concentrations (25, 50, 100, 200, 300, 400 mg/L) was added. Then, the flasks were placed in a rotary shaker for various contact times (i.e. 5, 10, 15, 30, 60, 120, 180 min) at 120 rpm at room temperature (27±1°C). The solution was filtered and the residual concentration of methylene blue was determined at 660 nm using a UV/vis spectrophotometer.

The amount of methylene blue adsorbed by biosorbent was calculated using the following equation:

$$q_e = \frac{(C_i - C_f)V}{W} \quad (5)$$

where the amount of methylene blue adsorbed per gram of biosorbent at equilibrium (mg/g) is represented by (q_e), the initial methylene blue concentration is represented by C_i (mg/L), the final or equilibrium methylene blue concentration is represented by C_f (mg/L), the volume of methylene blue solution in the flasks is represented by V (L) and the weight of biosorbent used is represented by W (g) [38].

The difference between methylene blue concentration before and after adsorption ($C_i - C_f$) to the initial concentration of methylene blue in the aqueous solution (C_i) is known as the percentage of methylene blue uptake, and it may be estimated using the following the equation:

$$Removal\ (\%) = \frac{(C_i - C_f)}{C_i} \times 100 \quad (6)$$

3. RESULTS AND DISCUSSION

3.1 Characterization of Adsorbent (BFS-ACs)

3.1.1 Effects of Activating Agents on the Yield of AC

The yield of prepared BFS-ACs was determined, and the results are shown in Table 1. Rendering to the results, the yield of BFS-ACs ranged from 30.3 % to 34.6 %, obtained under the same conditions (500 °C for 1 h). All the data presented are the average of three replicates. The highest yield was obtained by phosphoric acid (H_3PO_4) with 30.3 %, followed by zinc chloride ($ZnCl_2$) with 33.6 %, and potassium hydroxide (KOH), respectively. At a higher temperature, more volatiles and polymeric structures decompose such as cellulose, lignin, and hemicellulose are liberated causing a lower yield.

The existence of H_3PO_4 over the activation process favors the transformation of aliphatic to aromatic compounds, dehydration, promotes depolymerization, and redistribution of constituent biopolymers, consequently, raising the yield of activated carbon [34].

Table 1: Yield, moisture content, ash content, bulk density and iodine number of prepared BFS-ACs

Adsorbent	Yield (%)	Ash content (%)	Moisture content (%)	Bulk density (g/cm^3)	Iodine Number (mg/g)
$ZnCl_2$	33.663	17.3	1.15	0.3	1213.45
H_3PO_4	34.396	17.7	5.17	0.3	1248.35
KOH	30.377	14.3	1.63	0.3	1227.41

3.1.2 Ash and Moisture Content

Ash content and moisture content were determined due to their importance, and the results are also reported in Table 1. Ash content consists of inorganic substances that lead to a rise in hydrophilic nature and could cause catalytic effects leading in restructuring pending regeneration of AC [39]. Overall, ash content ranges from 1 to 20 %, which depends on the type of raw material, and the activating agents used in the activation process. High ash content affects adsorptive capacity and reduces the mechanical strength of carbon, therefore, it is unsuitable for the AC [40]. The highest value of ash content was that AC produced using H_3PO_4, whereas the lowest value of ash content was that AC produced using KOH.

The amount of water physically associated with activated carbon was determined by calculating the moisture content, which depends on various porous characteristics like the pore volume and surface area. Generally, when activated carbon is exposed to air, that with high surface areas absorb more moisture. Some studies reported that the rate of absorption of contaminant moisture increases with lower moisture content. The values of 1.15 %, 5.17 %, and 1.63 % were obtained for $ZnCl_2$, H_3PO_4, and KOH, respectively.

The bulk density is the mass of the activated carbon particles divided by the volume they occupy. It is an important characteristic of activated carbon for a densified bed or gas storage application. The bulk density of the prepared BFS-ACs is 0.3 g/cm^3.

3.1.3 Iodine Number

Iodine number, also called iodine adsorption, is considered one of the best methods to determine the capacity of adsorption and the quality of activated carbon. It is frequently utilized because of its simplicity and rapidity [36]. The higher iodine number signifies a great probability that activated carbon has a very large micropore structure and has a high surface area for the adsorption to take place [37]. Table 1 displays the variation of the iodine number for BFS-ACs prepared with different activating agents. H_3PO_4 impregnated activated carbon had the highest iodine number (1248 mg/g), while the lowest iodine number was obtained from activated carbon impregnated with $ZnCl_2$ (1213 mg/g). Therefore, the iodine number of carbons prepared by activation processes were in the order of H_3PO_4-AC>KOH-AC> $ZnCl_2$-AC.

3.1.4 Morphological Structures of the Activated Carbon

After activation in the tube furnace with the three different activating agents ($ZnCl_2$, KOH, and H_3PO_4), the samples were observed utilizing a scanning electron microscope (SEM) and the results are presented in Fig. 2.

The morphological structure of the raw BFS and prepared activated carbons were observed using the SEM technique. SEM reveals the raw material (Fig. 2a) have the lowest original pores structures. However, the treated-BFS with different activating agents ($ZnCl_2$, KOH, and H_3PO_4) contains more porous structures that are honeycomb-shaped with different pore sizes owing to the degradation of the cellulosic fraction, which is affected by the pyrolytic decomposition of biomass by influences of activating agents.

There is no difference in the material that has been used as raw material to prepare the AC, however, there are morphological differences between the prepared AC obtained which refer to tremendous differences in the activating agents that were used and their different reaction mechanisms. $ZnCl_2$ encouraged the elimination of water molecules from the lignocellulosic structures of the raw materials, while H_3PO_4 combined chemically into the lignocellulosic structures. Therefore, there is no selective elimination of carbon over the activation which ameliorates the yields. The KOH mechanism is more complex and contains the disintegration (almost explosively) of the structure following intercalation and some gasification by oxygen molecules of hydroxide as well [41].

Many mesopores with an irregular yet extremely porous structure were occupied by the KOH-AC structure (Fig. 2b). On the other hand, H_3PO_4 and $ZnCl_2$ activated carbons (Fig. 2c and 2d) clearly show developed honeycomb as strongly defined pores with regularity and cavities compared to raw BFS and $ZnCl_2$-AC. Zinc chloride activated carbon exhibited a high macropore size with an average of 44.21 µm, followed by KOH with an average of 42.24 µm, and H_3PO_4 with an average of 27.90 µm. Besides, the pores in the KOH micrograph are greater than in the H_3PO_4 and $ZnCl_2$ micrographs. Therefore, there is great potential to absorb and trap methylene blue.

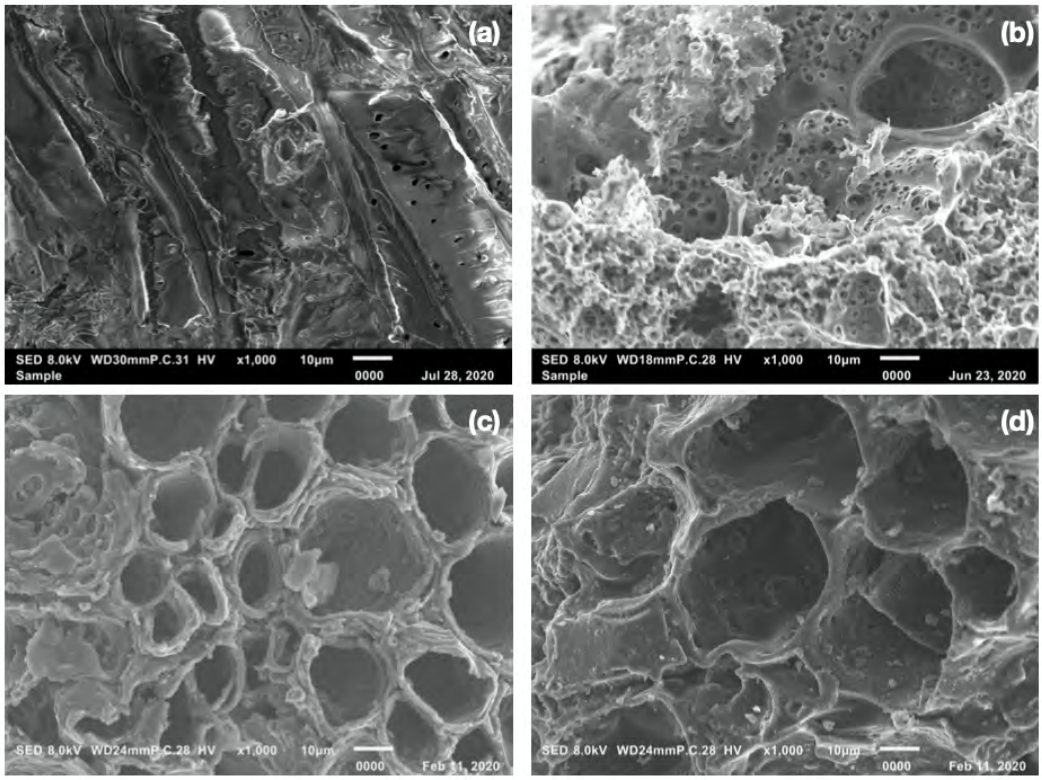


Fig. 2: SEM photomicrographs of (a) BFS; (b) AC produced from BFS using KOH; (c) AC produced from BFS using H_3PO_4; (d) AC produced from BFS using $ZnCl_2$ (magnification scale ×1000).

3.1.5 Surface Chemistry

The IR spectrum of raw BFS and BFS-ACs samples were recorded using a KBr disk in conjunction with a Perkin-Elmer infrared spectrophotometer. Figure 3 displays the results, and the graphs were taken between the wavenumbers of 4000 cm^{-1} and 400 cm^{-1}. As can be seen, the raw BFS spectrum exhibited more absorption peaks than the BFS-ACs spectrum, where several peaks that were present in the raw BFS disappeared once the activation step was performed, owing to the elimination of heat-sensitive functional groups at high temperatures and volatiles.

For the raw BFS, the broad band at 3285.18 cm^{-1} is characteristic of a stretching vibration of hydrogen-bonded (O-H) hydroxyl groups of cellulose, pectin, and lignin [18]. The peak appearing at 2889.65 cm^{-1} is described to (C–H) symmetric stretching and ($–CH_2$) alkyl groups. The small bands observed at 2160 cm^{-1} and 2031 cm^{-1} are due to the presence of (C≡N) stretching [18,34,36]. The brand at around 1613.83 cm^{-1} is the (C=O) stretching vibration due to bonds of esters and phenols [42], which disappeared in BFS-ACs. A comparatively low-intensity peak at 1367 cm^{-1} is owing to the presence of (C–O–C) stretching or C=C stretching that can be ascribed to the existence of ester, ether,and phenol. A peak situated at around 1033.37 cm^{-1} is the characteristics of anhydrides (C–O), which is observed also in $ZnCl_2$-AC and disappeared completely in KOH and H_3PO_4. The principal functional groups in BFS are hydroxyl groups, carbonyl groups and carboxyl groups.

FTIR spectrum of all prepared BFS-ACs shows broad weak peaks around 3900–3600 cm^{-1}. For H_3PO_4 and KOH activated carbons present a similar FTIR spectrum, which indicates the identical type of surface functional groups. The FTIR spectrum of BFS-ACs shows distinguished characteristic peaks as follows: 3333.13 cm^{-1} (vibrations of O–H

groups), 2886.24 cm^{-1} (stretching of C–H), 2161.38 cm^{-1} (stretching of C≡C), 1574.71 cm^{-1} (stretching of C=C), 1255.45 cm^{-1} (stretching of C–O), respectively. KOH activated carbon presents peaks with slighter intensity as contrasted to H_3PO_4, which demonstrates a lower quantity of functional groups in KOH. As stated by Puziy et al., the band obtained by H_3PO_4 at 1179 cm^{-1} can be assigned to the phosphorous-containing group P–O, O–C stretching vibrations in P–O–C linkage, and P=OOH [43]. On the other hand, the $ZnCl_2$ activated carbon, the very small bands around 2358 and 1087 cm^{-1} are related to the C≡N and the C–O stretching, respectively.

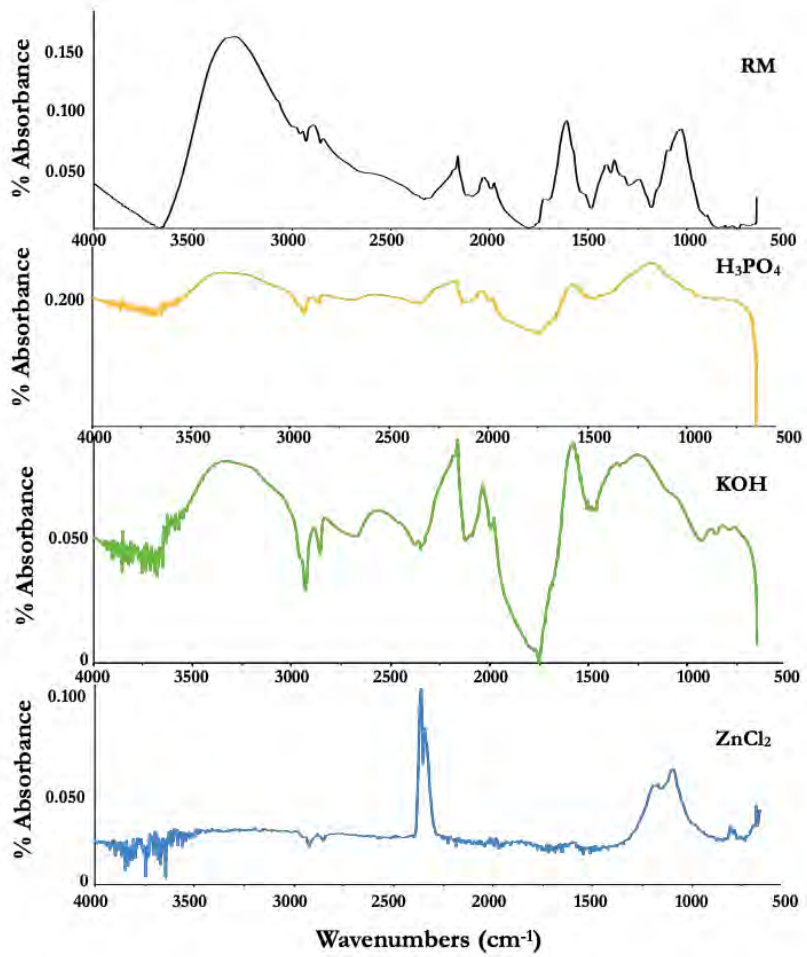


Fig. 3: Fourier transform infrared spectrum for raw Baobab fruit shells and BFS-derived activated carbons (BFS-ACs).

3.2 Adsorption Process

3.2.1 Effect of Initial Methylene Blue Concentration

For the adsorption process, the prepared BFS-ACs were employed as adsorbents to study the effect of chemical activating agents on the adsorption of methylene blue. The adsorption studies were performed using adsorbent of 3 g/L at agitation speed of 120 rpm at 27±1 °C. Figure 4 shows the effect of initial concentration on methylene blue adsorption, which were carried out at different initial dye concentrations (25 to 400 ppm). The results revealed that the actual amount of MB adsorbed (mg/g) increased with the increase in the initial dye concentration. The analysis of these results shows that the methylene blue uptake efficiencies of the prepared activated carbons were almost similar at low concentrations and significantly different at high concentrations.

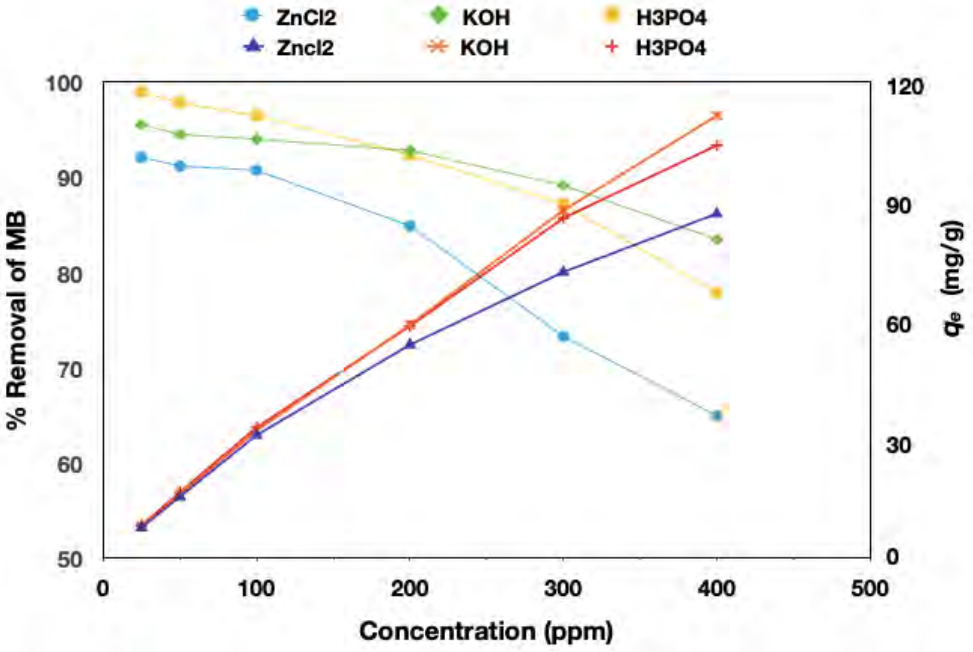


Fig. 4: Effect of initial concentration on methylene blue adsorption (adsorbent dose = 3 g/L, stirring speed = 120 rpm, temperature = 27±1 °C, volume of solution = 100 mL).

From the graph, it can be seen that the percentage of methylene blue removal decreased with the increase of the initial dye concentration. Where a high concentration of methylene blue contributes to higher interaction between methylene blue and adsorbent. Similar results were reported in the literature for the removal of methylene blue [44].

Generally, a high efficiency of removal of methylene blue was observed in KOH activated carbon, followed by H_3PO_4 and $ZnCl_2$ as have been revealed by another study [17]. However, at low concentration (25, 50, and 100 ppm), the efficiency of the removal of methylene blue by H_3PO_4 activated carbon reached to 98.97 %, 97.84 %, and 96.48 %, respectively, which is higher than the efficiency of the two other prepared activated carbons. Whereas, at high concentration (200, 300, and 400 ppm), the efficiency of dye removal by KOH activated carbon attained was 92.82 %, 89.1 %, and 83.48 %, respectively, which is higher than the efficiency of H_3PO_4 and $ZnCl_2$ activated carbons. Significant uptake of KOH compared to $ZnCl_2$ and H_3PO_4 could be explained by differences in the morphological structure of the activated carbon, as shown in Fig. 2.

3.2.2 Effect of Contact Time on Methylene Blue Adsorption

The adsorption data for the removal of methylene blue dye versus contact time at different initial concentrations is demonstrated in Fig. 5. The adsorption studies of MB into BFS-ACs were followed over 3 hours. In general, most of the methylene blue removal takes place during the first 30 minutes. The equilibrium adsorption was achieved after 1 hour and no remarkable changes being observed for longer contact time. The results show that the adsorption capacity of methylene blue on BFS-ACs dramatically augmented during the preliminary period and thereafter with slower speed. The adsorption capacity is highly prompted by the number of available active sites, which are numerous at first; consequently, the adsorbate extends the active sites easily. Afterward, the number of active sites reduces, and the adsorbent surface turns out to be saturated [45].

https://doi.org/10.31436/iiumej.v22i2.1682

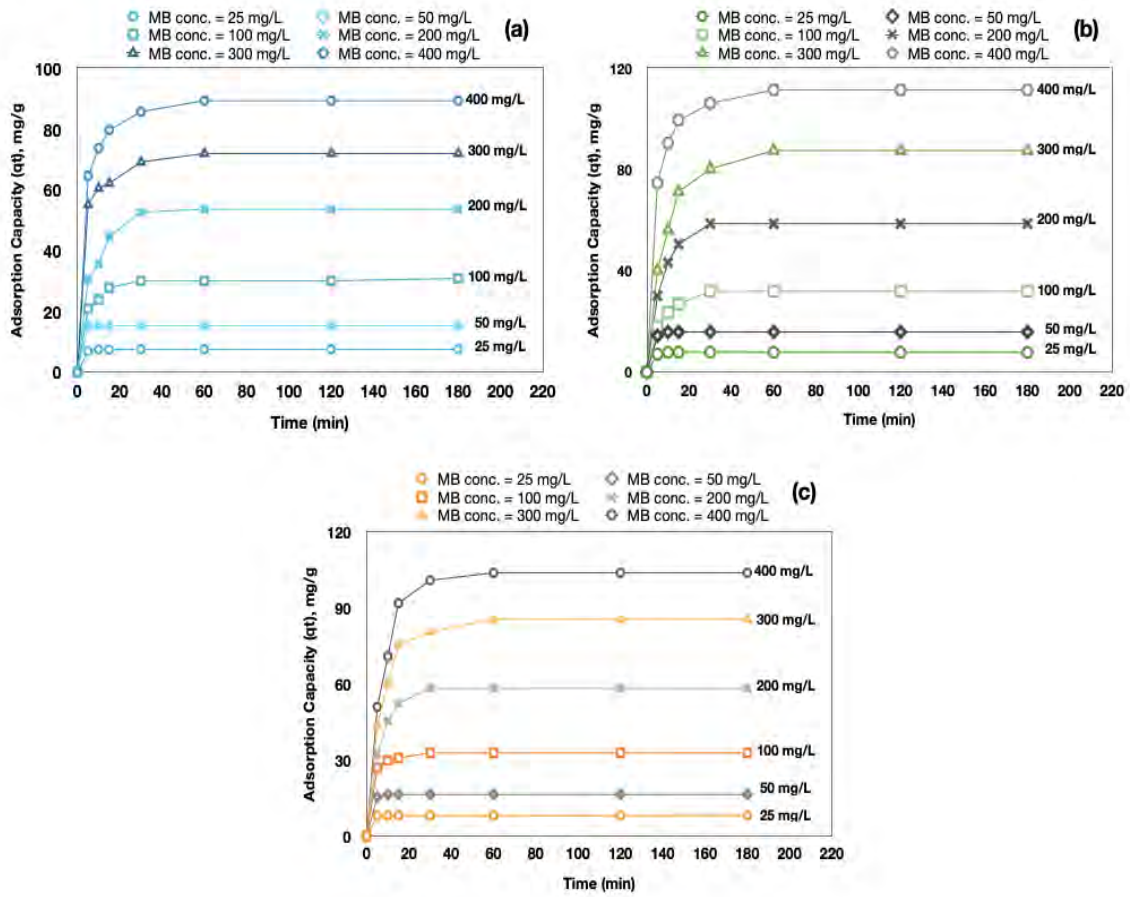


Fig. 5: Effect of contact time on methylene blue adsorption (a) $ZnCl_2$ -AC, (b) KOH-AC, (c) H_3PO_4-AC (adsorbent dose = 3 g/L, stirring speed = 120 rpm, temperature = 27±1 °C, volume of solution = 100 mL).

Figure 5a displays that the time to attain equilibrium was observed to be 10 min for 25 and 50 mg/L, whereas this increased to 60 min for a concentration of 300 and 400 mg/L. The increasing trend stopped once a state of equilibrium was achieved. At this stage, the quantity of MB was not significantly changed with time, which indicates that time is sufficient to accomplish equilibrium for the maximum elimination of dye from aqueous solutions by $ZnCl_2$-AC. A significant amount of dye was removed by KOH-AC (Fig. 5b) in the first 30 min of contact time due to the existence of a huge number of active sites on the surface of activated carbon. The adsorption equilibrium of methylene blue adsorption on KOH-AC increased from 7.75 to 111.36 mg/g as the initial concentration of methylene blue increased from 25 to 400 mg/L. Methylene blue adsorption onto the H_3PO_4-AC (Fig. 5c) followed a similar trend as methylene blue adsorption onto $ZnCl_2$-AC. At the equilibrium, the adsorption capacity increased from 8.02 to 103.86 mg/g, while the removal percentage dropped from 98.97 to 77.86 % with a rise in the initial methylene blue concentrations from 25 to 400 mg/L (see Fig. 4).

Generally, the equilibrium was achieved within approximately the first hour of adsorption, indicating a favorable interaction between the methylene blue dye and BFS-ACs. These results were similar to those reported by some previous studies [8,15,25,46].

https://doi.org/10.31436/iiumej.v22i2.1682

3.3 Adsorption Isotherms Models

Adsorption isotherms have been used to illustrate how the adsorbate molecules distribute between the liquid and solid phases. That is generally significant for the adsorption system design [12]. Adsorption isotherms results are typically presented as a plot of the adsorbed chemical concentration (mg/g) versus the remainder of the solution concentration (mg/L). In this study, the data of adsorption equilibrium was modeled to Langmuir and Freundlich isotherms.

3.3.1 Langmuir Isotherm

The Langmuir isotherm model suggests single-layer surface sorption with no transmigration of the sorbate on the surface without interplaying the sorbed molecules and the uniform energy of absorption [47]. Equation 7 describes the linear form of the Langmuir model and a plot of C_e/q_e against C_e is presented in Fig. 6. The slope and the interception are used to evaluate the maximum adsorption capacity (q_m) and the adsorption rate (K_L), respectively.

$$\frac{C_e}{q_e} = \frac{1}{q_m K_L} + \frac{C_e}{q_m} \tag{7}$$

where K_L (L/mg) and q_m (mg/g) are Langmuir constants related to rate of adsorption and adsorption capacity, respectively. C_e is the equilibrium concentration of the adsorbate (mg/L).

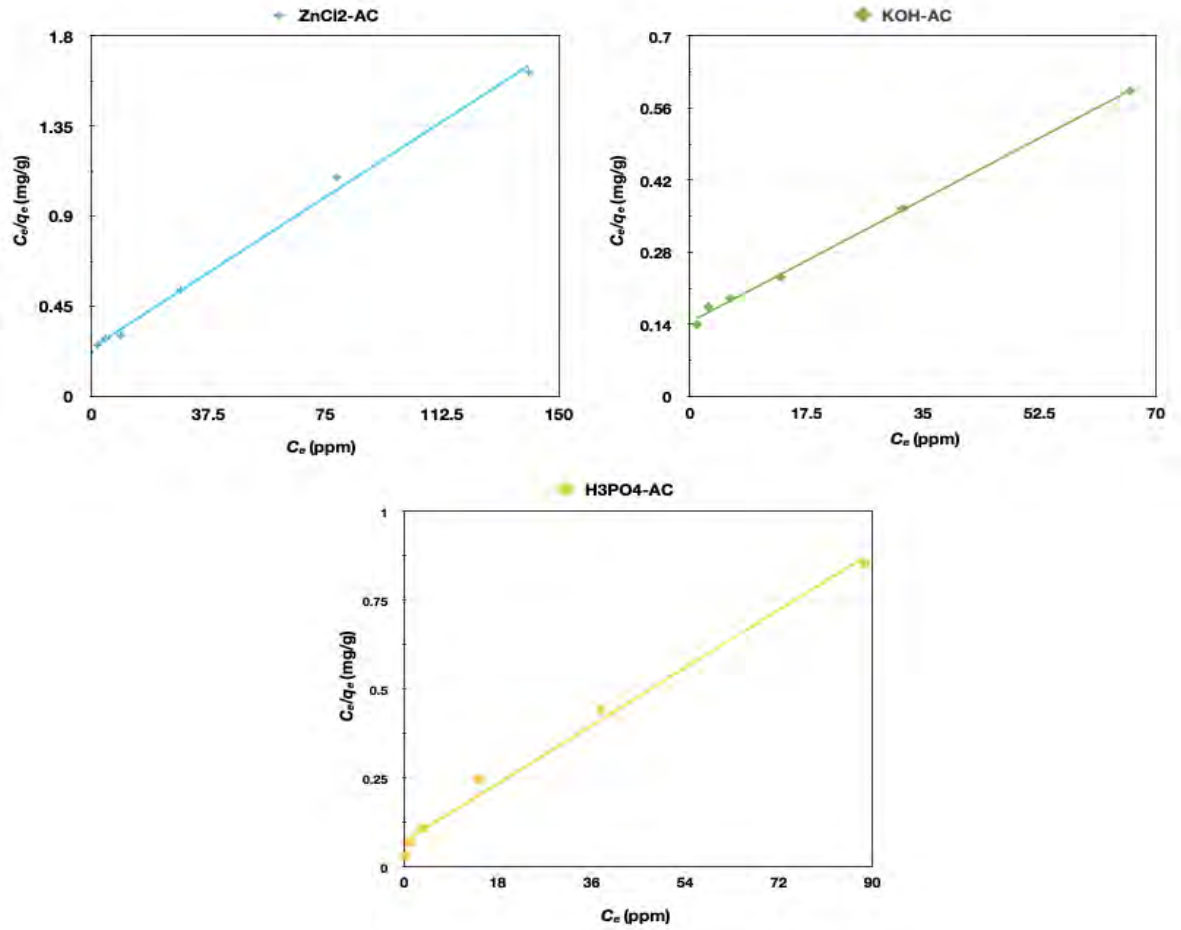


Fig. 6: Langmuir isotherm plot for the removal of methylene blue at 27±1 °C.

3.3.2 Freundlich Isotherm

The Freundlich isotherm is an empirical model based on the heterogeneous surface adsorption and on working sites with various energies [48]. The following equation indicates the linear form of the Freundlich isotherm:

$$log\ q_e = logK_F + \frac{1}{n} log\ C_e \tag{8}$$

where q_e is the amount of MB adsorbed at equilibrium (mg/g); C_e is the equilibrium concentration of the adsorbate (mg/L); K_F ((mg/g) (L/mg)$^{1/n}$) and n are Freundlich constants relating respectively to the adsorption capacity of the adsorbate and the favorability of adsorption process. If the value of $1/n$ is less than 1, the adsorption is favorable.

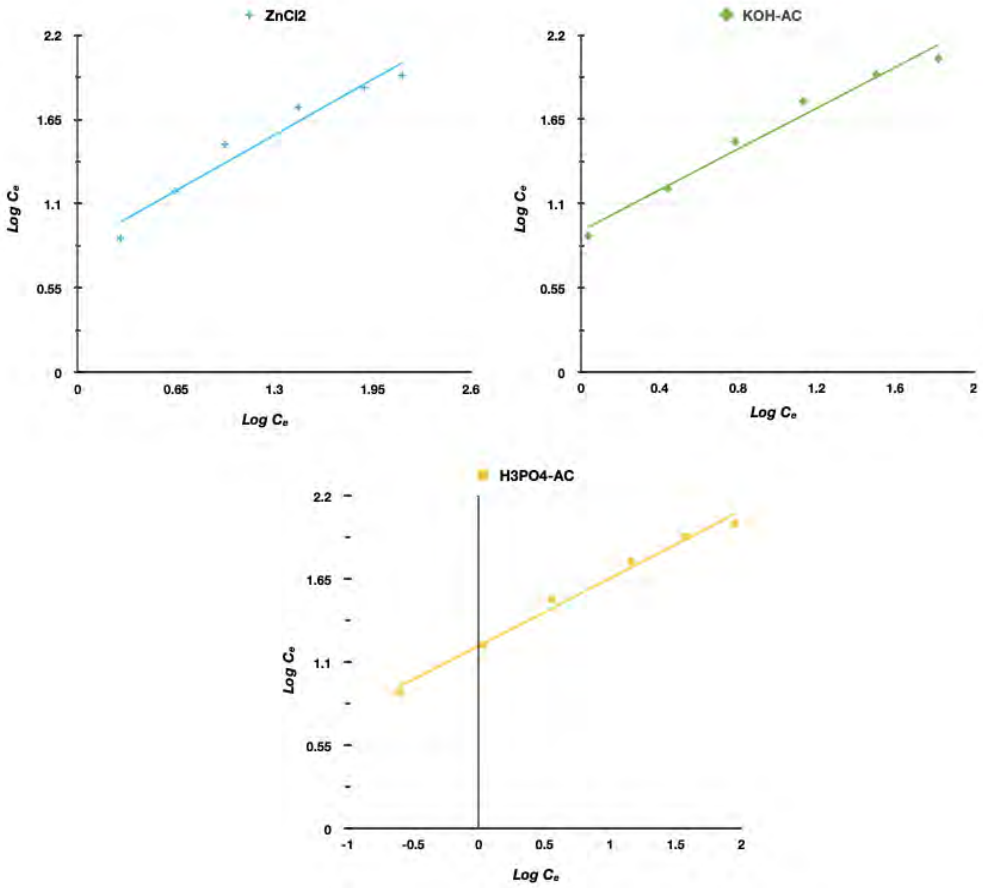


Fig. 7: Freundlich isotherm plot for removal of methylene blue onto synthesized BFS-ACs at 27±1 °C.

Table 2: Isotherm model parameters and correlation coefficients of MB adsorption onto synthesized BFS-ACs at 27±1 °C

Adsorbent	Langmuir parameters			Freundlich parameters			
	K_L (L/mg)	q_m (mg/g)	R^2	K_F (mg/g)(L/mg)$^{1/n}$	$1/n$ (L/mg)	n	R^2
$ZnCl_2$-AC	0.044	90.45	0.9963	6.663	0.557	1.793	0.9506
H_3PO_4-AC	0.1262	109.98	0.9891	13.188	0.446	2.238	0.9888
KOH-AC	0.060	113.63	0.9978	8.388	0.666	1.501	0.9770

https://doi.org/10.31436/iiumej.v22i2.1682

Table 2 lists all parameters and determination coefficients (R^2) of MB adsorption on BFS-ACs. For the Freundlich isotherm, the slope *1/n* of 0 to 1 tests the adsorption intensity or surface heterogeneity of the adsorption and is increasingly heterogeneous as its value reaches nil. A value of *1/n* lower than 1 suggests a normal Freundlich isotherm, whereas *1/n* higher than 1 suggests stronger sorption strength [49,50]. The results of adsorption processes showed that the values of *1/n* were 0.557, 0.446, and 0.666 for $ZnCl_2$, H_3PO_4, and KOH, respectively, indicating that the conditions of adsorption in this study are favorable. It could be observed that H_3PO_4 showed stronger affinity and greater heterogeneity for MB than $ZnCl_2$ and KOH.

Based on the R^2, the Langmuir model produced a better fit for the experimental data with higher R^2 for the three produced activated carbons, indicating that the adsorption process occurred on a uniform surface and there was no transmigration of the MB dyes. Therefore, the methylene blue adsorption onto the adsorbent tended to monolayer adsorption. Furthermore, the Langmuir model shows that the maximum adsorption capacities (q_m) were 90.45 mg/g, 109.98 mg/g, and 113.63 mg/g for $ZnCl_2$, H_3PO_4, and KOH, respectively. These agree well with the experimental results, which also showed that the adsorption process was in line with the Langmuir model. Table 3 shows a comparison of the maximum sorption capacities of the methylene blue with other reported values for some agricultural by-products based activated carbons. The findings from this table showed that BFS-ACs can effectively be used for the removal of MB dyes from the aqueous resolution with a great absorption capacity compared to many other sorbents.

Table 3: A comparison of maximum adsorption of methylene blue onto activated carbons produced from different precursors

Precursor	Activating agent	Adsorption capacity (mg/g)	References
Cashew nutshell	KOH	68.72	[51]
Pineapple waste biomass	$ZnCl_2$	288.34	[52]
Coconut (*Cocos nucifera*) Leaf	H_3PO_4	250	[53]
Wood	KOH	59.92	[54]
Corn cob residue	H_3PO_4	183.3	[55]
Baobab fruit shell	H_3PO_4	334.45	[18]
Baobab fruit shell	$ZnCl_2$	90.45	This Study
	H_3PO_4	109.98	
	KOH	113.63	

3.4 Adsorption Kinetic Studies

The time-dependent experimental data were evaluated in order to investigate the rate-limiting step by adapting them to different kinetic models, namely the pseudo-first order model [56] and the pseudo-second order model [57]. Linear forms of pseudo-first order and pseudo-second-order kinetic equations are represented by Equations (9) and (10), respectively.

$$\ln(q_e - q_t) = \ln q_e - k_1 t \tag{9}$$

$$\frac{t}{q_t} = \frac{1}{(k_2 q_e^2)} + \frac{1}{q_e} t \tag{10}$$

https://doi.org/10.31436/iiumej.v22i2.1682

where k_1 (min^{-1}) and k_2 (g/min/mg) are the rate constant of the pseudo-first order and pseudo second-order adsorption, q_e (mg/g) and q_t (mg/g) are the amounts of MB adsorbed onto BFS-ACs at the equilibrium and at any time t (min).

The values of first-order rate constants (k_1) and the equilibrium adsorption capacities (q_e) were determined from the slope and intercept of the plots of ln (q_e-q_t) against t, respectively. While the second-order sorption rate constants (k_2) and the equilibrium adsorption capacities (q_e) were determined from the slopes and intercepts of the plots of t/q_t versus t. Table 4 lists the derived kinetic parameters for the pseudo-first and pseudo-second order. From Table 4, the values of the rate constant pseudo-first-order model did not follow any particular pattern and the equilibrium adsorption capacities (q_e *cal*) were quite different and not in agreement with experimental data, and the correlation coefficients (R^2) of the pseudo-first order model were low values. On the other hand, the correlation coefficients (R^2) of the pseudo-second-order model are close or equal to 1 for all initial concentrations of the different carbons prepared, indicating that the experimental kinetic data fitted the pseudo-second-order model better (see Table 4). Furthermore, the equilibrium adsorption capacities calculated (q_e *cal*) values that were obtained from the second-order equations are very similar to equilibrium adsorption capacities experimental (q_e *exp*), suggesting that the second-order model is more suitable for the kinetics of the methylene blue and the rate-limiting step of MB onto BFS-ACs may be chemisorption. The rate constant (k_2) values decrease with the initial methylene blue concentration increasing owing to the lower competition for the sorption surface sites at lower concentrations [51].

Table 4: Kinetic model parameters and correlation coefficients for adsorption of MB onto $ZnCl_2$-AC, H_3PO_4-AC, and KOH-AC

Adsorbent	C_0 **(ppm)**	q_e **exp (mg/g)**	**Pseudo-first-order kinetics**			**Pseudo-second-order kinetics**		
			q_e **cal (mg/g)**	k_1**(/min)**	R^2	q_e **cal (mg/g)**	k_2 **(g/min/mg)**	R^2
$ZnCl_2$	25	7.472	0.00680	0.0297	0.51	7.4794	0.8392	1
	50	15.252	0.00023	0.0204	0.22	15.2671	1.3000	1
	100	30.805	35.1803	0.0306	0.74	30.8641	0.0187	0.9996
	200	53.583	36.8468	0.051	0.69	54.6448	0.0072	0.9993
	300	71.916	52.7351	0.0442	0.74	72.4637	0.0100	0.9998
	400	89.302	59.2788	0.0555	0.69	90.0900	0.0079	0.9999
KOH	25	7.750	0.000040	0.1015	0.2254	7.7579	1.3962	1
	50	15.805	0.0039	0.0217	0.33	15.8227	0.5547	1
	100	31.916	2.4027	0.0398	0.5143	32.3624	0.0148	0.9995
	200	58.583	0.2343	0.0886	0.5335	59.5238	0.0076	0.9993
	300	87.472	142.6921	0.1052	0.719	89.2857	0.0031	0.999
	400	111.361	85.9824	0.0926	0.719	112.3595	0.0056	0.9998
H_3PO_4	25	8.021	0.000037	0.0197	0.2182	8.0192	9.7188	1
	50	16.311	0.000110	0.0233	0.2252	16.3132	0.8946	1
	100	32.751	0.3016	0.0473	0.5372	32.8947	0.0486	1
	200	58.31	5.0944	0.0392	0.5014	59.1715	0.0095	0.9995
	300	85.521	106.1451	0.0626	0.696	86.9565	0.0044	0.9994
	400	103.861	120.0051	0.0620	0.6974	105.2631	0.0036	0.9992

4. CONCLUSION

The present study has put forward that baobab fruit shell is a promising precursor in generating high quality activated carbon at an affordable price. The results showed that KOH was found more effective than the other activating agents ($ZnCl_2$ and H_3PO_4) under the same conditions for the methylene blue removal. According to the results, properties of the activated carbon are significantly influenced by interaction with various chemical agents. Adsorption capacity of the adsorbent was affected by initial Methylene Blue concentration and contact time. SEM micrographs indicate that pores of varying sizes and shapes have been obtained from various active agents. Besides, chemical reagents had a significant effect on the nature of the surface functional groups as shown in FTIR results. The adsorption studies show that the pseudo second order model provided the best description of the kinetic uptake properties. On the other hand, the adsorption isotherms were well described by the Langmuir model. The results indicate that BFS-ACs is an effective adsorbent for MB adsorption from aqueous solutions.

ACKNOWLEDGEMENT

The authors are grateful to the Department of Biotechnology Engineering, Faculty of Engineering, International Islamic University Malaysia (IIUM) for providing continuous support to the research at hand. We are as well thankful to the International Institute for Halal Research and Training (INHART), IIUM for giving us access to FTIR instrument for characterization.

REFERENCES

[1] Özhan A, Şahin Ö, Küçük MM, Saka C. (2014) Preparation and characterization of activated carbon from pine cone by microwave-induced $ZnCl_2$ activation and its effects on the adsorption of methylene blue. Cellulose, 21: 2457-2467. https://doi.org/10.1007/s10570-014-0299-y

[2] Yang J, Qiu K. (2010) Preparation of activated carbons from walnut shells via vacuum chemical activation and their application for methylene blue removal. Chem Eng J, 165: 209-217. https://doi.org/10.1016/j.cej.2010.09.019

[3] Koo WK, Gani NA, Shamsuddin MS, Subki NS, Sulaiman MA. (2015) Comparison of wastewater treatment using activated carbon from bamboo and oil palm: an overview. J Trop Resour Sustain Sci, 3: 54-60

[4] Wang LG, Yan GB. (2011) Adsorptive removal of direct yellow 161dye from aqueous solution using bamboo charcoals activated with different chemicals. Desalination, 274: 81-90. https://doi.org/10.1016/j.desal.2011.01.082

[5] Gupta T. (2017) Carbon: The black, the gray and the transparent. Springer.

[6] Gupta VK, Tyagi I, Agarwal S, Singh R, Chaudhary M, Harit A, Kushwaha S (2016) Column operation studies for the removal of dyes and phenols using a low cost adsorbent. Glob J Environ Sci Manag, 2: 1-10. https://doi.org/10.7508/gjesm.2016.01.001

[7] Spagnoli AA, Giannakoudakis DA, Bashkova S. (2017) Adsorption of methylene blue on cashew nut shell based carbons activated with zinc chloride: The role of surface and structural parameters. J Mol Liq, 229: 465-471. https://doi.org/10.1016/j.molliq.2016.12.106

[8] Sogbochi E. (2017) Evaluation of Adsorption Capacity of Methylene Blue in Aqueous Medium by Two Adsorbents: The raw hull of lophira lanceolata and its activated carbon. Am J Phys Chem, 6(5): 76-87. https://doi.org/10.11648/j.ajpc.20170605.11

[9] Hameed BH, Din ATM, Ahmad AL. (2007) Adsorption of methylene blue onto bamboo-based activated carbon: Kinetics and equilibrium studies. J Hazard Mater, 141: 819-825. https://doi.org/10.1016/j.jhazmat.2006.07.049

[10] Islam MA, Sabar S, Benhouria A, Khanday WA, Asif M, Hameed BH. (2017) Nanoporousactivated carbon prepared from karanj (Pongamia pinnata) fruit hulls for methylene blue adsorption. J Taiwan Inst Chem Eng, 74: 96-104. https://doi.org/10.1016/j.jtice.2017.01.016

[11] Guo J, Song Y, Ji X, Ji L, Cai L, Wang Y, Zhang H, Song W. (2019) Preparation and characterization of nanoporous activated carbon derived from prawn shell and its application for removal of heavy metal ions. Materials, 12(2): 1-17. https://doi.org/10.3390/ma12020241

[12] Gao J, Kong D, Wang Y, Wu J, Sun S, Xu P. (2013a) Production of mesoporous activated carbon from tea fruit peel residues and its evaluation of methylene blue removal from aqueous solutions. BioResources, 8: 2145-2160. https://doi.org/10.15376/biores.8.2.2145-2160

[13] Gao JJ, Qin YB, Zhou T, Cao DD, Xu P, Hochstetter D, Wang YF. (2013b) Adsorption of methylene blue onto activated carbon produced from tea (*Camellia sinensis* L.) seed shells: Kinetics, equilibrium, and thermodynamics studies. J Zhejiang Univ Sci B, 14: 650-658. https://doi.org/10.1631/jzus.B12a0225

[14] Vicinisvarri I, Shanker Kumar S, Nor Aimi AW, Norain I, Nurul Izza H. (2014) Preparation and characterization of phosphoric acid activated carbon from *Canarium odontophyllum* (Dabai) nutshell for methylene blue adsorption. Res J Chem Environ, 18: 57-62

[15] Banat F, Al-Asheh S, Makhadmeh L. (2003) Preparation and examination of activated carbons from date pits impregnated with potassium hydroxide for the removal of methylene blue from aqueous solutions. Adsorpt Sci Technol, 21: 597-606. https://doi.org/10.1260/026361703771953613

[16] Ndi Nsami J, Ketcha Mbadcam J. (2013) The adsorption efficiency of chemically prepared activated carbon from cola nut shells by $ZnCl_2$ on methylene blue. J Chem, 2013: 1-7. https://doi.org/10.1155/2013/469170

[17] Idris-Hermann KT, Raoul TTD, Giscard D, Gabche AS. (2018) Preparation and characterization of activated carbons from bitter kola (*Garcinia kola*) nut shells by chemical activation method using H_3PO_4, KOH and $ZnCl_2$. Chem Sci Int J, 23: 1-15. https://doi.org/10.9734/csji/2018/43411

[18] Vunain E, Biswick T. (2018) Adsorptive removal of methylene blue from aqueous solution on activated carbon prepared from Malawian baobab fruit shell wastes: Equilibrium , kinetics and thermodynamic studies and thermodynamic studies. Sep Sci Technol, 54(1): 1-15. https://doi.org/10.1080/01496395.2018.1504794

[19] García JR, Sedran U, Zaini MAA, Zakaria ZA. (2018) Preparation, characterization, and dye removal study of activated carbon prepared from palm kernel shell. Environ Sci Pollut Res, 25: 5076-5085. https://doi.org/10.1007/s11356-017-8975-8

[20] Sahira J, Mandira A, Prasad PB, Ram PR. (2013) Effects of activating agents on the activated carbons prepared from Lapsi seed stone. Res J Chem Sci, 3: 19-24

[21] Bedin KC, Martins AC, Cazetta AL, Pezoti O, Almeida VC. (2016) KOH-activated carbon prepared from sucrose spherical carbon: Adsorption equilibrium, kinetic and thermodynamic studies for Methylene Blue removal. Chem Eng J, 286: 476-484. https://doi.org/10.1016/j.cej.2015.10.099

[22] Hong Tan Y, Xian Chin S, Lun Ang W, Mahmoudi E, Mohd Zainoodin A, Mohammad AW. (2018) Effect of H_3PO_4 and KOH as the Activating agents on the synthesis of low-cost activated carbon from duckweeds plants. J Kejuruter, 1(4): 37-43. https://doi.org/10.17576/jkukm-2018-si1(4)-05

[23] Baccar R. (2013) Removal of some water contaminants by adsorption on activated carbon prepared from olive-waste cakes and biological treatment using fungi. PhD thesis. Sfax University, Chemical Engineering Department.

[24] Hayashi J, Horikawa T, Muroyama K, Gomes VG. (2002) Activated carbon from chickpea husk by chemical activation with K_2CO_3: Preparation and characterization. Microporous Mesoporous Mater, 55: 63-68. https://doi.org/10.1016/S1387-1811(02)00406-7

[25] Ben Nasr J, Hamdi N, Elhalouani F. (2017) Characterization of activated carbon prepared from sludge paper for methylene blue adsorption. J Mater Environ Sci, 8: 1960-1967

[26] Alcañiz-Monge J, Illán-Gómez MJ. (2008) Insight into hydroxides-activated coals: Chemical or physical activation? J Colloid Interface Sci, 318: 35-41. https://doi.org/10.1016/j.jcis.2007.10.017

[27] Martins AC, Pezoti O, Cazetta AL, Bedin KC, Yamazaki DAS, Bandoch GFG, Asefa T,Visentainer JV, Almeida VC. (2015) Removal of tetracycline by NaOH-activated carbon produced from macadamia nut shells: Kinetic and equilibrium studies. Chem Eng J, 260: 291-299. https://doi.org/10.1016/j.cej.2014.09.017

[28] Dzigbor A, Chimphango A. (2019) Production and optimization of NaCl-activated carbon from mango seed using response surface methodology. Biomass Convers Biorefinery, 9: 421-431. https://doi.org/10.1007/s13399-018-0361-3

[29] Kabbashi NA, Mirghani MES, Alam MZ, Qudsieh SY, Bello IA. (2017) Characterization of the Baobab fruit shells as adsorption material. Int Food Res J, 24: 472-474.

[30] Ofomaja AE. (2007) Sorption dynamics and isotherm studies of methylene blue uptake on to palm kernel fibre. Chem Eng J, 126: 35-43. https://doi.org/10.1016/j.cej.2006.08.022

[31] Al-Khatib MFR, Munjid M. (2011) Production of activated carbon from palm oil empty fruit bunch by chemical activation. in: Current research and developments in biotechnology engineering at IIUM. IIUM Press, Kuala Lumpur, pp 209-216

[32] Milne T, Brennan AH, Glenn BH. (1990) Sourcebook of methods of analysis for biomass and biomass conversion processes. Springer Science & Business Media.

[33] American Society for Testing and Materials. (1991) Standard for method for moisture in activated carbon ASTM D 2867-91. PA: ASTM Committee on Standards, Philadelphia

[34] Prahas D, Kartika Y, Indraswati N, Ismadji S. (2008) Activated carbon from jackfruit peel waste by H_3PO_4 chemical activation: Pore structure and surface chemistry characterization. Chem Eng J, 140: 32-42. https://doi.org/10.1016/j.cej.2007.08.032

[35] J Hassen JH. (2017) Effect of KOH ratio on the formation of activated carbon from pressed wood residues. Int J Pharm Sci Res, 8: 4875-4880. https://doi.org/10.13040/IJPSR.0975-8232.8(11)4875-80

[36] Vunain E, Kenneth D, Timothy B. (2017) Synthesis and characterization of low-cost activated carbon prepared from Malawian baobab fruit shells by H_3PO_4 activation for removal of Cu (II) ions: equilibrium and kinetics studies. Appl Water Sci, 7: 4301-4319. https://doi.org/10.1007/s13201-017-0573-x

[37] Nunes CA. (2011) Estimation of surface area and pore volume of activated carbons by methylene blue and iodine numbers. Quim Nova, 34(3): 472-476. https://doi.org/10.1590/S0100-40422011000300020

[38] Mona S, Kaushik A, Kaushik CP. (2011) Biosorption of reactive dye by waste biomass of Nostoc linckia. Ecol Eng, 37: 1589-1594. https://doi.org/10.1016/j.ecoleng.2011.04.005

[39] Gottipati R. (2012) Preparation and characterization of microporous activated carbon from biomass and its application in the removal of chromium (VI) from aqueous phase. PhD thesis. Natl Inst Technol Rourkela Odisha, India. Department of Chemical Engineering.

[40] Abdullah AH, Kassim A, Zainal Z, Mohd Zobir H, Kuang D, Wooi SO, Ahmad F. (2001) Preparation and characterization of activated carbon from Gelam wood bark (*Melaleuca cajuputi*) Malaysian J Anal Sci, 7: 65-68

[41] Marsh H, Rodríguez F. (2006) Activated Carbon. Elsevier

[42] Jawad AH, Abdulhameed AS. (2020) Statistical modeling of methylene blue dye adsorption by high surface area mesoporous activated carbon from bamboo chip using KOH-assisted thermal activation. Energy, Ecol Environ, 5: 456-469. https://doi.org/10.1007/s40974-020-00177-z

[43] Puziy AM, Poddubnaya OI, Martínez-Alonso A, Suárez-García F, Tascón JMD. (2005) Surface chemistry of phosphorus-containing carbons of lignocellulosic origin. Carbon N Y, 43: 2857-2868. https://doi.org/10.1016/j.carbon.2005.06.014

[44] Kumar A, Jena HM. (2016) Removal of methylene blue and phenol onto prepared activated carbon from Fox nutshell by chemical activation in batch and fixed-bed column. J Clean Prod, 137: 1246-1259. https://doi.org/10.1016/j.jclepro.2016.07.177

[45] De La Luz-Asunción M, Sánchez-Mendieta V, Martínez-Hernández AL, Castaño VM, Velasco-Santos C. (2015) Adsorption of phenol from aqueous solutions by carbon nanomaterials of one and two dimensions: Kinetic and equilibrium studies. Journal of Nanomaterials. https://doi.org/10.1155/2015/405036

[46] Dural MU, Cavas L, Papageorgiou SK, Katsaros FK. (2011) Methylene blue adsorption on activated carbon prepared from Posidonia oceanica (L.) dead leaves: Kinetics and equilibrium studies. Chem Eng J, 168: 77-85. https://doi.org/10.1016/j.cej.2010.12.038

[47] Vimala R, Das N. (2009) Biosorption of cadmium (II) and lead (II) from aqueous solutions using mushrooms: A comparative study. J Hazard Mater, 168: 376-382. https://doi.org/10.1016/j.jhazmat.2009.02.062

[48] Kilic M, Apaydin-Varol E, Pütün AE. (2011) Adsorptive removal of phenol from aqueous solutions on activated carbon prepared from tobacco residues: Equilibrium, kinetics and thermodynamics. J Hazard Mater, 189: 397-403. https://doi.org/10.1016/j.jhazmat.2011.02.051

[49] Angin D, Altintig E, Köse TE. (2013) Influence of process parameters on the surface and chemical properties of activated carbon obtained from biochar by chemical activation. Bioresour Technol, 148: 542-549. https://doi.org/10.1016/j.biortech.2013.08.164

[50] Angin D. (2014) Utilization of activated carbon produced from fruit juice industry solid waste for the adsorption of Yellow 18 from aqueous solutions. Bioresour Technol, 168: 259-266. https://doi.org/10.1016/j.biortech.2014.02.100

[51] Kumar PS, Ramalingam S, Sathishkumar K. (2011) Removal of methylene blue dye from aqueous solution by activated carbon prepared from cashew nut shell as a new low-cost adsorbent. Korean J Chem Eng, 28: 149-155. https://doi.org/10.1007/s11814-010-0342-0

[52] Mahamad MN, Zaini MAA, Zakaria ZA. (2015) Preparation and characterization of activated carbon from pineapple waste biomass for dye removal. Int Biodeterior Biodegrad, 102: 274-280. https://doi.org/10.1016/j.ibiod.2015.03.009

[53] Jawad AH, Sabar S, Ishak MAM, Wilson LD, Ahmad Norrahma SS, Talari MK, Farhan AM. (2017) Microwave-assisted preparation of mesoporous-activated carbon from coconut (*Cocos nucifera*) leaf by H_3PO_4 activation for methylene blue adsorption. Chem Eng Commun, 204: 1143-1156. https://doi.org/10.1080/00986445.2017.1347565

[54] Danish M, Ahmad T, Hashim R, Said N, Akhtar MN, Mohamad-Saleh J, Sulaiman O. (2018) Comparison of surface properties of wood biomass activated carbons and their application against rhodamine B and methylene blue dye. Surfaces and Interfaces, 11: 1-13. https://doi.org/10.1016/j.surfin.2018.02.001

[55] Jawad AH, Bardhan M, Islam MA, M.A. Islam, Syed-Hassan SSA, Surip SN, Alothman ZA, Khan MR. (2020) Insights into the modeling, characterization and adsorption performance of mesoporous activated carbon from corn cob residue via microwave-assisted H_3PO_4 activation. Surfaces and Interfaces, 21:100688. https://doi.org/10.1016/j.surfin.2020.100688

[56] Budinova T, Ekinci E, Yardim F, Grimm A, Björnbom E, Minkova V, Goranova M. (2006) Characterization and application of activated carbon produced by H_3PO_4 and water vapor activation. Fuel Process Technol, 87: 899-905. https://doi.org/10.1016/j.fuproc.2006.06.005

[57] Ho YS, McKay G. (1999) Pseudo-second order model for sorption processes. Process Biochem, 34: 451-465. https://doi.org/10.1021/acs.oprd.7b00090

https://doi.org/10.31436/iiumej.v22i2.1687

CYTOTOXICITY EFFECT OF IONIC LIQUID-GRAVIOLA FRUIT (*Annona muricata*) EXTRACT TO HUMAN COLON CANCER (HT29) CELL LINES

DJABIR DADDIOUAISSA[1,2*], AZURA AMID[2], NASSERELDEEN AHMED KABBASHI[1], AHMED ADAM MOHAMMED ELNOUR[1,2] AND MOHAMAD ADIKA KHAIRY BIN MOHD SHAIFUDIN EPANDY[3]

[1]Biotechnology Engineering Department, Kulliyyah of Engineering,
[2]International Institute for Halal Research and Training (INHART),
International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia
[3]Adikafirdaus Resources, Lot 24, Jalan Klebang Selatan, 2/5 Kampung Tersusun,
Batu 6 Klebang Selatan, 31200 Ipoh, Perak, Malaysia

[*]Corresponding author: daddiouaissa.djabir@live.iium.edu.my

(Received: 3rd November 2020; Accepted: 31st March 2021; Published on-line: 4th July 2021)

ABSTRACT: The present study aimed to investigate the anti-proliferative effect of the ionic liquid-Graviola fruit (IL-GFE) extract on colon adenocarcinoma (HT29) cell lines and their kinetics behaviour to assess the Graviola fruit potential as a therapeutic alternative in cancer treatment. The phytoconstituents content of IL-GFE was identified using GC-TOFMS apparatus and measured its cytotoxicity on HT29 by tetrazolium bromide. Then the cytokinetic behaviour of the treated HT29 cells with IL-GFE was illustrated using the cells' growth curve. Besides, the cell cycle phase perturbation for the treated HT29 was applied using a flow cytometry technique. Qualitative identification of phytoconstituents of IL-GFE showed that Graviola fruit contains acetogenins, alkaloids, flavonoids, tannins and saponins compounds. IL-GF extract displayed a cytotoxicity effect on HT29 cells with the IC_{50} value of 10.56 µg/mL, while Taxol showed an IC_{50} value of 1.22 µg/mL. IL-GFE also decreased the cell generation number from 3.93 to 2.96 generations compared to Taxol-treated cells 2.01 generations. The microscope observation of the HT29 cells treated with the crude IL-GFE displayed loss of density and cell detachment. The extract's growth inhibition was related to the cell cycle arrest at the G0/G1 phase. IL-GFE inhibited colon adenocarcinoma HT29 cells' proliferation and affected their kinetic behaviour by lowering cell viability, inducing apoptosis, and arresting the cell cycle at the G0/G1 phase.

ABSTRAK: Kajian ini bertujuan untuk mengkaji kesan anti-proliferatif ekstrak buah-ion Graviola (IL-GFE) pada garis sel adenokarsinoma kolon (HT29) dan tingkah laku kinetik mereka untuk menilai potensi buah Graviola sebagai alternatif terapi untuk barah rawatan. Kandungan fitokonstituen IL-GFE dikenal pasti menggunakan alat GC-TOFMS dan mengukur sitotoksisitasnya pada HT29 oleh tetrazolium bromida. Kemudian tingkah laku sitokinetik sel HT29 yang dirawat dengan IL-GFE digambarkan menggunakan keluk pertumbuhan sel. Selain itu, gangguan fasa kitaran sel untuk HT29 yang dirawat diaplikasikan menggunakan teknik sitometri aliran. Pengenalpastian kualitatif fitokonstituen IL-GFE menunjukkan bahawa buah Graviola mengandungi asetogenin, alkaloid, flavonoid, tanin dan sebatian saponin. Ekstrak IL-GF memperlihatkan kesan sitotoksisiti pada sel HT29 dengan nilai IC_{50} 10.56 µg/mL, sementara Taxol menunjukkan nilai IC_{50} 1.22 µg/mL. IL-GFE juga menurunkan jumlah penjanaan sel dari 3.93 hingga 2.96 generasi berbanding sel yang dirawat Taxol 2.01 generasi. Pemerhatian mikroskop

sel HT29 yang dirawat dengan IL-GFE kasar menunjukkan kehilangan ketumpatan dan detasmen sel. Perencatan pertumbuhan ekstrak berkaitan dengan penangkapan kitaran sel pada fasa G0/G1. IL-GFE menghalang percambahan sel HT29 adenokarsinoma kolon dan mempengaruhi tingkah laku kinetik mereka dengan menurunkan daya maju sel, mendorong apoptosis, dan menghentikan kitaran sel pada fasa G0/G1.

KEYWORDS: *colon cancer; cell cycle; flow cytometry; Graviola (Annona muricata); growth kinetics; GC-TOFMS; ionic liquids*

1. INTRODUCTION

Fruit consumption is hugely recommended to preserve the healthy life of a human being. *Annona muricata* is known as Graviola, also called soursop in English, due to its sour and delicious fruit [1]. Graviola is predominately harvested in the tropical and subtropical regions of the world. Nowadays, Graviola fruit is vastly utilised to produce syrups, beverages and candies [2]. All Graviola tree parts have ethnobotanically reported being used in alternative medicine, mainly the fruit that is obtained to reduce worms and parasites [3], treat fever and enhance mother's milk [4], as well as renal, liver infections [5], hypertension [6] and cancer treatment [7]. The frequent use of Graviola in treating cancer may be related to its cytotoxic selectivity to cancer cells [4,8,9]. Also, studies have reported that the toxicity of Graviola to the cancer cells was more than healthy cells [10-12].

The potential anticancerous of the Graviola fruit extract was reported to be related to the phytochemicals called annonaceous acetogenins that are thought to be cytotoxic against different cancer cell lines, as highlighted in our previous study [13]. Recently, researchers are interested in the antitumor activity of acetogenins, as these are the main compounds isolated from the fruit [14]. The latest study confirmed that the IL-GFE showed selectivity toward cancer cells by affecting only the breast adenocarcinoma (MCF-7) cell lines but not the normal (VERO) cell lines, even with a high dose of 100 µg/mL [15]. This finding is in agreement with a study conducted by Dai et al. [11]. Hence, the Graviola fruit can be a promising substitute or complementary supplement to reduce tumour cell proliferation generally.

Colon adenocarcinoma is a frequently high incident rate in both sexes, with an estimation of 101,420 new cases and 51,020 colon cancer (CC) deaths in the United States in 2019 [16]. Therefore, it is hypothesised that the recent rapid declines in the colon cancer rate are due to the increased uptake of colonoscopy, which now is a useful screening test. Common used protocols in cancer treatment such as radiotherapy, chemotherapy and surgical intervention cause various side effects such as nausea, fatigue, hair loss, vomiting and weak immune system [17]. Hence, new approaches to treat cancer are highly recommended. Lately, our research team has successfully extracted bioactive compounds from Graviola fruit by ionic liquid-microwave-assisted extraction technique and showed a positive effect on breast cancer (MCF-7) cell line while it was safe toward normal VERO cell line [15]. Moreover, this extract did not cause any significant toxicity on the morphology of the treated zebrafish model [18].

Ionic liquid-based microwave-assisted extraction (IL-MAE) method has been applied successfully to the efficient extraction of bioactive compounds from medicinal plants. This is related to the high extraction efficiency, shorter extraction time and the unique properties of ILs which is proposed to be an environment-friendly technique in sample preparation. This may replace the well-known solvent used in extraction technique such as methanol (MeOH) in which, its toxicity makes it undesirable for eco-friendly approaches [19]. It is expected that the IL-MAE method would be able to extract new bioactive compounds from

Graviola fruit that are safe for human consumption to be used in cancer treatment. Many outcomes achieved by using ionic liquid-based processes in the extraction and isolation of several bioactive compounds such as small natural compounds from biomass, lipids, proteins, amino acids and pharmaceuticals [20].

Bogdanov [21] has reviewed the mechanism of ILs action in the process of solid-liquid extraction from the plant, showing that the ILs can attribute to the solvent-matrix interactions, which leads to the modification of permeability of the plant matrix. This is due to the disruption of the cell tissue and modifying the matrix permeability through the interaction of the H bonding with the carbohydrates from the cell walls. A previous study carried out by Cláudio et al. [22] reported that aqueous solution of [C4mim]Cl is the appropriate extractive solvent of caffeine from *Paullinia cupana* seeds (guaraná) in which, at the optimum extraction conditions of (2.34 M [C4mim]Cl, 70 °C, 30 min, s/l ratio 1:10) the extraction efficiency was enhanced by 50% compared with the one obtained by Soxhlet extraction method using dichloromethane solvent. Another study conducted by Wang and co-workers. [23], reported that the 1-hexyl-3-methylimidazolium tetrafluoroborate [C6mim][BF4] water solution was used to extract lipophilic and hydrophilic metabolites from chrysanthemums and showed better extraction performance for the desired product than methanol. The ILs extraction ability relies on their physicochemical characteristics, such as hydrophobicity, hydrogen bonding ability and viscosity, which may have promising applications in the extraction of the natural product from plant supplies [24]. Thus, this research aimed to explore the phytochemical content and the potential anti-proliferative of the ionic liquid extract of Graviola fruit (IL-GFE) toward colon adenocarcinoma cell lines using MTT assay, cytokinetic behaviour and the flow cytometry technique.

2. MATERIALS AND METHODS

2.1 Chemicals and Reagents

All experimental chemicals utilised in this study were laboratory grade. The 1-Butyl-3-Methylimidazolium Chloride [C4MIM][Cl] 96% solution was obtained from (Alfa, USA). The N,O bis(trimethylsilyl)-trifluoroacetamide (BSTFA) 99% was purchased from (Sigma-Aldrich, USA). Media for cell growth include Dulbecco's modified Eagle's medium (DMEM), fetal bovine serum (FBS) and penicillin/streptomycin were purchased from (GIBCO®, USA). For the anti-proliferative assay, the 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyl-tetrazolium bromide (MTT reagent) and the accutase solution (cell detachment) were obtained from (ICT, USA). In addition, the RNase A / Prepodium iodide kit of the cell cycle assay was purchased from (Beckman Coulter, USA). All these reagents were utilised to evaluate the anti-proliferative activity of IL-GF extract on colon adenocarcinoma cell lines.

2.2 Fruit Sample Preparation

Graviola fruit samples (DB3) were chosen randomly from the farm of Adikafirdaus (4.5921° N, 101.0901° E), Perak state, Malaysia. After that, the fruits were properly cleaned to eliminate all traces of insects and dust, removed the pericarp and seeds. The fruit pulp was then frozen at -20 °C for 72 hrs, then freeze-dried using a lyophiliser (Christ Alpha) based on the methodology of Hoeing [25]. Next, the dried sample was weighed and stored in falcon tubes at -20 °C for further analysis [26]. The fruit samples were stored in the herbarium with voucher number KAED/HBL/S1A047/2018/707 at the Kulliyyah of Architecture and Environmental Design, IIUM. The plant name was checked on www.theplantlist.org.

2.3 Fruit Sample Extraction

The dried Graviola fruit sample was extracted as described in our previous study [15]. In brief, the dried sample was mixed at a ratio of [1:30] g/mL of 0.5 mol/L of [C4MIM][Cl] reagent. The extract suspension was heated under a microwave oven with irradiation power of 700 W for 3 min. Then, the fruit extract (IL-GFE) was cooled down to room temperature and then filtered using a 3 mm Whatman filter paper. A back extraction to recover the ionic liquid was applied using phase separation with ethyl acetate at a ratio of 1:2 [27]. The extract was then lyophilised by freeze drier for 72 hrs. Finally, the dried extract was stored at 4 °C chiller until further analysis [28,29].

2.4 GC-TOFMS Sample Preparation and Analysis

The phytochemicals content of IL-GFE was identified using gas chromatography (GC–TOFMS) analysis; the dry IL-GF extract was silylated by using N,O bis(trimethylsilyl)-trifluoroacetamide (BSTFA) / acetonitrile (40:60; 1000 µL) and incubated at 60 °C for one hour. The sealed vials were vortexed at an interval of 1 min to achieve complete silylation. Silyl derivatisation is a common process to raise the volatility and detectability of the chemical compound in GCMS analysis [30]. In this research, GC-TOFMS analysis of the IL-GFE used a 7890A Agilent-Technologies GC system formed by a 7693-mass spectrometer system detector (Agilent Technologies, USA). The method of Muhamad et al. [31] was adopted with modification. The MS detector was run with an electron energy of 70 eV and a mass range of m/z 50–550. An HP-5 MS capillary column of (30 m × 0.25 mm, 0.25 µm film thickness) was utilised for separation. The GC oven temperature was set at 80 °C for 2 min, then raised to 240 °C at 5 °C/min and held for 5 min then finally increased to 300 °C at a rate of 3 °C/min and kept for 5 min. The temperature of the GC injector and MS transfer line was set at 225 °C and 300 °C, respectively. The sample was injected in the splitless mode at a ratio of 1:10. The carrier gas (helium) was utilised at a constant flow rate of 1.2 mL/min. Consequently, essential compounds were identified by comparing their retention time and mass spectra with those of the standard and authentic compounds at the National Institute of Standard (NIST) library.

2.5 Cell Culture

In This Study, the colon adenocarcinoma HT29 cell line (ATCC No: HTB-22™) was used as the experimental tumour cell. These cells were obtained from the American Type Collection Culture and stored in liquid nitrogen until further utilisation. After thawing the frozen cells, they were inoculated into 25 cm^2 T-flask with 5 mL of complete media, including DMEM complemented with 10% FBS and 1% pen/strip. Next, the culture cells were incubated at a 95% humidified incubator with 5% CO_2 at 37 °C. Then, the cells were used for further experimental analysis when they reached approximately 80% confluence.

2.6 The MTT Cell Viability Assay

The viable cell number was quantified using the tetrazolium-based colourimetric (MTT assay). The confluent cells were detached from the bottom of the T-flask by adding 2 mL of accutase, and then the cell was counted using a haemocytometer. Almost 5.0 x 10^4 cells were suspended in 100 µL of media and plated in each well of 96-well plate then incubated in a 5% CO_2 incubator for 24 hrs. After removing the media, 20 µL of serial dilutions (6.25-400 µg/mL) of the IL-GF extract in media free-FBS were added to the cells (triplicate wells per condition). The negative control received only DMSO, and the positive control treated with Taxol (0.39-6.25 µg/mL). The treated cells were incubated for 48 hrs, then added 20 µL of MTT reagent (5 mg/mL) into different wells. The incubation was then extended for another 3 hours then the media was carefully discarded. Next, the solubilisation liquid

(DMSO) (100 µL) was applied to each well to solubilise formazan then shake for 15 min at room temperature [32]. Finally, the formazan absorbance was measured at 570 nm by a microplate reader. The percentage of viable/dead cells were calculated, and the IC_{50} of IL-GF extract toward HT29 was determined using equation 1:

$$Percentage\ Cell\ Viability\ (\%) = \frac{Abs\ of\ treated\ cells}{Abs\ of\ untreated\ cells} x\ 100 \quad (1)$$

2.7 The 50% Inhibitory Concentration Estimation

The inhibitory effect of the IL-GF extract on the colon cancer cell growth was shown as the inhibitory concentration that induced 50% inhibition of the cell population (IC_{50}). This IC_{50} value was obtained graphically by plotting the percentage of viable cells against the associated different concentrations of the IL-GF extract used. Findings were presented as the mean ± one standard error (SE), and statistically, the $P \leq 0.05$ level was considered acceptable.

2.8 Growth Kinetics Study

The growth kinetics study was performed to establish a model that predicted the IL-GFE effect on treated and non-treated colon adenocarcinoma HT29 cell lines and compared it to Taxol-treated cells as a positive control. First, HT29 cells were counted and inoculated into 25 cm^2 T-flask at 2.0×10^5 cells/mL and incubated for 24 hrs. There were three groups of flasks in this study (triplicate flasks per condition): untreated HT29, Taxol-treated HT29 at the IC_{50} value of 1.22 µg/mL and HT29 treated with IL-GF extract at the IC_{50} value of 10.56 µg/mL [33]. The treated cells in 20 flasks per group were counted every 8 hrs interval (from 0 to 144 hrs). After incubation, the media was discarded, and the cells received fresh media containing the samples, then incubated from 0 to 144 hrs. Cell images were observed and recorded every 8 hours using a 10X magnification microscope. After that, cells were harvested from a monolayer and quantified using a trypan blue dye exclusion assay [34]. Finally, the cell viability number was determined by using Eq. 2:

$$C = n\ x\ 2\ x\ 10^4 \quad (2)$$

Where n presents the average of the quantified cells, the number 10^4 is the conversion of volume 0.1 mm^3 to 1 millilitre, and 2 is the dilution factor. The growth kinetics graph was plotted using the resulted data to calculate the number of cell generation, the specific growth rate and doubling time by using Eqs. (3), (4) and (5).

$$X = \frac{Log_{10}N - Log_{10}N_0}{Log_{10}2} \quad (3)$$

$$Log_{10}N = Log_{10}N_0 + \mu t \quad (4)$$

$$t_d = \frac{Log_{10^2}}{\mu} = \frac{0.301}{\mu} \quad (5)$$

Where (X) is the number of cell generation; (N) is the number of cells at the end; (N_0) is the primary number of cells; (μ) is the specific growth rate; (t) is the duration of treatment; (t_d) is the doubling time.

2.9 Cell Cycle Assay

The cell cycle phase perturbation in IL-GF extract-treated HT29 was analysed in a time-dependent manner using flow cytometry. First of all, approximately 5.0×10^4 of colon HT29 cell lines were inoculated into 25 cm^2 T-flask and incubated for 24 hrs. After that, the media

was discarded and replaced with new media contains IL-GF extract at the IC_{50} concentration value, then incubated progressively for 24, 48 and 72 hrs. After harvesting the cells with accutase, cells were thoroughly washed using 1X PBS and centrifuged at 1500 rpm for 5 min. The cells were fixed for one hour with 2 mL of 70% ethanol then centrifuged at 4000 rpm for 10 min at room temperature. Then removed the supernatant and washed the cells with PBS. Next, the treated cells were stained with 500 µL of 0.05 mg/mL of RNase A and 0.05 mg/mL of propidium iodide reagent in PBS [35]. The perturbation in cell cycle phases between the treated and untreated HT29 cells (10,000 events/group) was determined by the flow cytometry (CytoFLEX S, Beckman Coulter, USA). The cells in phases S and G2M were considered as proliferative cells and in the sub-G1 phase as apoptotic cells [36].

2.10 Statistical Analysis

The experiments of cell viability and cytokinetic growth were performed in triplicates, and the findings were displayed by the overall mean of three independent experiments ± standard deviation (SD). The statistical analysis of the data was carried out using GraphPad Prism software (Version 7.00). One-way (ANOVA) analysis of variance and Tukey's test were applied using Minitab software (version 17; PA, USA) to show the significant differences. The statistical analysis was considered acceptable when $p \leq 0.05$.

3. RESULTS

3.1 GC-TOFMS Analysis

The phytochemical content of the IL-GF extract was identified by using the GC-TOFMS instrument. The secondary metabolites detected in the extracts were validated by measuring the similarity of the pattern obtained from the GC-TOFMS reader with the existing database pattern. Graviola fruit extract (GFE) constituents generally belong to the class of alkaloids, flavonoids, phenols, and acetogenins. According to many kinds of literature, these compounds have antidiabetic, anti-oxidant and anti-cancer activities [37]. The result of the GC-TOFMS analysis of IL-GF extract was shown in Table 1, where 55 compounds were identified as having greater than 70% similarity with the standard mass spectroscopy in the NIST library. The major components identified by GC-TOFMS were D-psicofuranose, pentakis ether (13.53%), propyldecyl cyclopropane carboxylate (7.86%), tri-ruthenium dodecacarbonyl (7.86%), N-acetylimino dimethylsulfurane (7.86%), mannopyranose (7.74%), pyranone (7.74%), carbohydrazide (5.97%), nickel tetracarbonyl (5.57%), benzoic acid (5.46%), formic acid, ethenyl ester (4.45%) and propanetriol, 1-acetate (3.68%). Table 1 presents the complete list of bioactive components from the IL-GF extract.

3.2 The IC_{50} of IL-GF Extract Against HT29 Cell Lines

After plotting the graph of viable cells (%) vs different doses of the sample (µg/mL), The estimation of the IC_{50} values for IL-GF extract and Taxol were 10.56 µg/mL (Fig. 1) and 1.22 µg/mL (Fig. 2), respectively. These findings confirmed that the IL-GF extract has significantly inhibited the HT29 cell proliferation in a dose-dependent manner.

https://doi.org/10.31436/iiumej.v22i2.1687

Table 1: Compounds present in the ionic liquid extract of Graviola fruit identified by GC-TOFMS

Peak No	Area %	Compound Name	R. time (s)	Molecular Weight g/mol
1	0.17	1-Chloropropylene	180.232	76.01
3	0.15	Butanedione, monooxime	182.125	101.05
4	0.55	Nitrous acid, methyl ester	182.457	61.02
2	0.44	Alpha-Carotene	194.044	536.44
5	0.53	Molybdenum carbonyl	205.697	662.12
6	0.80	Propanediol, trimethylsilyl ether	206.062	148.09
7	1.70	Pentanedione, 3-diazo	206.626	126.04
8	1.18	Cyclopentane, 1-acetyl-1,2-epoxy	206.925	126.07
9	0.18	Acetophenone, 4'-nitro-, thiosemicarbazone	208.585	238.05
10	0.69	Propenoic acid, oxiranylmethyl ester	213.964	128.05
11	0.20	Mercaptamine	220.039	77.03
12	0.13	Nitroso-3-pyrroline	232.921	98.05
13	0.16	Oxalic acid	256.061	90.00
14	0.92	Acetic acid, (aminocarbonyl)	264.295	132.02
15	0.17	Acronize	270.105	478.11
16	5.97	Carbohydrazide	275.251	90.05
17	6.05	Pyranone	275.45	144.04
18	4.45	Formic acid, ethenyl ester	276.014	72.02
19	1.69	Trifluoromethanesulfonic acid ethyl ester	276.313	177.99
20	0.44	Ala-gly, trimethylsilyl ester	289.593	218.11
21	1.04	Furanone, dihydro-4-hydroxy	290.755	102.03
22	0.18	β-Phorbol	296.565	686.48
23	0.75	Cyclobutane	344.44	56.06
24	0.15	Butanoic acid, methyl ester	344.672	185.03
25	0.58	Aminourea	345.9	75.04
26	0.23	Cobalt, tetracarbonylsilyl	357.321	201.91
27	0.40	Valeric acid hydrazide	363.994	116.10
28	0.19	Oxotetrahydrofuran-2-carboxylic acid	365.256	130.03
29	0.30	à-D-Glucopyranoside	390.853	331.16
30	3.68	Propanetriol, 1-acetate	397.46	134.06
31	0.21	n-Propyl fluoride	443.741	62.05
32	5.57	Nickel tetracarbonyl	451.41	169.92
33	0.14	L-Proline	456.755	187.31
34	7.86	Propyldecyl cyclopropane carboxylate	493.375	268.24
35	7.86	Tri-ruthenium dodecacarbonyl	494.172	641.65
36	7.86	N-Acetylimino dimethylsulfurane	494.271	119.04
37	0.18	Carbamic acid, ethylnitroso, ethyl ester	537.564	146.07
38	0.28	Nitroethane	573.486	75.03
39	0.20	Iron, hexacarbonyl	704.128	762.16
40	0.21	Apigenin 8-C-glucoside	706.386	432.11
41	0.96	Manganese, acetylpentacarbony	769.632	237.93
42	0.23	Cyclobutanone	939.317	70.04
43	0.91	Butanedioic acid	945.625	350.63
44	0.13	Diethylene glycol, monotrimethylsilyl ether	996.853	178.10
45	0.20	Trimethylene oxide	1012.82	58.04
46	0.53	Silanol trimethyl, propanoate	1098.31	146.07
47	0.12	Nitrohexane	1105.82	131.09
48	0.22	β-Xylopyranose, tetrakis-TMS	1334.63	438.21
49	13.54	D-Psicofuranose, pentakis ether	1396.25	541.06
50	7.74	Mannopyranose	1397.34	452.88
51	1.10	Erythro-pentonic acid	1435.26	438.85
52	0.22	Gluconic acid, ç-lactone	1435.46	421.71
53	0.11	Quinazolin-2(3H)-thione	1974	392.16
54	0.19	Difluoroacetylene	2278.77	62.00
55	0.74	Aucubin, hexakis	2353.17	779.42

https://doi.org/10.31436/iiumej.v22i2.1687

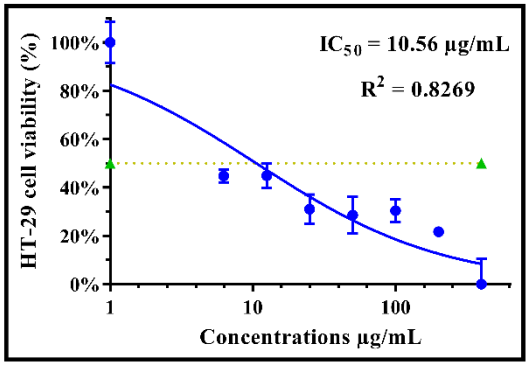


Fig. 1: The percentage of cell viability of colon adenocarcinoma HT29 cells vs different concentrations of IL-GF extract.

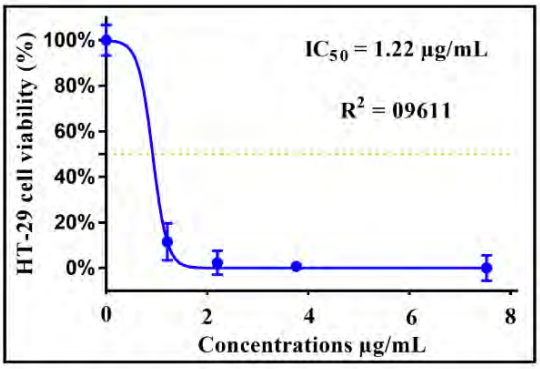


Fig. 2: The percentage of cell viability of colon adenocarcinoma HT29 cells vs different concentrations of Taxol.

3.3 Effect of IL-GF Extract on HT29 Cell Growth Kinetics

The growth kinetics graph of untreated HT29 cells (control), the IL-GF extract-treated HT29 (10.56 µg/mL) and Taxol-treated HT29 (1.22 µg/mL) were plotted and compared. The changes in the growth behaviour of the treated HT29 can be seen in (Fig. 3). First, in the lag phase, the untreated and treated HT29 cells' growth was similar until 16 hrs of treatment. After that, a clear log phase was seen when the cells started to proliferate at 24 hrs. The IL-GF extract-treated HT29 cells took 72 hrs to reach the peak, while the Taxol treated HT29 cells showed less exponential phase and a smaller number of cells than control cells.

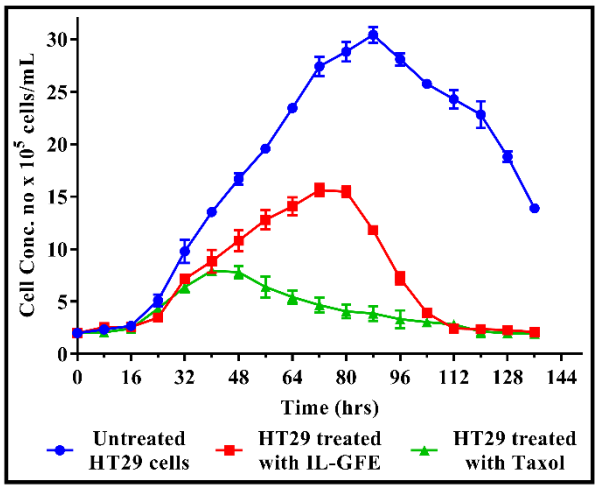


Fig. 3: Growth profiles of untreated HT29 cells and the treated with IL-GF extract (IC_{50}=10.56 µg/mL) and Taxol (IC_{50}=1.22 µg/mL).

From the growth curve, it can be seen that the untreated HT29 showed significant proliferation compared to the treated cells with IL-GF extract and Taxol. The cell number at the peak of the stationary phase for the treated HT29 cells with IL-GF extract and Taxol were only $15.6 \times 10^5 \pm 0.592$ and $7.93 \times 10^5 \pm 0.411$ cells/mL, respectively while the number reached $30.4 \times 10^5 \pm 0.753$ for the control. Based on the growth curve (Fig. 3) and the fourth equation, cell growth at the log phase (Fig. 4A), the death phase (Fig. 4B), and the number of cell generations (Table 2) were obtained for all treatments.

Table 2: The number of cell generations (X), specific growth rate (μ) and duplication period (td) for HT29 control cells and the treated cells with IL-GF extract and Taxol

Cell growth	Maximum cells volume (N) (cells/mL)	Number of generations (X)	Specific growth rate μ (h^{-1})	Doubling time t_d (h)
Untreated HT29 cells	30.4 x 10^5 ± 0.753[a]	3.93	0.0134	22.46
IL-GFE treated HT29 cells (IC_{50}= 10.56 μg/mL)	15.6 x 10^5 ± 0.592[b]	2.96	0.0124	24.27
Taxol treated HT29 cells (IC_{50}= 1.22 μg/mL)	7.93 x 10^5 ± 0.411[c]	2.01	0.012	25.08

*For all experiments, the initial number of cells (N0) was fixed at 2.0 x 10^5 cells/mL. Different superscripts (a, b, c) display a significant difference ($p \leq 0.05$).

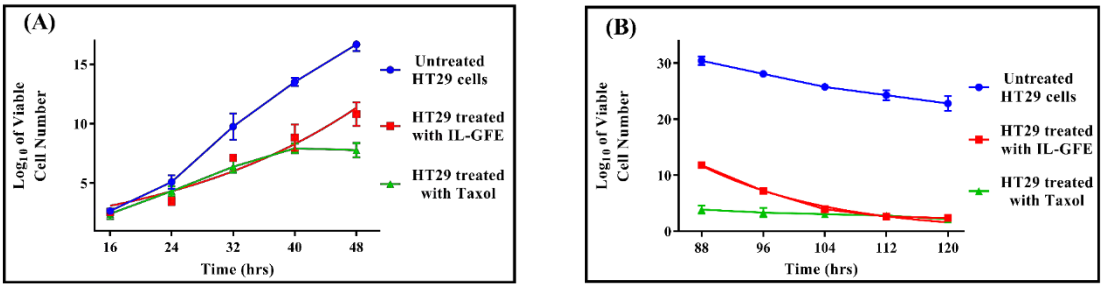


Fig. 4: HT29 Cells Growth kinetic of the control and the treated cells with IL-GF extract and Taxol at A: exponential phase and B: death phase.

The number of cell generation (X) was decreased from 3.93 generations in control to 2.96 generations in the IL-GF extract-treated cells and 2.01 generations in the Taxol-treated cells. This finding showed a significant difference when compared the number of cell generations between the control and different treatments.

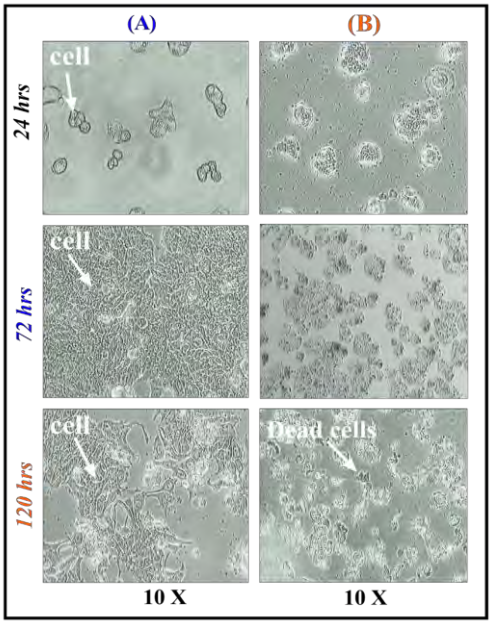


Fig. 5: The HT29 cells photographs of A: control cells and B: treated cells with IL-GF extract at a specified time point.

The typical photographs of the control cells and the treated HT29 with IL-GF extract at 24, 72 and 120 hrs were shown in Fig. 5. The density of cells was reduced in the IL-GFE treated HT29 cells. That is probably because of the anti-proliferation effect of the IL-GF extract against HT29 cell growth.

The volume of cancer can be estimated in some types of tumours using the specific growth rate [38]. Thus, the IL-GF extract effect on the specific growth of HT29 cells was assessed by plotting the log number of cell viability versus treatment time, and the specific growth rate (μ) was taken from the regression slope. The growth rate of the treated HT29 cells with IL-GF extract and Taxol has significantly reduced by 0.0124 h^{-1} and 0.012 h^{-1}, respectively compared to the control 0.0134 h^{-1}.

3.4 Cell Cycle Phase Distribution

The distribution of cell cycle phases in the treated HT29 with IL-GF extract was analysed by flow cytometry combined with RNase A / Propidium iodide. The flow cytometry result showed that in the treated HT29 cells with IL-GF extract for 24, 48 and 72 hrs, the percentage of cells that were arrested at the G0/G1 phase has increased to 76.08% compared to the control 69.56% at 48 hrs then reduced. In contrast, the percentage of cells has declined in phases S and G2M compared to the control cells (Fig. 6).

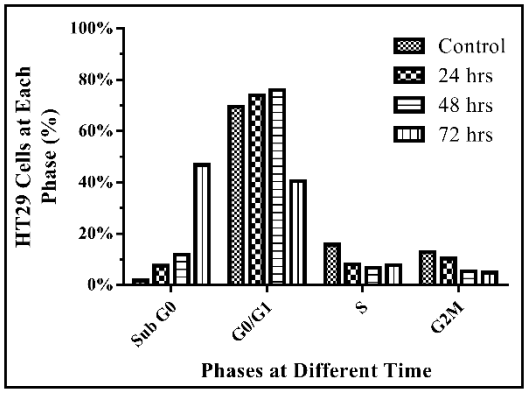


Fig. 6. Distribution of cell cycle phases of the treated HT29 cells with the corresponding IC_{50} of IL-GF extract in a time-dependent manner.

Furthermore, the treated HT29 cells showed an increase of cells arrested at the sub G0 phase in a time-increased treatment as compared with untreated HT29, considering the cells at the phase of sub-G0 as apoptotic cells [36]; there was an apparent increase of the apoptotic cells particularly after 48 hrs treatment.

4. DISCUSSION

Fruit and vegetables are related to a healthy life and reduce the risk of diseases. That is undoubtedly due to the different phytochemicals and secondary metabolites, including alkaloids, phenols, carotenoids, vitamins and flavonoids that are important in strengthening immunity [39]. Many anti-cancer agents are used in cancer treatment, such as Taxol, topotecan, cisplatin, etoposide and doxorubicin [35]. However, cancer patients suffer from the severe side effect of these drugs related to their capacity to damage healthy cells. Hence, new anti-cancer drugs with fewer health risks are highly recommended to be investigated. The National Cancer Institute (USA) reported that any plant extracts showing an IC_{50} value under 20 μg/mL after a duration of incubation of 48 to 72 hrs are considered as a potent cytotoxic substance [40].

Different identified phytochemicals in this study have been found to possess various biological activities. It was reported a long time ago that the Annonaceae family contained a considerable number of phytochemicals that exhibit anti-cancer activities such as acetogenins, alkaloids, terpenoids and flavonoids [41]. The claimed anti-cancer activity of several identified compounds have been reported, such as α-D-mannopyranose in cloves of *Allium sativum* were showed cytotoxicity against breast cancer cells [42], tri-ruthenium dodecacarbonyl also was reported to have a highly selective anti-angiogenic against ovarian carcinoma A2780 cell lines [43,44]. Silanol trimethyl, D-psicose pentakis, and butanedioic acid have been reported to contribute to drug delivery system, anti-proliferation, apoptosis and cell cycle arrest in different cancer cells lines [45-47]. Another study reported that dihydro-dihydroxy-methyl-pyranone (DDMP) from onions had inhibited the growth of colon adenocarcinoma (HCT116 and SW620) cells by inducing apoptotic cell death, modulation of tumour necrosis factor-α (TNF-α) and DNA binding activity [48].

Moreover, Kameue et al. [49] reported that gluconic acid could induce apoptosis by stimulating butyrate production in the large intestine in a p53-independent manner and inhibiting cancer cell proliferation. These findings confirm the value of traditional use of the plant for medicinal purposes as well as pharmaceutical development. The use of Graviola fruit in herbal medicine has been supported by the existence of these bioactive compounds with proven health benefits and hence an interest for further research on this plant.

The anti-proliferative effect of the IL-GF extract was evaluated on adherent colon adenocarcinoma HT29 cell line. The IL-GF extract showed an anti-proliferative effect toward cancer cells as presented by MTT assay. Many previous studies have reported the anti-cancer activity of Graviola fruit toward many cancer cells, including human lung carcinoma A-549, breast adenocarcinoma MCF-7 cells, hepatoma HepG2 cell, prostate cancer PC-3 cells, and pancreatic tumour PACA-2 cell lines with comparable outcomes to our research findings [14,15,50-52]. Furthermore, the typical photographs of HT29 cells after treatment with IL-GFE and the control at designated time points displayed cell density reduction after treatment. The increase of apoptotic cells can be indicated by the changes in the morphology and biochemical of the HT29 cells, including DNA fragmentation and chromatin condensation, cell shrinkage and membrane blebbing [53]. Moreover, a recent study conducted by Daddiouaissa et al. [54] reported that after treating colon cancer HT29 cell line with IL-GF extract, the pathway analysis of metabolomic profiles showed an alteration of different metabolic pathways such as aerobic glycolysis, amino acid metabolism, ketone bodies metabolism and urea cycle that contribute to cancer cells proliferation and energy metabolism.

The growth kinetics model was investigated to estimate the inhibitory and cytotoxicity effects of IL-GF extract on HT29 cell growth. A general model was developed for colon adenocarcinoma HT29 cells in response to the treatment of IL-GF extract. Generally, efficient therapies decrease the cells' proliferation rate (cytostatic effect) and increase the death rate of the cells (cytotoxic effect). Many chemotherapy drugs disrupt the mechanism of cell division and dysregulate the cell cycle either through interfering or destroying DNA synthesis [55]. This can be confirmed when the doubling time has expanded from 22.46 hrs in control to 24.27 hrs and 25.08 hrs for IL-GF extract and Taxol-treated HT29, respectively. The cytostatic effects can be observed at the log phase (Fig. 4A), whereas cytotoxic effects were clearly demonstrated at the death phase (Fig. 4B).

Although a biopsy sample is usually used to define the growth and spread of cancer growth and prescribe an effective therapy for it, the kinetics model in terms of simple

formulas may help predict the tumour size in a function of time. The first-order linear equation for the untreated colon adenocarcinoma HT29 cell growth was given as follows:

$$log_{10}N = log_{10}N_0 + 0.0134t \quad (6)$$

After treating HT29 cell lines with IL-GF extract, the equation of the first order linear presented as follows:

$$log_{10}N = log_{10}N_0 + 0.0124t \quad (7)$$

Moreover, after treating HT29 cell lines with Taxol, the formula can be stated as follows:

$$log_{10}N = log_{10}N_0 + 0.012t \quad (8)$$

However, if we neglect the tumour growth characteristics, the treatment's effectiveness can be estimated [38]. Therefore, the final equation of the kinetic model in response to IL-GF extract treatment was given as follows:

$$log_{10}\mathrm{Nf} = (log_{10}N_0 + 0.0134t) - (log_{10}N_0 + 0.0124t) \quad (9)$$

Furthermore, the final equation of the kinetic model in response to Taxol treatment was given as follows:

$$log_{10}\mathrm{Nf} = (log_{10}N_0 + 0.0134t) - (log_{10}N_0 + 0.012t) \quad (10)$$

where, (log_{10} Nf) means the ultimate size of cancer after treatment, (log_{10} N_0) means the primary size of a tumour and (t) means the treatment duration ($t_{final} - t_{initial}$).

With this model, it is possible to identify the presence and the size of the early-stage tumours (less than 1 cc). The interpolation and extrapolation of tumour volume can be predicted at any time points. Cancer is classified as a cell cycle dysregulation disease. Thus, an effective drug can block cancer cell division [56]. After treatment, the cell cycle distribution of the HT29 cells was investigated using the flow cytometry technique to see whether IL-GF extract induced cell growth inhibition by cell cycle arrest. This study showed that IL-GF extract arrested the cell cycle of the treated HT29 at the G1 phase of the growth-static, explaining the anti-proliferative effect of IL-GF extract. This confirmed the result of Moghadamtousi and co-workers [57] when treated colon cancer HCT-116 with ethyl acetate extract of *A. muricata* leaves.

Moreover, in our previous study, the flow cytometry result showed that IL-GF extract had induced apoptosis in breast adenocarcinoma MCF-7 cell lines [15]. This was also validated by Moghadamtousi et al. [57], and Chamcheu et al. [58] after treating non-melanoma skin cancer NMSC and colorectal cancer HCT-116 and HT-29 with *A. muricata* leaves extract. On the other hand, Torres et al. [59] reported the induction of necrosis in pancreatic cancer PC cells when treated with the Graviola extract. This study concluded that Graviola extracts inhibit cell metabolism through numerous signalling pathways, metastatic properties and arrests the cell cycle machinery in pancreatic cancer cells.

Many investigated studies of the phytochemicals contained in the leaves and fruits of Graviola have extracted and identified a large number of secondary metabolites with a potential therapeutic such as anti-proliferative, apoptotic and cytotoxic towards many human cancer cells [57,60]. The synergistic effect of different metabolites in the Graviola fruit extract such as acetogenins, flavonoids, terpenes, and alkaloids may be responsible for its anti-proliferative effect. Thus, more investigations are still needed to detect the responsible active compounds for the anti-proliferation and inhibition effects of the Graviola fruit.

5. CONCLUSION

The GC-TOFMS analysis of Graviola fruit extracted using IL-MAE confirmed the existence of various chemical groups of compounds like acetogenins, terpenoids, alkaloids, fatty acids, sugars and lactones. IL-GF extract presents anti-proliferative activity toward colon adenocarcinoma HT29 cell lines by inducing cell death through morphological changes, apoptosis, and the arrest of the cell cycle at the G0/G1 phase. This cytotoxicity showed selectivity toward the proliferation of colon cancer cells, indicating selective antitumor properties in IL-GF extract against tumour cells. Moreover, IL-GF extract showed a great potentiality anticancerous when compared to the positive control, Taxol. IL-GF extract has inhibited the growth of HT29 cells by decreasing the number of cell generations and increasing duplication duration compared with untreated cells. These findings suggest that IL-GF extract can be established as a new supplementary agent for preventing and curing colon cancer. Additional research to explain the mechanism attached to the therapeutic effects and define the responsible bioactive compounds for the fruit extract's cytotoxicity.

ACKNOWLEDGEMENTS

The authors are grateful to the International Islamic University Malaysia for offering a research grant (P-RIGS18-065-0075) and laboratory facilities.

REFERENCES

[1] Patel MS, Patel JK. (2016) A review on a miracle fruits of *Annona muricata*. Journal of Pharmacognosy and Phytochemistry, 5(1): 137-148.

[2] Ioannis P, Anastasis S, Andreas Y. (2015) Graviola: A systematic review on its anti-cancer properties. Am. J. Cancer Prev, 3(6): 128-131. https://doi.org/10.12691/ajcp-3-6-5

[3] Yajid AI, Ab Rahman HS, Wong MPK, Zain WZW. (2018) Potential benefits of *Annona muricata* in combating cancer: A Review. The Malaysian Journal of Medical Sciences, 25(1): 5-15. https://dx.doi.org/10.21315%2Fmjms2018.25.1.2

[4] Badrie N, Schauss AG. (2010) Soursop (*Annona muricata* L.): composition, nutritional value, medicinal uses, and toxicology. Bioactive Foods in Promoting Health, Elsevier, pp. 621-643. https://doi.org/10.1016/B978-0-12-374628-3.00039-6

[5] Coe FG. (2008) Rama midwifery in eastern Nicaragua. Journal of Ethnopharmacology, 117(1): 136-157. https://doi.org/10.1016/j.jep.2008.01.027

[6] Samuel AJSJ, Kalusalingam A, Chellappan DK, Gopinath R, Radhamani S, Husain HA, Muruganandham V, Promwichit P. (2010) Ethnomedical survey of plants used by the Orang Asli in Kampung Bawong, Perak, West Malaysia. Journal of Ethnobiology and Ethnomedicine, 6(1): 1-6. https://doi.org/10.1186/1746-4269-6-5

[7] Tisott G, Molin D, Colet C. (2013) The use of medicinal plants and herbal medicines for patients in chemotherapy in an oncology center of Ijuı/RS. O Mundo da Saude, 39(3): 287-298.

[8] George VC, Kumar D, Rajkumar V, Suresh P, Kumar RA. (2012) Quantitative assessment of the relative antineoplastic potential of the n-butanolic leaf extract of *Annona muricata* Linn. in normal and immortalised human cell lines. Asian Pacific Journal of Cancer Prevention, 13(2): 699-704. https://doi.org/10.7314/APJCP.2012.13.2.699

[9] Monigatti M, Bussmann RW, Weckerle CS. (2013) Medicinal plant use in two Andean communities located at different altitudes in the Bolívar Province, Peru. Journal of Ethnopharmacology, 145(2): 450-464. https://doi.org/10.1016/j.jep.2012.10.066

[10] Betancur-Galvis LA, Saez J, Granados H, Salazar A, Ossa J. (1999) Antitumor and antiviral activity of Colombian medicinal plant extracts. Memórias do Instituto Oswaldo Cruz, 94(4): 531-535. http://dx.doi.org/10.1590/S0074-02761999000400019

[11] Dai Y, Hogan S, Schmelz EM, Ju YH, Canning C, Zhou K. (2011) Selective growth inhibition of human breast cancer cells by graviola fruit extract *in vitro* and *in vivo* involving downregulation of EGFR expression. Nutrition and Cancer, 63(5): 795-801. https://doi.org/10.1080/01635581.2011.563027

[12] Gavamukulya Y, Abou-Elella F, Wamunyokoli F, AEl-Shemy H. (2014) Phytochemical screening, anti-oxidant activity and *in vitro* anti-cancer potential of ethanolic and water leaves extracts of *Annona muricata* (Graviola). Asian Pacific Journal of Tropical Medicine, 7(13): S355-S363. https://doi.org/10.1016/S1995-7645(14)60258-3

[13] Daddiouaissa D, Amid A. (2018) Anti-cancer activity of acetogenins from *Annona muricata* fruit. International Medical Journal Malaysia, 17(3): 103-112. https://doi.org/10.31436/imjm.v17i3.236

[14] Sun S, Liu J, Sun X, Zhu W, Yang F, Felczak L, Dou QP, Zhou K. (2017) Novel Annonaceous acetogenins from Graviola (*Annona muricata*) fruits with strong anti-proliferative activity. Tetrahedron Letters, 58(19): 1895-1899. https://doi.org/10.1016/j.tetlet.2017.04.016

[15] Daddiouaissa D, Amid A, Kabbashi NA, Fuad FA, Elnour AM, Epandy MA. (2019) Anti-proliferative activity of ionic liquid-Graviola fruit extract against human breast cancer (MCF-7) cell lines using flow cytometry techniques. Journal of Ethnopharmacology, 236: 466-473. https://doi.org/10.1016/j.jep.2019.03.003

[16] Siegel RL, Miller KD, Jemal A. (2019) Cancer statistics, 2019. CA: A cancer journal for clinicians 69(1), 7-34. https://doi.org/10.3322/caac.21551

[17] Griffin A, Butow P, Coates A, Childs A, Ellis P, Dunn S, Tattersall M. (1996) On the receiving end V: patient perceptions of the side effects of cancer chemotherapy in 1993. Annals of Oncology, 7(2): 189-195. https://doi.org/10.1093/oxfordjournals.annonc.a010548

[18] Daddiouaissa D, Amid A, Syahida A, Elnour AM. (2020) Phytochemical analysis of ionic liquid-Graviola (*Annona muricata*) fruit extract and its acute toxicity on zebrafish early-life stages. Asia Pacific Journal of Molecular Biology and Biotechnology, 28(2): 113-124. https://doi.org/10.35118/apjmbb.2020.028.2.10

[19] Koel M, Kuhtinskaja M, Vaher M. (2020) Extraction of bioactive compounds from *Catharanthus roseus* and *Vinca minor*. Separation and Purification Technology, 252: 117438. https://doi.org/10.1016/j.seppur.2020.117438

[20] Ventura SPM, e Silva FA, Quental MV, Mondal D, Freire MG, Coutinho JAP. (2017) Ionic-liquid-mediated extraction and separation processes for bioactive compounds: Past, present, and future trends. Chemical Reviews, 117(10): 6984-7052. https://doi.org/10.1021/acs.chemrev.6b00550

[21] Bogdanov MG. (2014) Ionic liquids as alternative solvents for extraction of natural products, alternative solvents for natural products extraction, Springer, pp 127-166.

[22] Cláudio AFM, Marques CF, Boal-Palheiros I, Freire MG, Coutinho JA. (2014) Development of back-extraction and recyclability routes for ionic-liquid-based aqueous two-phase systems. Green Chemistry, 16(1): 259-268. https://doi.org/10.1039/C3GC41999A

[23] Wang T, Wang Q, Li P, Yang H. (2019) Temperature-responsive ionic liquids to set up a method for the simultaneous extraction and in situ preconcentration of hydrophilic and lipophilic compounds from medicinal plant matrices. Green Chemistry, 21(15): 4133-4142. https://doi.org/10.1039/C9GC00995G

[24] Fan Y, Li W, Zhang S, Sun S, Yang L. (2020) Vitamin B3-based protic ionic liquids as green solvents for the isolation of astilbin from rhizoma smilacis glabrae. Industrial Crops and Products, 152: 112563. https://doi.org/10.1016/j.indcrop.2020.112563

[25] Hoenig M. (2001) Preparation steps in environmental trace element analysis—facts and traps. Talanta, 54(6): 1021-1038. https://doi.org/10.1016/S0039-9140(01)00329-0

[26] Gyamfi K, Sarfo D, Nyarko B, Akaho E, Serfor-Armah Y, Ampomah-Amoako E. (2011) Assessment of elemental content in the fruit of graviola plant, *Annona muricata*, from some selected communities in Ghana by instrumental neutron activation analysis. Elixir Food Sci, 41: 5671-5675.

[27] Rahman NRHA, Idris A, Yunus NA, Mustaffa AA. (2019) Optimisation of ionic liquid-based microwave extraction of flavonoid and phenolic acid from *Labisia pumila*. AIP Conference Proceedings (Vol. 2124, No. 1, p. 020027), AIP Publishing LLC. https://doi.org/10.1063/1.5117087

[28] Bhan M, Satija S, Garg C, Dureja H, Garg M. (2017) A novel approach towards green extraction for glycyrrhitinic acid by ionic liquid based microwave assisted extraction and optimisation through response surface methodology. Pharmacognosy Journal, 9(6): 866-872. https://doi.org/10.5530/pj.2017.6.136

[29] Zhang Y, Liu Z, Li Y, Chi R. (2014) Optimisation of ionic liquid-based microwave-assisted extraction of isoflavones from *Radix puerariae* by response surface methodology. Separation and Purification Technology, 129: 71-79. https://doi.org/10.1016/j.seppur.2014.03.022

[30] Jang Z-H, Chung H-C, Ahn YG, Kwon Y-K, Kim J-S, Ryu J-H, Kim C-H, Hwang G-S. (2012) Metabolic profiling of an alcoholic fatty liver in zebrafish (*Danio rerio*). Molecular BioSystems, 8(7): 2001-2009. https://doi.org/10.1039/C2MB25073J

[31] Muhamad S, Jamilah B, Russly A, Faridah A. (2017) *In vitro* antibacterial activities and composition of *Carica papaya* cv. Sekaki/Hong Kong peel extracts. International Food Research Journal, 24(3): p976-984.

[32] Liu N, Yang HL, Wang P, Lu YC, Yang YJ, Wang L, Lee SC. (2016) Functional proteomic analysis revels that the ethanol extract of *Annona muricata* L. induces liver cancer cell apoptosis through endoplasmic reticulum stress pathway. Journal of Ethnopharmacology, 189: 210-217. https://doi.org/10.1016/j.jep.2016.05.045

[33] Chik WW, Amid A, Jamal P. (2010) Purification and cytotoxicity assay of tomato (*Lycopersicon esculentum*) leaves methanol extract as potential anti-cancer agent. Journal of Applied Sciences, 10(24): 3283-3288. https://dx.doi.org/10.3923/jas.2010.3283.3288

[34] Fouz N, Amid A, Hashim Y. (2013) Cytokinetic study of MCF-7 cells treated with commercial and recombinant bromelain. Asian Pacific Journal of Cancer Prevention, 14(11): 6709-6714. https://doi.org/10.7314/APJCP.2013.14.11.6709

[35] Magadi VP, Ravi V, Arpitha A. (2015) Evaluation of cytotoxicity of aqueous extract of Graviola leaves on squamous cell carcinoma cell-25 cell lines by 3-(4, 5-dimethylthiazol-2-Yl)-2, 5-diphenyltetrazolium bromide assay and determination of percentage of cell inhibition at G2M phase of cell cycle by flow cytometry: An *in vitro* study. Contemporary Clinical Dentistry, 6(4): 529–533. https://dx.doi.org/10.4103%2F0976-237X.169863

[36] Agu KC, Okolie NP, Falodun A, Engel-Lutz N. (2018) *In vitro* anti-cancer assessments of *Annona muricata* fractions and *in vitro* anti-oxidant profile of fractions and isolated acetogenin (15-acetyl guanacone). Journal of Cancer Research and Practice, 5(2): 53-66. https://doi.org/10.1016/j.jcrpr.2017.12.001

[37] Suhendar U. (2018) Geographical Effect on the Cytotoxic Activity of *Annona muricata* L. Leaves Extract Against MCF7 Cancer Cell. Fitofarmaka, 8(2): 12-19.

[38] Mehrara E, Forssell-Aronsson E, Ahlman H, Bernhardt P. (2007) Specific growth rate versus doubling time for quantitative characterisation of tumor growth rate. Cancer Research, 67(8): 3970-3975. https://doi.org/10.1158/0008-5472.CAN-06-3822

[39] Wang H, Oo Khor T, Shu L, Su Z-Y, Fuentes F, Lee J-H, Tony Kong A-N. (2012) Plants vs. cancer: a review on natural phytochemicals in preventing and treating cancers and their druggability. Anti-Cancer Agents in Medicinal Chemistry (Formerly Current Medicinal Chemistry-Anti-Cancer Agents), 12(10): 1281-1305. https://doi.org/10.2174/187152012803833026

[40] Boik J. (2001) Natural compounds in cancer therapy (Vol. 851). Oregon Medical Press, 1st ed.

[41] Coria-Tellez AV, Montalvo-Gónzalez E, Yahia EM, Obledo-Vázquez EN. (2018) *Annona muricata*: A comprehensive review on its traditional medicinal uses, phytochemicals, pharmacological activities, mechanisms of action and toxicity. Arabian Journal of Chemistry, 11(5): 662-691. https://doi.org/10.1016/j.arabjc.2016.01.004

[42] Nazeer AA, Veeraiyan S, Vijaykumar SD. (2017) Anti-cancer potency and sustained release of phytosomal diallyl disulfide containing methanolic *Allium sativum* extract against breast cancer. International Research Journal of Pharmacy, 8(8): 34-40. https://doi.org/10.7897/2230-8407.088141

[43] Nazarov A, Nosova YN, Mikhalev O, Kovaleva O, Dyson P, Milaeva E. (2016) Anti-proliferative activity of ruthenium and osmium clusters with phosphine ligands. Russian Chemical Bulletin, 65(2): 546-549. https://doi.org/10.1007/s11172-016-1335-x

[44] Nowak-Sliwinska P, Zava O, Van Beijnum JR, Groessl M, Chisholm DM, Ahmadi Z, McIndoe JS, Griffioen AW. (2013) Synthesis and characterisation of a new class of anti-angiogenic agents based on ruthenium clusters. Scientific Reports, 3: 1485. https://doi.org/10.1038/srep01485

[45] Belayachi L, Aceves-Luquero C, Merghoub N, de Mattos SF, Amzazi S, Villalonga P, Bakri Y. (2017) Induction of cell cycle arrest and apoptosis by *Ormenis eriolepis* a Morrocan endemic plant in various human cancer cell lines. African Journal of Traditional, Complementary and Alternative Medicines, 14(2): 356-373. https://doi.org/10.21010/ajtcam.v14i2.37

[46] Cheng W, Liang C, Xu L, Liu G, Gao N, Tao W, Luo L, Zuo Y, Wang X, Zhang X. (2017) TPGS-functionalized polydopamine-modified mesoporous silica as drug nanocarriers for enhanced lung cancer chemotherapy against multidrug resistance. Small, 13(29): 1-12. https://doi.org/10.1002/smll.201700623

[47] Sui L, Dong Y, Watanabe Y, Yamaguchi F, Hatano N, Tsukamoto I, Izumori K, Tokuda M. (2005) The inhibitory effect and possible mechanisms of D-allose on cancer cell proliferation. International Journal of Oncology, 27(4): 907-912. https://doi.org/10.3892/ijo.27.4.907

[48] Ban JO, Hwang IG, Kim TM, Hwang BY, Lee US, Jeong H-S, Yoon YW, Kim DJ, Hong JT. (2007) Anti-proliferate and pro-apoptotic effects of 2, 3-dihydro-3, 5-dihydroxy-6-methyl-4H-pyranone through inactivation of NF-κB in human colon cancer cells. Archives of Pharmacal Research, 30(11): 1455. https://doi.org/10.1007/BF02977371

[49] Kameue C, Tsukahara T, Ushida K. (2006) Alteration of gene expression in the colon of colorectal cancer model rat by dietary sodium gluconate. Bioscience, Biotechnology, and Biochemistry, 70(3): 606-614. https://doi.org/10.1271/bbb.70.606

[50] Consolacion Y, Geneveve S, Ming-Jaw D, Chien-Chang S. (2012) Acetogenins from *Annona muricata*. Pharmacognosy Journal, 4(32): 32-37. https://doi.org/10.5530/pj.2012.32.7

[51] Sun S, Liu J, Kadouh H, Sun X, Zhou K. (2014) Three new anti-proliferative Annonaceous acetogenins with mono-tetrahydrofuran ring from graviola fruit (*Annona muricata*). Bioorganic & Medicinal Chemistry Letters, 24(12): 2773-2776. https://doi.org/10.1016/j.bmcl.2014.03.099

[52] Sun S, Liu J, Zhou N, Zhu W, Dou QP, Zhou K. (2016) Isolation of three new annonaceous acetogenins from Graviola fruit (*Annona muricata*) and their anti-proliferation on human prostate cancer cell PC-3. Bioorganic & Medicinal Chemistry Letters, 26(17): 4382-4385. https://doi.org/10.1016/j.bmcl.2015.06.038

[53] Pieme CA, Kumar SG, Dongmo MS, Moukette BM, Boyoum FF, Ngogang JY, Saxena AK. (2014) Anti-proliferative activity and induction of apoptosis by *Annona muricata* (Annonaceae) extract on human cancer cells. BMC Complementary and Alternative Medicine, 14(1): 516. https://doi.org/10.1186/1472-6882-14-516

[54] Daddiouaissa D, Amid A, Sani MSA, Elnour AA. (2021) Evaluation of metabolomics behavior of human colon cancer HT29 cell lines treated with ionic liquid graviola fruit pulp extract. Journal of Ethnopharmacology, 270: 113813. https://doi.org/10.1016/j.jep.2021.113813

[55] O'Reilly MS, Boehm T, Shing Y, Fukai N, Vasios G, Lane WS, Flynn E, Birkhead JR, Olsen BR, Folkman J. (1997) Endostatin: an endogenous inhibitor of angiogenesis and tumor growth. Cell, 88(2): 277-285. https://doi.org/10.1016/S0092-8674(00)81848-6

[56] Mantena SK, Sharma SD, Katiyar SK. (2006) Berberine, a natural product, induces G1-phase cell cycle arrest and caspase-3-dependent apoptosis in human prostate carcinoma cells. Molecular Cancer Therapeutics, 5(2): 296-308. https://doi.org/10.1158/1535-7163.MCT-05-0448

[57] Moghadamtousi SZ, Karimian H, Rouhollahi E, Paydar M, Fadaeinasab M, Kadir HA. (2014) *Annona muricata* leaves induce G1 cell cycle arrest and apoptosis through mitochondria-mediated pathway in human HCT-116 and HT-29 colon cancer cells. Journal of Ethnopharmacology, 156: 277-289. https://doi.org/10.1016/j.jep.2014.08.011

[58] Chamcheu J, Rady I, Chamcheu R-C, Siddique A, Bloch M, Banang Mbeumi S, Babatunde A, Uddin M, Noubissi F, Jurutka P. (2018) Graviola (*Annona muricata*) exerts anti-proliferative, anti-clonogenic and pro-apoptotic effects in human non-melanoma skin cancer UW-BCC1 and A431 cells *in vitro*: Involvement of hedgehog signaling. International Journal of Molecular Sciences, 19(6): 1791. https://doi.org/10.3390/ijms19061791

[59] Torres MP, Rachagani S, Purohit V, Pandey P, Joshi S, Moore ED, Johansson SL, Singh PK, Ganti AK, Batra SK. (2012) Graviola: A novel promising natural-derived drug that inhibits tumorigenicity and metastasis of pancreatic cancer cells *in vitro* and *in vivo* through altering cell metabolism. Cancer Letters, 323(1): 29-40. https://doi.org/10.1016/j.canlet.2012.03.031

[60] Gajalakshmi S, Vijayalakshmi S, Devi Rajeswari V. (2012) Phytochemical and pharmacological properties of *Annona muricata*: A review. International Journal of Pharmacy and Pharmaceutical Sciences, 4(2): 3-6.

https://doi.org/10.31436/iiumej.v22i2.1589

INFLUENCE OF DIFFERENT PERCENTAGE BOILER ASH-BASED GEOPOLYMER IN LATERITE SOIL

ATIQAH NAJWA ZAINUDDIN[1*], MAZIDAH MUKRI[2], DIANA CHE LAT[1], ROSLIZAYATI ROSLI[1] AND NOOR HIDAYU ABDUL RANI[3]

[1]Faculty of Civil Engineering,
Universiti Teknologi MARA (UiTM) Johor Branch,
Kampus Pasir Gudang, Johor, Malaysia
[2]Faculty of Civil Engineering,
Universiti Teknologi MARA (UiTM) Selangor, Malaysia
[3]Faculty of Chemical Engineering,
Universiti Teknologi MARA (UiTM) Johor Branch,
Kampus Pasir Gudang, Johor, Malaysia

[*]Corresponding author: atiqa387@uitm.edu.my

(Received: 23rd August 2020; Accepted: 9th January 2021; Published on-line: 4th July 2021)

ABSTRACT: The waste generation of palm oil boiler ash has been one of its big problems as it is less used and deposited in landfills as a by-product. Geopolymer is a new green technology that has been intensively studied in concrete applications. However, few studies on geopolymers have been conducted in soil applications. Thus, this study investigated the influence of palm oil boiler ash-based geopolymer in laterite soil strength. Different percentages, 5, 10, 15, and 20% of geopolymer mixtures, were added to laterite soil. The process of producing a geopolymer binder was performed by sieving boiler ash (150 µm), then mixing with sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3) as an alkaline activator at a ratio of 1:2. This material effectiveness was tested through compaction test using a standard proctor, unconfined compressive strength, and the scanning electron microscope (SEM). Fifteen percent (15%) of geopolymer in laterite soil indicated the best-mixed design with a maximum dry density of 2.23 Mg/m^3 with a moisture content of 13.58%. The unconfined compressive strength test at curing times of zero, seven, and twenty-eight days on the LS-15.0 GPOBA sample, show a slightly increased strength of 47, 58, and 76 kPa, respectively. The SEM images proved that the geopolymer gel's development stabilized the soil structure from a loose structure to a denser soil structure. This study aims to investigate the influence of geopolymer in laterite soil. Boiler ash as an alternative material in geotechnical applications was studied to understand and develop new green alternative materials to sustain the environment from industrial waste and to enhance laterite soil properties.

ABSTRAK: Abu kelapa sawit adalah salah satu sisa utama yang terhasil dari industri kelapa sawit tetapi penggunaannya kurang dimanfaatkan dan dibuang ke tempat pembuangan sampah. Teknologi hijau baru yang dikenali sebagai geopolimer telah dikaji secara intensif dengan kekuatan konkrit tetapi hanya sedikit kajian telah dibuat dalam penggunaan tanah. Tujuan kajian ini adalah bagi mengesan geopolimer berasaskan abu kelapa sawit terhadap kekuatan tanah laterit. Peratusan campuran geopolimer yang berbeza (0, 5, 10, 15 dan 20%) dicampur pada tanah laterit. Bagi menghasilkan geopolimer, saiz 150 µm abu kelapa sawit disintesis dengan kombinasi bahan kimia natrium hidroksida (NaOH) dan natrium silikat (Na_2SiO_3) pada nisbah 1:2 bagi semua campuran sebagai pengaktif alkali. Ujian terhadap keberkesanan bahan adalah melalui proses ujian pemampatan menggunakan proktor standard, kekuatan pemampatan tidak terbatas, dan Pengimbas Mikroskop Elektron (SEM). Berdasarkan dapatan ujian

pemadatan, 15% geopolimer di tanah laterit menunjukkan campuran terbaik dengan memberikan kepadatan pengeringan maksimum 2.23 Mg/m^3 pada kelembapan 13.58%. Ujian kekuatan mampatan tidak terbatas pada masa pempolimeran sebanyak 0, 7 dan 28 hari diuji pada sampel LS-15.0GPOBA bagi menguji kekuatan campuran. Dapatan menunjukkan kekuatan geopolimer sedikit meningkat pada 47, 58 dan 76, masing-masing. Imej SEM membuktikan pengembangan gel geopolimer menstabilkan struktur tanah daripada struktur lopong kepada struktur tanah padat. Oleh itu, abu kelapa sawit berasaskan geopolimer dan tanah laterit berpotensi sebagai alternatif bagi merawat tanah dalam aplikasi geoteknik dan berpotensi mengurangkan kadar kebolehtelapan.

KEYWORDS: *palm oil boiler ash; geopolymer; laterite soil; compaction compressive strength*

1. INTRODUCTION

Increased population is leading to higher demand in the construction industry and the growth of infrastructure. Malaysia is the largest exporter of palm oil mill. The production of palm oil in Malaysia has had a positive impact on sustained economic growth in the global market over the past four decades, and the annual production of palm oil in Malaysia in the period of 2016–2020 has reached 15.4 million tons [1]. Despite the apparent benefits, oil palm mills' production significantly contributes to environmental degradation [2].

Production from palm oil processing for oil extraction generates waste such as palm kernel, nutshell, palm fiber, and empty fruit brunch. These wastes are incinerated and produce ashes at the boiler's lower compartment, known as boiler ash. It has now been revealed that the palm oil industry will produce 657,000 tons of waste at the end of December 2020 [3]. The wastes are either dumped to landfill or eventually exposed and contaminate the soil, creating a significant disposal issue and harming the environment [4].

Nowadays, Ordinary Portland Cement (OPC) and lime have become the most common soil stabilizers used as additives to enhance soil properties. However, the use of Ordinary Portland Cement (OPC) is held responsible for some CO_2 emission during its production and has thus polluted the environment. Several efforts are in progress to reduce Ordinary Portland cement as an additive material in soil stabilization to address global warming issues. Studies are being conducted to find alternative and eco-friendly soil stabilizers from waste material to reduce CO_2 emission and decrease construction costs. Therefore, in this study, boiler ash was introduced as an alternative by-product material in soil stabilizer by synthesizing it into geopolymer form.

Daviddovits [5] introduced a new green technology known as geopolymer to enhance mechanical properties and material strength in the construction industry. All the materials that contain Si and Al can be used as a constituent in geopolymer to be activated with alkali solution presence [6]. According to Van Jaarsveld et al. [7], the SiO_2 and Al_2O_3 in the ash provide a rich source of Si and Al atoms as a source of aluminosilicate materials for producing geopolymers. Xu et al. [8] investigated 16 natural Al-Si minerals and concluded that all the minerals are, to some extent, soluble in concentrated alkaline solution.

According to Daviddovits, geopolymerization is a reaction that consists of Si-O-Al and Si-O-Si bonds depending on the ability of alumina ion (6-fold or 4-fold coordination) to induce the crystallographer and chemical changes in silica backbone. In geopolymerization, a reaction three-dimensional techto-aluminosilicate framework called polysialate is shown in Equation (1), where M is a cation (K, Na, Ca), n is the degree of polycondensation, and z is 1, 2, 3 or ≥3. In the sialates network, Si-O-Al bridge forms form the chain and ring of

SiO_4 and AlO_4 tetrahedra's cross-linked together with oxygen atoms. Positive ions (Na^+, K^+, Ca^{2+}) must balance the negative charge if Al is in 4-fold coordination [7].

$$Mn[-(SiO_2)z - AlO_2]n - wH_2O \quad (1)$$

Previous researchers have investigated the application of geopolymer by using different waste materials, including sludge, fly ash, kaolin, risk husk, and slag [9-11]. The boiler ash contained the sum of SiO_2 and Al_2O_3 about 44 to 55% which is the most required composition to produce geopolymer [12-13]. The researchers have systematically studied the effect of using different alkaline solutions, sodium hydroxide concentrations, curing times, temperatures, and sodium hydroxide and sodium silicate ratios [14-15]. Duchesne et al. [16] in their study, proved that NaOH presence in alkaline solution activated the reaction, dissolved Si^{4+} and Al^{3+}, and produced smooth gel to fly ash. The presence of sodium silicate in the alkaline solution resulted in the rapid process of geopolymerization.

Khan et al. [17] agreed that the higher the percentage of fly ash-based geopolymer added, the higher the value of compressive strength. In one of his samples, 70% of compressive strength was achieved using fly ash geopolymer in cement compared to the use of ground granulated blast furnace slag (GGBS) in cement with only a 30% increase. The concentration of sodium hydroxide played an important role in dissolving the silica-alumina from the source material. In cement application, most of NaOH's molarity is 8 to 16 M at curing days of seven to twenty-eight days. As agreed by Emdadi et al. [18] in their study, compressive strength was intensively ranging from 10-14 M NaOH. Different molarity of NaOH gives a different effect on geopolymer. In Puertas et al. [19], a mixture of fly ash and NaOH of 10 M was added to slag, resulting in compressive strength of approximately 50 MPa. Investigation on the potential use of geopolymer of low calcium fly ash with different ratios of alkali activator, M1(0.5), M2(1.0), M3(1.5), M4(2.0), and M5(2.5), at curing time of three days gave compressive strength of 34.7, 61.6, 40.4, 40.5, and 22.3 MPa respectively. As agreed by Skvara et al. [20] who observed the increasing ratio from 2 to 2.5, the compressive strength decreased due to the excess sodium silicate hindering water evaporation and structure formation.

Soil stabilization is a typical soil remediation method to improve the physical and chemical properties of low strength soil by mixing the potential raw materials as additives or processes. Laterite soil is known to be very resistant to erosion. Laterite soil can be demonstrated by the nature of its properties such as high shear strength, low infiltration capacity, and low clay content, making it an excellent geotechnical material [21]. Ahmad Sujeeth et al. [22] mentioned that laterite soil properties can easily change as it is exposed to air and water during construction. Laterite soil is full of cavities and pores; therefore, it is crucial to select sufficient treatment in laterite soil to enhance soil strength and properties. Marto et al. [23] stated that laterite is widely used and economically convenient in construction. Hence, information on laterite properties is essential to design the geotechnical process's strength and durability in construction.

In recent years, the stabilization technique has been studied using geopolymerization. The geopolymer has become an effective method because the geopolymer source is cheaper since it uses industrial waste and reduces environmental pollution. Varied types of soil such as red mud soil, laterite soil, loess deposit, and sandblasting grit have been studied using a geopolymer to improve the soil properties [24-27].

Thiha et al. [28] studied the shear parameter of three soil types at curing times of zero and seven days. NaOH and Na_2SiO_3 were used as alkaline solutions and mixed with fly ash to produce a geopolymer. The ratio of fly ash-geopolymer to the soil was at 0.1. In the study,

soil with silty sand and high plasticity clay gave a higher soil strength than high plasticity silt at a curing time of seven days. Moreover, geopolymer curing at seven days was higher at 30% at day zero. This was supported by Wisam et al. [29] who made an extension of research to investigate NaOH's different molarity (4, 6, 8, 10 and 12 M) in geopolymer technique. Control soil samples displayed a UCS value of 77.9 kPa. Alkali activation for soil-olivine mixture increased the strength of the mixture to 392.8, 421.4, 583.9 1082.8 and 904.3 kPa at 7 days, respectively. As a conclusion on this research, geopolymer reaction gives higher soil strength over time. Alkaline activator dissolves the bonding of Si-Al-Si from the source material and gains stronger atom bonding. Further study on the long-term performance should be carried out on boiler ash to compare other raw materials' effectiveness.

Nik et al. [30] studied the effectiveness of different percentages of class C fly ash geopolymer on the compaction effort used in a soil liner application. In their study, an investigation of compaction and a hydraulic conductivity test were conducted. Based on the compaction test, 15% of the geopolymer gave optimum MDD, which led to a good hydraulic conductivity test. MDD 20% of the geopolymer decreased as the excess water in the geopolymer failed the geopolymerization.

A few works have studied the use of geopolymer binder as material in soil improvement such as sludge, fly ash, kaolin, risk husk, and slag [9-11]. However, no Malaysian works have used boiler ash as geopolymer source material. It is essential to carry out a study to understand and develop new green alternative materials to sustain the environment from industrial waste and convert it into alternative raw materials to enhance soil properties. This study investigates the potential of boiler ash as source material in producing a geopolymer to enhance laterite soil properties and sustain the environment.

2. METHODOLOGY

Laboratory work was carried out in the Faculty of Civil Engineering in UiTM Johor Kampus Pasir Gudang, Johor, Malaysia. The total number of samples in this study was nine samples for chemical properties, sixteen samples for physical properties, five samples for compaction value, and three samples for unconfined compression test conductivity with a total of thirty-three samples. The sample number and sample definition of this analysis are shown in Table 1. The best percentage of boiler-ash geopolymer added to laterite soil was investigated using five different percentages, (0, 5, 10, 15, and 20%). This percentage was determined based on the previous soil stabilization literature review [31-32].

Table 1: Soil descriptions

No	Sample Number	Soil Description
1	LS - P	Laterite Soil + 0% POBA geopolymer
2	LS – 5.0 GPOBA	Laterite Soil + 5.0% POBA geopolymer
3	LS – 10.0 GPOBA	Laterite Soil + 10.0% POBA geopolymer
4	LS – 15.0 GPOBA	Laterite Soil + 15.0% POBA geopolymer
5	LS – 20.0 GPOBA	Laterite Soil + 20.0% POBA geopolymer

Three stages of laboratory work were conducted to determine the physical properties. The Atterberg test, pH test, particle density test, and shrinkage test followed the British Standard BS 1377-2:1990. Scanning electron machine (SEM) was utilized at the Faculty of Chemical Engineering, UiTM Johor Kampus Pasir Gudang, and X-Ray Fluorescence at Jabatan Kimia UTM Skudai Johor, covering the chemical properties test to determine the

https://doi.org/10.31436/iiumej.v22i2.1589

soil structure of laterite soil mixed with geopolymer and the chemical composition in palm oil boiler ash. Compaction and unconfined compression strength (UCT) engineering tests were performed to investigate the reaction of soil strength to geopolymer. Fig.1 shows the study flow chart in this study.

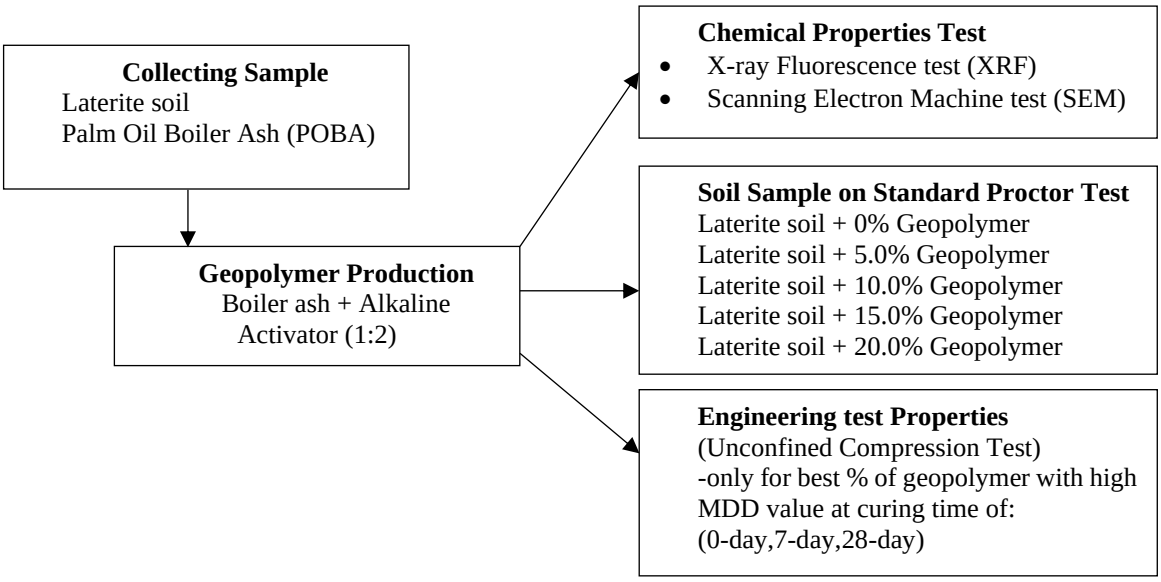


Fig 1: Study flow chart.

2.1 Materials

In this study, laterite soil was mixed with a geopolymer to produce a good binder to enhance the properties of laterite soil. To generate a geopolymerization reaction, raw material from industrial waste known as palm oil boiler ash and alkaline activator (sodium hydroxide and sodium silicate) were mixed with a suitable mixed design as discussed below, thus generating a new green technology binder known as geopolymer. Data of properties and soil strength at different percentages of boiler ash- geopolymer mixed in laterite soil were recorded.

2.1.1 Laterite Soil

The laterite soil used was collected at a depth of 1.5 m, 100 m away from Tanjung Langsat Landfill, Pasir Gudang, Johor, Malaysia. The gradation test was conducted using BS 1377: Part 2:1990. Fig. 2 shows the gradation chart of the laterite soil. From the sieve analysis test, the laterite soil used in this study consisted of about 25.98% of gravel, 35.55% of sand, and 38.47% of fine size grain and was classified under very “silty sand” as shown in Table 2. The laterite soil had a specific gravity of 2.62. An Atterberg limit test that was conducted on the laterite soil of sieve size 425 µm showed the result of the liquid limit (LL) of the soil was 58.61% and the plastic index (PI) was 6.15%. According to the Unified Soil Classification system (USCS), this soil is classified as high plasticity (MH). The laterite soil in this study consisted of a high portion of silt compared to clay resulting in a low shear strength or low compressibility. This type of soil was therefore chosen in this study to investigate the effect of geopolymer palm oil (boiler ash) to enhance the mechanical and properties of laterite soil.

https://doi.org/10.31436/iiumej.v22i2.1589

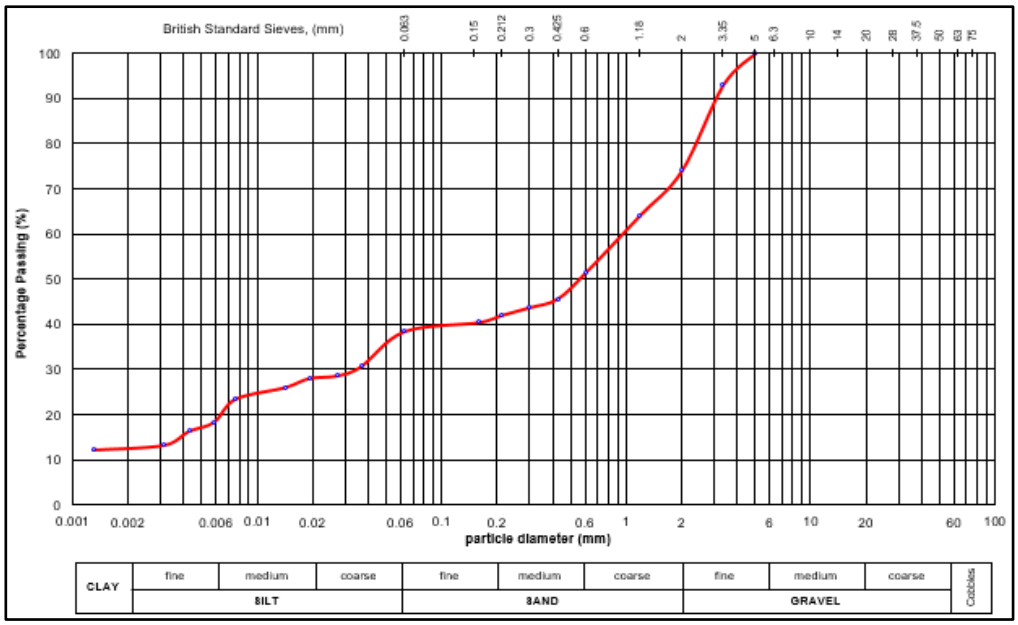


Fig. 2: Grain size distribution of laterite soil.

Table 2: Grading characteristics of the laterite soil

Composition of soil	
Gravel %	25.98
Sand %	35.55
Silt %	25.81
Clay (%)	12.66
Gradation	
D_{60} (mm)	0.995
D_{50} (mm)	0.558
D_{30} (mm)	0.033

2.1.2 Palm Oil Boiler Ash (POBA)

Palm oil boiler ash (POBA) was sourced from Teluk Sengat palm oil mill, Kota Tinggi, Johor. The palm oil boiler ash, as shown in Fig.3, consisted of boiler ash that was obtained from the lower compartment of a palm oil boiler burning what consisted of large particles including nutshells and kernels. Palm oil boiler ash is one of the industrial wastes from the burning of palm oil that creates disposal problems and causes soil contaminants as the waste is stored in landfills. The chemical composition of the palm oil boiler ash was determined using x-ray fluorescence (XRF), as shown in Table 3. The higher content of silica could be the potential of a geopolymer.

Fig. 3: Palm oil boiler ash (POBA) in Teluk Sengat Palm Mill, Johor.

https://doi.org/10.31436/iiumej.v22i2.1589

Table 3: Chemical composition of laterite soil and boiler ash

Oxides	**SiO_2**	**Al_2O_3**	**CaO**	**Fe_2O_3**	**Na_2O**	**MgO**	**P_2O_5**	**K_2O**
Laterite Soil (LS-P)	56.0	30.3	0.381	9.54	-	-	-	0.290
Palm Oil Boiler Ash (POBA)	43.9	4.55	13.1	4.11	-	3.27	2.45	12.5

Based on Table 3, the percentage of boiler ash was suitable to be a geopolymer source material as the percentage of silica was more than 40% [5]. In laterite soil, the silica-alumina comprised 30 percent to 60 percent content. However, the properties of laterite soil were complex such as shrinkage and low plastic index due to the presence of water. Therefore, it was not sufficient for laterite soil alone without additives to be the source material for geopolymerization production. Fig. 4 shows the size sample of the palm oil boiler ash at 150 µm used in this study.

Fig. 4: Palm oil boiler ash passing 150 µm sieve size.

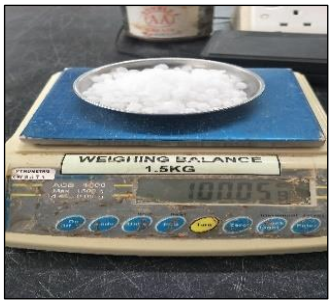
2.1.3 Alkaline Solution

The alkaline activator used in this study was sodium hydroxide (NaOH) in pellet form with 97% purity and sodium silicate (Na_2SiO_3), as shown in Fig. 5. Referring to Thiha et al. [28], the concentration of sodium hydroxide (NaOH) in this study was chosen to be 5 M. Sodium silicate (Na_2SiO_3) was a liquid solution that consisted of 9.4% Na_2O, 30% SiO_2, and 60.5% H_2O taken from chemical suppliers in Johor. The ratio of sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3) in the study was 1:2.

2.2 Sample Preparation

The palm oil boiler ash was collected and ground to finer particles and sieved using 150 µm sieves and put in an oven at 80 ^oC for 24 hours to activate the silica content before being mixed with an alkaline activator. The fine particles of palm oil boiler ash were used to produce the geopolymerization reaction between palm oil boiler ash and alkaline solution at different mixed designs. The purpose of using an alkaline activator was to synthesize the chemical structure of palm oil boiler ash to produce a geopolymer. Sodium hydroxide (NaOH) and sodium silicate (Na_2SiO_3) were selected at a ratio of 1:2. The sodium hydroxide (NaOH) pellets were prepared by diluting them with distilled water at a concentration of 5 M (molarity) to create an alkaline solution. The concentration of the sodium hydroxide solution (NaOH) used was 5 M. For preparation of 1 M solution, 40 g NaOH flakes were dissolved in water to prepare 1 liter of sodium hydroxide solution. Thus, for preparation of 5 M, 200 g of NaOH flakes were used.

https://doi.org/10.31436/iiumej.v22i2.1589



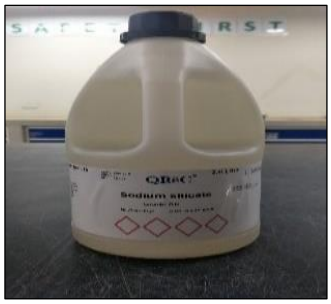


Fig. 5: (left) Sodium hydroxide in pellet form; (right) Sodium silicate in liquid form.

During the preparation of NaOH, precautions were taken to put the pellets into the distilled water and not the other way around as sodium hydroxide could react with moisture from the air and generate heat while dissolving and cause a fire near flammable material, as well as possible sparks that could come into contact with eyes or skin. The sodium silicate was poured into the NaOH solution at a ratio of 1:2, stirred for 10 minutes, and left at room temperature for 24 hours to produce a better strength of mortar [12]. Lastly, to produce the geopolymer binder, 80% of boiler ash and 20 % of alkaline solution were carefully mixed before being added to laterite soil at a different percentage of mixed design, as shown in Fig. 6.

After the geopolymer curing at ambient temperature for 24 hours, the geopolymer was mixed with laterite soil at a different percentage of 0, 5, 10, 15, and 20% as shown in Fig. 7 The samples were then tested to determine the physical properties, chemical properties, compaction test, and soil strength using an unconfined compression test (UCT) at zero, seven, and twenty-eight days of curing.

Fig 6: Alkaline solution mixed with POBA.

Fig 7: Preparation of geopolymer mixed with laterite soil.

2.3 Experimental Design

2.3.1 Compaction Test

Five samples at different percentages of geopolymer (0, 5, 10, 15, and 20%) were prepared. The percentage of water required was added and mixed until a uniform consistency was achieved [30]. The standard proctor test as shown Fig.9 was conducted in this study to determine the maximum dry density (MDD) and optimum moisture content (OMC). The soil sample was dynamically compacted with a release of a steel hammer with a weight of 2.5 kg and 27 blows per each of three layers, falling freely 300 m with a 50 mm diameter of rammer. After compaction was conducted, samples were then trimmed to remove the excess soil, molded, and weighed. Equation (2) and Equation (3) show the calculation of the standard proctor test:

$$\rho = \frac{m_2 - m_1}{v} \; Mg/m^3 \quad (2)$$

where:

m_1 = mass of mold and base

m_2= mass of soil and mold

V = volume of mold

Then, dry density was calculated from the equation below:

$$\rho = (\frac{100}{100+w})\rho \quad Mg/m^3 \quad (3)$$

The dry density was recorded to determine the optimum moisture content (OMC) at the maximum dry density (MDD) to conduct an unconfined compression test (UCT).

2.3.2 Unconfined Compression Test (UCT)

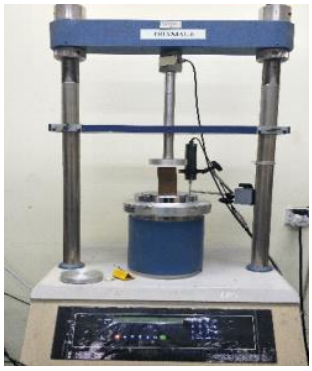


Fig. 8: UCT machine.

Unconfined compressive strength is based on ASTM Designation D2166. The size of samples in this test was 50 mm in diameter with 100 mm height based on the procedure in ASTM D2166. In this study, the investigation of unconfined strength was conducted for sample LS-15.0 GPOBA measured at curing times of the compacted sample under ambient temperature. Curing times of zero, seven, and twenty-eight days were chosen for this study. These samples were selected to determine the before and after reaction of the geopolymer towards laterite soil. Also, LS-15.0 GPOBA was chosen due to the best compaction obtained from MDD and OMC results.

The sample was mixed with water, based on moisture content resulting from the compaction test for LS- 15.0 GPOBA. Then, the samples were compacted in the mold with three (3) layers 10 tamping's of compaction for each layer. The compressed samples at zero day were immediately diverted and installed for the UCT test as shown in Fig. 8. Samples of seven and twenty-eight days were kept in a plastic bag and tightly sealed under ambient temperature based on curing times. When the sample reached its curing time, it was taken out and immediately run on the UCT machine. In this test, all samples were sheared with the dry condition without submerging the sample into water.

3. RESULTS AND DISCUSSION

3.1 Physical Properties of Samples

This study conducted laboratory investigations on different percentages of geopolymers to enhance the properties of laterite soil. This study aimed to determine the soil compaction and shear strength of geopolymer as additives in laterite soil. Table 4 shows the physical properties of laterite soil. The properties and characteristics of laterite soil are important to be investigated for the future development of laterite soil.

Table 4: Physical properties of different percentages of geopolymer in laterite soil

Soil Description	LS-P
Liquid Limit (%)	58.61
Plastic Index (%)	6.15
Specific Gravity	2.62
Ph	6.49
Shrinkage	11.49

Based on the experimental data collected, with the right percentage of boiler ash-based geopolymer added to laterite soil, it improves the properties of soil and enhances the soil strength. Further investigation in the compaction test and soil strength was conducted as follows.

3.2 Compaction Test

Fig. 10 shows the compaction curve indicating maximum dry density (MDD) and optimum moisture content (OMC) of all geopolymer percentages in laterite soil. Based on the compaction curve produced, 0% of geopolymer in laterite soil gave a maximum dry density of 1.80 Mg/m^3 at 14.07% optimum moisture content. The adding of geopolymer at 5% and 10% showed increasing maximum dry densities of 1.88 Mg/m^3 and 1.90 Mg/m^3 and decreasing optimum moisture contents at 13.88% and 13.70%, respectively. The increase of moisture content was probably because of the reaction between laterite soil, and geopolymer led to the loss of water [27]. However, drastic increase can be seen of maximum dry density of 2.23 Mg/m^3 and decrease of optimum moisture content at 13.58% at 15.0 % of geopolymer in laterite which give positive results of geopolymerization reactions. 15.0 % of geopolymer in laterite soil indicated the optimum mixed design of geopolymer as a mixed design. The bonding created by alkaline solution and the subsequent isomorphous substitution of silica (*SI*) has indeed increased the packing between the grains [32]. This result is supported by Nik et al. [30] in their study on laterite soil and fly ash geopolymer mentioned, whereby geopolymer added in laterites soil showed an increment in dry density due to the action of alkaline attack which changed the soil mineral. In 20% of geopolymer, results gave a decreasing value of dry density of 1.71 Mg/m^3 and increasing optimum moisture content at 16.38%. This result happened due to the optimum and peak reaction achieved in 15.0 % of geopolymer in laterite soil. The structure of silica-alumina in geopolymerization was lost due to increased moisture content because of the rapid reaction between geopolymer and laterite soil, which led to the loss of water. This result is in good agreement with the Noushini et al. [35] study of the behavior in laterite soil with sodium silicate-liquid stabilizer. Based on the result obtained from this study in compaction parameters, the increased percentage of geopolymer added to laterite soil gives increasing MDD and decreasing OMC. It was observed that the presence of a geopolymer in laterite

soil gave a strong relationship between moisture content and maximum dry density shows the compaction curve of tested samples.

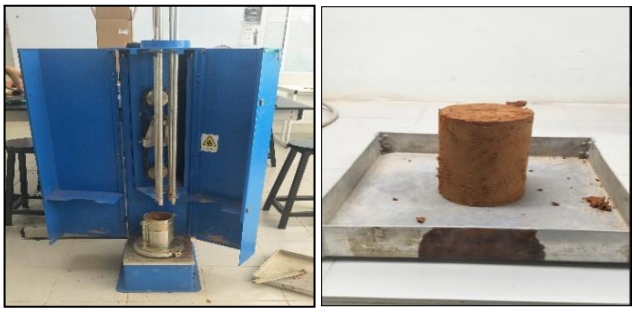


Fig. 9: Standard Proctor Machine and compaction sample.

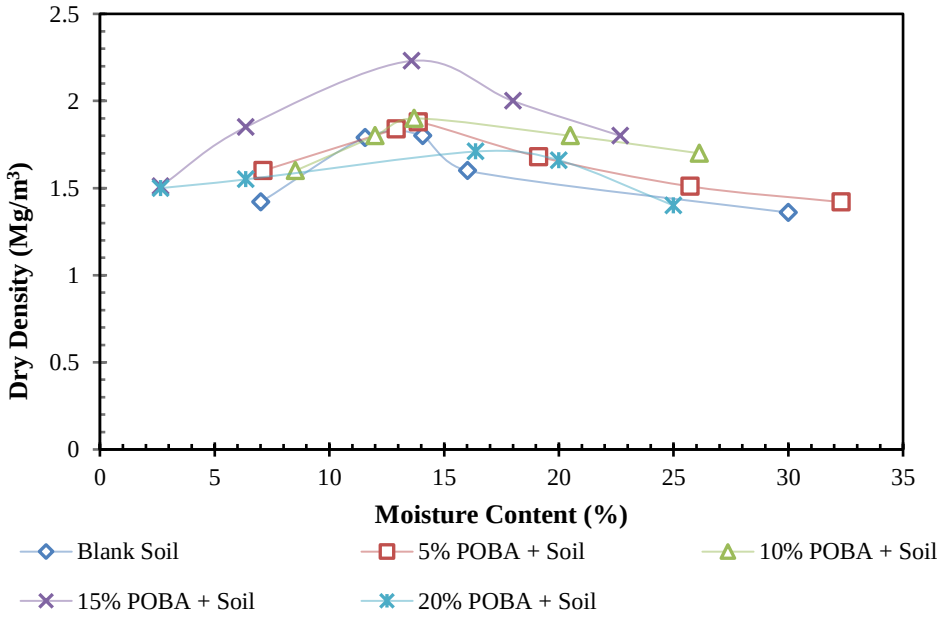


Fig. 10: Compaction Curve of Laterite soil with different percentages of geopolymer**.**

3.3 Unconfined Compression Test (UCT)

Soil compaction has direct effects on soil physical properties such as bulk density, strength, and porosity [36]. Therefore, the LS-15.0 GPOBA sample that gave the best value of maximum dry density (MDD) of 2.23 Mg/m^3 was directly chosen as the optimum percentage to determine the soil strength using an unconfined compression test (UCT). In this study, the effect of compressive force on sample LS–15.0 GPOBA was investigated at different curing times of zero, seven, and twenty-eight days. The samples were compacted under moisture content based on the compaction parameter of LS- 15.0 GPOBA, which was 13.70%. The pieces were cured at ambient temperature (27°C to 30°C) reflecting the field conditions. The soil sample took on a hardened state after compacting. During the curing time, chemical reactions were expected to happen between the soil grain and the geopolymer molecules. Table 5 shows the samples from lowest strength at zero day of curing (47 kPa) after seven days of curing (58 kPa).

https://doi.org/10.31436/iiumej.v22i2.1589

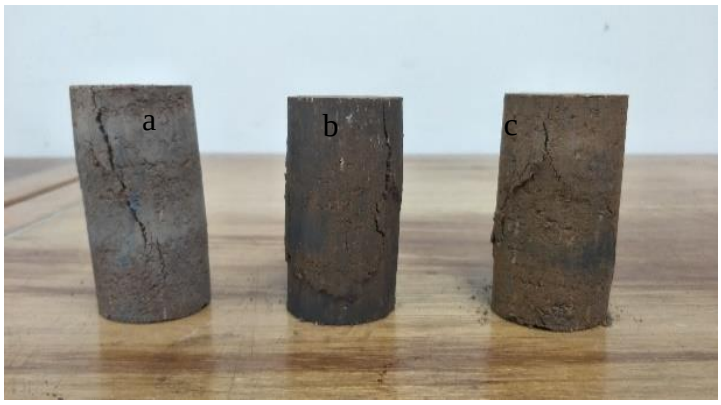


Fig. 11: (a) zero days, (b) seven days, (c) twenty-eight days of UCT.

The soil sample obtained after twenty-eight days achieved the highest compressive strength of 76 kPa. Fig.12 showed a slight increase in compressive strength in increasing curing times of zero to twenty-eight days. The result is similar to previous studies [30,37-38], where it shows the increment of compressive strength that proves the improvement of the geopolymerization process along the curing time. According to Emdadi et al. [18] the form of aluminosilicate and alkaline influences the effect of compressive force in geopolymer. Based on the result in this study, palm oil boiler ash (POBA) that contains a high percentage of silica can be used as pozzolanic materials to enhance the shear strength of soils. The geotechnical technology using a geopolymer can be a potential study and practice in applying soil strength such as soil slope, embankment, and landfill soil liner.

Table 5: Shear parameters of laterite soil with 15% of geopolymer at curing times of zero, seven, and twenty -nine days condition

Curing Time (Days)	Shear Strength (kPa)	UCS (kPa)
0	47	94
7	58	132
28	76	153

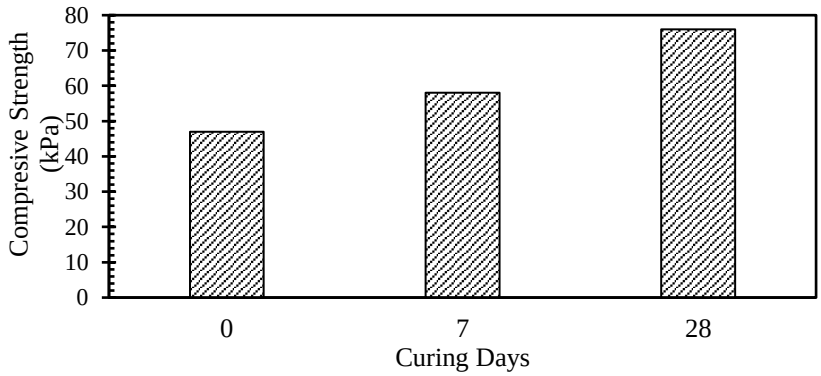


Fig. 12: The comparison of soil strength (LS-15.0 GPOBA) at zero, seven, and twenty-eight days of curing time.

3.2 Microstructure of Materials

Fig 13 shows the morphological features of plain laterite soil (LS-P) and Fig 14 shows the laterite soil with 15% of geopolymer (LS.15.0 GPOBA) at a curing time of zero days. The test was examined using Scanning Electron Machine analysis using magnifying power 5000x at the Faculty of Chemical Engineering in UiTM Johor Campus, Pasir Gudang, Johor, Malaysia. As seen in Fig. 13, the structure layered with larger voids led to low compressive strength. The microstructure of the laterite soil without geopolymer can be found to be loose and not uniform; cracks and pores are observed between soil particles. Meanwhile, in Fig. 14 some particles show popcorn type and spherical type. Moreover, the structure of laterite soil with a geopolymer demonstrated a denser geopolymer matrix after the geopolymerization process with less pores compare to blank laterite. In the study by Temujun et al. [39] the soil structure becomes compact causing higher packing density, low porosity in microstructure, and leading to higher strength.

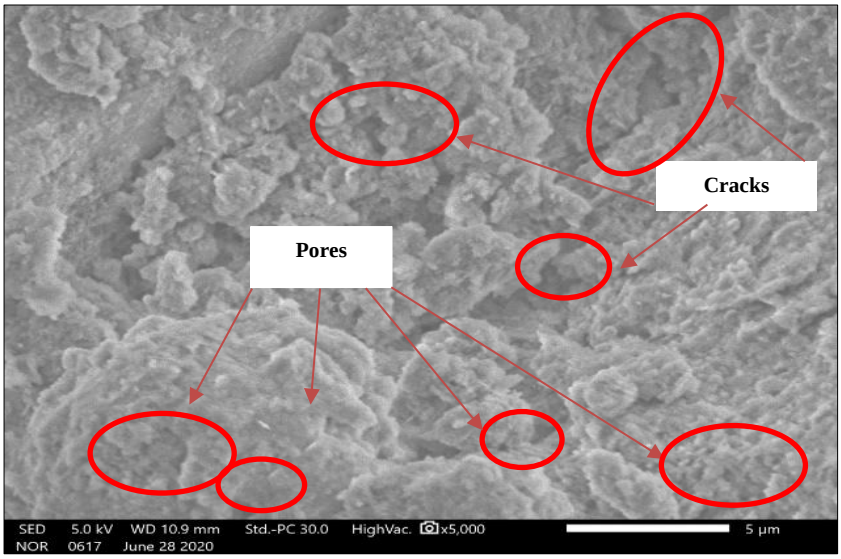


Fig. 13: SEM Samples of Plain Laterite (LS-P).

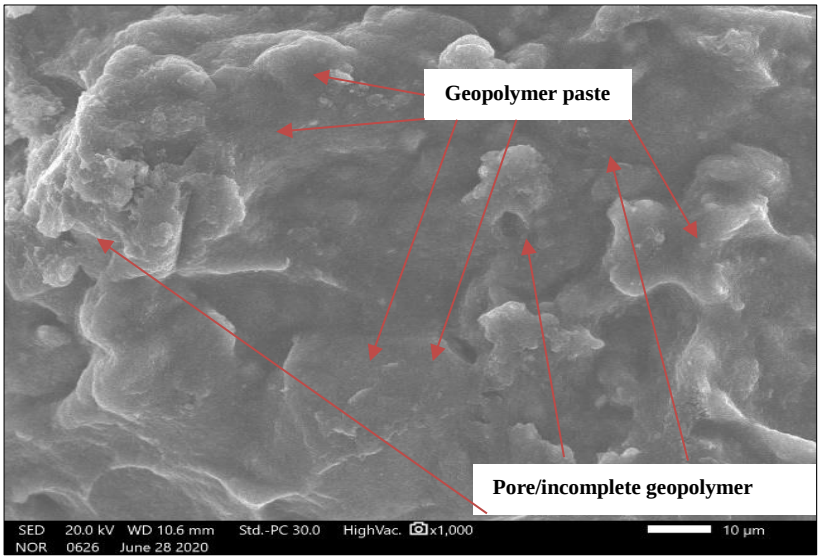


Fig. 14: SEM Samples of Laterite soil with 15% of Geopolymer (LS.15.0 GPOBA).

4. CONCLUSION

This paper examines the influence of geopolymers on the properties of laterite soil to be applied in the field of geotechnical application. The palm oil boiler ash combined with an alkaline solution at a constant ratio to form a geopolymer binder has shown to have a beneficial effect on compaction and high compressive strength by improving the properties of the laterite soil. Based on the results of the study, the effective percentage of geopolymer in laterite soil was found at 15% that generated a maximum dry density of 2.23 Mg/m^3 at 13.58% optimum moisture content which gave a good compaction parameter leading to the high strength of soil properties. The UCT was only conducted on the sample of LS.15.0 GPOBA, which offered the highest MDD. Samples were cured in a plastic bag at ambient temperature to suit field conditions at zero, seven, and twenty-eight days accordingly. Based on the experimental data, the compressive strength of the geopolymer slightly increased as curing days increased on zero, seven, and twenty-eight and gave the strength of 47, 58, and 76kPa, respectively. Based on SEM test, it is apparent that it is due to the geopolymerization process that makes the sample harder and denser, which improves the soil strength. The SEM images have proven that the development of geopolymer gel in laterite soil has stabilized the soil structure from a loose structure to a denser soil structure. Thus, geopolymer-based palm oil boiler ash and laterite soil may be a potential by-product as an alternative to soil treatment in geotechnical applications that minimize porosity and potentially reduce permeability.

ACKNOWLEDGEMENT

The author would like to acknowledge the help of Associate Professor Dr. Ts. Mazidah Mukri as a post-graduate supervisor in this study. This work is also supported by the Faculty of Civil Engineering and Faculty of Chemical Engineering, UiTM Johor Campus, Pasir Gudang for their facilities in the testing laboratory.

REFERENCES

[1] Hai TC, Ng A, Prudente C, Pang C, Tek Choon Yess, J. (2001) Balancing the Need for Sustainable Oil Palm Development and Conservation: The Lower Kinabatangan Floodplains Experience. Strategic Directions for The Sustainability of the Oil Palm Industry, 53 pp. Retrieved From http://Assets.Panda.Org/Downloads/Balancingtheneed.pdf

[2] Abdullah N, Sulaim F. (2013) The Oil Palm Wastes in Malaysia. in Biomass Now - Sustainable Growth and Use. Intech. https://doi.Org/10.5772/55302

[3] Mangi,SA, Ibrahim MHW, Jamaluddin N, Arshad MF, Memon, FA, Jaya RP, Shahidan S. (2018) A Review on Potential Use of Coal Bottom Ash as a Supplementary Cementing Material in Sustainable Concrete Construction. International Journal of Integrated Engineering, 10(9): 127-135. https://doi.Org/10.30880/Ijie.2018.10.09.006

[4] Deraman LM, Al Bakri Abdullah,MM, Ming LY, Hussin K, Yahya Z. (2015) A Review on Processing and Properties of Bottom Ash Based Geopolymer Materials. in Key Engineering Materials, 660: 3-8. https://doi.Org/10.4028/www.Scientific.Net/KEM.660.3

[5] Davidovits J. (2005) Geopolymer, Green Chemistry and Sustainable Development Solutions. in J. Davidovits (Ed.), Proceedings of the World Congress Geopolymer 2005 (p. 236). Institut Geopolymere.

[6] Wattimena OK, Antoni, Hardjito D. (2017) A review on the effect of fly ash characteristics and their variations on the synthesis of fly ash based geopolymer. in AIP Conference Proceedings, 1887: 020041. https://doi.org/10.1063/1.5003524

[7] H Xu, Van Deventer JSJ. (2000) The geopolymerisation of alumina-silicate minerals. Int. J. Miner. Process., 59(3): 247-266.

[8] Van Jaarsveld JGS, Van Deventer JSJ, Lorenzen L. (1997) Potential use of geopolymeric materials to immobilize toxic metals: Part I. Theory and applications, Miner. Eng., 10(7): 659-669.

[9] Kanesan D, Ridha S, Rao P. (2017) Formulation of geopolymer cement using mixture of slag and class f fly ash for oil well cementing. In IOP Conference Series: Materials Science and Engineering, 201: 012014. https://doi.org/10.1088/1757-899x/201/1/012014

[10] Nanavati S, Chetan Nanavati S, Jaywant Lulla S, Ranjeet Singh A, Bharat Mehta D, Mohan Patel A, Author C. (2017) A review on fly ash based geopolymer concrete. IOSR Journal of Mechanical and Civil Engineering, 14(4): 12-16. https://doi.org/10.9790/1684-1404071216

[11] Kabir SMA, Alengaram UJ, Jumaat MZ, Sharmin A, Islam A. (2015) Influence of molarity and chemical composition on the development of compressive strength in POFA based geopolymer mortar. Advances in Materials Science and Engineering, 2015: 647071 https://doi.org/10.1155/2015/647071.

[12] Yahya Z, Abdullah MMAB, Hussin K, Ismail KN, Razak RA, Sandu AV. (2015) Effect of solids-to-liquids, Na_2SiO_3-to-NaOH and curing temperature on the palm oil boiler ash (Si + Ca) geopolymerisation system. Materials, 8(5): 2227-2242. https://doi.org/10.3390/ma8052227.

[13] Lopez E., Vega-Zamanillo A., Calzada Perez M.A., & Hernandez-Sanz A. (2015). Bearing capacity of bottom ash and its mixture with soils. Soils and Foundations, 55, 529–535. https://doi.org/10.1016/j.sandf.2015.04.005

[14] Jindal B. B. (2019). Investigations on the properties of geopolymer mortar and concrete with mineral admixtures: A review. Construction and Building Materials. Elsevier Ltd. https://doi.org/10.1016/j.conbuildmat.2019.08.025.

[15] Ma C. K, Awan, A. Z., & Omar W. (2018). Structural and material performance of geopolymer concrete: A review. Construction and Building Materials. Elsevier Ltd. https://doi.org/10.1016/j.conbuildmat.2018.07.111.

[16] Duchesne J., Duong L., Bostrom T., & Frost R. (2010). Microstructure study of early in situ reaction of fly ash geopolymer observed by environmental scanning electron microscopy (ESEM). Waste and Biomass Valorization, 1(3), 367–377. https://doi.org/10.1007/s12649-10-9036-4.

[17] Khan, H., Purohit, D., Bagara, D., & Pahadiya, H.S. (2016). Geopolymer Concrete with Replacement of Cement. International Journal of Engineering Research & Technology (IJERT) 4-23.

[18] Emdadi, Z., Asim, N., Amin, M. H., Yarmo, M. A., Maleki, A., Azizi, M., & Sopian, K. (2017). Development of Green Geopolymer Using Agricultural and Industrial waste Materials with High Water Absorbency. Applied Sciences (Switzerland), 7(5). https://doi.org/10.3390/app7050514.

[19] Puertas, F. & Martınez-Ramırez, S & Alonso, S & Vazquez, T. (2000). Alkali-Activated Fly Ash/Slag Cement Strength Behavior and Hydration Products. Cement and Concrete Research. 301. 1625-1632.

[20] Skvara, F., Kopecky, L., Nemecek, J., & Bittnar, Z. (2006). Microstructure of geopolymer materials based on fly ash. Ceramics - Silikaty, 50(4), 208–215.

[21] Rosli, Ros, Nadiah. (2017). Properties of Laterite Soils from Sources Near Nibong Tebal, Malaysia. Conference: ICONTES Vistana Hotel Kuala Lumpur https://www.researchgate.net/publication/324720622

[22] Sujeeth, Ahmed. (2015). An Investigation into The Geotechnical Engineering Properties of Laterite Soils in Nilai, Malaysia. 10.13140/RG.2.2.20036.04489.

[23] Marto, A., Latifi, N., & Sohaei, H. (2013). Stabilization of laterite soil using GKS soil stabilizer. Electronic Journal of Geotechnical Engineering, 18 C, 521–532.

[24] Parikshith, M. V., & Sekhar, D. C. (2019). Feasibility of fly-ash based geopolymer for soil stabilization. International Journal of Innovative Technology and Exploring Engineering, 9(1), 4348–4351. https://doi.org/10.35940/ijitee.A5019.119119

[25] Mukri, M., Aziz, N. N. S., & Khalid, N. (2018). Geopolymer Effect in Modelling Hydraulic Conductivity for Designing Soil Liner of Laterite Soil. International Journal of Civil Engineering and Technology, 9(7), 1962–1974.

[26] Liu, Z., Cai, C. S., Liu, F., & Fan, F. (2016). Feasibility study of loess stabilization with fly ash-based geopolymer. Journal of Materials in Civil Engineering, 28(5). https://doi.org/10.1061/(ASCE)MT.1943-5533.0001490.

[27] Nikolic, I., Zejak, R., Radmilovic, V., Blecic, D., & Tadic, M. (2012). Geopolymerization of fly ash as a possible solution for stabilization of used sandblasting grit.

[28] Thiha, S., Lertsuriyakul, C., & Phueakphum, D. (2018). Shear Strength Enhancement of Compacted Soils Using High-Calcium Fly Ash-Based Geopolymer. International Journal of GEOMATE, 15(48), 1–9. https://doi.org/10.21660/2018.48.35692

[29] Dheyab, W., Ismael, Z. T., Hussein, M. A., & Huat, B. B. K. (2019). Soil stabilization with geopolymers for low cost and environmentally friendly construction. International Journal of GEOMATE, 17(63), 271–280. https://doi.org/10.21660/2019.63.8159

[30] Nik Ab Aziz, N. N. S., Mukri, M., Hashim, S., & Khalid, N. (2015). Influence of Compaction Effort for Laterite Soil Mix with Geopolymer In Designing Soil Liner. Electronic Journal of Geotechnical Engineering, 20(22), 12353–12364.

[31] Ahmad, F. H., Mukri, M., & Azmi, N. A. C. (2017). Effectiveness of Different Percentage of Bentonite in Soil Liner on Interface Shear Strength with Geosynthetic. International Journal of Applied Engineering Research, 12(7), 1360–1364.

[32] Ashraf, M. A., Hossen, M. A., Ali, M. A., & Chakaraborty, B. P. (2018). Stabilization of Soil by Mixing with Different Percentages of Lime, 1–10. The 4th International Conference on Civil Engineering for Sustainable Development

[33] Zhou, S. Q., Zhou, D. W., Zhang, Y. F., & Wang, W. J. (2019). Study on Physical-Mechanical Properties and Microstructure of Expansive Soil Stabilized with Fly Ash and Lime. Advances in Civil Engineering. https://doi.org/10.1155/2019/4693757

[34] Fattah, M., Y., Al-Saidi, A., And Jaber, M. (2014) Consolidation Properties of Compacted Soft Soil Stabilized with Lime-Silica Fume Mix. International Journal of Scientific & Engineering Research, 5, 1675-1682.

[35] Noushini, A., & Castel, A. (2016). The Effect of Heat-Curing on Transport Properties of Low-Calcium Fly Ash-Based Geopolymer Concrete. Construction and Building Materials, 112, 464–477. https://doi.org/10.1016/j.conbuildmat.2016.02.210.

[36] Nawaz, M. F., Bourrié, G., & Trolard, F. (2013). Soil Compaction Impact and Modelling. A Review. Agronomy for Sustainable Development, 33(2), 291–309. https://doi.org/10.1007/s13593-011-0071-8.

[37] Khale, D., & Chaudhary, R. (2007). Mechanism of Geopolymerization and Factors Influencing Its Development: A Review. Journal of Materials Science, 42(3), 729–746. https://doi.org/10.1007/s10853-006-0401-4.

[38] Alvarez-Ayuso, E., Querol, X., Plana, F., Alastuey, A., Moreno, N., Izquierdo, M., Barra, M. (2008). Environmental, Physical and Structural Characterization of Geopolymer Matrixes Synthesized from Coal (Co-)Combustion Fly Ashes. Journal of Hazardous Materials, 154(1–3), 175–183. https://doi.org/10.1016/j.jhazmat.2007.10.008

[39] Temuujin, J., & Van Riessen, A. (2009). Effect of Fly Ash Preliminary Calcination on The Properties of Geopolymer. Journal of Hazardous Materials, 164(2–3), 634–639. https://doi.org/10.1016/j.jhazmat.2008.08.065.

https://doi.org/10.31436/iiumej.v22i2.1599

CHARACTERIZATION OF MORTAR WITH *PENNISETUM PURPUREUM* ASHES AS CEMENT REPLACEMENT MATERIAL

MOHD HAZIMAN WAN IBRAHIM[1*], MOHAMMED YAHYA AL-FASIH[2], NIK NADIA AMIRA NIK PA[1], RAMADHANSYAH PUTRA JAYA[3], AND MUHAMMAD IKHSAN SETIAWAN[4]

[1]Jamilus Research Center, Faculty of Civil Engineering and Built Environment, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor, Malaysia
[2]Forensic Engineering Centre (FEC), Institute of Smart Infrastructure and Innovative Construction (ISIIC), School of Civil Engineering, Faculty of Engineering, Universiti Teknologi Malaysia, 81310 Johor Bahru, Johor
[3]Department of Civil Engineering, College of Engineering, Universiti Malaysia Pahang, 26300 Gambang, Pahang, Malaysia.
[4]Department of Civil Engineering, Faculty of Engineering, Narotama University, Surabaya, Jawa, Indonesia

[*]Corresponding author: haziman@uthm.edu.my

(Received: 27th August 2020; Accepted: 30th January 2021; Published on-line: 4th July 2021)

ABSTRACT: In this study, the properties of mortar such as standard consistency, setting time, compressive strength, and water absorption rate were investigated. The cement was replaced with *Pennisetum purpureum* ashes (PPA) in different particle sizes and dosages. PPA was produced in greyish-white ash with total reactive oxides ranging from 37% to 41.1%. *Pennisetum purpureum* grass (PPG) was burned with a controlled process at 350 °C for the first 3 hours and 600 °C for another 3 hours at a heating rate of 10 °C/min. Then, PPG was ground in three different grinding durations (1, 3, and 6 hours) producing ashes with particle sizes of 10.58 µm, 10.25 µm, and 9.30 µm, respectively. The physical, chemical, and microstructural properties of PPA were evaluated through several tests; particle size analysis, X-ray diffraction analysis, loss on ignition, and chemical composition. Results indicated that PPA is more suitable for use as filling material as a substitute for cement than pozzolanic material as its reactive oxides are less than 50%. The 15% 6H-PPA at 28 days was found to be the optimum PPA replacement dosage and grinding time with cement as it achieved the highest strength and lower water absorption rate among all samples at 7 and 28 days.

ABSTRAK: Kajian ini berkenaan sifat mortar seperti konsistensi standard, masa penyediaan, kekuatan menyeluruh, dan kadar penyerapan air. Simen ditukar dengan abu *Pennisetum purpureum* (PPA) dalam saiz partikel dan sukatan berbeza. PPA pula dihasilkan melalui habuk putih-kekelabuan dengan total reaktif oksida dengan julat purata 37% ke 41.1%. Rumput *Pennisetum purpureum* (PPG) dibakar dengan proses kawalan pada 350 °C selama 3 jam pertama dan 600 °C pada 3 jam berikutnya pada kadar pemanasan 10 °C/min. Kemudian, PPG dikisar dalam tiga tempoh kisaran berbeza (1, 3, dan 6 jam) menghasilkan abu dengan saiz partikel 10.58 µm, 10.25 µm, dan 9.30 µm, masing-masing. Fizikal, larutan kimia dan sifat struktur mikro PPA telah dikaji melalui pelbagai ujian; analisis saiz partikel, analisis pembelauan X-ray, kehilangan semasa penyalaan dan kandungan kimia. Dapatan kajian menunjukkan PPA adalah lebih sesuai digunakan sebagai material pengisian ganti kepada simen berbanding material pozzolanik kerana reaktif oksida adalah kurang daripada 50%. PPA adalah maksimum pada 15% 6H-

PPA selama 28 hari, didapati lebih sesuai sebagai dos pengganti dan masa kisaran bersama simen, kerana kekuatan menyeluruh adalah paling tinggi dan kadar penyerapan air paling kurang antara semua sampel pada 7 dan 28 hari.

KEYWORDS: *Pennisetum purpureum ashes; PPA mortar; replacement dosage; grinding time*

1. INTRODUCTION

The cement industry is one a strategic industry acting as a contributor to several other industries such as real-estate, construction, and infrastructure, where it plays an important role in the world economy [1]. Concrete materials should not only possess good workability, excellent mechanical properties, and durability, but also offer environmental and economic benefits. However, the cement production of the world, which is estimated at around 1.6 billion tons, contributes about 7% of the global carbon dioxide emissions due to the burning of limestone and clay. Thus, widespread cement usage in the industry has been considered unsustainable. Therefore, an alternative need to be found to help the cement industry, protect the environment, and also to promote sustainable development for the construction industry and the country. The sustainability concept in engineering practice is focused on the development of new materials for construction purposes and environmentally friendly materials.

Recently, the search for alternative binders, natural admixtures, or cement replacement materials has been the subject of many studies. The waste raw materials have been used as cement replacement materials to reduce the carbon footprint and enhance sustainability. The biomass waste ashes that contain a high amount of silica content such as sugar cane bagasse ashes [2], rice husk ashes [3], palm oil fuel ash [4], and corn cob ash [5] can be used as supplementary cementitious materials (SCMs). Biomass is defined as a sustainable energy source with carbon dioxide (CO_2) neutralization potential. Ghiasvand and Ramezanianpour [1] reported that the energy consumption and carbon dioxide (CO_2) emissions can be reduced through using SCMs. Furthermore, a portion of cement in a concrete system replaced by SCMs will participate in primary or secondary hydration reactions, increasing the volume fraction of binding phases and decreasing porosity in the microstructure of the concrete.

Other than traditional cementitious materials such as rice husk ash, silica fume, and fly ash, one of the alternatives is the ashes extracted from elephant grass or Napier grass (*Pennisetum purpureum*). The leaves of elephant grass are currently used as a food source for ruminants, but its trunks are dumped in the landfill. The trunks from the grass could be recycled into silica-based products. Napier grass can be harvested four times a year and it has the potential to produce about 100 barrels of oil equivalent per hectare [6]. Elephant grass is used now in Malaysia as a sustainable resource for the electricity supply system due to its advantages of yield capability, wide propagation, and high dry material [7]. However, a large quantity of elephant grass ashes was generated during the direct burning process of these activities. These ashes will finally be deposited in landfills with the consequent environmental, technical, economic, and social problems [8]. The output energy from Napier grass is estimated as 25 times higher than the input energy consisting of a high amount of lignocellulosic material.

However, elephant grass contains a considerable amount of amorphous silica due to the extraction of ortho-silicic acid from groundwater by the plant and its ashes can be economically attractive to the production of a pozzolan, which is similar to some by-products used as biomass such as sugar cane bagasse and rice husk [8]. Saruchera et al. [9]

found that elephant grass was suitable for ash production as all results of chemical composition meet the minimum requirement to be classified as a pozzolan. Nakanishi et al. [10] found that the pozzolanic behavior of elephant grass paste was similar to silica fume paste where the main hydrated phases in EGC/$Ca(OH)_2$ consisted of C-S-H gels similar to SF/$Ca(OH)_2$ paste.

Therefore, using Napier grass as an alternate binder resource may protect the environment and enable the building industry to grow sustainably. Several studies have already reported the potential activity of PPA and its effect on the mechanical properties of concrete or mortar. Cordeiro and Sales [8] found that the 28-day compressive strength, Young's Modulus, and water absorption did not change due to the replacement of 20% of the cement with EGA compared to the reference concrete. Cordeiro and Sales [11] found that the compressive strength increased after 180 days of curing compared to the reference concrete. However, Erika et al. [12] reported that the compressive strength decreased by 12% due to 20% cement replacement with EGA.

In the context of pozzolanic ash, Ramadhansyah et al. [13] reported that the use of finer particles of the ash produced better strength and affected the workability, creep, shrinkage, and water-cement ratio properties. As the fineness of the material relies on their grinding times, it was proven that the strength of concrete increased with the increase of the grinding time of the Rice Husk ash until the optimum value is reached [13]. Since the ashes from the pozzolanic material are expected to behave similarly, the grinding time of the PPA might also improve the mechanical properties of the mortar.

Therefore, the objective of this study is to investigate the effects of grinding times and dosages of PPA as a partial replacement for cement on the mechanical properties of the blended-cement pastes. The compressive strength, standard consistency, setting time, and water absorption rate of the cement blended paste were considered. The pozzolanic character of PPA was determined through the particle size analysis, scanning electron microscope (SEM), loss on ignition, and chemical composition.

2. ORGANIZING EXPERIMENTAL PROGRAM

2.1 Materials

The cement used in this study was Ordinary Portland Cement (OPC) conforming to EN 197-1: 2000 standard [14] and manufactured by Tasek Corporation Berhad (ISO 9001, ISO 14001 and OHSAS 18001 certified company). Physically, the Portland cement particle size of 63 μm and the fine aggregates that passed 2.36 mm sieve were used for cement mortar. *Pennisetum purpureum,* also known as Napier grass of Taiwan type, was used. The raw materials of *P. purpureum* grass were supplied from Hole Grass Farm Pagoh, Muar, Johor, Malaysia.

2.2 *Pennisetum purpureum* Ash (PPA) Production

The production of PPA included four steps: first, the *P. purpureum* grass was cut into small pieces, and washed with tap water to remove undesirable residue. Then, the grass was oven-dried at 110 °C for 24 hours to ensure the grass was completely dried. The oven-dried grass then burned using Lenton Atmospheric control furnace at 350 °C for the first 3 hours and 600 °C for another 3 hours to produce final ash with a heating rate of 10 °C/min. The ash was allowed to cool down inside the furnace. The PPA was sieved and ground in Fritsch ball mill for 1, 3, and 6 hours to produce ash with three different particle size ranges. The production process is shown in Fig. 1 and Fig. 2.

(a) Cut and washed

(b) 24 hours' oven-dried at 110 °C

Fig. 1: Pieces of *P. purpureum* plant.

(a) At Lenton Atmospheric control furnace

(b) After 6 hours of the burning process

(c) Ground PPA (final ashes)

Fig. 2: Burning process of PPA.

2.3 Characterizations and Microstructural Analysis of PPA

The particle size distributions of the PPA and OPC were determined using the laser particle size analyzer (Fritsch Analysette 22) as shown in Table 1. It can be observed that PPA particle size decreased or specific surface area increased with the increase in grinding duration. The average particle size (D_{50}) of the PPA was less than the cement particle size (18 µm) so that it can be classified as fillers as stated by Abu Bakar et al. [15]. Also, it can participate actively with the mixes through pozzolanic reaction as stated by Venkatanarayanan and Rangaraju [12].

Table 1: Granulometric value for D10, D50, and D90 of materials

Sample	Grinding duration (hours)	Specific surface area (cm^2/cm^3)	*D_{10} (µm)	*D_{50} (µm)	*D_{90} (µm)
1H-PPA	1	10949	3.41	10.58	20.58
3H-PPA	3	11008	3.11	10.25	21.03
6H-PPA	6	12671	2.55	9.30	21.97
OPC	-	3350	1.2	18	67

*D_{10}: particle size which lies below 10% of the material; D_{50}: particle size which lies below 50% of the material; D_{90}: particle size which lies below 90% of the material.

The chemical composition analysis of raw materials was carried out using X-ray fluorescence (XRF). Table 2 shows that PPA mostly consists of silica, iron, and alumina with small quantities of sulfate, magnesium, and calcium, etc. Despite different grinding durations, there are no significant variations in the chemical composition of the PPA.

Summations of the percentage of SiO_2, Al_2O_3, and Fe_2O_3 present in PPA were 37.98%, 37.02%, and 41.15% for 1H-PPA, 3H-PPA, and 6H-PPA, respectively. Thus, the PPA does not achieve a desirable amount of reactive oxides to be classified as Class F or Class C pozzolan where the total reactive oxides (SiO_2, Al_2O_3, and Fe_2O_3) must be above 50% and 70%, respectively.

Table 2: Chemical composition of PPA by weight percent

Oxide	1H-PPA (*wt.%)	3H-PPA (*wt.%)	6H-PPA (*wt.%)	OPC (*wt.%)
SiO_2	31.7	30.4	33.8	17.9
Al_2O_3	4.54	4.70	5.45	4.69
Fe_2O_3	1.74	1.92	1.81	2.97
SO_3	5.52	5.23	4.95	-
MgO	0.74	0.75	0.78	1.62
CaO	20.8	21.0	22.2	-
K_2O	1.62	1.87	1.62	0.51
TiO_2	0.25	0.23	0.27	-

*wt% means weight percent

The microstructural analysis for PPA was conducted using a JEOL scanning electron microscope (SEM) for all grinding periods as shown in Fig. 3. It was found that the unground PPA (UG-PPA) sample has an irregular shape with many sizes as shown in Fig. 3a. The UG-PPA grains have a porous structure that was quite similar to that found by [10,11]. The skeleton of a honeycomb-like particle and vesicular shape in the UG-PPA, whereas the particles are in highly irregular shapes and porous structures as shown in Fig. 3b. After being ground as shown in Fig. 3 c, d, and e, the particle sizes become smaller, which indicates a significant degradation of the PPA structure due to the grinding process. The particle shape of the ground PPA had better dispersion than UG-PPA which proved that the mechanical grinding improved the particle surface and morphology of PPA.

2.4 Mix Proportions

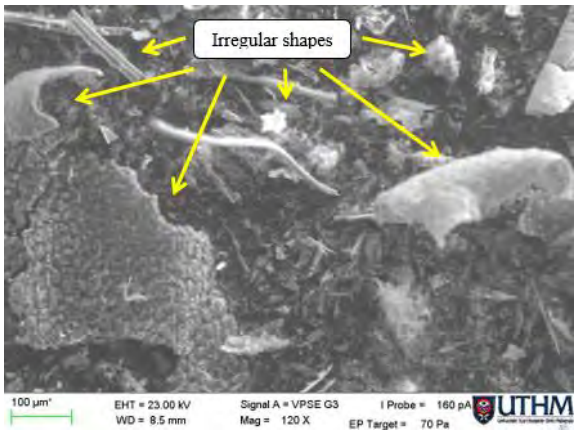
Five different percentages of PPA with cement by weight of 0%, 5%, 10%, 15%, and 20% were prepared. Three different particle sizes of PPA from three different grinding times (1, 3, and 6 hours) were considered in this study. The cement to sand ratio was 1:2.75 by weight while the water-cement ratio for all samples was 0.485. The mixed composition of the mortar cube was shown in Table 3. A total of 234 sample cubes was prepared in which 9 samples were prepared for each mixing batch.

2.5 Testing

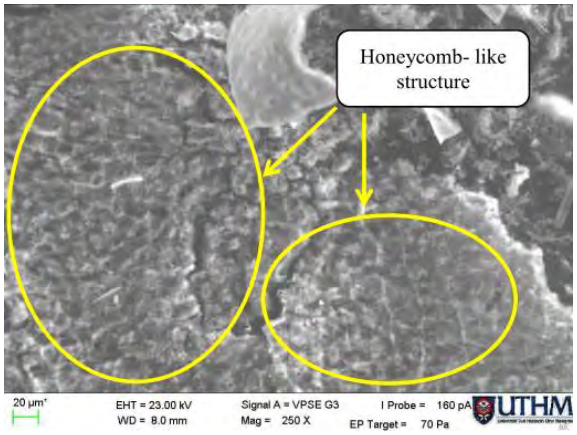
2.5.1 Loss on Ignition Test

The loss on ignition is defined as the percent mass loss of a constant weight of a sample as ignited at a high temperature. The loss on ignition test was conducted for both the OPC and PPA as per ASTM C114-07 (Part16) [16]. An empty platinum crucible was weighed and marked as W_1. About 1.5 grams of the sample were weighed in the platinum crucible and marked as W_2. Then, the platinum crucible was covered and ignited in a muffle furnace at a temperature of 950°C±50°C for 15 minutes. The sample was taken out and cooled for 20 minutes, then weighted and marked as W_3. The percentage of loss on ignition was calculated to the nearest 0.1 using Eq. 1.

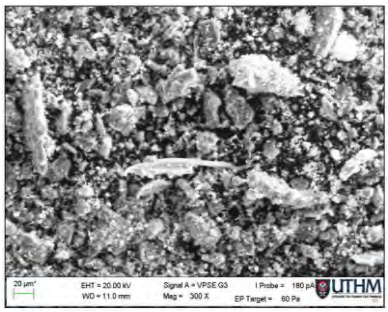
$$\text{Loss on ignition (\%)} = [(W_2-W_3) / (W_2-W_1)] \times 100\% \quad (1)$$



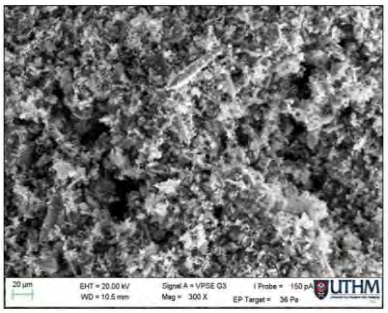
(a) UG -PPA contained irregular shapes



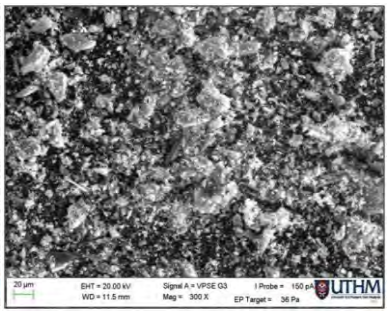
(b) UG-PPA contained a honeycomb-like structure



(c) 1H-PPA



(d) 3H-PPA



(e) 6H-PPA

Fig. 3: Microstructural analysis of *P. purpureum* ashes.

Table 3: The mix composition of mortar cube

Sample	Water-to-cement ratio	Water (ml)	Cement (g)	PPA (g)	Sand (g)
Control	0.485	359	740	-	2035
5%	0.485	359	703	37	2035
10%	0.485	359	666	74	2035
15%	0.485	359	629	111	2035

2.5.2 Standard Consistency and Setting Time Tests

The standard consistency test was applied to the Ordinary Portland Cement and blended cement paste with 5%, 10%, 15%, and 20% of replacement by PPA as per ASTM C187 [17]. The Vicat apparatus with a plunger of diameter 10 mm was used for the test. The amounts of water for normal consistencies of all samples were calculated using Eq. 2.

$$\text{Normal consistency (\%)} = \frac{M_W}{M_C} \times 100\% \quad (2)$$

where M_W is the mass of water (g) and M_C is the mass of cement (g).

On the other hand, the setting time test was conducted as per ASTM C191 [18] to determine the value of the Vicat initial and final time of setting of OPC and blended paste with 5%, 10%, 15%, and 20% cement replacement with PPA. A paste that was proportioned and mixed to normal consistency, as described in the test method ASTM C187 [17], was molded and placed in a moist cabinet and allowed to start setting. Periodic penetration tests were performed on this paste by the Vicat apparatus with a needle of diameter 1 mm.

The Vicat initial setting time was the time elapsed between the initial contact of cement and water and the time when the penetration is measured to be 25 mm. The Vicat final setting time was the time elapsed between initial contact of cement-water and the time when the needle does not leave a complete circular impression on the paste surface. The Vicat setting times were calculated to the nearest 1 minute according to Eq. 3.

$$\text{Setting time} = \left(\left(\frac{(H-E)}{(C-D)} \right) \times (C - 25) \right) + E \quad (3)$$

2.5.3 Compressive Testing

The samples were cast into cubical molds with a size of 50 mm × 50 mm × 50 mm. The top of the sample was smoothened and leveled off with a straight edge trowel. The samples were placed in a moist place immediately upon completion of molding. After 24 hours, the samples were removed from the molds and immersed in the water storage tanks until the testing day. Compressive strength tests were conducted as per ASTM C109 [19] for each replacement percentage and grinding time at the age of 3, 7, and 28 days curing age. The load rate with a range of 900 to 1800 N/second was applied to the specimen.

2.5.4 Water Absorption Testing Procedure

Mortar cube samples of size 50 mm × 50 mm × 50 mm were cast for water absorption testing. After 24±1 hour of curing, the mortar cube samples were cured in a moisture tight plastic bag at 24±8 °C until it reaches the age of 3, 7, and 28 days. Then, the mortar cube samples were dried in a ventilated oven at 110±5 °C for around 24 hours. The successive weights of the samples were checked at an interval of 2 hours until two successive weights showed an increment of loss not greater than 0.2%. The mortar cube samples were then removed from the oven and were cooled until the samples achieved ambient temperature. Then after 24 hours, the water absorption test was conducted by following the Average Absorption Standard Test Method for Rate of Water Absorption of Mortars ASTM C1403 [20]. The water absorption rate at 24 hours ± 15 minutes was calculated following Eq. 4, and recorded as "A_T". The A_T (Avg.) was calculated for each set of three cube samples.

$$AT = (W_T - W_0) \times 10000 / (L_1 \times L_2) \quad (4)$$

where WT (g) is the weight of cube sample at time T, W_0 (g) is the initial weight of cube sample, L_1 is the average length (mm) of test surface of cube sample, and L_2 (mm) is the average width of test surface of cube sample.

The percentage of water absorption A_T (%) also was calculated using Eq. 5.

$$A_T = [(W_T - W_0) / (W_0)] \times 100\% \quad (5)$$

3. RESULTS AND DISCUSSION

3.1 Loss on Ignition

The results of loss on ignition (LOI) for all PPA samples were shown in Table 4. Based on the results, the LOI of all PPA samples was below 6%, ranging from 5.27% to 5.47%. The lack of significant variations between samples means that the grinding durations do not affect the LOI of PPA. Meanwhile, the LOI for OPC was 2.25%, which was lower than PPA. However, Cordeiro and Sales [9] stated that it is important to emphasize that all ashes presented LOI lower than 5%, indicating that the burning procedures were appropriate to produce pozzolan from PPA in the context of carbon content. However, LOI lower than 5% could not be achieved in this study, possibly due to untreated PPA with hydrochloric acid

pretreatment. However, ASTM C618-3 [21] limits the LOI of pozzolanic materials to less than 6%, largely because higher LOI levels commonly result in discoloration, poor air entrainment, segregation, and low compressive strength of the mixed components. Note that this standard is specified principally for fly ashes. For other calcined natural pozzolans, the ASTM C618-3 [21] limits the LOI at 10%, since these materials are produced at lower burning temperatures (500 °C to 800 °C) and contain higher LOI levels.

Table 4. Loss on ignition of PPA

Sample	1H PPA	3H PPA	6H PPA	OPC
Weight of crucible, W_1 (g)	24.383	24.385	24.383	24.387
Initial Weight of crucible + PPA, W_2 (g)	25.883	25.885	26.183	25.887
Final Weight of crucible + PPA, W_3 (g)	25.804	25.803	26.087	25.853
Loss on ignition (%)	5.27	5.47	5.33	2.25

Legend: 1H-PPA: 1-hour ground PPA; 3H-PPA: 3-hour ground PPA; 6H-PPA: 6-hour ground PPA.

3.2 Standard Consistency of Cement Paste Containing PPA

According to ASTM C191 [18], the standard consistency test is related to the quantity of water that should be able to flow in the fresh mixture. The plasticity of cement capacity to flow will increase if excessive water is added, while the cement will be stiff with a lower quantity of water. Fig. 4 shows the standard consistency results of cement paste containing PPA compared to the OPC paste. From the results, the standard consistency of OPC was 38%. Meanwhile, the standard consistency of cement paste containing PPA increased with the increase of replacement level percentage of PPA.

The standard consistency of the samples with 5%, 10%, 15%, and 20% 1H-PPA were 39%, 41%, 50%, and 53%, respectively. For the cement paste containing 5%, 10%, 15%, and 20% of 3H-PPA, the standard consistency was 40%, 53%, and 57%, respectively, which does not differ much from the standard consistency of 1H-PPA. The same pattern applies to cement paste containing 5% to 20% of 6H-PPA with normal consistency ranges from 42% to 57%. All of the above results show that the cement pastes containing PPA have higher standard consistency than OPC due to the properties of PPA that absorbed the water during the early hydration process of cement and PPA mixtures. The standard consistency of cement paste containing PPA also increased with the increase of PPA sizes. As the specific surface area of PPA increased, the water demand would be increased [22]. However, the standard consistency of cement paste containing 15% and 20% of 3H-PPA was the same as 15% and 20% of 6H-PPA. The increased water demand for the PPA paste to achieve the same level of consistency may be due to the ability of the PPA particles to quickly absorb the mixed water from the paste due to their cellular microstructure. At any constant replacement level and particle size of PPA, the standard consistency of the PPA paste was higher than the control OPC paste. This is because the rate of water absorption of the PPA particles is higher than that of the OPC particles due to the presence of larger pore sizes and wider pore size distribution of the former. Venkatanarayanan and Rangaraju [23] proved that the grinding of PPA may not significantly reduce the absorption of water by its grains and hence, affected the workability characteristics of concrete and mortars. The OPC paste registered lower standard consistency probably due to the lower specific surface area and non-cellular shape of the OPC particles compared to the ground PPA paste.

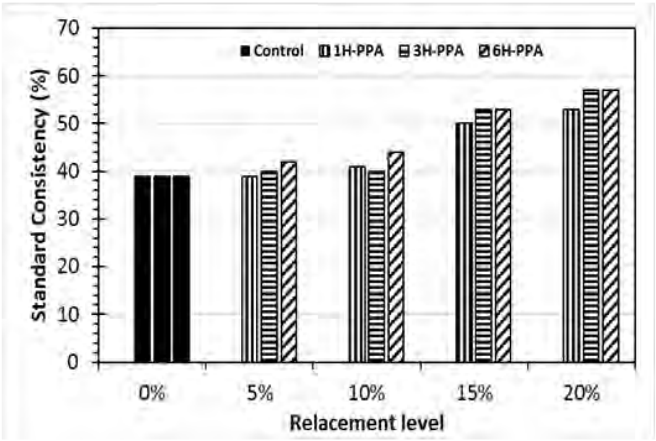


Fig. 4: Standard consistency of cement paste containing PPA at different replacement rate and control paste.

3.3 Setting Time of Cement Paste Containing PPA

The results of the setting time of OPC paste and paste containing PPA were shown in Fig. 5. The initial setting time (I-ST) for the samples containing PPA started from 226 minutes for 5% 1H-PPA to the longest I-ST of 298 minutes for 20% 6H-PPA while the I-ST of the cement paste was 269 minutes.

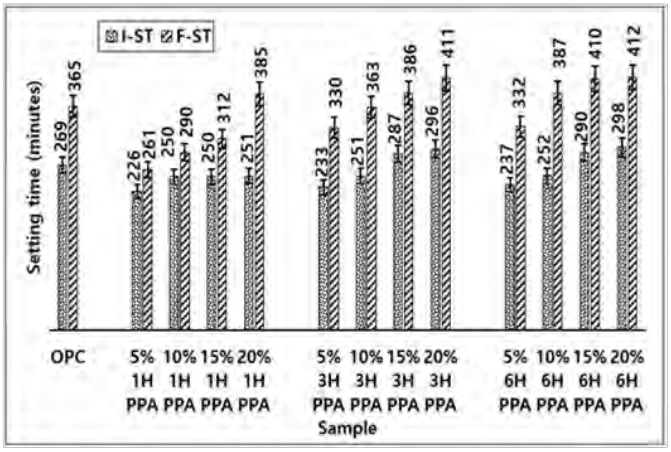


Fig. 5: Setting time of cement paste containing PPA at different replacement rates and control paste.

The final setting time (F-ST) started from 261 minutes for 5% 1H-PPA to the longest F-ST of 412 minutes for 20% 6H-PPA while the F-ST of cement paste was 365 minutes. Based on the results, it was found that the mortar containing PPA achieved lower I-ST and F-ST than the control mortar except for those containing 15% and 20% PPA. However, both I-ST and F-ST increased with the increase of cement replacement percentage with PPA. Also, both I-ST and F-ST slightly increased with the increase of grinding time in which 3H-PPA and 6H-PPA have longer initial and final setting times than 1H-PPA pastes. The setting time behavior is a function of the rate of stiffening of the paste influenced by the effective water to cement ratio [23]. The effective water to cement ratio is the remaining water available to react with Portland cement after the water was absorbed by the pozzolan grains. At any percentage of replacement of PPA, the lower grinding time of PPA was shown to absorb more water due to greater porosity resulting in lower effective water to cement ratio in pastes. Thus, the final setting time would be shorter for the specimens containing larger

particle sizes of PPA due to the reduction of porosity with continued hydration, hence affecting their stiffening rate of the paste.

3.4 Water Absorption Rate of Mortar Containing PPA

The effects of PPA replacement with cement at different percentages and particle sizes on the water absorption rate of cement mortar mixes are shown in Fig. 6. The results show that all samples containing PPA initially exhibited a high-water absorption rate for 3 days. The water absorption rate significantly slowed down for 7 days and 28 days as most of the pores available had become saturated. It is found that the absorption rate of the mortar containing PPA was higher than the control samples at all testing ages and replacement levels. However, 15% 6H-PPA samples at 7 days and 28 days displayed a lower water absorption rate than the control sample. It was observed that the water absorption rate increased as the replacement level of 1H-PPA and 3H-PPA increase, while it remains unchanged for 6H-PPA.

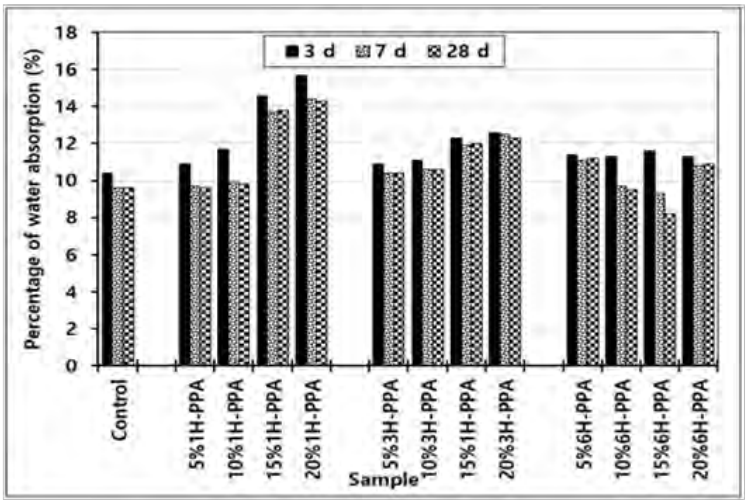


Fig. 6: Percentage of water absorption of 1H-PPA mortar after 24 hours of immersion.

Based on Fig. 6, the finer samples (3H-PPA and 6Hh-PPA) exhibited a lower water absorption rate than 1H-PPA samples. This is due to the filler effect, in which the voids between the cement grains are filled with a micro size of silica particles available in pozzolan to enhance its structural properties. A higher packing density resulted in lower water demand of the mixture with the right composition of pozzolans, which also contributed to the strength enhancement due to the reduced capillary porosity [24]. The higher the percentage of PPA replacement, the denser the particle packing in cement mortar, which also applies to the fineness of PPA. Also, the incorporation of suitable fine pozzolan refined the pore structure and reduces the porosity of cement-based materials due to filler and nucleation effects [25].

3.5 Compressive Strength with Activity Index of Mortar Containing PPA

The results of the compressive strength of cement mortar containing PPA for all percentages based on grinding durations at the testing age of 3 days, 7 days, and 28 days were illustrated in Fig. 7. Based on the plot, the strength of mortar increased linearly with the curing ages for all samples.

The results also displayed that the compressive strength of the mortar sample increased with the increase of cement replacement of up to 15% with PPA in the mortar mix, but then the strength mostly decreased after 15% of PPA replacement. However, the strength

obtained was still higher than the control mortar. The filler supplemented the dense-pack structure in cement mortar and the filler dispersed very well. Filler effect or particle size effect is defined as the proper arrangement of small particles that fill the voids and contribute to the increment of compressive strength without any chemical reaction [26]. The compressive strength of cement mortar may be increased due to pore refinement from the pozzolanic reaction of PPA with the cement particles. Furthermore, finer PPA could react more quickly with $Ca(OH)_2$ and increase the compressive strength of mortar due to the pozzolanic reaction.

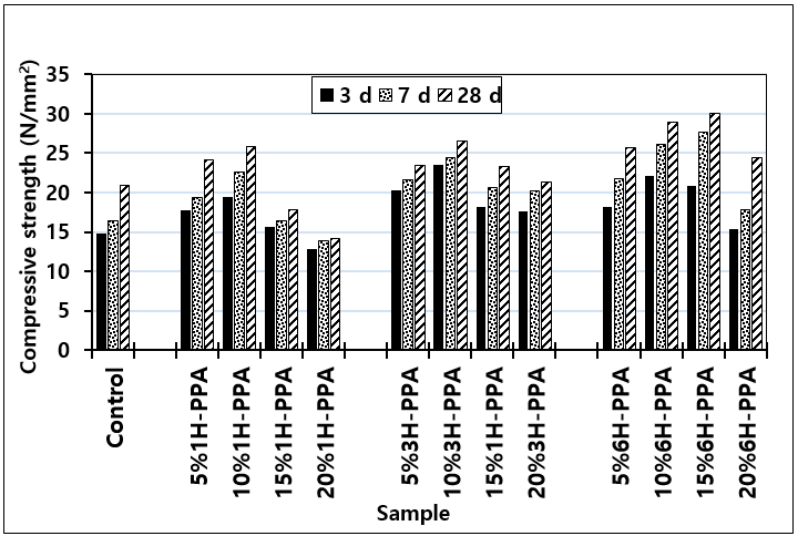


Fig. 7: Compressive strength of mortar containing PPA at 3 days, 7days, and 28 days.

Overall, the compressive strength of cement mortar decreases with increased PPA dosage starting from 10% of replacement. It is observed that the optimum compressive strength is approximately 10% of PPA replacement. However, the mortar containing 15% of 6H-PPA at 28 days has the highest strength among all samples in which it achieved a strength of 30 MPa. In the context of grinding time, the 3 hours grinding time achieved the highest compressive strength at 3 days followed by 6 hours, and 1 hour, respectively. However, the compressive strength at 7 days and 28 days increased with the increase of the grinding time of the PPA in which the 6-hours grinding time achieved the highest compressive strength followed by 3 hours and 1 hour, respectively. Therefore, the 6-hours grinding was found to be the optimum grinding time for the PPA as a replacement of cement for the mortar.

The compressive strength activity indices for all mortars containing PPA are listed in Table 5. At each testing time, the strength index is defined as the ratio of the compressive strength of the mortar containing PPA to the strength of the control sample. It is mostly depending on the hydration rate. The minimum requirement of 75% of the strength activity index was achieved by mortar containing ash as specified in ASTM C618 [21]. Based on Table 5, all strength activity indices at the ages of 3 days, 7 days, and 28 days were higher than the minimum requirement of 75% as stated in ASTM C618 [21]. It was found that the strength activity index decreased compared to the control sample due to replacing 10%, 15%, and 20% 1H-PPA at the age of 28 days, as well as 20% 1H-PPA at the ages of 3 days and 7 days. This is due to the mitigation influence and delay of the pozzolanic reaction of PPA/$Ca(OH)_2$. Altwair et al. [27] found the same result for a 20% replacement of OPC with POFA at the age of 3 and 7 days. However, the samples with a high fineness (3H-PPA and

6H-PPA) achieved strength indices higher than the control samples for all replacement percentages. This is due to the amorphous aluminous and siliceous minerals that still actively react with $Ca(OH)_2$, producing C-S-H and hydrated calcium aluminates, improving interfacial bonding between the sand and pastes. The outcome of this study is in agreement with Altwair et al. [27] in which the POFA cement mortar recorded a strength activity index of 101.72 %.

Table 5: Strength activity index (%)

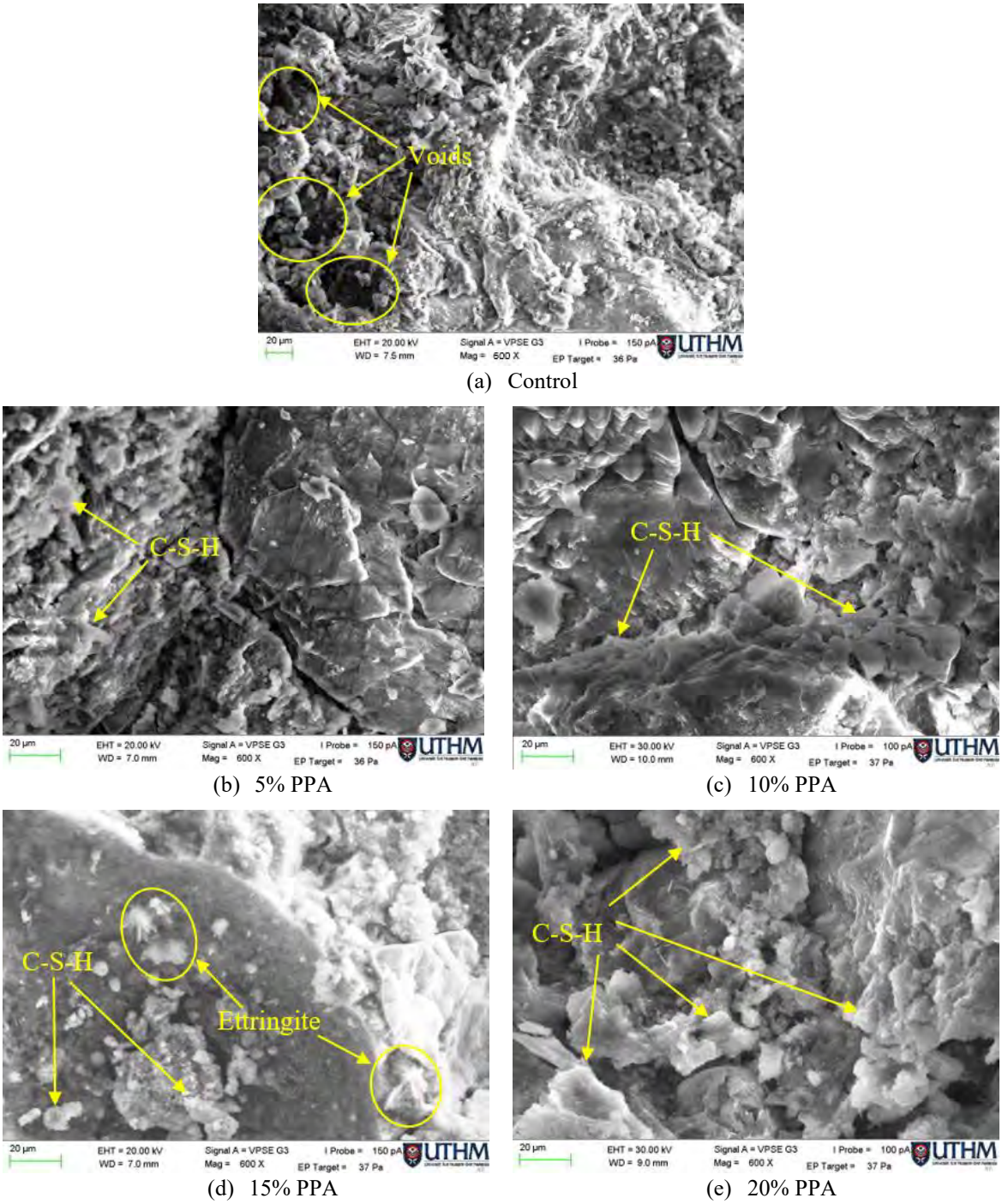
Sample	Testing age (days)		
	3	**7**	**28**
Control	100	100	100
5%1H-PPA	120	119	115
10%1H-PPA	131	138	81
15%1H-PPA	106	100	76
20%1H-PPA	87	84	78
5%3H-PPA	150	132	91
10%3H-PPA	159	149	127
15%3H-PPA	123	127	112
20%3H-PPA	118	124	97
5%6H-PPA	123	133	123
10%6H-PPA	149	160	138
15%6H-PPA	141	169	144
20%6H-PPA	104	108	117

Furthermore, the impact of replacing cement with PPA in the mortar was examined using the microstructure images analysis through SEM at 28 days as shown in Fig. 8. The microstructure of mortar containing 15% of PPA has hydrated products consisting of ettringite and C-S-H particle as shown in Fig. 8d. Ettringite formed in needle-like or club shape crystal, while C-S-H was in fine layer particle, both found in a mortar containing 15% PPA at 28 days curing period but did not exist in the control mortar. This is illustrated the highest strength for 28 days curing period of 15% of PPA.

4. CONCLUSION

Based on the results, the following conclusions can be drawn:

- The characterization of PPA material based on chemical composition, particle size distribution, and X-ray diffraction showed that the PPA material is adequate to produce pozzolan; however, it did not achieve 50% and 70% of reactive oxides to be classified as Class F or Class C pozzolan, respectively. Therefore, the PPA is more suitable as filling material than the pozzolanic material as a substitute for OPC, which helps in reducing the environmental pollution during the disposal of excess elephant grass wastes. Although the particle shape of the ground PPA had better dispersion than UG-PPA, the grinding durations do not affect the chemical composition and LOI of PPA.



(a) Control

(b) 5% PPA

(c) 10% PPA

(d) 15% PPA

(e) 20% PPA

Fig. 8: Scanning Electron Micrographs of mortar mixtures at 28 days.

- The mortar containing PPA achieved a higher standard consistency value than the control sample while it gained a lower setting time except for those containing 15% and 20% PPA. However, the standard consistency and setting time of cement paste containing PPA increased with the increase of PPA grinding time.
- The compressive strength of cement mortar increased with the increase of PPA until 10% for all grinding times, then it was decreased with the increase of PPA percentage.
- The mortar containing PPA achieved a water absorption rate higher than the control samples at all testing ages and replacement levels. The water absorption rate

increased as the replacement level of 1H-PPA and 3H-PPA increased, while it remained unchanged for 6H-PPA.

- The 6-hours grinding was found to be the optimum grinding time for the PPA as a replacement of cement for the mortar as it achieved the highest compressive strength at 7 and 28 days.
- The 15% of 6H-PPA at 28 days was found to be the optimum PPA replacement dosage and grinding time with cement as it achieved the highest strength and lower water absorption rate among all samples at 7 and 28 days.
- Finally, based on the results of chemical analysis and strength activity index, the characterization displayed the potential pozzolan of PPA, although it is not Class F or Class C pozzolan. Therefore, the PPA can be considered as filler and source of pozzolan production in tropical regions.

ACKNOWLEDGEMENT

The research described in this paper was financially supported by Research Management Centre (RMC), Universiti Tun Hussein Onn Malaysia (UTHM), under grant GPPS Vot No. U45.

REFERENCES

[1] Ghiasvand E, Ramezanianpour A. (2017) Effect of grinding method and particle size distribution on long term properties of binary and ternary cements. Construction and Building Materials, 134: 75-82. https://doi.org/10.1016/j.conbuildmat.2016.12.122

[2] Cordeiro GC, Kurtis KE. (2017) Effect of mechanical processing on sugar cane bagasse ash pozzolanicity. Cement and Concrete Research, 97: 41-49. https://doi.org/10.1016/j.cemconres.2017.03.008

[3] Ambedkar B, Alex J, Dhanalakshmi J. (2017) Enhancement of mechanical properties and durability of the cement concrete by RHA as cement replacement: Experiments and modeling. Construction and Building Materials, 148: 167-175. https://doi.org/10.1016/j.conbuildmat.2017.05.022

[4] Thomas BS, Kumar S, Arel HS. (2017) Sustainable concrete containing palm oil fuel ash as a supplementary cementitious material–A review. Renewable and Sustainable Energy Reviews, 80: 550-561. https://doi.org/10.1016/j.rser.2017.05.128

[5] Adesanya DA, Raheem AA. (2009) Development of corn cob ash blended cement. Construction and Building Materials. 23(1): 347-352. https://doi.org/10.1016/j.conbuildmat.2007.11.013

[6] Mohammed I, Abakr Y, Kabir F, Yusup S. (2015) Effect of aqueous pretreatment on pyrolysis characteristics of napier grass. Journal of Engineering Science and Technology, 10(11): 1487-1496. http://jestec.taylors.edu.my/Vol%2010%20Issue%2011

[7] Ghani WWAK, Salleh MAM, Adam SN, Shafri HM, Shaharum SN, Lim KL, Tapia JF. (2019) Sustainable bio-economy that delivers the environment-food-energy-water nexus objectives: the current status in Malaysia. Food and Bioproducts Processing, 118: 167-186. https://doi.org/10.1016/j.fbp.2019.09.002

[8] Cordeiro GC, Sales CP. (2015) Pozzolanic activity of elephant grass ash and its influence on the mechanical properties of concrete. Cement and Concrete Composites, 55: 331-336. https://doi.org/10.1016/j.cemconcomp.2014.09.019

[9] Saruchera L, Mupa M, Muchanyereyi-Mukuratirwa N, Witzleben S. (2018) A study on the pozzolanic properties of *Pennisetum Purpureum* ash. Journal of Interdisciplinary Academic Research, 1: 1-10. https://doi.org/10.32476/b9691748-2ac0-4d2fab85-6660b03aeac5

[10] Nakanishi EY, Frías M, Martínez-Ramírez S, Santos SF, Rodrigues MS, Rodríguez O, Savastano Jr H. (2014) Characterization and properties of elephant grass ashes as supplementary cementing material in pozzolan/$Ca(OH)_2$ pastes. Construction and Building Materials, 73: 391-398. https://doi.org/10.1016/j.conbuildmat.2014.09.078
[11] Cordeiro GC, Sales CP. (2016) Influence of calcining temperature on the pozzolanic characteristics of elephant grass ash. Cement and Concrete Composites, 73: 98-104. https://doi.org/10.1016/j.cemconcomp.2016.07.008
[12] Nakanishi EY, Frías M, Santos SF, Rodrigues MS, de la Villa RV, Rodriguez O, Junior HS. (2016) Investigating the possible usage of elephant grass ash to manufacture the eco-friendly binary cements. Journal of Cleaner Production, 116: 236-243. https://doi.org/10.1016/j.jclepro.2015.12.113
[13] Ramadhansyah PJ, Bakar BA, Azmi MM, Ibrahim MW. (2011) Engineering properties of normal concrete grade 40 containing Rice husk ash at different grinding times. International Journal of Technology, 2(1): 10-19.
[14] EN 197-1 (2000) Cement: Part 1. Compositions and conformity criteria for common cements.
[15] Abu Bakar BH, Ramadhansyah PJ, Megat Azmi MJ. (2011) Effect of rice husk ash fineness on the chemical and physical properties of concrete. Magazine of Concrete Research, 63(5): 313-320. https://doi.org/10.1680/macr.10.00019
[16] ASTM C114 (2007) Standard Test Methods for Chemical Analysis of Hydraulic Cement. ASTM International. West Conshohocken, PA.
[17] ASTM Designation C187-04 (2004) Standard test method for normal consistency of hydraulic cement. West Conshohocken, PA.
[18] ASTM C191-08 (2008) Standard Test Method for Time of Setting of Hydraulic Cement by Vicat Needle, West Conshohocken, PA, USA.
[19] ASTM C109-08 (2008) Standard test method for compressive strength of hydraulic cement mortars. ASTM International, West Conshohocken, PA.
[20] ASTM C1403-14 (2014) Standard Test Method for Rate of Water Absorption of Masonry Mortars, ASTM International, West Conshohocken, PA.
[21] ASTM C618 (2008) Standard specification for coal fly ash and raw or calcined natural pozzolan for use as a mineral admixture in concrete Cement; Lime; Gypsum, Annual Book of ASTM Standards, vol. 04.01, American Society for Testing and Materials, Philadelphia, PA.
[22] Sathawane SH, Vairagade VS, Kene KS. (2013) Combine effect of rice husk ash and fly ash on concrete by 30% cement replacement. Procedia Engineering, 51: 35-44. https://doi.org/10.1016/j.proeng.2013.01.009
[23] Venkatanarayanan HK, Rangaraju PR. (2015) Effect of grinding of low-carbon rice husk ash on the microstructure and performance properties of blended cement concrete. Cement and concrete composites, 55: 348-363. https://doi.org/10.1016/j.cemconcomp.2014.09.021
[24] Rai S, Tiwari S. (2018) Nano silica in cement hydration. Materials Today: Proceedings, India, 5(3): 9196-9202. https://doi.org/10.1016/j.matpr.2017.10.044
[25] Wang D, Shi C, Farzadnia N, Shi Z, Jia H, Ou Z. (2018) A review on use of limestone powder in cement-based materials: Mechanism, hydration and microstructures. Construction and Building Materials, 181: 659-672. https://doi.org/10.1016/j.conbuildmat.2018.06.075
[26] Jaturapitakkul C, Tangpagasit J, Songmue S, Kiattikomol K. (2011) Filler effect and pozzolanic reaction of ground palm oil fuel ash. Construction and Building Materials, 25(11): 4287-4293. https://doi.org/10.1016/j.conbuildmat.2011.04.073
[27] Altwair NM, Johari MAM, Hashim SFS. (2011) Strength activity index and microstructural characteristics of treated palm oil fuel ash. International Journal of Civil & Environmental Engineering, 11: 100-107

https://doi.org/10.31436/iiumej.v22i2.1464

WSN-BASED MONITORING SYSTEMS FOR THE SOLAR POWER STATIONS OF TELECOMMUNICATION DEVICES

KHUJAMATOV KHALIMJON ERGASHEVICH[1], KHASANOV DOSTON TURAYEVICH[1*], FAYZULLAEV BAYRAM ARTIKBAYEVICH[2] AND REYPNAZAROV ERNAZAR NURJAMIYEVICH[1]

[1]Department of "Data Communication Networks and Systems" of TUIT named after Muhammad al-Khwarizmi, 100200 Tashkent, Uzbekistan.
[2]Department of "Telecommunication engineering" of Nukus Branch of TUIT named after Muhammad al-Khwarizmi, Nukus, Karakalpakstan

[*]Corresponding author: dhasanov0992@gmail.com

(Received: 27th April 2020; Accepted: 11th November 2020; Published on-line: 4th July 2021)

ABSTRACT: Renewable energy sources are an increasingly popular way to generate electrical energy for telecommunications systems. The use and effectiveness of remote monitoring systems has led to growing interest relative to telecommunications systems. As with any new technology, new problems have arisen in using renewable energy sources in telecommunications systems and the use of IoT-based remote monitoring systems in them. In particular, one of these problems is the decentralization of remote monitoring systems for renewable power plants for telecommunication systems. The paper details the stages of development and modelling of open-source centralized monitoring systems for solar power stations for telecommunication systems. In this paper, a real-time remote monitoring system for solar power sources was modelled and investigated by wireless sensor networks for telecommunications devices. Proteus software environment was obtained for modelling. Before modelling the system, a system structure and a block diagram were developed, where each of the elements that are part of the system are described. Software for the system was created. The designed structure and software were tested by modelling. The results of the modelling were presented in a virtual terminal, an oscillography, and a local web browser.

ABSTRAK: Sumber tenaga boleh baharu semakin meningkat popular dalam menghasilkan tenaga elektrik bagi sistem telekomunikasi. Penggunaan dan keberkesanan sistem pemonitoran jarak jauh dalam jaringan ini telah membawa kepada peningkatan minat kepada tenaga boleh baharu. Walau bagaimanapun, melalui sistem teknologi baru ini, masalah baru timbul dalam penggunaan punca tenaga boleh baharu dalam sistem telekomunikasi dan penggunaan bersama sistem pemonitoran berdasarkan IoT. Khususnya dalam desentralisasi sistem pemonitoran jarak jauh bagi loji kuasa sistem telekomunikasi. Kajian ini merincikan peringkat pembinaan dan model sumber terbuka berpusat sistem pemonitoran bagi sistem janakuasa solar sistem telekomunikasi. Kajian ini turut memodelkan sistem pemonitoran jarak jauh secara langsung bagi sumber kuasa solar dan dikaji dengan rangkaian sensor tanpa jaringan bagi alatan telekomunikasi. Perisian Proteus telah dihasilkan bagi model ini. Sebelum model sistem dibina, struktur sistem dan gambar rajah blok dibina. Setiap unsur dalam sistem ini diperihalkan. Perisian bagi sistem ini turut dibina. Struktur yang dibina dan perisian ini diuji melalui model. Hasil dapatan model dibentangkan dalam pangkalan maya, oscilograf dan pelayar web tempatan.

KEYWORDS: *real-time monitoring system; solar power stations; IoT; WSN; ZigBee; Proteus software*

1. INTRODUCTION

Presently, Smart Grids - smart energy systems based on IoT - are used to provide a stable and high-quality power supply to telecommunication facilities. This leads to solving many problems, including remote monitoring, two-way information systems, energy consumption controls, increasing energy demands, and providing reliability and security [1, 2].

Real-time remote monitoring of solar power sources for Telecommunication devices through IoT based on wireless sensor networks are the basis of Smart Grid systems [1], [2]. Real-time remote monitoring of these solar power sources requires remote monitoring of solar panel voltage, current, light illumination, dusting, environment and battery temperature, battery current, battery voltage, as well as other types of measurements of telecommunication devices.

In traditional remote monitoring of solar power sources based on IoT, the controller devices unit (CDU) receives relevant data from each sensor (how it is used in real-time, in hours, in days, in weeks, etc.) [3]. The resulting data is transferred to the database (DB) and will be provided to personnel by mobile communication via the Arduino Ethernet module or Arduino GSM module, and viewed as SMS, web pages, and other forms (Fig.1) [4], [1].

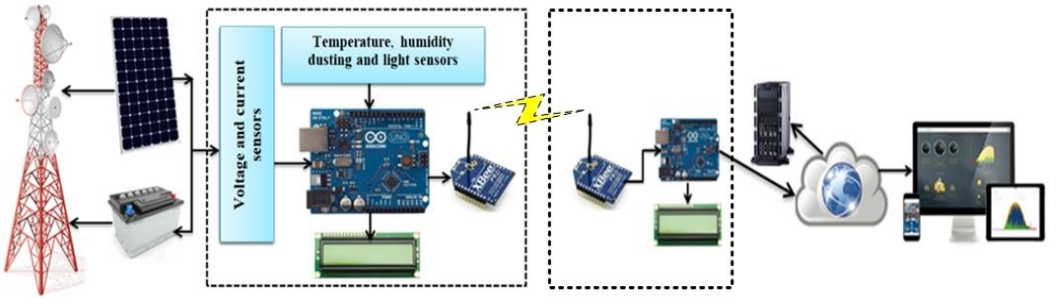


Fig. 1: Scheme of construction of monitoring system of the solar power station for telecommunication devices.

The collected data about the operating mode in the monitoring will be sent to the monitoring center, where the responsible personnel can promptly make a decision to correct the situation. Operational and maintenance personnel will be able to quickly identify and fix the cause of the malfunction based on detailed information from a set of technical tools installed in the monitoring system [2]. In this regard, by reducing the cost of recovery and downtime during preventive maintenance, the use of monitoring systems will increase the reliability of telecommunications systems, which can be evaluated with the coefficient of technical maintenance.

A positive economic effect is achieved with the use of monitoring systems, many actions aimed at checking the technical condition, at maintaining the system's performance, and at identifying the causes of malfunctions are performed remotely automated or automatically [5-6].

2. LITERATURE REVIEW

Several leading scientists have researched remote monitoring of energy supply sources, including remote monitoring of solar power plants, and the use of wireless sensor networks in remote monitoring systems. We have analyzed some of these studies that are relevant to the research topic.

In 2012, Chunming et al. [4] described the monitoring system for wind power generation based on a wireless sensor network. Aimed at meeting the demands of a wind power generation monitoring system, this thesis describes a system that builds up a wireless, real-time, multi-object, distant monitoring system for wind power generation.

In 2014, Nagalakshmi et al. [7] described the design and development of a remote monitoring and maintenance of a solar plant supervisory system. In this paper, authors created a prototype of the solar PV monitoring and optimization included a data acquisition system, control station, and supervisory monitoring at the plant level and Decision Support System (DSS) at the Central Management Station.

In 2015, Prakash et al. [8] described wireless wind turbine monitoring using Arduino. This paper described the implementation and design of a wireless monitoring system for turbines using the Arduino platform. There they employed an Arduino Uno-based WSN that measured and transmitted the various mechanical and electrical parameters of the turbine.

In 2015, Parikh et al. [9] described a solar panel condition monitoring system based on a wireless sensor network. They developed hardware and software discovery for fault detection and continuous monitoring system for solar panels in the distant areas.

In 2017, Kabalci et al. [3] presented the design and implementation studies of a remote monitoring system. The proposed modelling was improved in terms of smart measuring infrastructure based on voltage, current, and power metering.

In 2017, Jumaat et al. [10] described solar energy measurement using Arduino Uno. This project aimed to develop a measurement of solar energy using Arduino Uno board. In this paper, four parameters were measured: light intensity, temperature, current, and voltage.

In 2018, Madhubala et al. [11] described solar power based remote monitoring and control of industrial parameters using IoT. The main aim of this project was to develop a system for remote control and monitoring of important industrial parameters such as measurement of speed, gas, pressure, temperature, voltage, and current for prevention of human fatalities, diagnosis, automated range based access and two-way control, and extension of life of machinery using green technology in a friendly and secure manner using biometrics and the Internet.

In 2018, Parveen et al. [12] described an IoT-based solar tracking system for efficient power generation. The main goal of this paper was to design a very obvious tracker for solar panels and transmit the information through IoT. In this model, sensors of the sun's position was accomplished in two stages, primary and secondary.

In 2019, Vijayaragavan et al. [13] proposed a windmill monitoring system using IoT and WIFI. This paper described considerations of making use of WSN to scrutinize status within wind farms, taking into consideration the exact queries.

Finally, in 2019, Parveen et al. [14] described the simulation, fabrication, and monitoring of solar power plants using Arm7 and Proteus Software. It proposed an

autonomous intellectual system. LDR and rain sensors were used to prevent the excess errors by solar cells during inappropriate situations. In this project, the modeling had to be done all the way to maintain high efficiency irrespective of environmental conditions.

However, this research work did not investigate the complex monitoring system of solar power plants of telecommunications facilities. Therefore, this research was implemented based on the energy consumption demands of telecommunications facilities.

3. STATEMENT OF THE PROBLEM

The ability to remotely monitor the production capacity of solar power stations and the status of batteries is important in the long-term maintenance of such power systems [15]. Various means of communication and IoT are used in remote monitoring systems of optional power sources [16-17]. Such means of communication are selected based on the principle of construction of the monitoring system.

For now, remote monitoring systems of existing solar energy sources is not an optimal solution for real-time monitoring of energy sources of telecommunications facilities. Existing systems have a number of technical limitations [8]. The main reasons for this are:

- The use of optical fiber technologies in remote monitoring systems is not always desirable, as telecommunications facilities are often located far away from the monitoring center, in deserts, mountains, forests, and plains.
- The monitoring system of energy sources of telecommunications facilities must promptly transmit data in real-time, implementation of a preventive system aimed at troubleshooting and diagnostics. This requires a secure and reliable connection to avoid serious failures and interruptions.
- It is important to keep the monitored parameters on the local monitoring center's server (database). This data collection allows us to study and predict failures and their timelines.

This work is focused on developing an optimal system for real-time remote monitoring of solar power sources of telecommunication facilities through wireless sensor networks. It aims at designing and simulating the system on the Proteus environment, with a focus to the technical constraints arising from the above.

4. MAIN PART

When developing a Solar Power Plant Remote Monitoring System, it is important to take into account all the parameters that apply to it. The main parameters for a solar power plant can be set as follows: solar panel power; the amount of solar radiation; environment temperature; environment humidity; voltage and current at controller output; controller temperature; voltage and current on the battery; battery temperature; accumulator charge; voltage and current at load; and required power of the load. They define the basic characteristics of the system [7], [9], [11-12]. The development of a monitoring system of all parameters for a solar power plant was initially designed as a block diagram.

4.1 Structure of Monitoring System

The solar power plant monitoring system has a complex construction. Figure 2 shows the overall structure of the monitoring system. It's built on a wireless sensor network (WSN) technology or ZigBee (IEEE Std 802.15.4) technology.

ZigBee Networks consist of three types of devices (base stations): ZigBee Coordinator (Node Coordinator), ZigBee Router (Router Node), ZigBee End Device (End Node). ZigBee coordinator sets up, to forms and is a wireless network control center [18-19].

ZigBee router expands network coverage built on this technology. It uses dynamic routing protocols to avoid various problems and obstacles (such as distance). It restores the route direction when a network load increases or when a device fails. ZigBee uses two types of routing protocols: the distance vector (AODV) and the link state (Cluster-Tree) [18].

The ZigBee End device can receive or transmit the message, but it does not have the ability to retransmit or route packets. The End device can be connected to a router or coordinator and does not support other devices.

The structural scheme of the system (Fig. 2) consists of a solar power plant, sensors, three types of ZigBee devices, a monitoring center, and Internet network interfaces. Solar power plant parameters are measured by sensors. One CDU and one ZigBee End device (End Node) are installed for each sensor that is used at the solar power plant. CDU detects sensor values and transmits them to the ZigBee End device as a message. The ZigBee End device transmits the message in packet form to the ZigBee router (Router Node). The ZigBee router detects the route direction of packets to the ZigBee coordinator (Coordinator Node) and transmits them on the optimum path. Each of the four ZigBee routers can transmit received packets to the ZigBee coordinator through the remaining routers, or one of the optional routers can transmit the packets directly to the coordinator [20]. The remaining routers can be intermediate routers. As this system consists of four routers, it can have between one, two, or three intermediate routes.

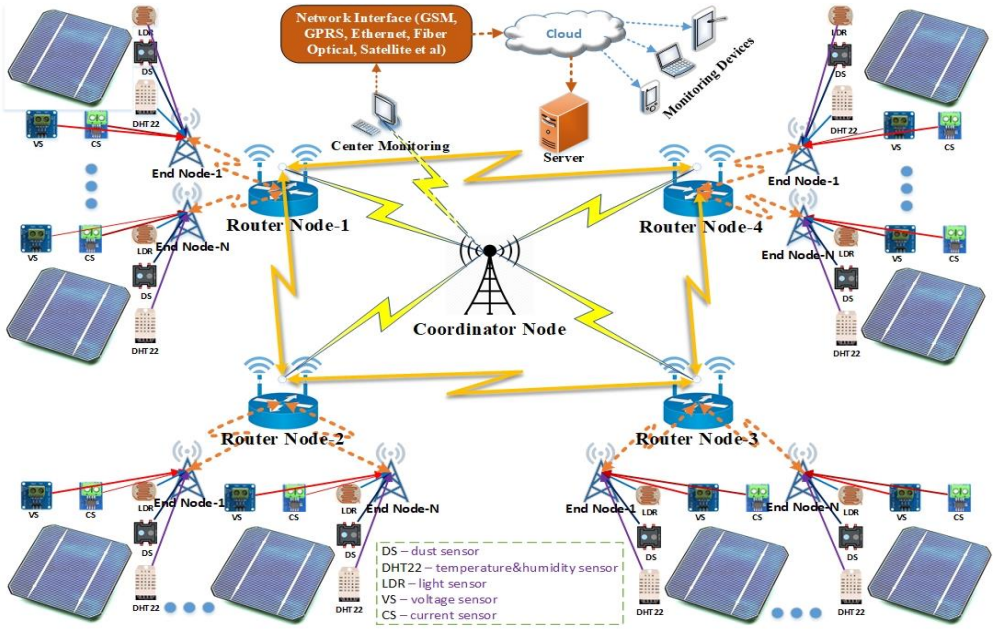


Fig. 2: Structure scheme of monitoring system based WSN.

The main task of the ZigBee coordinator in the system is to control the wireless network built on the sensors and transfer the data that are transmitting from the sensors to the monitoring center. In the monitoring center, the values measured by the sensors are controlled by administrator. At the same time, their information is transmitted to the server through Internet network interfaces (GSM, GPRS, Ethernet, Fiber optical network,

Satellite et al.) [21]. The server saves hourly, daily, weekly, monthly, and yearly reports of values measured by sensors. This information can be tracked with authorization, using any devices that have access to the Internet network.

4.1 Block Diagram of the Monitoring System of Solar Station

The remote monitoring system of solar power sources through wireless sensor networks of the telecommunication facilities was conditionally divided into two parts: the data collection unit from the solar power station and the monitoring center (Fig. 3). The data collection unit from the solar power plant consists of the following blocks (Fig. 3a):

- solar panel;
- accumulator battery;
- voltage sensor;
- current sensor;
- humidity sensor;
- temperature sensor;
- light sensor;
- dust sensor;
- controller devices unit (Arduino UNO);
- wireless network module (ZigBee/Xbee);
- display;
- telecommunication objects (load).

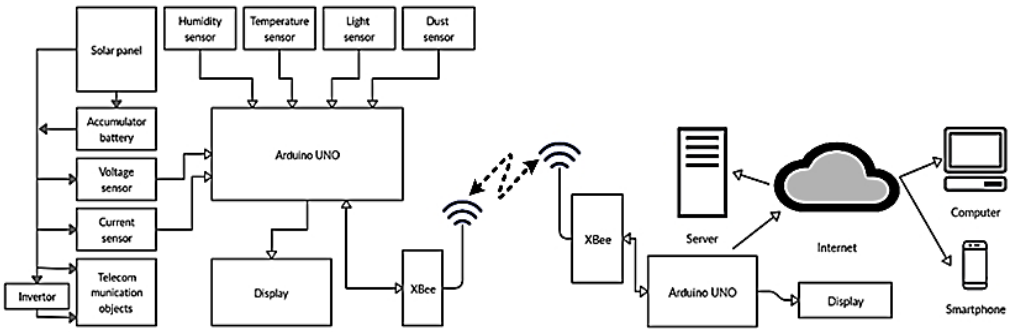


Fig. 3: Block diagram of the monitoring system of the solar station.

The monitoring center consists of the following blocks (Fig. 3.b):

- controller device unit (Arduino MEGA);
- wireless network module (ZigBee/Xbee);
- display;
- local server.

The monitoring center is also connected to the Internet, where the operator has access to the rights monitoring system assigned to it through any Internet-connected device [12].

https://doi.org/10.31436/iiumej.v22i2.1464

5. HARDWARE OF THE MONITORING SYSTEM OF SOLAR STATION

This section provides basic information about the devices and modules used for designing and modeling on the Proteus software environment as described above [15].

5.1 Solar Panel

Solar panels are semiconductor devices that convert solar energy directly into electricity and are widely used as an alternative source of energy for telecommunications facilities. Depending on the number of photovoltaic converters in solar panels, they can produce unchanged currents of varying voltage and power, and may vary in size.

5.2 Controller Device Units (Arduino UNO/MEGA)

The Arduino Uno and the Arduino MEGA boards are taken as CDUs. The Arduino Uno was used to measure the parameters of the solar power supply from the control unit and transfer it to the central control unit. Arduino Uno is a board that is based on an open source microcontroller ATmega328P (Fig. 4.a) [22, 23].

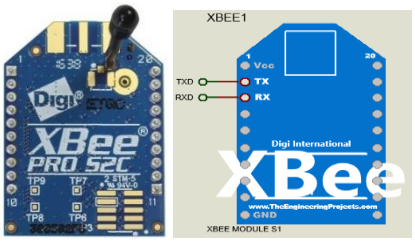
Arduino MEGA was used as a centralized control unit. This control unit is designed to receive, process, and transmit signals from distributed solar panel control units of telecommunications facilities to the server. The Arduino MEGA board is integrated with the Atmega2560 microcontroller (Fig. 4.b) [24]. These boards are equipped with a set of digital and analogue input/output pins, which allow us to connect to various expansion boards and other circuits.

(a) (b)

Fig. 4: CDUs: (a) Arduino Uno, (b) Arduino MEGA.

5.3 Wireless Network Module (ZigBee/Xbee)

ZigBee is a standard for high-level wireless communication protocols that support small, low-power digital transmitters based on IEEE 802.15.4 standard [25].

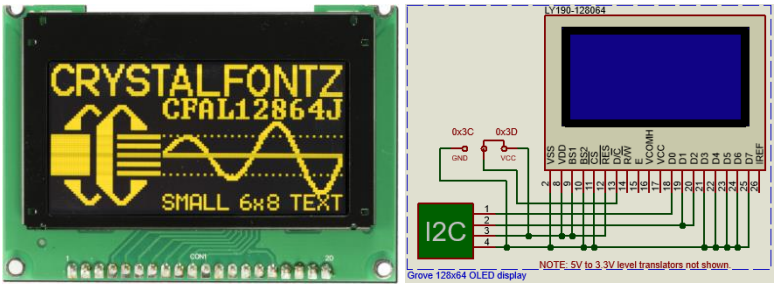


(a) (b)

Fig. 5: ZigBee/Xbee wireless network module: a) real view of the module, (b) environmental view of the module in Proteus software.

5.4 Display

The LCD display is used to reflect the parameters being monitored. LCD displays are the most common type of TV and monitor screens, as well as phone and other device displays. We used a 128x64 OLED graphical display in this experimental model. Its real and Proteus software environmental views are illustrated in Fig. 6a and 6b. The display connects to the microcontroller unit through the I2C interface. This enables efficient use of the microcontroller ports [7], [10], [14].

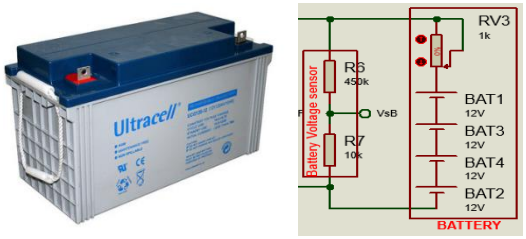


(a) (b)

Fig. 6: 128x64 OLED graphical display: (a) real view of the module, (b) environmental view of the module in Proteus software.

5.5 Accumulator Battery

A battery is a chemical power source; a reusable electrical driving force. Its main unique feature is the repetition of internal chemical processes, which allows it to accumulate energy and undergo multiple charge-recharge cycles for the autonomous energy supply of various electrical appliances and equipment [26]. In the model, four 12 V batteries (48 V total) were used (Fig. 7b). But in this case, there is one problem: it is impossible to show the battery charging or discharging during the ANIMATING time of the Proteus environment in the model. Therefore, an RV3 pot resistor was installed to show this condition. By changing this resistance, the battery voltage is increased or decreased which increases or decreases its charge. A voltage sensor consisting of R6 and R7 resistors was used to detect the voltage in the battery [22]

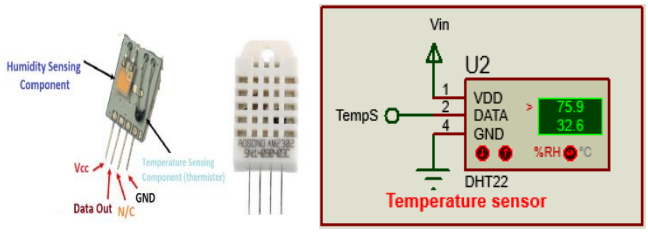


(a) (b)

Fig. 7: Accumulator battery: (a) real view of the module, (b) environmental view of the module in Proteus software.

5.6 Temperature and Humidity Sensor

The DHT22 series sensor was used to measure the temperature and humidity of the area where the solar panel was installed (Fig. 8). The DHT22 sensor is a well-established electronic temperature and humidity sensor that connects the microcontroller and the DHT22 sensor through a single data-exchange line. The single-line data exchange time is 5ms, with the measured temperature and humidity data being transmitted simultaneously [11], [30].



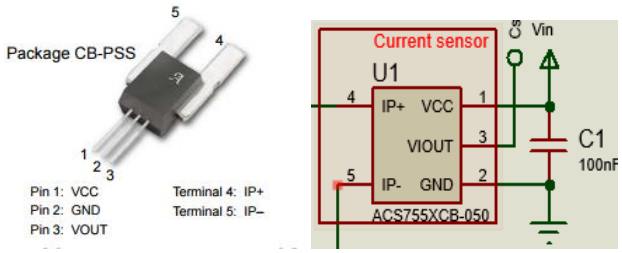
a) b)

Fig. 8. Temperature and humidity sensor: a) real view of the module b) environmental view of the module in Proteus software.

5.7 Current Sensor

The AC power sensor is a device used to measure and control alternating current, direct current, and pulse current. Current sensors are widely used in electrical engineering to create contrariwise communication systems. There are various types of current sensors such as resistive current sensors, current transformers, and Hall effect sensors. Hall effect sensors can measure and control DC current, AC current, and pulse current. In addition, they have reliable electrical insulation, stand out with a wide range of frequencies, and do not emit heat [5], [22], [24], [27-28].

Presently, there are many sensors that detect current. They differ from each other by the power measurement limit. Because the system we are building is high-power, we used ACS755XCB-050 as a 50 A current sensor [3], [7-8], [12]. Its real and virtual views are shown in Fig. 9.

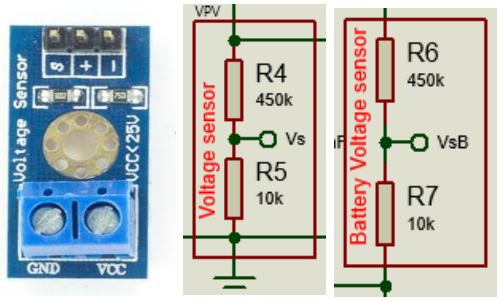


(a) (b)

Fig. 9: ACS755XCB-050 current sensor: (a) real view of the module, (b) environmental view of the module in Proteus software.

5.8 Voltage Sensor

The voltage sensor is a module that measures the voltage in the network. There are several types: voltage sensors of transformers, resistors, capacitors and other types. The most common type of voltage sensor is made with two resistors. It has a very simple structure. The two resistors are connected in series (such as R4/R5 and R6/R7) and their loose ends are connected to the power supply (as shown in Fig. 10). At the same time, a high voltage is distributed over these resistors. As a result, a voltage of less than 5 V is generated on the point where the two resistors connected. To convert high voltage to low voltage, the Atmega microcontroller control unit receives a voltage signal of up to 5 V. From their common endpoint, a small voltage signal is transmitted to the control unit of the Arduino board. The small voltage signal is converted to numerically accurate value using a special formula in the program [5], [22], [24], [27], [29].



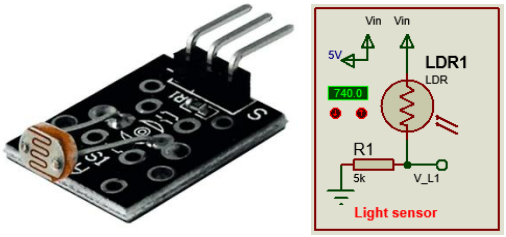
(a) (b)

Fig. 10: Voltage sensor: (a) real view of the module,
(b) environmental view of the module in Proteus software.

The imitation model uses two voltage sensors. One of them determines the amount of voltage in the common mains and the other determines the amount of voltage in the battery. The amount of charge on the battery is calculated with the second sensor.

5.9 Light Sensor

The light sensor is a device for measuring the level of ambient light. In the model, light sensors are used to monitor the degree of illumination of the telecommunications facilities in the area where solar energy sources are located, to determine whether it is sunny or cloudy and whether the area's illumination is sufficient for energy production. The range of light sensors is formed at different intervals (luxe). They can be adjusted depending on the location and conditions of installation [7], [9-10]. A photoresistor was used to detect the illumination in this system (Fig. 11). The photoresistor resistance varies linearly as the ambient light changes. In this case, at the same time as the amount of light that comes on the photoresistor surface, its resistance decreases linearly, and with the decrease in the amount of light, its resistance increases. Therefore, the amount of ambient light is measured depending on the photoresistor resistance [22].



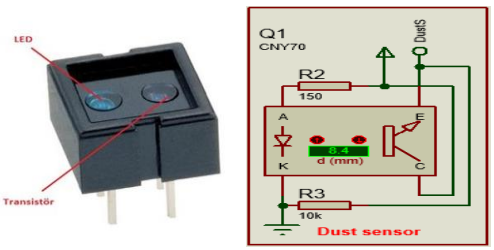
(a) (b)

Fig. 11. Light sensor: (a) real view of the module,
(b) environmental view of the module in Proteus software.

5.10 Dusting Sensor

The dust sensor can measure the dust value by transmitting a signal between 4 and 20 mA, or report that the dust level exceeds the set values. Many dust sensors operate on the principle of quantum particle measurement and detect high-resolution particles larger than one micron. These particles include cigarette smoke, household dust, spores, mold and more. Typically, sensors such as SM-PWM-01C, GP2Y1010AU0F are used to determine the level of dust. These sensors determine the concentration of dust in the air using an optical sensitivity method. The infrared light-emitting diode (IR LED) and photo-sensor

(Photo Transistor) are used as optical sensors. The photosensor receives infrared light passing through dust particles [9].



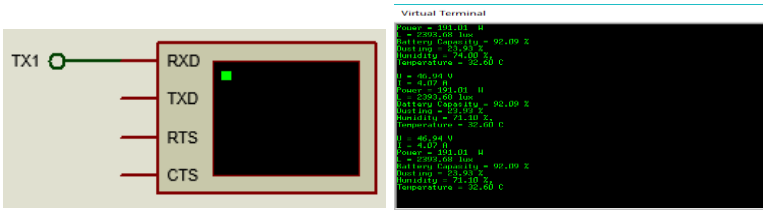
(a) (b)

Fig. 12: Dusting sensor: (a) real view of the module,
(b) environmental view of the module in Proteus software.

The CNY70 element was used as a dust sensor in the Proteus software environment (Fig. 12). In fact, the CNY70 element is used for other purposes. However, experience has shown that this element can also be used as a dust sensor in real conditions, and that is appropriate for our experimental sample necessitating only some minor experiments.

5.11 Virtual Terminal

In the modeling, we used the additional capabilities of the Proteus software environment to monitor and analyze transmitting data in real-time.



(a) (b)

Fig. 13: Virtual terminal: (a) method of connection to the system,
(b) appearance of values in the simulation.

One of them is the Virtual Terminal. It communicates with Arduino through the UART interface [13]. It is possible to print optional information in the virtual terminal based on the parameters entered in the program. In this model, the measured constants shown in Fig. 13 can be monitored in real-time.

6. SOFTWARE ALGORITHM OF THE MONITORING SYSTEM OF SOLAR STATION

The solar power station monitoring system has been modelled in two parts: solar power station management and its measurement system and monitoring center. They can be viewed as a distributed system. In other words, these two systems are at a certain distance from each other and have related and unrelated control processes [4]. Therefore, a separate software was written for each system model in the Arduino IDE environment. Their algorithm is illustrated in Fig. 14 as a block diagram.

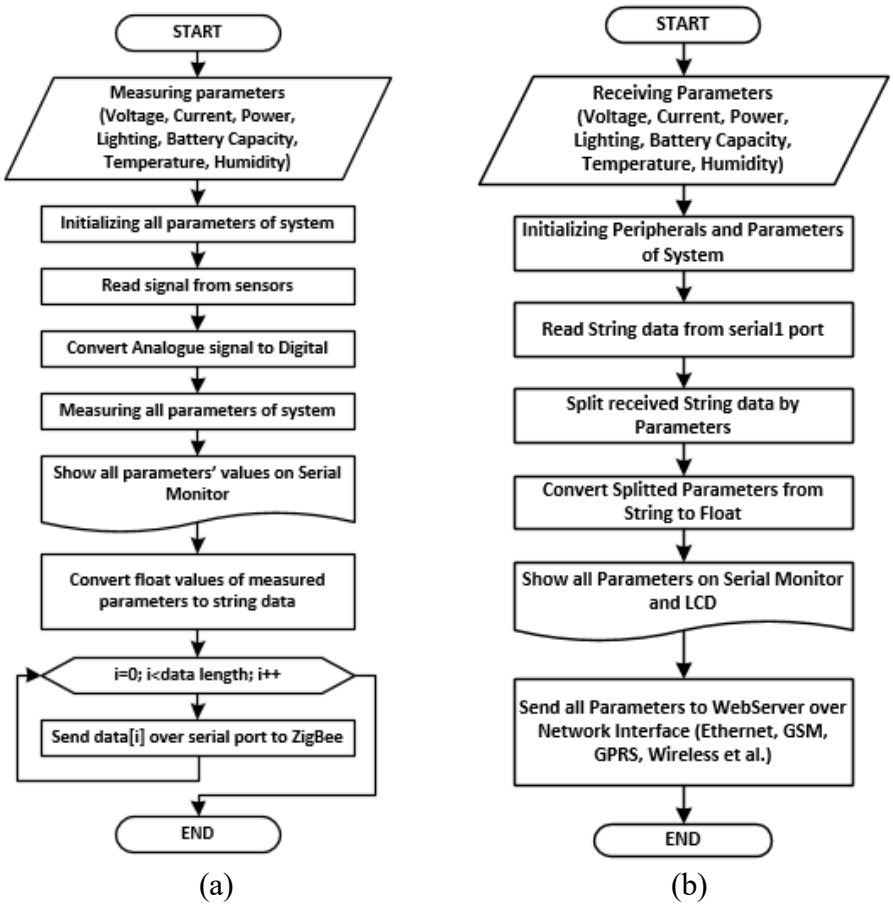


Fig. 14: Software algorithm for monitoring system: (a) software algorithm for transmission part, (b) software algorithm for receiving part.

Figure 14a illustrates the solar power station management and its algorithm for the measuring software system. As shown in this algorithm, the incoming signals from the sensors located in the object are converted to digital and the value is determined. All defined parameters are first displayed in the Serial Monitor [22].

At the same time, the values of this parameter are passed through the serial ports to the XBee wireless network module. However, because the measured values are multiple and the value type is float, they cannot be written directly to the serial port. Otherwise, the values may be mixed and the monitoring center will receive error values. Therefore, the value of each parameter (value type - float) is initially converted to a specified amount or byte sequence (Value type - String). Then, in order to distinguish each parameter from the other, the characters are appended to the front of the values and formed in the form of a common data package. Each symbol of the formed data is passed through serial ports to the XBee wireless network module. The data is transmitted to the monitoring center using a wireless network module.

Figure 14b illustrates the algorithm of the management system of the solar power station's monitoring center. In this case, after the system is fully loaded, the serial port determines the data flow or byte sequence (value type - String) from the XBee wireless network module. Then each character in this information is scanned and the value corresponding to the character is divided by the measured parameters. The value of each allocated parameter is initially specified as a String type of value. On such a value, it is impossible to perform mathematical calculations and display on the screen thus, they are

converted to float type of values (type of value - float). Defined parameters are displayed on the Serial Monitor and LCD. At the same time, these values are transmitted to Web Server using Internet network interfaces.

7. DESIGNED STRUCTURES OF THE MONITORING SYSTEM

The structure of the power monitoring system of any facility consists of three parts: the measuring part, the server part, and the web page. Specific sensors (such as current, voltage, etc.) are installed in the measurement unit for energy sources and consumer loads. These sensors are connected to the control unit, which is a small computing system.

The sensors detect changes in the object and transmit the signal to the control unit as voltage or current. The control unit converts the voltage or current signal into digital form and transmits it via the wireless network to the central control unit (or Monitoring Center). There may be more than one such system. The information transferred from them is consolidated in the central control unit (Monitoring Center) and transmitted via the Internet to the server. At the same time, it is possible to monitor real-time values from sensors using a web browser [3], [11].

The process of designing remote monitoring of solar power sources of telecommunications facilities in the Proteus software environment was carried out in three stages:

- Equipping the solar power station with sensors and IoT technology;
- Equipping the monitoring center with IoT technology
- Creating a website for monitoring.

Typically, this system is designed to remotely monitor solar power stations located in different areas. As an example, in this experiment, we implemented a monitoring system on only one solar power plant in a virtual environment. But our idea is to monitor all the power plants in the designated area. There may be more than 100 such solar power stations.

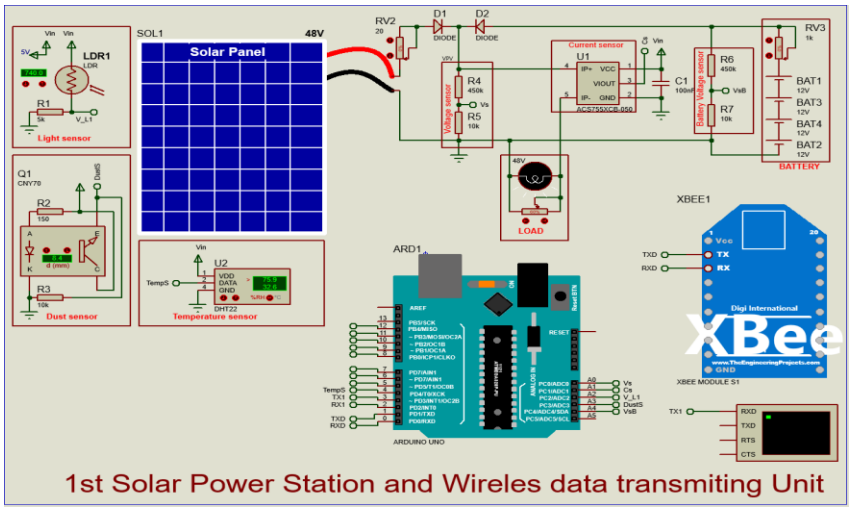


Fig. 15: Solar power plant equipped with IoT technology.

Figure 15 shows the first step of the system we designed in the virtual environment. Several sensors are used for the area where solar panels are located and telecommunications facilities: light sensors, dust sensors, temperature and humidity

sensors, voltage sensors, and current sensors. The characteristics of these sensors are mentioned above. They usually serve to determine the condition of the solar panel's installed location (such as the temperature of the solar panel, dust on the surface, and ambient humidity), the amount of energy the solar panel produces, the amount of battery charge, and energy consumed by telecommunications facilities [2], [22-24]. Defined parameters are transmitted to the monitoring center using the ZigBee/Xbee wireless network module.

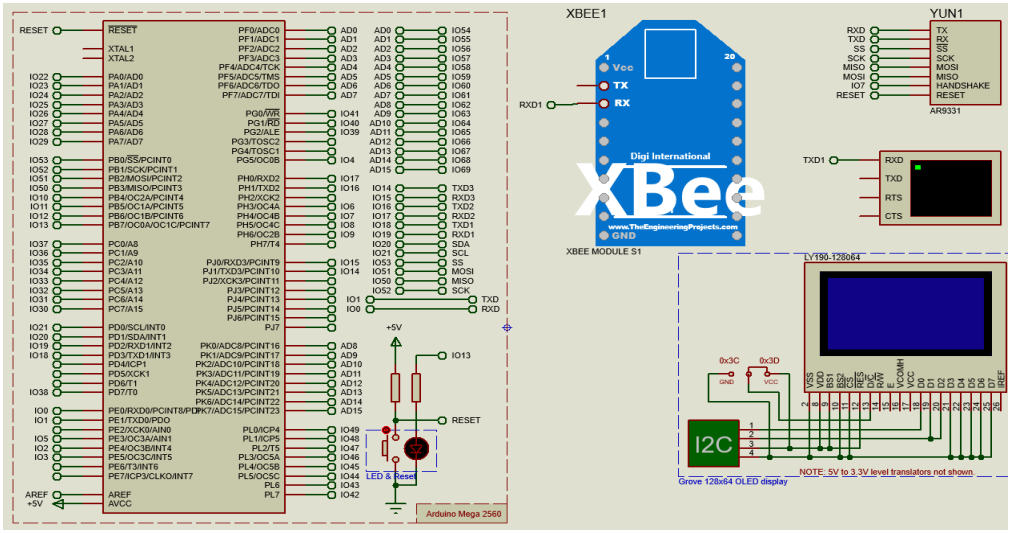


Fig. 16: Monitoring center equipped with IoT technology.

In the next section, the Monitoring Center is designed to monitor and analyze the parameters measured at solar power plants (Fig. 16). The monitoring center is equipped with Arduino MEGA, virtual terminal, ZigBee/Xbee wireless network module, LCD display, as well as Internet network module (may include YUN WIFI & Ethernet module, as well as GSM, Ethernet, WIFI). The monitoring center receives the parameters measured and transmitted at the solar power plant using the ZigBee/Xbee wireless network module. The received data is displayed using a virtual terminal and LCD display. Simultaneously, these parameters are passed to the webserver. The measured parameters are stored in the webserver memory.

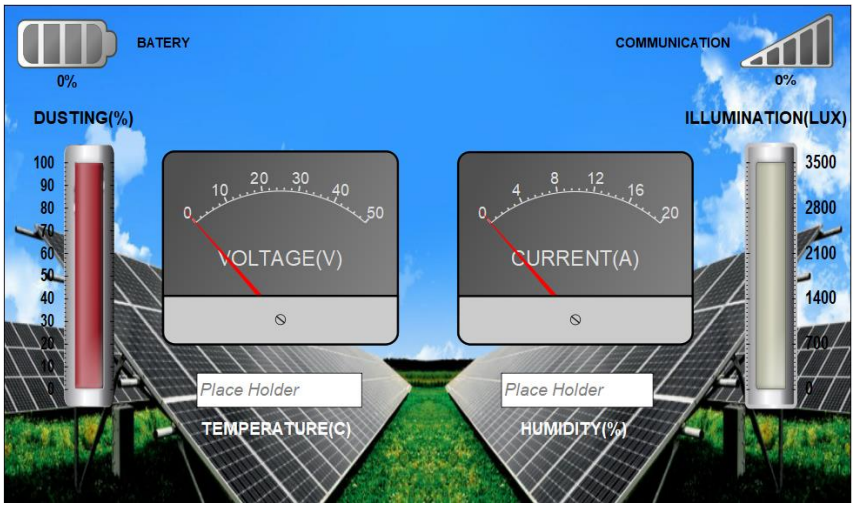


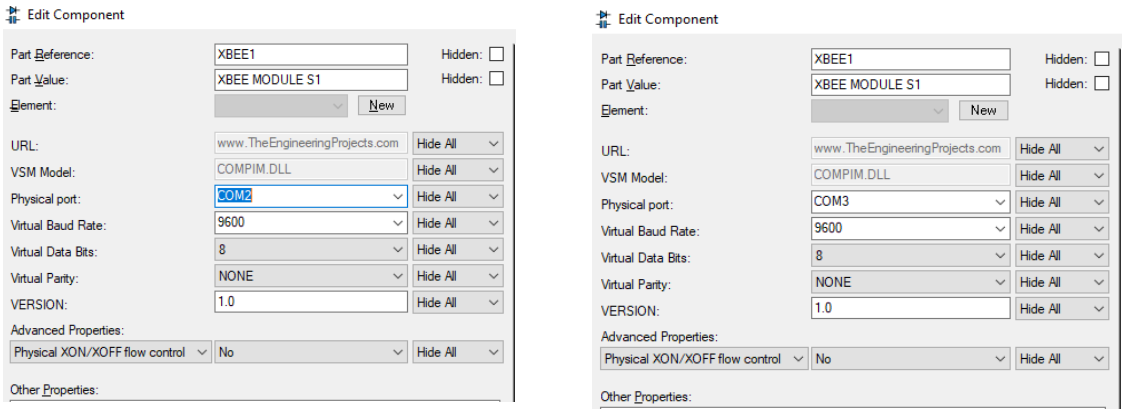
Fig. 17: Web page view for the monitoring system.

It is well known that the monitoring system requires a system of remote control, management and evaluation of the distributed system parameters to stay at a discretionary distance. A simple web page was created to fulfil the tasks set out in the third stage of monitoring system modelling (Fig. 17).

The webpage can be referenced by typing a web address in any web browser. As a result, a web browser will display the page shown in Fig. 17. On this page, the values of the solar power plants' measured parameters (solar panel dust, voltage, current, battery charge, ambient light, temperature, humidity) are presented in various graphical images and figures. Every change can be seen in real-time using this webpage. This allows us to monitor system performance, quickly identify problems, and use optimal solutions to address them, to evaluate and improve the system's performance.

8. SIMULATION PROCESS OF THE MODEL

The solar power station monitoring system of telecommunications facilities modelled in the Proteus software environment is a virtual system. It is required to create a wireless sensor network between the ZigBee/Xbee network modules to transfer from the station to the monitoring center. The exchange of data between wireless sensor networks is carried out by electromagnetic waves in the environment [31-32]. However, the simulation environment is embedded in a single PC software that cannot generate signals using electromagnetic waves. The Xbee network module in the Proteus software environment (Fig. 5b) is built on the RS232 serial communication interface. Information exchange between RS232 serial communication interfaces is provided through a physical communication port (COM Port - Communication Port). This means that a virtual communication port should be created on the PC to provide communication between the Xbee network modules in the Proteus software environment. A lot of software is available on the PC to develop virtual communication ports. For this model, VSPE (Virtual Serial Port Emulator) software was used to create a virtual COM port on the PC, and virtual COM2 and COM3 ports were created. They are configured for the Xbee network module in the Proteus software environment (Fig. 18).



(a) (b)

Fig. 18: Xbee network module configuration window in Proteus software environment: (a) for the transmitter Xbee network module at the station, (b) for the receiving Xbee network module at the monitoring center.

As mentioned above, data received by the monitoring center is transmitted to the Internet using the YUN WIFI & Ethernet network interface. To view the transmitted

values in a PC web browser, you need to configure the local host port number for the YUN WIFI & Ethernet network interface. In this model, the 8181 localhost port number is configured for the YUN WIFI & Ethernet network interface (Fig. 19).

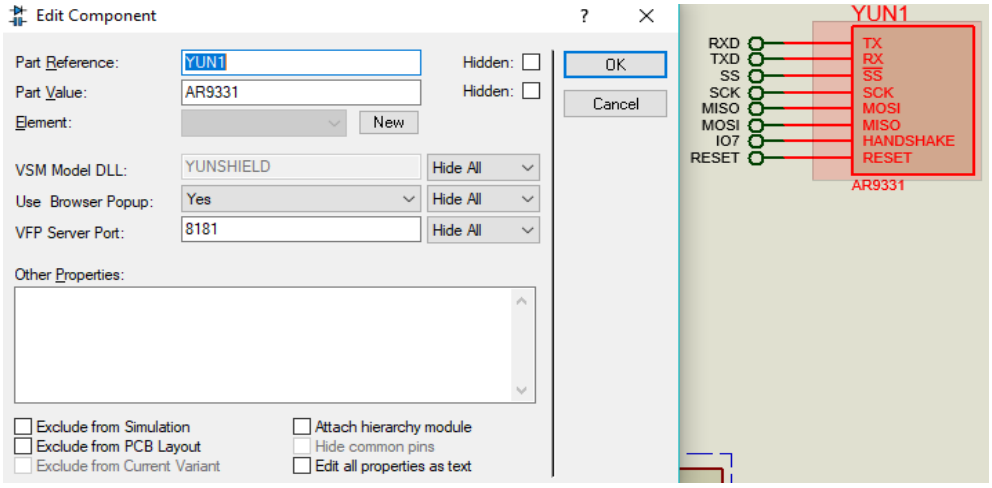
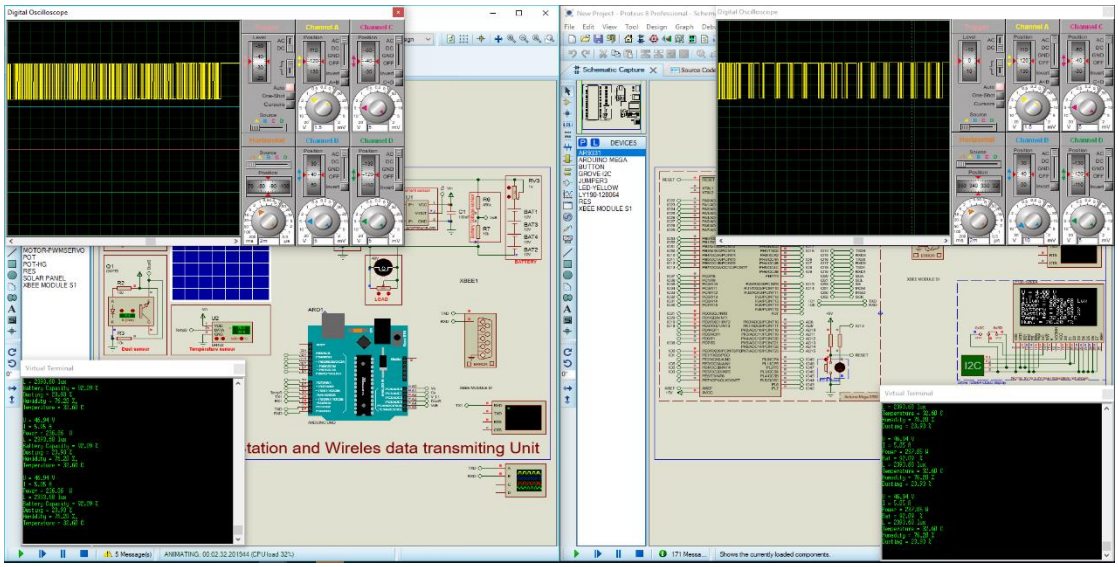


Fig. 19: Configuration window the YUN WIFI & Ethernet network interface in Proteus software environment.

After the settings have been fully set up, by simultaneously launching the simulation models of the solar power plant wireless transmission system and the monitoring center, the simulation results can be verified, the values of sensors located on the solar station (Fig. 20a) can be changed, and the results can be monitored on the LCD display and virtual terminal located in the monitoring center (Fig. 20b). Also, the monitoring center transmits the measured values to the webserver using the Internet network interface. Figure 20 shows the simulation result of the monitoring system. The signal view of transmitted and received information is displayed through the digital oscilloscope.

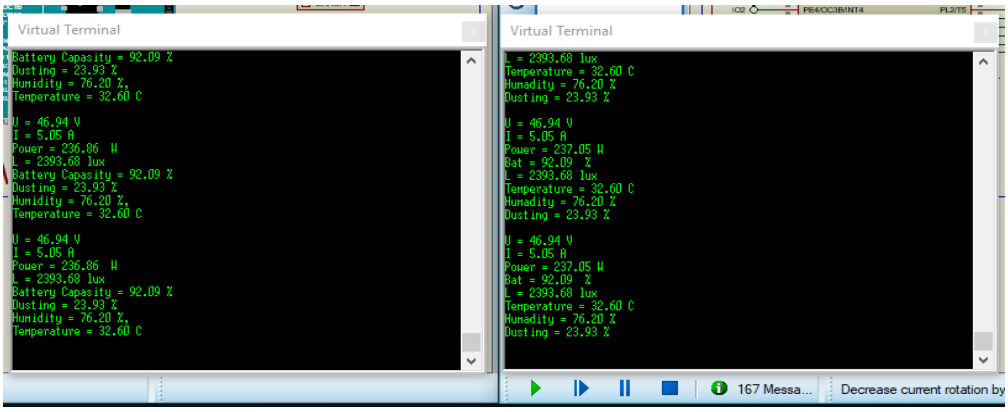


(a) (b)

Fig. 20: The window of the simulation process with transmitting (a) and receiving, (b) data.

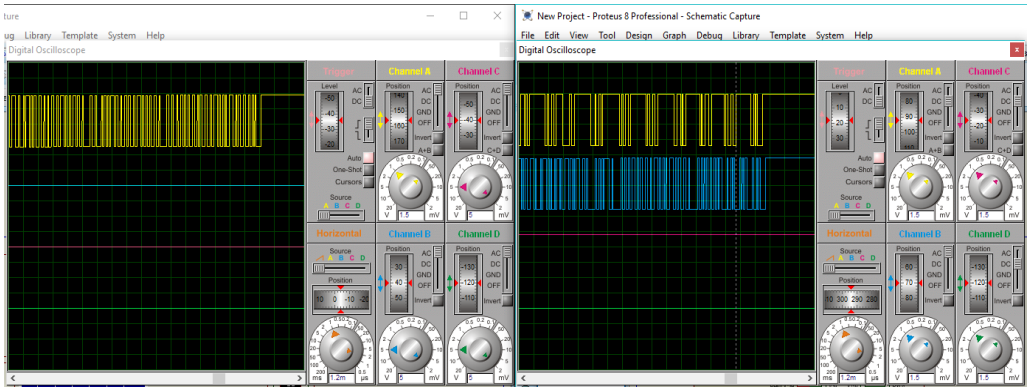
9. RESULTS

The results of the system model simulation process were usually analyzed using the Virtual Terminal and a digital oscilloscope. The Virtual Terminal and Digital Oscilloscope are connected to the signal channels transmitted by the solar power station to the Xbee network module and received through the Xbee network module located in the monitoring center, as well as the Internet wireless interface (YUN WIFI & Ethernet). The initial check was performed for transmitted and received signals, and the results are illustrated in Fig. 21 and 22. Figure 21a belongs to the transmitting values and Fig. 21b belongs to the receiving values. In this Fig., the values transmitted by the station and received by the monitoring center are the same. This means that the measured values are passed and received through the Xbee network module without any errors.



(a) (b)

Fig. 2: The view window of values transmitted and received by the Xbee network module during the simulation in the Virtual Terminal.



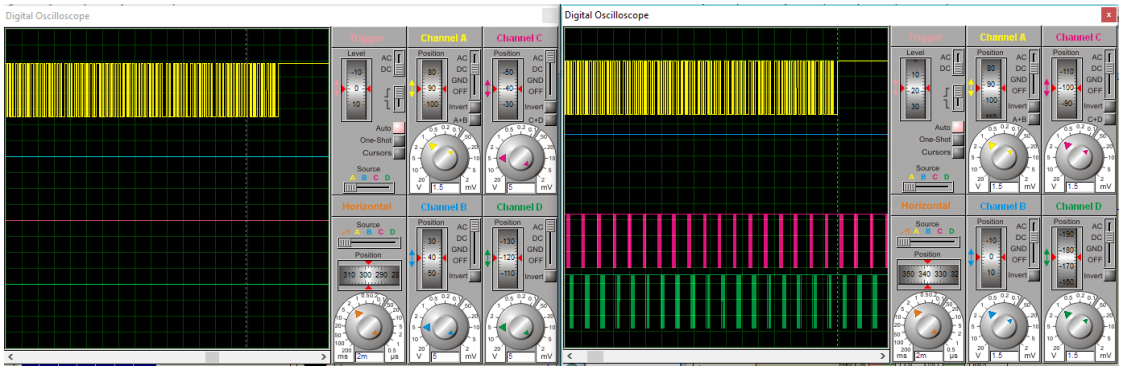
a) b)

Fig. 22: The electrical signal graph of the transmitting and receiving values by the Xbee network module during the simulation.

Using the Digital Oscilloscope, the electrical signal view of the values shown above gave the following results (Fig. 22). All channels (channels A, B, C, D) of the Digital Oscilloscope are set to the same setting to compare the signal status. In the graph, the yellow line represents the signal being transmitted and received. In this case, the electrical signal of the transmitted data differs from that of the received signal. This is because additional information is being added to ensure the reliability and accuracy of the transmitted data. The blue line signal graph in Fig. 22b represents an unexpectedly

generated signal. This can be seen as a signal that the information has been received. As you can see from these pictures (Fig. 21 and 22), the parameters measured at the solar power station and reach the monitoring center in the form of data using a wireless sensor network.

The next step was to analyze the signal graphs transmitted and received by the monitoring center to the webserver (Fig. 23b). As you know, web servers are based on the client-server architecture. The client first sends a request to the server, and the webserver responds to the request based on the received request. Therefore, this image also shows the transmitting information (or request) signal by the monitoring center and the receiving information (or confirmation response) signal from the server. In this case, the transmitted signal graphic to the webserver is a red line, and the received signal graphic is green.



(a) (b)

Fig. 23: The electrical signal graphics of data transmitted and received by the Xbee network module (a) and the Internet network module (b) during the simulation.

In this figure, the graph shows the transmitted and received signals appearing at the same time. We place them on top of each other to see the difference between the transmitted and received signals (Fig. 24), the difference is obvious.

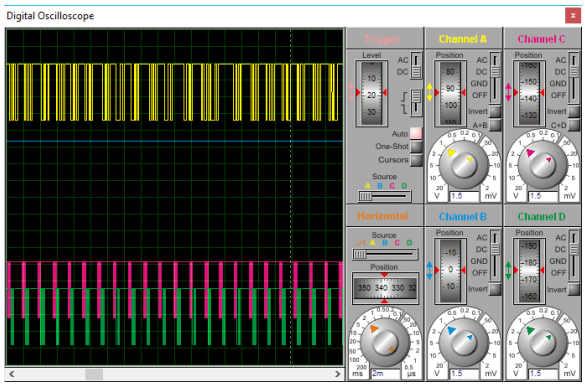


Fig. 24: Comparative electrical signal graph of data transmitted and received by the network module in the simulation process.

We can see the simulation results of the telecommunications facilities solar power station monitoring system model in any web browser installed on PCs or smartphones. The result of the simulated system was verified using a Microsoft Edge web browser installed on the PC (Fig. 22). To view the results, open the web browser and enter "localhost: 8181". As a result, the web page shown in Fig. 25 will open. With this web page, the system status can be monitored in real-time

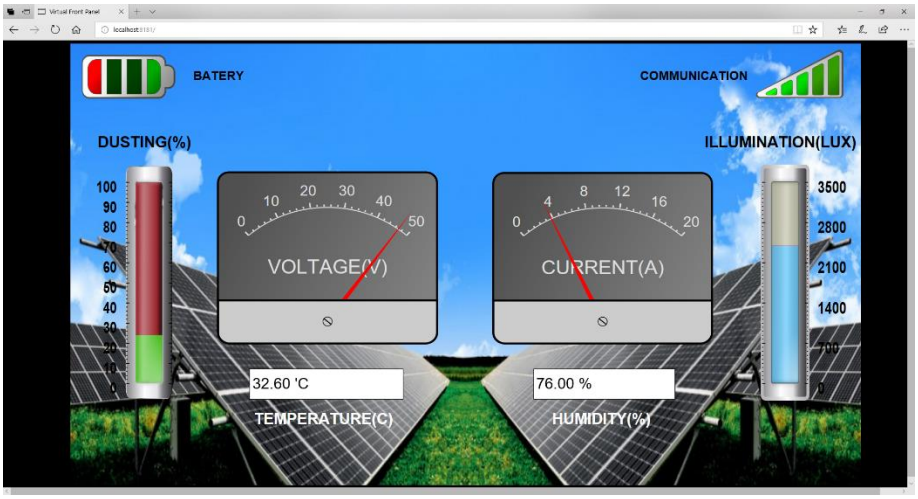


Fig. 25: Web page view of solar power station monitoring system.

10. CONCLUSION

Today, the increasing demand for renewable energy sources has led to the development of research and improved the efficiency of such systems used in various facilities. In particular, the use of solar power stations for telecommunications facilities is a crucial factor in energy conservation. Therefore, the design of this system was developed and modelled. A software product for this system has been created. There was research on all parameters of the solar power plant. We can be watching when these settings were changed. The monitoring center and webpage settings were changed. In this case, when the optional parameters of the solar power station were changed, the monitoring center and website were updated at the same time. The value has been changed to replace the previous one. This system is a crucial factor for the accurate and reliable operation of the system.

REFERENCES

[1] Pansare J. Sonavane SS. (2016) Efficient Energy Management in Smart Grids based on Raspberry Pi & Web of Things. International Journal of Science Technology & Engineering, 2(12).

[2] Tanwar S, Tyagi S, Kumar S. (2018) The role of Internet of things and smart grid for the development of a smart city. Intelligent Communication and Computational Technologies: January 2018; pp 23-33. https://doi.org/10.1007/978-981-10-5523-2_3

[3] Kabalci E, Kabalci Y. (2017) Remote Monitoring System Design for Photovoltaic Panels. The 10th International Symposium on Advanced Topics in Electrical Engineering: 23-25 March 2017, Bucharest, Romania.

[4] Chunming W, Liang Ch. (2012) The Monitoring System for the Wind Power Generation Based on the Wireless Sensor Network. International Conference on Future Electrical Power and Energy Systems: 2012; Hainan University, China.

[5] Siddikov IX, Sattarov KhA., Khujamatov KhE. (2017) Modeling of the Transformation Elements of Power Sources Control. International Conference on Information Science and Communications Technologies: 2-4 November 2019; Tashkent, Uzbekistan.

[6] Muradova AA, Khujamatov H. (2019) Results of Calculations of Parameters of Reliability of Restored Devices of the Multiservice Communication Network. International Conference on Information Science and Communications Technologies: 2-4 November 2019; Tashkent, Uzbekistan.

[7] Nagalakshmi R, Babu BK, Prashanth D. (2014) Design and Development of a Remote Monitoring and Maintenance of Solar Plant Supervisory System. International Journal of Engineering and Computer Science, 3(12).
[8] Prakash S, Vijayaragavan SP. (2015) Wireless Wind Turbine Monitoring Using Arduino. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 4(7).
[9] Parikh A, Pathan F, Rathod B, Shah S. (2015) Solar Panel Condition Monitoring System based on Wireless Sensor Network. International Journal of Science, Engineering and Technology Research, 4(12).
[10] Jumaat SA, Othman MH. (2017) Solar Energy Measurement Using Arduino. MATEC Web of Conferences 150: 2017, MUCET.
[11] Madhubala S, Nachammai R, Nandhini I, Preethisha AM, Paulin JJ. (2018) Solar power based remote monitoring and control of industrial parameters using IoT. International Research Journal of Engineering and Technology, 5(3).
[12] Parveen R, Mohammed AM, Ravinder K. (2018) IoT based solar tracking system for efficient power generation. International Journal of Research and Analytical Reviews, 5(4).
[13] Vijayaragavan SP, Karthik B, Sriram M. (2019) Wind Mill Monitoring System using IOT and WI-FI. International Journal of Recent Technology and Engineering, 8(1).
[14] Farzhana PM, Neelavathy R, Reshma SB, Saravanakumar U. (2019) Simulation, Fabrication and Monitoring of a Solar Power Plants using Arm7 and Proteus Software. International Journal of Engineering Science and Computing, 9(5).
[15] Davronbekov DA, Matyokubov UK. (2020) The Impact of Mobile Communication Power Supply Systems on Communication Reliability and Viability and Their Solutions. International Journal of Advanced Science and Technology, 29(5): 374-385.
[16] Davronbekov DA, Aliev UT, Isroilov JD. (2017) Using the energy of electromagnetic radiation as a source of power. International Conference on Information Science and Communications Technologies, Applications, Trends and Opportunities: 2-4 November 2016; Tashkent, Uzbekistan https://doi.org/10.1109/ICISCT.2017.8188565
[17] Pulatov ShU, Aliev UT, Isroilov JD. (2017) Energy harvesters wireless charging technology. International Conference on Information Science and Communications Technologies, Applications, Trends and Opportunities: 2-4 November 2017; Tashkent, Uzbekistan. https://doi.org/10.1109/ICISCT.2017.8188566
[18] IEEE Computer Society. (2015) IEEE Standard for Low - Rate Wireless Networks. IEEE Std 802.15.4™-2015.
[19] Khujamatov Kh, Khaleel A, Reypnazarov E, Khasanov D. (2020) Markov Chain Based Modeling Bandwith States of the Wireless Sensor Networks of Monitoring System. International Journal of Advanced Science and Technology, 29(4): 4889-4903. http://sersc.org/journals/index.php/IJAST/article/view/24920
[20] Khujamatov Kh, Khasanov D, Reypnazarov E, Akhmedov N. (2020) Networking and Computing in Internet of Things and Cyber-Physical Systems. International Conference Application of Information and Communication Technologies: 7-9 October 2020; Tashkent, Uzbekistan.
[21] Isroilov JD. (2016) Linearization spectral characteristics through passage by means of akusto-optical reconstructed filters. International Conference on Information Science and Communications Technologies, Applications, Trends and Opportunities: 2-4 November 2016; Tashkent, Uzbekistan. https://doi.org/10.1109/ICISCT.2016.7777386
[22] Khujamatov Kh.E., Khasanov D.T., Reypnazarov E.N. (2019) Modeling and Research of Automatic Sun Tracking System on the bases of IoT and Arduino UNO. International Conference on Information Science and Communications Technologies: 4-6 November 2019; Tashkent, Uzbekistan.
[23] Davronbekov DA. (2016) Features measurement parameters and control functioning of integrated chips. International Conference on Information Science and Communications Technologies: 2-4 November 2016; Tashkent, Uzbekistan.

[24] Khujamatov Kh.E., Khasanov D.T., Reypnazarov E.N. (2019) Research and Modelling Adaptive Management of Hybrid Power Supply Systems for Object Telecommunications based on IoT. International Conference on Information Science and Communications Technologies: 4-6 November 2019; Tashkent, Uzbekistan.
[25] Davronbekov DA, Aliev UT, Isroilov JD, Alimdjanov XF. (2019) Power providing methods for wireless sensor. International Conference on Information Science and Communications Technologies: 4-6 November 2019; Tashkent, Uzbekistan. https://doi.org/10.1109/ICISCT47635.2019.9011850
[26] Skyllas-Kazacos M, Menictas C. (2013) Electricity Transmission, Distribution and Storage Systems. Woodhead Publishing Series in Energy.
[27] Siddikov I, Sattarov Kh, Khujamatov Kh, Dexkhonov O, Agzamova M. (2018) Modeling of Magnet Circuits of Electromagnetic Transducers of the Three-Phases Current. International Scientific-Technical Conference on Actual Problems of Electronics Instrument Engineering: 2-6 October; 2018; Novosibirsk; pp 419-422.
[28] Siddikov I, Sattarov Kh, Khujamatov H. (2016) Research of the Influence of Nonlinear Primary Magnetization Curves of Magnetic Circuits of Electromagnetic Transducers of the Three-phases Current. Universal Journal of Electrical and Electronic Engineering. Horizon Research Publishing Corporation, 4(1): 29-32
[29] Siddikov I, Sattarov Kh, Khujamatov H. (2018) Modeling and research circuits of intelligent sensors and measurement systems with distributed parameters and values. Chemical technology control and management. International Scientific and Technical Journal, 4(5): 50-55.
[30] Aosong Electronics Co., Ltd. Digital-output relative humidity & temperature sensor/module DHT22 (also named as AM2302), DHT22 datasheet.
[31] Arsheen S, Wahid A, Khaleel A, Khujamatov Kh. (2020) Flying Ad hoc Network Expedited by DTN Scenario: Reliable and Cost-effective MAC Protocols Perspective. International Conference Application of Information and Communication Technologies: 7-9 October 2020; Tashkent, Uzbekistan.
[32] Siddikov IKh, Sattarov KhA, Khujamatov KhE, Dekhkonov OR. (2016) Modeling the processes in magnetic circuits of electromagnetic transducers. International Conference on Information Science and Communications Technologies: 2-4 November 2016; Tashkent, Uzbekistan.

https://doi.org/10.31436/iiumej.v22i2.1541

DEVELOPING A PARALLEL CLASSIFIER FOR MINING IN BIG DATA SETS

AHAD SHAMSEEN[1], MORTEZA MOHAMMADI ZANJIREH [1*], MAHDI BAHAGHIGHAT [1], QIN XIN[2]

[1]*Computer Engineering Department, Imam Khomeini International University, Qazvin, Iran*
[2]*Faculty of Science and Technology, University of the Faroe Islands, Tórshavn, Faroe Islands*

**Corresponding author: Zanjireh@eng.ikiu.ac.ir*

(Received: 10th July 2020; Accepted: 12th November 2020; Published on-line: 4th July 2021)

ABSTRACT: Data mining is the extraction of information and its roles from a vast amount of data. This topic is one of the most important topics these days. Nowadays, massive amounts of data are generated and stored each day. This data has useful information in different fields that attract programmers' and engineers' attention. One of the primary data mining classifying algorithms is the decision tree. Decision tree techniques have several advantages but also present drawbacks. One of its main drawbacks is its need to reside its data in the main memory. SPRINT is one of the decision tree builder classifiers that has proposed a fix for this problem. In this paper, our research developed a new parallel decision tree classifier by working on SPRINT results. Our experimental results show considerable improvements in terms of the runtime and memory requirements compared to the SPRINT classifier. Our proposed classifier algorithm could be implemented in serial and parallel environments and can deal with big data.

ABSTRAK: Perlombongan data adalah pengekstrakan maklumat dan peranannya dari sejumlah besar data. Topik ini adalah salah satu topik yang paling penting pada masa ini. Pada masa ini, data yang banyak dihasilkan dan disimpan setiap hari. Data ini mempunyai maklumat berguna dalam pelbagai bidang yang menarik perhatian pengaturcara dan jurutera. Salah satu algoritma pengkelasan perlombongan data utama adalah pokok keputusan. Teknik pokok keputusan mempunyai beberapa kelebihan tetapi kekurangan. Salah satu kelemahan utamanya adalah keperluan menyimpan datanya dalam memori utama. SPRINT adalah salah satu pengelasan pembangun pokok keputusan yang telah mengemukakan untuk masalah ini. Dalam makalah ini, penyelidikan kami sedang mengembangkan pengkelasan pokok keputusan selari baru dengan mengusahakan hasil SPRINT. Hasil percubaan kami menunjukkan peningkatan yang besar dari segi jangka masa dan keperluan memori berbanding dengan pengelasan SPRINT. Algoritma pengklasifikasi yang dicadangkan kami dapat dilaksanakan dalam persekitaran bersiri dan selari dan dapat menangani data besar.

KEY WORDS: *Data mining, Big data, Decision tree, Parallel classifier, SPRINT*

1. INTRODUCTION

Recently, too much data is generated and stored on servers per day. This data has great amounts of essential and useful information and therefore, searching through this data can be one of the most essential topics in data science. The process of extracting information from large data sets and discovering patterns is termed *data mining* and it involves methods from artificial intelligence, machine learning, and database systems. One of the main goals of the data mining process is to build a logical structure from the extracted information and prepare

this information for further use. The data mining process involves many sub-processes like data collection, database management, data pre-processing, and online updating [1].

In data mining, classification is one of the learning models, and the classifications are used to build a model that classifies the unclassified data in the dataset. These models are rules or patterns that connect the input data with the output target and used to predict the future output of the system that the data was taken from [2] and can be used in several fields such as marketing, advertisement, and medical science [3].

The data mining process steps that were explained by Berry and Linoff in 2004 [2] are in Figure 1.

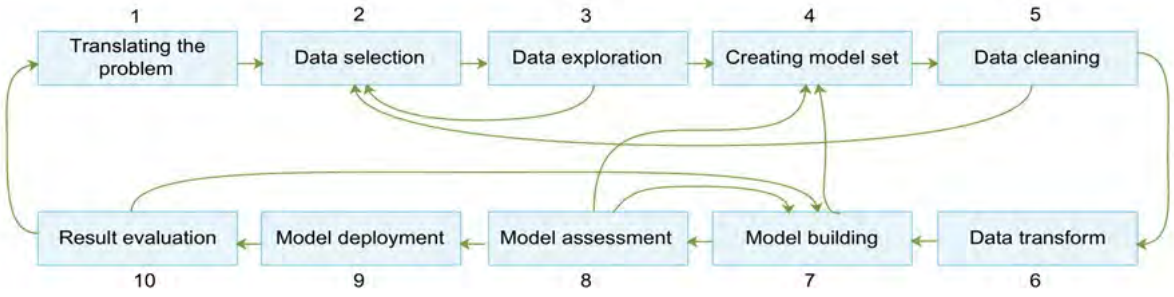


Fig. 1. Data mining process [2].

To create data mining models, we can use data mining algorithms. First, the algorithm starts by analysing the data set and extracting trends, results, and specific patterns, and these will be used next to determine the priorities and the parameters of the data mining model. Finally, these parameters are applied to the data set to find the patterns and statistical needs [4]. Selecting the best algorithm for the application depends on the data mining tasks, available data, data kind, and the size of data [5].

The decision tree algorithm is one of the most popular data mining algorithms because it has high accuracy, fast training performance, and is easy to understand. Dividing the data into several subsets is the primary purpose of this algorithm. The first step is to evaluate the attributes to determine how to divide the data between the classes. Then, this process is applied over all subsets until the decision tree is formed. The decision tree algorithm is a hybrid algorithm because it supports classification and regression tasks. The classification task can be used to predict continuous and categorical variables where the regression task can be used to predict only continuous variables. Decision tree algorithms can also be applied for association analysis [4]. The decision tree algorithm has many advantages over all data mining algorithms, which are easy to understand, quick to build the tree, and efficient to a high degree [6].

Naive Bayes algorithm is another data mining algorithm that achieves the learning phase based on evidence. This algorithm learns the evidence by counting the connection between the critical variables and all other variables [7]. Naive Bayes can detect the factors in a production line or how the products are related. However, this algorithm cannot classify nonlinearly separable classes correctly [4].

The neural network algorithm is a data mining algorithm that works by creating relationships between the input and output data and then uses these relationships to create

prediction patterns [4]. The algorithm result is difficult to explain, thus, a data mining process should start with decision trees, and then use the neural network algorithm [8].

The clustering algorithm finds the variable that classifies the data, and it is often used as a step in a big data analysis project. The clustering algorithm uses two methods to assign the clusters. The first method is the K-means method. In this method, the centre is moved to the mean of all assigned objects after the assigned object's phase. The second method is the EM method (Expectation-Maximization method), that uses a probabilistic measure to assign objects to the clusters and uses a bell curve for the dimension from the low centre and a standard deviation [4, 9].

An association algorithm is a rule-based method in data mining. This method is used to discover interesting relations between the attributes in big data sets [10]. This algorithm uses impressive measures to identify the discoverable rules in the data set.

The sequence clustering algorithm is a combination of two techniques, the clustering technique and sequence technique [4]. A time-series algorithm is a series of the data collection, which is collected in a timeframe. This algorithm is used to forecast the future based on history. This algorithm can provide many services and the most commonly used specify the seasonality and multiple periods. This algorithm has an excellent advantage, and it can do a great job with minimum information [4].

The data classification process divides the data into classes where each class has a unique symbol. The data in each class have common properties. To classify the data, we have first to determine how many tasks we want, and we have to determine the methods for breaking data into ranges, this method is related to the data that we are classifying and the project that we are working on. Data classification has some challenges, such as how to determine the best method to break and split the input data, and the challenge of determining the number of classes and what these classes are.

The main goal of classifying data in data mining is to generate a predicted model, by learning from input data. This model can predict and divide the new input data between the classes. Predicted models discover relationships between the attributes that would make it possible to predict the outcome [11].

The first step in classification is prepared data in the training set, which is done in three parts [2]:

- Clean the data: we have to use smoothing techniques to remove the noise from data, and we have to solve the missing data problem by replacing it with the commonly occurring value for that attribute.
- Irrelevant attribute problem: in some databases, the attributes are not related. Therefore, correlation analysis is used to know if there is any connection between attributes.
- Database normalization: The process of organizing (such as the attributes and tables) for reducing data redundancy and improve data integrity is called normalization.

A decision tree is a compact data structure that can be built quickly and provides significant descriptions of the relationships between the input variables and the target attributes. The decision trees can provide accurate predictions if applied to the problems that can be represented with contrapuntal expressions. It can handle continuous and categorical attributes and is very powerful with separate attributes. The decision tree classifier has many advantages over all other classifiers [6]. The first advantage is that the distribution in the decision tree is free, consequently, there is no prior assumption about the previous data distributions. The

second advantage is that the decision tree does not require as much training as the other classifiers. Finally, the most important advantage is that the decision tree is easy to understand. These trees can be a set of rules that describe the connection between the input and target attributes.

Sometimes the decision tree generates some unwanted and meaningless rules while it is growing deeps. This problem called the overfitting problem. The overfitting problem increases the tree size and data noise in the tree, and reduces the efficiency and accuracy of the data. To solve the overfitting problem, we have to use pruning methods to avoid a large tree size. There are two kinds of pruning methods in the decision tree. Post sorting is a kind of pruning that first builds the tree, and then reduces unwanted branches in the decision tree. The second kind of pruning method is pre-sorting. In this kind, the tree keeps on checking in each step of the building phase [12].

Each kind of pruning method has many techniques. For post sorting, there are many methods such as reduced error, error complexity, minimum error, and cost-based. For pre-sorting, there are a minimum number of object and chi-square pruning [13].

Nowadays, data mining is a critical technology in the world because all the information and data in different fields of work such as business and science are loaded and stored in servers and have much valuable information that needs to be instructed. The classification task in data mining is useful in many fields of life, and classifying data and creating the predictive model allows us an opportunity to have advanced knowledge of events. In the classification process, we are dealing with big data sets. Therefore, we have to improve the classification algorithms to be able to handle this size of data and to be more comfortable for implementation. Data size is growing very fast every day, faster than the growth in hardware abilities. To analyse and classify this size of data, we have to make classification algorithms more effective in both the time that they take to classify the data and the data structure of the algorithms.

The data mining process has many tasks like association, classification, clustering, regression, anomaly detection, and summarization. These tasks can formulate the problem and discover knowledge from data. Classification is the most common task in data mining. This technique is used to create a predictive model for future behaviour. Decision tree is the most common classification algorithm in data mining and is easy to implement, has a high degree of accuracy, and has easily recognizable patterns.

There are many algorithms to build the decision tree; each one of them has advantages and disadvantages. However, in data mining, the size of data is very big, thus in the implementation phase, all of these algorithms struggle with the amount of data that must stay in active memory until the mining process ends.

Due to the strain on computing assets caused by the massive amounts of data involved, the ideal classifier should take less time to complete the classification process, thus requiring less computer memory. In this research, we propose a new classifier to enhance classification tasks in data mining, and we will compare our proposed classifier with the SPRINT classifier in parallel computations.

The goal of this research is to design and implement a new classifier to build the decision tree for mining in big data sets. The aim is to make the classification process easier and quicker in both parallel and serial implementation and to decrease the data size that should stay in the memory. In this proposed classifier, we solve the problems in the other classifiers such as runtime and memory requirements, and also we fix the incorrect results obtained by some classifiers that are proposed to improve the runtime and memory requirements.

https://doi.org/10.31436/iiumej.v22i2.1541

In our research, we have to implement two decision tree algorithms to test their runtime and the memory requirements. The training data have to be datasets consisting of several attributes. If we download a data set from the internet, we have to make many changes to it to make it more suited to the parallel programs. Consequently, we created a dataset with three attributes and use it as a training set for the implementation phase.

The rest of this article is as follows. In Section 2, we review the literature survey of related researches in the data classification algorithms. The methodology of our proposed classifier method and a comparison with the best previously proposed parallel classifier have been presented in Section 3. The simulations and experimental results of the proposed algorithm have been presented in Section 4. Finally, in Section 5, we discuss the conclusion of this research and future work.

2. RELATED WORK

Decision trees are one of the most crucial classification algorithms that can be constructed quickly and possess simple and easy-to-use models [6, 14]. The most common decision trees builders are CHAID (CHI-squared Automatic Interaction Detection), ID3 (Iterative Dichotomies 3), C4.5, CART (Classification and Regression Tree) and MARS (Multivariate Adaptive Regression Splines). Each one of them can be implemented in serial and parallel situations, but SLIQ (Scalable Classifier for Data Mining) and SPRINT (Scalable Parallel Classifier for Data Mining) algorithms have a perfect performance in parallel implementation.

CHAID classifier is a decision tree builder for dependent variables proposed in [15]. This classifier uses CHI-squared Automatic Interaction Detection (CHAID), which is a technique for analysing a large amount of data by dividing it into separated subsets. The classifier then detects the interactions between categorized variables of a data set that are the dependent variables.

The CART algorithm is the classification and regression tree published in [16]. This algorithm is used to build a prediction model from data. The prediction model is created by partitioning the data recursively and fitting a simple prediction model in each partition, then, a decision tree is built to represent and describe the partition graphically.

In [17], the authors presented a framework called CLS (Concept Learning System). This system builds a decision tree with the minimum cost of classifying. The cost of classifying consists of two components: the cost of determining the property of an object and the cost of deciding. In [6], Ross Quinlan presented ID3, which is one of the serial algorithms developed from the CLS algorithm, which is used to build the decision tree invented. It handles category attributes only, begins with the original set as the root node, and all the attributes in the set belong to the same class. ID3 algorithm is suffering from an overfitting problem. To solve the overfitting problem, Ross Quinlan created a new algorithm in 1993 and named it the C4.5 algorithm.

In [18], the MARS algorithm, which is a group of techniques implemented together, was presented. This algorithm solves the regression-type problems and has the same purpose as the classification algorithms for predicting the values of the dependent (outcome values) and independent (income values) values.

Most of the classification algorithms are designed for resident memory data, and this issue limits the ability to mine in large data sets. In [19], the SLIQ algorithm was proposed to build the decision tree to solve the resident memory data problem by using a new data structure. In [20], the authors designed SPRINT. This algorithm has the same advantage of SLIQ, but a

different data structure. SPRINT data structure is an attribute list with three rows, the first row is the attribute value, the second row is the class, and the third row is the index of the record. This list is sorted in the first step and only one time.

In [21], the ScalParC algorithm (Scalable Parallel Classifier algorithm) proposed for enhancing the runtime and the memory required in SPRINT. ScalParC has the same data structure of SPRINT, and it creates a separate list for each attribute consisting of the attribute, the class, and the index. It then creates another data structure to stay in the memory during execution time that determines to which node each datum belongs to. The difference between the two algorithms is that ScalParC distributes the data structure, which stays in memory between the processors to increase the memory request and execute the split by level not by not to increase the communication time [21]. However, in 2000, Kevin Bowyer has proved that ScalParC gives incorrect results in some situations [22].

CLOUDS algorithm is a breadth-first strategy to build the decision tree and uses the Gini index for evaluating the split points [23]. It uses either sampling the splitting method or sampling the splitting points with estimation method to determine the splitter at each node of the tree, and it evaluates the split points for categorical attributes as in SPRINT.

In [24], the authors presented a Random Forest classifier. It consists of numbers of simple trees, each one of these trees is capable of producing a predicted response. For the classification tasks, the response of these trees has a class membership form. This task classifies the independent variables in categories. For the regression task, the response is an estimation of the dependent variable, given the predictors. Each tree in the forest has an individual responsibility for the queries. This response depends on the predictor values which have the same distribution for all the trees and are selected independently.

The public classifier is developed from SPRINT and proposed in [25]. It has the same data structure, the same steps, and the phases. The main goal of the public is dealing with data that has much noise and suffers from missing and incorrect values, the kind of problems that can be fixed using the pruning method.

3. METHODOLOGY

The new algorithm has the same data structure for SPRINT. It creates an attribute list for each attribute. This list consists of three rows: the attribute, the class, and the rids. The sorting process is done in the first phase, and it is done once. Our algorithm also uses the hash table to determine the rids in each node and use the Gini index to found the best split point in each attribute list.

In the first step, the parallel pattern attribute records are distributed equally over all the processors in SPRINT and the proposed algorithm. Thus, in this case, each processor processes the same number of records. Figure 2 shows parallel data placement in the proposed algorithm in the first step.

In our algorithm, we distribute the hash table between the processors and create a new data structure called the node table. This data structure determines the number of classes in each node because in the calculation phase, each processor calculates (Gini) for its rids. In this phase, the processor needs information about the rids in each node and the number of each class in the node; thus, this new data structure determines the number of classes in nodes in each level of the tree. Figure 3 shows the new data structure in the proposed algorithm.

https://doi.org/10.31436/iiumej.v22i2.1541

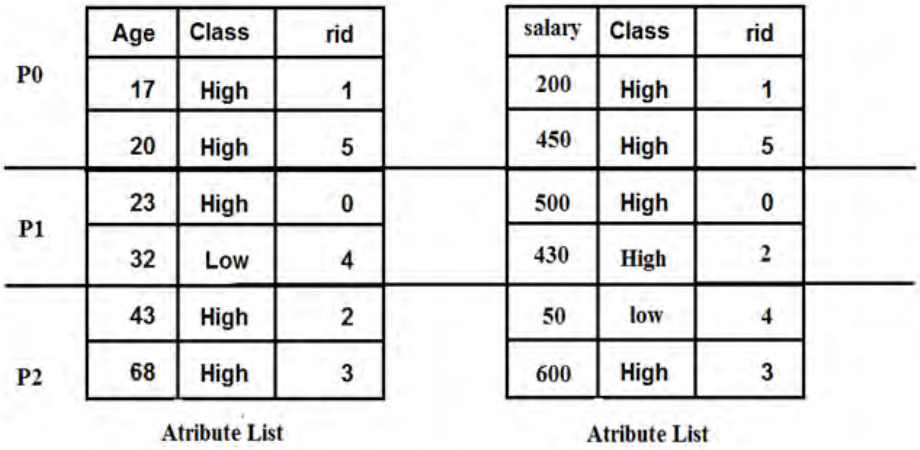


Fig. 2. The parallel data placement in the proposed algorithm.

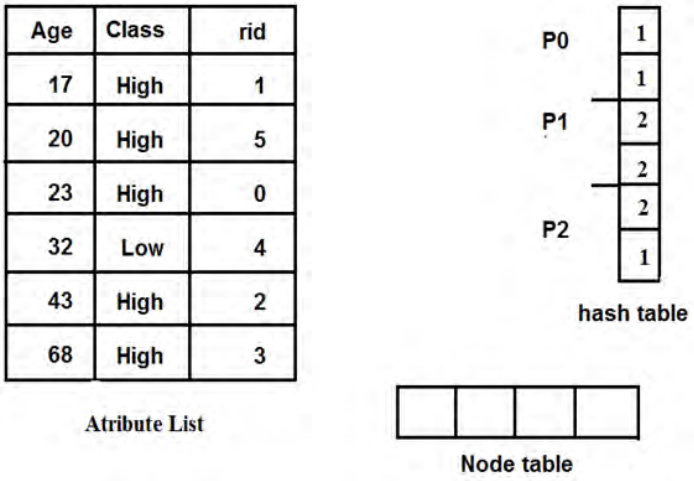


Fig. 3. The new data structure in the proposed algorithm.

The proposed algorithm has many advantages over SPRINT, and it enhances many weak points in SPRINT. In SPRINT, the size of the hash table is equal to the size of the data set, and the hash table is memory resident. Each processor has to frequently connect to the hash table to determine the rids in each node of the tree, and the hash table must be updated after each split.

The total communication overhead (per processor) in SPRINT is O(N), where N is the number of records in the dataset, and the total memory requirement (per processor) is O(N), as the size of the hash table is the same order as the size of the training dataset for the upper levels of the decision tree, and it resides on every processor. Therefore, SPRINT is unscalable in runtime and memory requirements.

The communication cost in SPRINT is very high, especially when the data set is vast. In this case, the size of hash table will be a problem because it should stay in memory and the frequent connection with the hash table increases the runtime of the algorithm. Figure 4 shows the communication cost in SPRINT.

https://doi.org/10.31436/iiumej.v22i2.1541

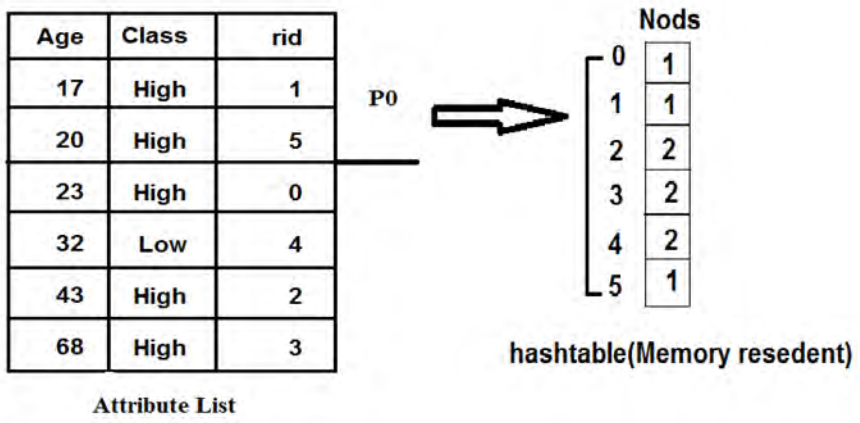


Fig. 4. Communication cost in SPRINT.

In the proposed algorithm, the hash table is distributed between the processors. In the calculation phase, each processor will connect with part of the hash table to determine the rids in the node and will connect to the node table to determine the number of classes in the node. Each processor will update its part of the hash table.

In this way, the total communication; over each processor is O(N/P+M/2), P is the number of processors, and M is the size of the node table (M=number of classes* 2), in this case, we decrease the communication cost and the time of the update. Figure 5 shows the communication cost in the proposed algorithm.

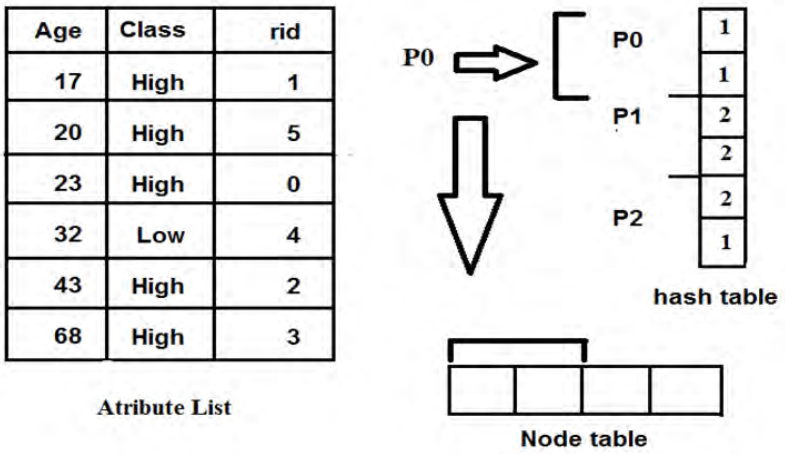


Fig. 5. Communication cost in the proposed algorithm.

In SPRINT, the records in the attribute list are distributed between the processors equally in the first step after the sorting process, and this distribution continues until the last step. Each processor processes the records that were assigned to it in the first step.

This approach caused a communication problem when the records that were assigned to one processor are located in more than one node. In this case, the fast processor will wait until the slower processor finishes its work, increasing the runtime of the classification process.

The reason for the problem is that the process is done per level. Thus, the records in more than one node are evaluated to determine the best split in each node, in this case, waiting time will be considerable because each processor is waiting for another processor to finish its work

https://doi.org/10.31436/iiumej.v22i2.1541

and each processor is evaluating records for more than one node; thus, the creation of a new node will take much more time. Figure 6 shows the communication problem in SPRINT.

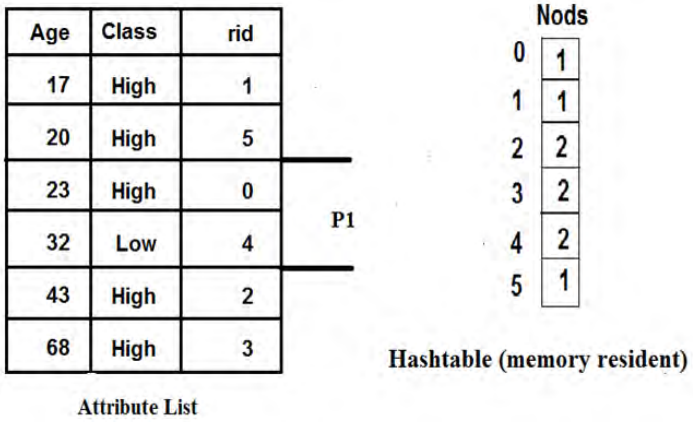


Fig. 6. Communication problem is SPRINT.

In our algorithm, distributing the records between the processors is based on two rules. The first rule is that the distribution is done per node, not per level. Assigning the record is done in each node, therefore, there is no previous record assignation in the nodes and each processor connect with its part of hash table to determine how many records are in its responsibility in each node and the max number of records that each process handles is (N/P).

In this way, the split process in each node is done entirely before going to the next node, and the processors evaluate the split point only for one node, and this will decrease the waiting time in each processor.

To determine the best attribute for a split in the node, each processor connects with its part of the hash table and determine the rids that belong to the node. After determining the rids in the node, each processor performs calculation operations on these rids for the records in remaining attributes to select the best attribute for the next split step. Figure 7 shows the communication per node in the proposed algorithm.

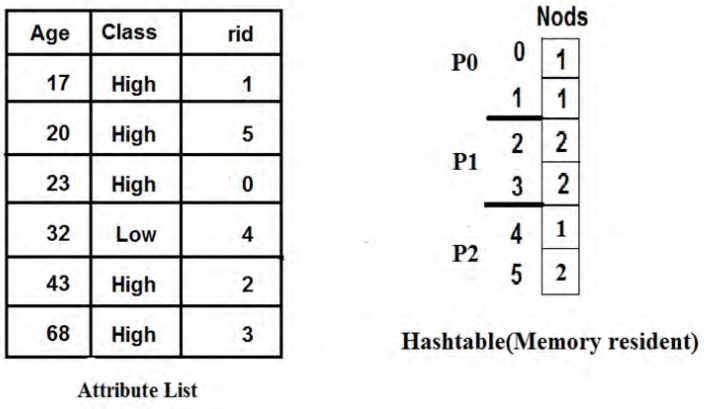


Fig. 7. Communication per node in the proposed algorithm.

The size of the hash table is a critical issue when dealing with a big data set because, in the high level of the tree, the size of the hash table is in the same order as the size of the dataset.

In SPRINT, the entire hash table stays in the memory of each processor, and in each evaluation step, the processor needs to connect to all the elements in the hash table to determine

https://doi.org/10.31436/iiumej.v22i2.1541

the rids in each node. In our proposed algorithm, the hash table is distributed among all the processors; thus, the memory-resident in each processor is (N/P), and the new data structure also stay in memory; thus, the memory-resident of our algorithm is (N/P + M), where M is the size of the node table, and in each evaluate step, the processor connects only with its part of the hash table. Figure 8 shows the memory requirements in both SPRINT and the proposed algorithm.

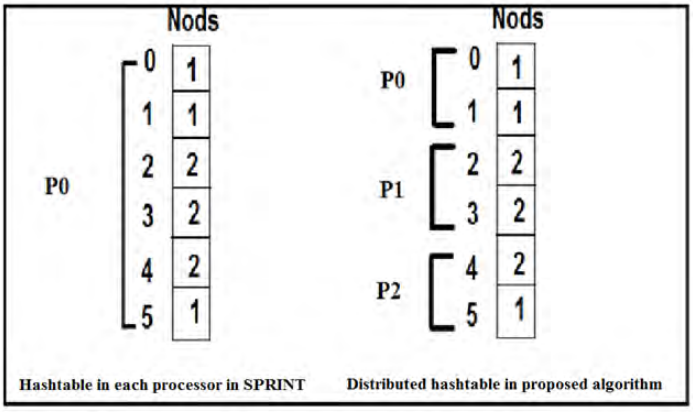


Fig. 8. Hash table (resident memory) in SPRINT and the proposed algorithm.

In SPRINT, each processor has to stay in contact with all the hash tables in each step to determine the rids in each node of the tree. Therefore, all the hash tables should stay in memory until the build process is completed. However, in the proposed algorithm, there is no need to keep the entire hash table in memory, the hash table is divided between the processors, and each one of them connects with its part of the hash table to determine the rids in the tree node.

Processors update the hash table in each level of the tree in the proposed algorithm after the split step. Because the hash table is distributed between the processors, each processor updates its part of the table. Figure 9 shows how hash table parts are updated in each processor.

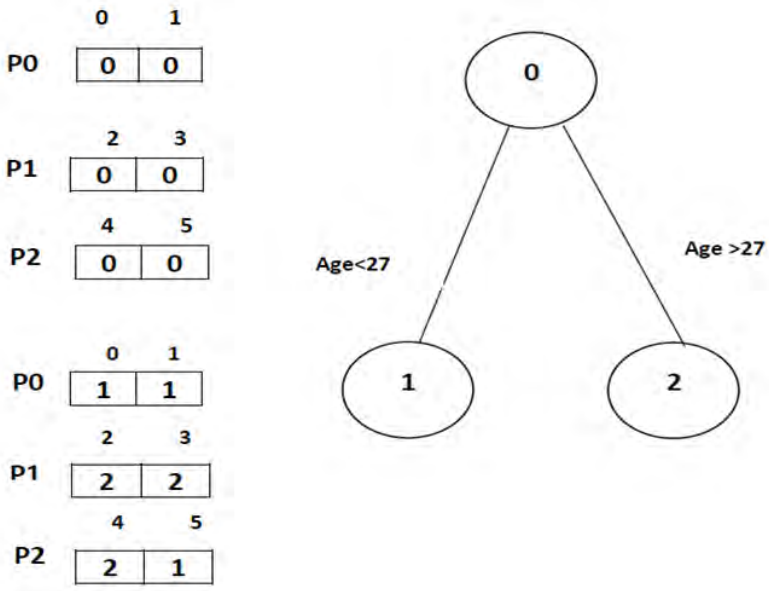


Fig. 9. Update of the hash table in the proposed algorithm.

https://doi.org/10.31436/iiumej.v22i2.1541

Communication cost is one of the most critical issues in the classification process when we are dealing with big data sets. The changes that we have made in the SPRINT algorithm will decrease this cost and make the algorithm more scalable in runtime and memory size. The new data structure will make the algorithm more efficient for massive data sets, and the connection with the hash table will be at the minimum level.

To execute the evaluation step, the split step per node will let the tree be constructed faster, and there will much less waiting time. This approach of work distribution between the processors will make them focus on finishing the tree node by node.

The proposed algorithm data structure and the per-node build strategy will enhance the decision tree building process. The communication problem that is happening when the processor evaluates the records in more than one node in one time will be solved.

4. EXPERIMENTAL RESULTS

4.1. Evaluating the execution time

We have implemented the SPRINT algorithm and the proposed algorithm using MPI, and data sets in up to (10^5) and we used six, nine, and twelve processors in a cluster with total processing hardware of 80 cores/10 nodes and Ram 16 GB, we got the attribute value from the database and created an attribute list for each attribute, then we sorted the attributes.

After the sorting phase, we executed the steps of the algorithm. First, we found the Gini for each attribute, then selected the best split and the split point. Finally, we executed the split and repeated this circle for the remaining attributes. Figure 10 shows the execution time in both SPRINT and the proposed algorithm.

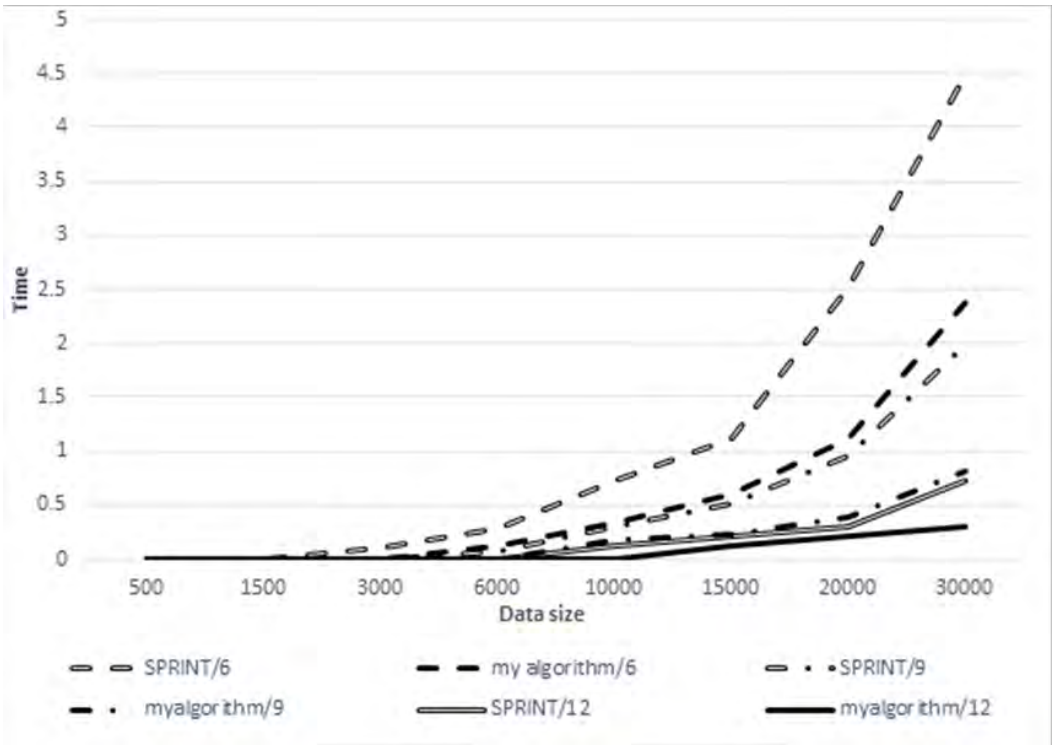


Fig. 10. The execution time (in second).

https://doi.org/10.31436/iiumej.v22i2.1541

From Figure 2, we can divide the different runtime between two algorithms into three parts, the first part is when data is between (500-3000), the proposed algorithm gives a better result than SPRINT but the range of execution time is not significant, the second part when data is from (3000-6000), proposed algorithm also gives a better result than SPRINT and the range in execution time is bigger than the first part. We can see that every time the size of the data gets bigger, the difference range in runtime between the algorithm gets better and the proposed algorithm gives a better result.

In SPRINT, when the data size gets bigger, the hash table does not fit in memory, so we have to do many rounds to update the hash table, these rounds take much time and increase the communication cost because the next step is dependent on the hash table rids to determine the node for each attribute value.

In our algorithm, a hash table is distributed between the processors, so even if the data size gets bigger, the hash table in each processor fits in memory, and we need no more rounds. The results of implementation show that the proposed algorithm is faster than SPRINT in small and big data sets. Table 1 shows the results of implementation.

Table 1: Comparing the runtime of SPRINT and the proposed algorithm (in second).

Data Size (records)	**SPRINT 6**	**SPRINT 9**	**SPRINT 12**	**My algorithm 6**	**My algorithm 9**	**My algorithm 12**
500	0.00	0.00	0.00	0.00	0.00	0.00
1500	0.00	0.00	0.00	0.00	0.00	0.00
3000	0.10	0.00	0.00	0.00	0.00	0.00
6000	0.29	0.08	0.00	0.12	0.00	0.00
10000	0.72	0.3	0.12	0.34	0.18	0.00
150000	1.12	0.51	0.22	0.6	0.24	0.12
20000	2.52	0.96	0.3	1.12	0.39	0.22
30000	4.48	1.98	0.72	2.28	0.82	0.31

4.2. Evaluating the memory requirement

In SPRINT, the size of the hash table is equal to the order of the size of the training dataset for the upper levels of the decision tree. It is resident in all processors and the number of processors does not have any effect on the size of the hash table. Thus, when we are dealing with massive data sets, the hash table size is a series problem.

In the proposed algorithm, the hash table is distributed between the processors, therefore, if the data set order is O(N), then the hash table in each processor is O(N/P); by this solution, we solved the memory requirement problem. Figure 11 shows memory requirements in each processor for both algorithms.

https://doi.org/10.31436/iiumej.v22i2.1541

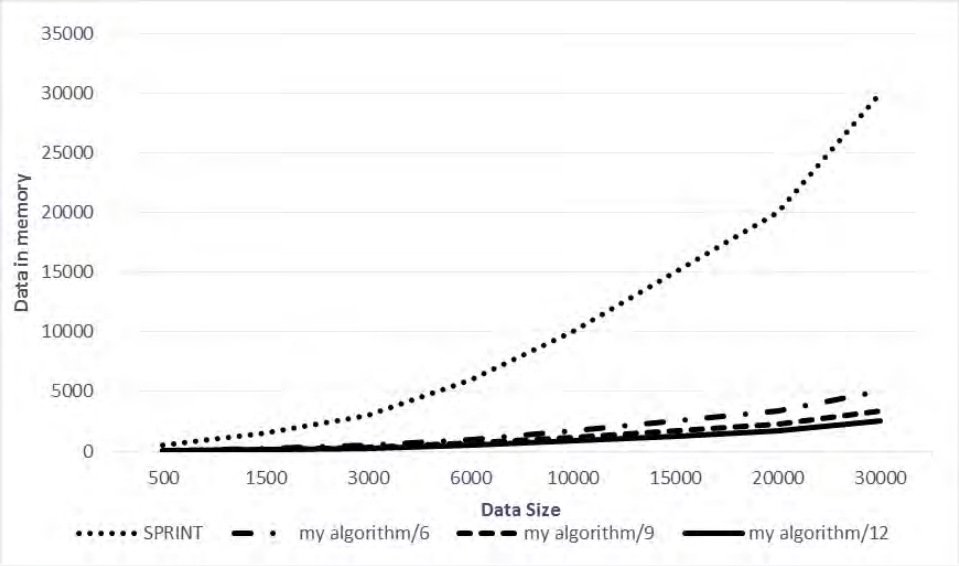


Fig. 11. The memory requirement in each processor.

Data size is a severe problem because hardware cannot handle the fast growth of data size, therefore, developing the existing classification algorithms and making them able to handle a huge data size, without the need of potent hardware, is very important. Table 2 shows the memory requirement in SPRINT and our algorithm.

Table 2: Comparing the memory requirement by the record in each processor.

Data Size (records)	My algorithm 6	My algorithm 9	My algorithm 12	SPRINT
500	88	60	46	500
1500	254	171	129	1500
3000	504	338	254	3000
6000	1004	671	504	6000
10000	1671	1116	838	10000
15000	2504	1671	1254	15000
20000	3338	2227	1671	20000
30000	5004	3338	2504	30000

In the previous experiment, the proposed algorithm showed a better result than SPRINT. It needs a smaller memory size per processor than SPRINT. In our algorithm, the hash table should stay in memory in all the steps, because it is essential to determine the record rids in each node. In SPRINT, the hash table in each node has the same size order as the data set, thus, if the dataset is huge, SPRINT suffers from a memory requirement problem and needs more time to complete the classification process. However, in our algorithm, hash table size is

https://doi.org/10.31436/iiumej.v22i2.1541

distributed equally among the processors. Thus, each processor has (N/P) records which stay in memory, and the node table also stays in memory and has size M which is the number of classes, therefore, the memory requirement for our algorithm is (N/P +M).

5. CONCLUSION AND FUTURE WORK

A decision tree is one of the most popular algorithms in data mining. It has many advantages such as being easy to understand, having a high degree of accuracy, and building the tree very quickly. The decision tree algorithm can handle both categorical and continuous attributes. This algorithm classifies the data set records and builds a predictive model that can be used to predict future outputs.

There are many techniques to build the decision tree such as C4.5, ID3, CART, and CHAID. Each one of them has a unique method and has many advantages, which make the building process easier and more efficient. In data mining, the decision tree building algorithms deal with very big data sets; thus, these algorithms must be able to handle this size of data and consider the ability of the computer's hardware.

Most of the decision tree algorithms are designed for resident memory data, this approach decreases the ability of these algorithms when the data sets are significant. Therefore, many studies have been done in this area, and many algorithms have been proposed to solve the limited ability of the existing algorithms to handle big data sets. SPRINT algorithm is designed to solve the resident memory problem using a new data structure, which decreases the resident memory data size and makes the parallel designing of the decision tree easier.

In this paper, we proposed a new parallel builder for the decision tree, which solves the runtime problem and the memory requirement problem in other algorithms. We tested the new algorithm using MPI and compared the results with the best parallel algorithm, SPRINT.

Our algorithm showed a better result than SPRINT with different sizes of data sets. It also showed a better result in runtime because the communication cost of the split step was decreased from O(N) to O(N/P +M/2) after we distributed the hashtable among the processors.

Distributing the hashtable also solved the memory requirement problem because all the processors share the size of the dataset equally. In this case, the algorithm will not suffer from overfitting problems when the size of the hashtable is too big and there will be no need for many rounds to update the hashtable.

Also in SPRINT, the initial distribution of the attribute lists records between the processors was not changed. This approach caused communication problems when the rids for one processor existed in more than one node. However, in the proposed algorithm, we solved this problem by building the tree node by node; thus, in the evaluation step, each processor evaluated its rids in one node, in this case, the process completed the node and went to another node.

By evaluating the implementation results for SPRINT and the proposed algorithm, we can see that the proposed algorithm is better than the SPRINT in runtime and memory requirements. These advantages make the proposed algorithm more useful when the size of the dataset is significant. Selecting the best algorithm to build the decision tree is based on many rules, the size of the dataset is one of the critical issues that should be considered in the selection process. There are many algorithms that are very useful and powerful when the amount of data is not very big such as C4.5 and CART, but, when the size of dataset is big, SPRINT and the proposed algorithm are better than other algorithms. The new data structure of the proposed algorithm helps to handle this big size, and it is even better than the SPRINT data structure.

In parallel computing, load balancing between the processors is essential, and the data load must be divided equally between the processors, otherwise, some processors will work more than others. In the decision tree, some nodes reach a dead end, and this happens when all the records in the node are related to one class. In this case, there is no need to evaluate and process the records in these nodes. Our algorithm suffers from an unbalancing problem when some nodes reach a dead end, and some processors are processing more than the others, which increases the runtime of the algorithm. For the future work, we will solve the unbalancing problem when some nodes reach a dead end and divide the work equally between the processors to decrease the runtime. This step will enhance the building process and make the proposed algorithm more powerful and more suitable for massive data sets.

Nowadays, data scientists are more interested in deploying emerging and strong deep learning (DL) algorithms [26-30]. DL can perform big data analytics efficiently. It can be widely utilized for handling some traditional problems in Big Data Analytics such as data labelling, semantic indexing, probing complex and nonlinear patterns from giant volumes of data, fast and efficient information retrieval, and simplifying discriminative tasks. As a result, it would be a progressive solution to our work in the future.

REFERENCES

[1] Fayyad, U., G. Piatetsky-Shapiro, and P. Smyth (1996). From data mining to knowledge discovery in databases. AI magazine. 17(3):37-37. https://doi.org/10.1609/aimag.v17i3.1230

[2] Berry, M.J. and G.S. Linoff (2004). Data mining techniques: for marketing, sales, and customer relationship management. John Wiley & Sons.

[3] Tan, P.-N., M. Steinbach, and V. Kumar (2016). Introduction to data mining. Pearson Education India.

[4] Aggarwal, C.C. (2014). An Introduction to Data Classification. Data Classification: Algorithms and Applications. Chapman and Hall/CRC.

[5] Zaki, M.J. and W. Meira (2014). Data mining and analysis: fundamental concepts and algorithms. Cambridge University Press.

[6] Quinlan, J.R (1986). Induction of decision trees. Machine learning. 1(1):81-106. https://doi.org/10.1007/BF00116251.

[7] Rennie, J.D., et al (2003). Tackling the poor assumptions of naive Bayes text classifiers In Proceedings of the Twentieth International Conference on Machine Learning: August 21-24, 2003; Whasington DC. https://dl.acm.org/doi/10.5555/3041838.3041916.

[8] Hagan, M., et al (2014). Neural Network Design. 2nd Edtion. Oklahoma. Martin Hagan.

[9] Jain, A.K. and R.C (1988). Dubes, Algorithms for clustering data. Prentice-Hall, Inc.

[10] Zhang, C. and S. Zhang (2003). Association rule mining: models and algorithms. Springer.

[11] Aggarwal, C (2014). Data Classification: Algorithms and Applications, ser. Frontiers in physics. Chapman and Hall/CRC.

[12] Esposito, F., et al (1997). A comparative analysis of methods for pruning decision trees. IEEE transactions on pattern analysis and machine intelligence. 19(5): 476-491. https://doi.org/10.1109/34.589207.

[13] Mingers, J (1989). An empirical comparison of pruning methods for decision tree induction. Machine learning. 4(2):227-243. https://doi.org/10.1023/A:1022604100933.

[14] Rokach, L. and O.Z. Maimon (2008). Data mining with decision trees: theory and applications. World scientific.

[15] Kass, G.V (1980). An exploratory technique for investigating large quantities of categorical data. Journal of the Royal Statistical Society: Series C (Applied Statistics). 29(2):119-127. https://doi.org/10.2307/2986296.

[16] Breiman, L., et al (1984). Classification and regression trees. CRC press.

[17] Hunt, E.B , J. Marin, and P.J. Stone (1966). Experiments in induction. Academic Press.

[18] Friedman, J.H (1991). Multivariate adaptive regression splines. The annals of statistics, 19(1):1-67. https://doi.org/10.1214/aos/1176347963

[19] Mehta, M., R. Agrawal, and J. Rissanen (1996). SLIQ: A fast scalable classifier for data mining. In Proceedings of the International conference on extending database technology: 25-29 March 1996; Avignon, France; pp 18-32. https://doi.org/10.1007/BFb0014141.

[20] Shafer, J., R. Agrawal, and M. Mehta (1996). SPRINT: A scalable parallel classifier for data mining. In Proceedings of the 22nd VLDB Conference. 3-6 September 1996; Mumbai(Bombay), India; pp 544-555.

[21] Joshi, M.V., G. Karypis, and V. Kumar (1998). ScalParC: A new scalable and efficient parallel classification algorithm for mining large datasets. In Proceedings of the First Merged International Parallel Processing Symposium and Symposium on Parallel and Distributed Processing. 30 March-3 April 1998; Orlando, FL, USA. https://doi.org/10.1109/IPPS.1998.669983.

[22] Bowyer, K.W., et al. A parallel decision tree builder for mining very large visualization datasets. In Proceedings of the ieee international conference on systems, man and cybernetics.'cybernetics evolving to systems, humans, organizations, and their complex interaction. 8-11 Oct 2000; Nashville, TN, USA. https://doi.org/10.1109/ICSMC.2000.886388.

[23] Ranka, S. and V. Singh (1998). CLOUDS: A decision tree classifier for large datasets. In Proceedings of the 4th Knowledge Discovery and Data Mining Conference. 27 – 31 August 1998; New York, USA. https://dl.acm.org/doi/10.5555/3000292.3000294.

[24] Liaw, A. and M. Wiener (2002). Classification and regression by random Forest. R news. 2(3):18-22.

[25] Rastogi, R. and K. Shim (2000). PUBLIC: A decision tree classifier that integrates building and pruning. Data Mining and Knowledge Discovery. 4(4):315-344.

[26] Bahaghighat, M., et al (2020). Estimation of wind turbine angular velocity remotely found on video mining and convolutional neural network. Applied Sciences. 10(10):35-44. https://doi.org/10.3390/app10103544.

[27] Bahaghighat, M., et al (2020). ConvLSTMConv network: A deep learning approach for sentiment analysis in cloud computing. Journal of Cloud Computing: Advances, Systems and Applications. 9(16). https://doi.org/10.1186/s13677-020-00162-1.

[28] Abedini, F., et al (2019). Wind turbine tower detection using feature descriptors and deep learningFacta Universitatis, Series: Electronics and Energetics. 33(1):133-153. https://doi.org/10.2298/FUEE2001133A.

[29] Bahaghighat, M., et al (2019). Vision Inspection of Bottle Caps in Drink Factories Using Convolutional Neural Networks. In Proceedings of the IEEE 15th International Conference on Intelligent Computer Communication and Processing (ICCP). 5 - 7 September 2019; City Plaza, Cluj-Napoca, Romania. https://doi.org/10.1109/ICCP48234.2019.8959737.

[30] Bahaghighat, M., et al (2019). A machine learning based approach for counting blister cards within drug packages. IEEE Access. 7: 83785-83796. https://doi.org/10.1109/ACCESS.2019.2924445.

https://doi.org/10.31436/iiumej.v22i2.1559

INFLUENCE OF LIGHT ABSORPTION PROFILE ON THE PERFORMANCE OF ORGANIC PHOTOVOLTAICS

ABDUL HALIM IKRAM MOHAMED[1] AND MOHD LUKMAN INCHE IBRAHIM[2*]

[1]Department of Electrical and Computer Engineering,
[2]Department of Science in Engineering,
Kulliyyah of Engineering, International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia.

[*]Corresponding author: lukmanibrahim@iium.edu.my

(Received:23rd June 2020; Accepted: 16th January 2021; Published on-line: 4th July 2021)

ABSTRACT: We investigate how an enhanced light absorption at a specific position inside the active layer affects the performance of organic photovoltaic cells (OPVs), namely the short-circuit current density (J_{sc}), the open-circuit voltage (V_{oc}), the fill factor (FF), and the power conversion efficiency (PCE). The performance is calculated using an updated version of a previously published analytical current-voltage model for OPVs, where the updated model allows the light absorption profile to be described by any functions provided that analytical solutions can be produced. We find that the light absorption profile affects the performance through the drift current. When the mobility imbalance is not very high (when the ratio of the mobility of the faster carrier type to the mobility of the slower carrier type is less than about 10^3), the PCE is maximized when the light absorption is concentrated at the center of the active layer. When the mobility imbalance is very high (when the ratio of the mobility of the faster carrier type to the mobility of the slower carrier type is more than approximately 10^4), the PCE is maximized when the light absorption is concentrated near the electrode collecting the slower carrier type. Therefore, it is important to ensure that the light absorption profile is properly tuned so that the performance of OPVs is maximized. Moreover, any efforts that we make to improve the performance should not lead to a light absorption profile that would actually impair the overall performance.

ABSTRAK: Kajian ini menilai bagaimana penyerapan cahaya yang tinggi pada bahagian tertentu lapisan aktif mempengaruhi prestasi sel fotovoltaik organik (OPV), iaitu ketumpatan arus litar pintas (Jsc), voltan litar terbuka (Voc), faktor pengisian (FF), dan kecekapan penukaran kuasa (PCE). Prestasi dikira mengguna pakai model terkini yang diperbaharui dari model asal analitikal OPV voltan-arus, di mana model ini membenarkan mana-mana profil penyerapan cahaya digunakan asalkan penyelesaian analitikal terhasil. Dapatan kajian mendapati profil penyerapan cahaya mempengaruhi prestasi berdasarkan arus hanyut. Apabila ketidakseimbangan pergerakan caj tidak begitu tinggi (di mana nisbah pergerakan pembawa caj laju kepada perlahan adalah kurang daripada 10^3), PCE menjadi maksimum jika penyerapan cahaya bertumpu pada tengah lapisan aktif. Apabila ketidakseimbangan pergerakan caj sangat tinggi (di mana nisbah pergerakan pembawa caj laju kepada perlahan adalah lebih daripada 10^4), PCE menjadi maksimum jika penyerapan cahaya bertumpu pada elektrod yang mengutip pembawa caj perlahan. Oleh itu, kedudukan talaan profil penyerapan cahaya yang tepat adalah sangat penting bagi menentukan prestasi OPV dimaksimumkan. Tambahan, apa sahaja usaha penambahbaikan prestasi seharusnya tidak menyebabkan pengurangan keseluruhan prestasi profil penyerapan cahaya.

https://doi.org/10.31436/iiumej.v22i2.1559

KEYWORDS: *carrier mobility; fill factor; open-circuit voltage; power conversion efficiency; short-circuit current*

1. INTRODUCTION

Organic photovoltaic cells (OPVs) have several unique advantages such as lightweight, high mechanical flexibility, good transparency, and low fabrication cost [1]. Hence, OPVs are viewed as a promising alternative to conventional photovoltaic technologies. For example, OPVs have a great potential to be used for powering various devices such as wearable devices [1], next generation self-driven biomedical devices [2], and off-grid devices for the Internet of Things [3]. The recorded power conversion efficiency (PCE) of OPVs has been improving steadily from 11% [4] several years ago to 17.3% currently [5]. However, the PCE of OPVs is still quite low compared with traditional photovoltaics, and even when compared with another emerging photovoltaic technology, namely perovskite solar cells [4]. Therefore, improving the performance has been a focus of OPV research.

There are many factors that influence the performance of OPVs such as the carrier mobility [6,7], the work function of the electrodes [8,9], the permittivity of the active layer [7,10] and the thickness of the active layer [11-13]. The light absorption profile inside the active layer has been reported to be another factor that influences the performance [14-16]. The light absorption profile is determined by several factors such as the optical properties of the OPV components and the thickness of the active layer [13].

An improvement in the PCE by optimizing the light absorption profile can be achieved in practice since the absorption profile can be controlled to a certain extent. For example, optical spacers can be used to alter the absorption profile [14,15]. Another method that could be used to control the absorption profile is by employing plasmonic nanoparticles, which can be located outside or embedded inside the active layer [17]. Light intensity increases near the plasmonic nanoparticles [17], and hence, the light absorption should be higher near the nanoparticles. Therefore, one could control the absorption profile to a certain extent by choosing the embedment positions of the nanoparticles.

Although understanding the effect of the light absorption profile is clearly important, there is still a gap concerning our understanding on this matter. It has been shown that a high light absorption at a specific position inside the active layer improves the performance [14-16]. In more detail, Mescher et al. [14] and Islam et al. [16] concluded that a higher light absorption at the center of the active layer gives a better performance but did not conclude on the possible effect of carrier mobility imbalance. On the other hand, Tress et al. [15] concluded that the fill factor (FF) improves if more light is absorbed near the contact that collects the slower carrier type even when the mobility imbalance is quite low. Hence, there are still uncertainties regarding this matter. For example, from previous studies, we still have no clue on how the optimum light absorption profile evolves as the carrier mobilities evolve from balanced to highly imbalanced. Furthermore, there is a disagreement between the results of previous studies. When the mobility imbalance is quite low (when the ratio of the mobility of the faster carrier type to the mobility of the slower carrier type is of one order of magnitude), Islam et al. [16] found that the best performance is produced when the light absorption is concentrated at the center of the active layer whereas Tress et al. [15] found that the best performance is produced when the light absorption is concentrated near the electrode collecting the slower carrier type.

In this paper, we will investigate the effect of the light absorption profile on the performance of OPVs with the aim of providing insights on how an enhanced light absorption at a specific position inside the active layer affects the OPV performance at different levels of mobility imbalance, and enlightening the disagreement between the results of previous studies.

2. METHODOLOGY

2.1 Analytical Model

The current-voltage model for OPVs presented in this section is employed for all calculations in this article. The model is essentially the same as a previously published model [18]. Therefore, finer details of the model presented here can be obtained in the published work [18]. The only difference between the model presented here and the previous model is that the charge-transfer (CT) state generation profile here is open to any functions provided that the continuity equations can be solved analytically whereas the CT state generation profile in the previous model has an exponential function. The device structure of the model is illustrated in Fig. 1.

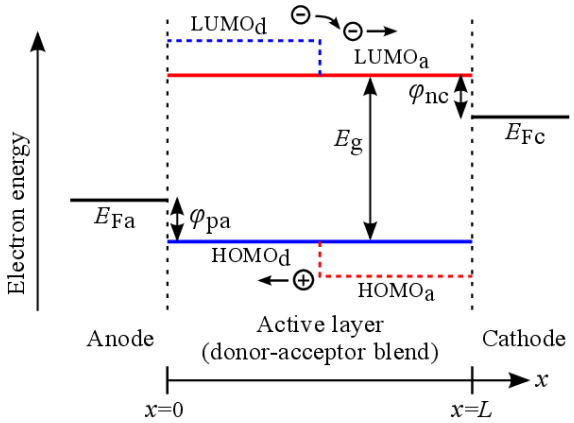


Fig. 1: A schematic showing the energy levels and the device structure. LUMOd and HOMOd denote the lowest unoccupied molecular orbital (LUMO) and the highest occupied molecular orbital (HOMO) of the donor, respectively. LUMOa and HOMOa denote the LUMO and the HOMO of the acceptor, respectively. E_g denotes the effective band gap. E_{Fa} and E_{Fc} denote the Fermi levels of the anode and the cathode, respectively. L is the thickness of the active layer. φ_{pa} and φ_{nc} denote the hole injection barrier at anode and the electron injection barrier at cathode, respectively. In the operation of OPVs, anode is the electrode that collects free holes and cathode is the electrode that collects free electrons.

Light absorption by the active layer generates excitons. When the excitons reach the donor-acceptor interface, CT states may be produced. Free charge carriers are generated from the dissociation of the CT states. When free electrons and free holes recombine bimolecularly, CT states are produced back. Therefore, CT states can be generated due to exciton relaxations at the donor-acceptor interface (through photon absorption) and due to bimolecular recombination. Since bulk heterojunction design is used to construct the active layer, the photogenerated CT states are produced at approximately the same location where the excitons are photogenerated (or where the photons are absorbed). Therefore, the profile of the light absorption can be assumed to have the same equivalent shape as the profile of the CT state photogeneration rate.

https://doi.org/10.31436/iiumej.v22i2.1559

According to the previous model, when the non-geminate recombination is not considered in the electron (hole) continuity equation, the resulting electron (hole) density that arises from the solution of the equation is called the maximum electron (hole) density n_{max} (p_{max}) [18]. The electron and the hole continuity equations at steady state without the non-geminate recombination are given by

$$D_n \frac{\partial^2 n_{max}}{\partial x^2} + \mu_n F \frac{\partial n_{max}}{\partial x} + P_d G_{CT} = 0 \quad (1)$$

$$D_p \frac{\partial^2 p_{max}}{\partial x^2} - \mu_p F \frac{\partial p_{max}}{\partial x} + P_d G_{CT} = 0 \quad (2)$$

where D_n (D_p) is the electron (hole) diffusion coefficient, μ_n (μ_p) is the electron (hole) mobility, F is the electric field, P_d is the dissociation probability of CT states, and G_{CT} is the CT state photogeneration rate per unit volume. The CT state dissociation probability is defined as

$$P_d = \frac{k_d}{k_d + k_f} \quad (3)$$

where k_d is the CT state dissociation rate coefficient, and k_f is the CT state decay rate coefficient. As in the previous model [18], k_d given by the work of Inche Ibrahim [19] is used in this article, which is an improvement to the Onsager-Braun model [20].

Solving Eq. (1) and Eq. (2) are the same as solving non-homogenous second-order linear ordinary differential equations with constant coefficients where D_n, D_p, μ_n, μ_p, F, and P_d are independent of x [18] with $P_d G_{CT}$ being the non-homogenous term. Any functions can be assigned to the non-homogenous term (i.e. to G_{CT} since P_d is a constant) provided that Eq. (1) and Eq. (2) can be solved analytically. The boundary conditions as given in the previous model [18] are used and MATLAB is employed to analytically solve Eq. (1) and Eq. (2) to obtain n_{max} and p_{max}.

Furthermore, according to the previous model [18], when the non-geminate recombination is considered in the electron (hole) continuity equation, the resulting electron (hole) density that arises from the solution of the equation is called the net electron (hole) density n_{net} (p_{net}). The electron and the hole continuity equations at steady state with the non-geminate recombination considered are therefore given by

$$D_n \frac{\partial^2 n_{net}}{\partial x^2} + \mu_n F \frac{\partial n_{net}}{\partial x} + P_d G_{CT} - (1 - P_d) R_b - R_{mn} = 0 \quad (4)$$

$$D_p \frac{\partial^2 p_{net}}{\partial x^2} - \mu_p F \frac{\partial p_{net}}{\partial x} + P_d G_{CT} - (1 - P_d) R_b - R_{mp} = 0 \quad (5)$$

where R_b is the non-geminate bimolecular recombination rate per unit volume, and R_{mn} (R_{mp}) is the non-geminate monomolecular recombination rate per unit volume for electrons (holes). R_b is given by

$$R_b = \gamma k_L n_{max} p_{max} \quad (6)$$

where γ is the bimolecular recombination reduction coefficient, and k_L is the Langevin recombination coefficient [18]. R_{mn} and R_{mp} are given by

$$R_{mn} = k_{mn} n_{max} \quad (7)$$

$$R_{mp} = k_{mp} p_{max} \quad (8)$$

where k_{mn} (k_{mp}) is the monomolecular recombination coefficient for electrons (holes). It is important to note that $R_{mn} = R_{mp}$ as explained in the previous model [18]. Solving Eq.

https://doi.org/10.31436/iiumej.v22i2.1559

(4) and Eq. (5) are also the same as solving second-order linear ordinary differential equations with constant coefficients. Again, the boundary conditions as given in the previous model [18] are used and MATLAB is employed to analytically solve Eq. (4) and Eq. (5) to obtain n_{net} and p_{net}.

The electron current density J_n and the hole current density J_p are given by

$$J_n = q\mu_n F n_{net} + qD_n \frac{\partial n_{net}}{\partial x} \tag{9}$$

$$J_p = q\mu_p F p_{net} - qD_p \frac{\partial p_{net}}{\partial x} \tag{10}$$

The total current density J is the sum of J_n and J_p. The first terms on the right sides of Eq. (9) and Eq. (10) are the electron drift current density $J_{n,dr}$ and the hole drift current density $J_{p,dr}$, respectively, whereas the second terms on the right sides of Eq. (9) and Eq. (10) are the electron diffusion current density $J_{n,diff}$ and the hole diffusion current density $J_{p,diff}$, respectively.

2.2 Light Absorption Profiles and Parameter Values

To achieve the objective of this paper, we need to use the light absorption profile (and hence, the G_{CT} profile) that has an enhanced absorption (and hence, an enhanced CT state photogeneration rate) at a specific position inside the active layer. To produce such a profile, we use G_{CT} given by

$$G_{CT} = G_0 + G_1(x - C)^2 \tag{11}$$

where G_0 and G_1 are the CT state photogeneration rate parameters, and C is the location of the peak CT state photogeneration rate (or the location of the peak light absorption). The CT state photogeneration rate per unit area of the active layer is $G_{CT,area} = \int_0^L G_{CT}\, dx$. Hence,

$$G_{CT,area} = G_0 L + \frac{1}{3} G_1 [(L - C)^3 + C^3] \tag{12}$$

To ensure a fair comparison, each of the studied G_{CT} profile must have the same the value of $G_{CT,area}$. When $C \leq L/2$, we impose the condition $G_{CT}(x = L) = 0$. Therefore, when $C \leq L/2$, we have

$$G_1 = \frac{3G_{CT,area}}{C^3 + (L-C)^3 - 3L(L-C)^2} \tag{13}$$

$$G_0 = -G_1(L - C)^2 \tag{14}$$

When $C > L/2$, we impose the condition $G_{CT}(x = 0) = 0$. Therefore, when $C > L/2$, we have

$$G_1 = \frac{3G_{CT,area}}{C^3 + (L-C)^3 - 3LC^2} \tag{15}$$

$$G_0 = -G_1 C^2 \tag{16}$$

In this paper, four different values of μ_n/μ_p (ratio of the electron mobility to the hole mobility) are considered. Table 1 shows the values of μ_n/μ_p together with the corresponding values of μ_n and μ_p used in this study. The values of the other parameters are shown in Table 2. The parameter values in Table 2 are typical for OPVs [11,18], particularly OPVs based on P3HT:PCBM blend [21]. Figure 2 illustrates several G_{CT} profiles used in this study.

https://doi.org/10.31436/iiumej.v22i2.1559

Table 1: Values of the electron mobility μ_n and the hole mobility μ_p, and the resulting μ_n/μ_p used in the calculations

μ_n ($m^2V^{-1}s^{-1}$)	μ_p ($m^2V^{-1}s^{-1}$)	μ_n/μ_p
1×10^{-7}	1×10^{-7}	1
9×10^{-7}	9×10^{-9}	10^2
9×10^{-6}	9×10^{-10}	10^4
3×10^{-5}	3×10^{-10}	10^5

Table 2: Parameter values used in the calculations unless otherwise specified. The symbols in this table are the same as in the previous works [11,18]

Symbols	Parameter Description	Value
E_g	Effective band gap	1.1 eV
N_c, N_v	Density of states	2×10^{26} m^{-3}
μ_{na}	Actual electron mobility	$200\mu_n$
μ_{pa}	Actual hole mobility	$200\mu_p$
ε	Effective permittivity of the active layer	3×10^{-11} $F \cdot m^{-1}$
φ_{na}, φ_{nc}	Injection barriers	0.05 eV
k_f	CT state decay rate coefficient	1×10^8 s^{-1}
a	Electron-hole separation of the CT state	1.8×10^{-9} m
T	Temperature	300 K
λ	Donor-acceptor morphology parameter	0.15
L	Active layer thickness	100 nm
$G_{CT,area}$	CT state photogeneration rate per unit area	1×10^{21} $m^{-2}s^{-1}$
γ	Bimolecular recombination reduction coefficient	0.002
k_{mn}	Monomolecular recombination coefficient for electrons	200 s^{-1}

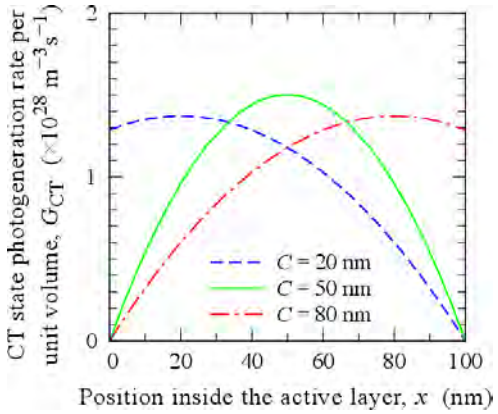


Fig. 2: Several profiles of the CT state photogeneration rate per unit volume G_{CT} that are used in this study. The expression for G_{CT} is given by Eq. (11). G_{CT} profile with light absorption peaks C at 0 nm, 40 nm, 60 nm and 100 nm are also used but are not shown in the figure.

3. RESULTS AND DISCUSSION

In this section, we present and discuss the results of our calculations. Different values of the carrier mobilities affect the OPV performance as shown by previous studies [6,7,18]. Therefore, it is imperative that we do not compare between performances of an OPV with a given μ_n/μ_p and an OPV with a different μ_n/μ_p. What we should analyze

https://doi.org/10.31436/iiumej.v22i2.1559

here is how OPVs with different values of μ_n/μ_p respond to different G_{CT} profiles. Furthermore, for clarity and brevity of this study, we only confine our analysis from the theoretical perspective. We cannot fully control all device parameters when conducting experiments. For example, every time we use a different active layer material in order to change the carrier mobilities, we would also inevitably change a few other device parameters (e.g. the light absorption properties, and thus the value of $G_{CT,area}$). Therefore, the inclusion of experimental data into our analysis would complicate our explanation and understanding on this matter.

3.1 Effect on Short-Circuit Current Density

Figure 3 shows the magnitude of the short-circuit current density ($|J_{sc}|$) as a function of C for each of the studied μ_n/μ_p.

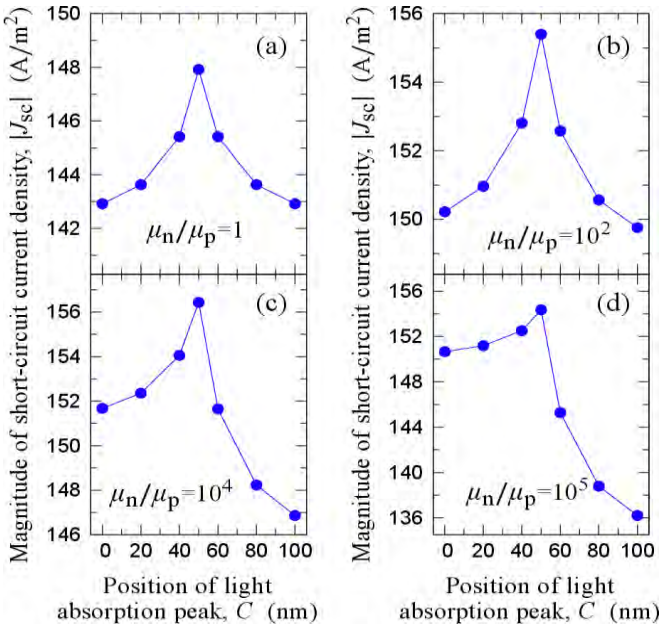


Fig. 3: Magnitude of the short-circuit current density $|J_{sc}|$ as a function of the position of the light absorption peak C when (a) $\mu_n/\mu_p = 1$, (b) $\mu_n/\mu_p = 10^2$, (c) $\mu_n/\mu_p = 10^4$, and (d) $\mu_n/\mu_p = 10^5$. If μ_n/μ_p is increased to a value significantly above 10^5, $|J_{sc}|$ produced by the light absorption profile with $C < L/2$ (e.g. $C = 20$ nm) would even surpass $|J_{sc}|$ produced by the light absorption profile with $C = L/2$.

To explain the trend shown in Fig. 3, let us explore the role of the drift and the diffusion currents. We define the average electron drift current density $\langle J_{n,dr} \rangle$ and the average electron diffusion current density $\langle J_{n,diff} \rangle$ as

$$\langle J_{n,dr} \rangle = q\mu_n F \int_0^L \frac{n_{net}}{L} dx = q\mu_n F \langle n_{net} \rangle \tag{17}$$

$$\langle J_{n,diff} \rangle = qD_n \int_0^L \left(\frac{\partial n_{net}}{\partial x} \right) dx/L = qD_n \langle \frac{\partial n_{net}}{\partial x} \rangle \tag{18}$$

where $\langle n_{net} \rangle$ is the average n_{net} and $\langle \partial n_{net}/\partial x \rangle$ is the average $\partial n_{net}/\partial x$.

Table 3 shows $\langle J_{n,dr} \rangle$'s and $\langle J_{n,diff} \rangle$'s at short-circuit for the balanced mobility case ($\mu_n/\mu_p = 1$). As seen in Table 3, as the light absorption is concentrated farther from the cathode (i.e. as C is lowered), $\langle J_{n,diff} \rangle$ remains unchanged because $\langle \partial n_{net}/\partial x \rangle$ is

unchanged, but the magnitude of $\langle J_{n,dr} \rangle$ increases because $\langle n_{net} \rangle$ increases. Even when we assume there is no non-geminate recombination in our calculations (by using $\gamma = 0$ and $k_{mn} = 0$ s^{-1}), we find that $\langle n_{net} \rangle$ still increases if the light absorption is concentrated farther from the cathode. Therefore, the non-geminate recombination is not the deciding factor why $\langle n_{net} \rangle$ increases as C is lowered.

Table 3: The average electron drift current density $\langle J_{n,dr} \rangle$ and the average electron diffusion current density $\langle J_{n,diff} \rangle$ at short-circuit for the balanced mobility case ($\mu_n/\mu_p = 1$) calculated using G_{CT} with different light absorption peak positions C's.

C (nm)	$\langle J_{n,dr} \rangle$ (A·m^{-2})	$\langle J_{n,diff} \rangle$ (A·m^{-2})
20	-1.19848×10^5	1.19762×10^5
50	-1.19836×10^5	1.19762×10^5
80	-1.19821×10^5	1.19762×10^5

To rationalize why $\langle n_{net} \rangle$ increases as the light absorption is concentrated farther from the cathode (the electrode collecting free electrons), we consider the following two basic facts. First, the profile of the free electrons per unit volume at steady state (which is n_{net} profile) is attained after some of the generated electrons are redistributed within the active layer (where some electrons may be (accidentally) extracted from the active layer during the redistribution process) and some of the electrons are involved in the non-geminate recombination. However, we can ignore the effect of recombination since we have shown that it is not the deciding factor. Second, the cathode has a higher Fermi level than the anode, and thus the resulting boundary conditions [18] mean that the free electrons (holes) must have a significantly higher concentration near the cathode (anode) than near the anode (cathode).

Now consider the $C = 20$ nm and $C = 80$ nm cases as shown in Table 3. In both cases, the electric field magnitudes are the same since both cases are at the same applied voltage V_a (i.e. at short-circuit). For the $C = 20$ nm case, there should be less free electrons extracted from the active layer in attaining n_{net} than for the $C = 80$ nm case since majority of the electrons are generated farther from the cathode, and thus are less likely to be extracted during the redistribution process. This means more of the generated electrons are retained inside the active layer, thus causing a higher n_{net} for the $C = 20$ nm case than for the $C = 80$ nm case. Therefore, the $C = 20$ nm case would give a higher magnitude of $\langle J_{n,dr} \rangle$ than the $C = 80$ nm case due to a higher magnitude of n_{net} (see Eq. (17)).

From our analysis above, we can conclude that the farther the light absorption is concentrated from the cathode (anode), the more the free electrons (holes) are retained inside the active layer, and this gives a higher $|J_{n,dr}|$ ($|J_{p,dr}|$), but $|J_{n,diff}|$ ($|J_{p,diff}|$) is effectively unchanged because the gradient of n_{net} (p_{net}) is effectively unchanged. Therefore, the farther the light absorption is concentrated from the cathode (anode), the higher the $|J_n|$ ($|J_p|$).

Based on the conclusion above, the results in Fig. 3(a) can be rationalized as follows. When $C = 0$ nm (i.e. when C is the closest to the anode or the farthest from the cathode), $|J_n|$ is at its highest whereas $|J_p|$ is at its lowest. As we start increasing C from $C = 0$ nm, $|J_p|$ increases whereas $|J_n|$ decreases, where the increase in $|J_p|$ is stronger than the decrease in $|J_n|$, and hence $|J_{sc}|$ (i.e. $|J|$ at open-circuit) increases. Then, at one point, the

decrease in $|J_n|$ equals to the increase in $|J_p|$, and at this optimum C, we have the peak $|J_{sc}|$. Then, as we further increase C beyond the optimum C, $|J_{sc}|$ decreases since the decrease in $|J_n|$ is stronger than the increase in $|J_p|$. For $\mu_n/\mu_p = 1$ case, the contribution from electrons (therefore $|J_n|$) and holes (therefore $|J_p|$) towards $|J_{sc}|$ are equal (since electrons and holes have the same mobilities), and therefore $|J_{sc}|$ as a function of C is symmetric about $C = L/2$ (i.e. the optimum C).

When $\mu_n/\mu_p = 10^2$, the contribution of $|J_n|$ on $|J_{sc}|$ is stronger than the contribution of $|J_p|$ on $|J_{sc}|$ due to the fact that the electron mobility is higher now, and since $|J_n|$ is higher if C is closer to the anode, therefore $|J_{sc}|$ produced by the profile with $C < L/2$ (i.e. C closer to the anode) is higher than $|J_{sc}|$ produced by the profile with $C > L/2$. This explains the results shown in Fig. 3(b).

As μ_n/μ_p is increased and becomes more imbalanced, the contribution of $|J_n|$ on $|J_{sc}|$ becomes even more significant, and this widens the difference between $|J_{sc}|$ produced by the profile with $C < L/2$ and $|J_{sc}|$ produced by the profile with $C > L/2$. This explains the trends as we increase μ_n/μ_p from 1 to 10^5 as shown in Fig. 3. Due to these trends, we can expect that when μ_n/μ_p reaches a threshold value (when $\mu_n/\mu_p \gg 10^5$), $|J_{sc}|$ produced by the profile with $C < L/2$ would even surpass $|J_{sc}|$ produced by the profile with $C = L/2$. It is worth noting that the explanation presented in section 3.1 is applicable at any given applied voltage V_a in general, and not just at the short-circuit.

3.2 Effect on Open-Circuit Voltage

Figure 4 shows the open-circuit voltage V_{oc} as a function of C for each of the studied μ_n/μ_p. Note that at any given μ_n/μ_p, the trend of V_{oc} as a function of C is determined by the trend of $|J|$ (the magnitude of the total current density) as a function of C at applied voltage V_a near but lower than V_{oc}. This is because, at a given V_a that is near but lower than V_{oc}, a higher $|J|$ means a higher V_{oc} is expected since a higher extra V_a is required in order to push and make $|J| = 0$. For example, for the $\mu_n/\mu_p = 1$ case at $V_a = 0.7$ V as shown in Fig. 5(a), the profile with $C = 50$ nm has a higher $|J|$ than the $|J|$ for the profile with $C = 20$ nm, and hence the profile with $C = 50$ nm has a higher V_{oc} than the V_{oc} for the profile with $C = 20$ nm. Therefore, the trend of V_{oc} shown in Fig. 4(a) can be explained by the trend of $|J|$ shown in Fig. 5(a), where the results shown in Fig. 5(a) can be rationalized in the same way as we rationalize the results shown in Fig. 3(a).

At high V_a (near but below V_{oc}), the electric field magnitude is significantly lower and has significantly less influence in transporting and extracting the carriers than at short-circuit. Hence, at high V_a, the light absorption profile (and thus C) becomes significantly more important in determining n_{net} and p_{net}, and thus in determining $|J_n|$, $|J_p|$ and $|J|$, than at short-circuit. Consequently, $|J|$ at high V_a is more sensitive to C compared with $|J|$ at short-circuit (i.e. $|J_{sc}|$), and therefore the trend of $|J|$ at high V_a as a function of C for a given μ_n/μ_p (where $\mu_n/\mu_p > 1$) is equivalent to the trend of $|J_{sc}|$ as a function of C but at a significantly higher μ_n/μ_p. That is why at a high V_a, the $|J|$ produced by the profile with $C < L/2$ already surpasses the $|J|$ produced by the profile with $C = L/2$ when $\mu_n/\mu_p = 10^4$ (see Fig. 5(c)) whereas at $V_a = 0$ V (i.e. short-circuit), $|J_{sc}|$ produced by the profile with $C < L/2$ still does not surpass $|J_{sc}|$ produced by the profile with $C = L/2$ when $\mu_n/\mu_p = 10^4$ (see Fig. 3(c)). Figure 5 shows the trends of $|J|$ at high V_a as a function of C for each of the studied μ_n/μ_p, and the results shown in Fig. 5 give rise to the trends of V_{oc} as a function of C for each of the studied μ_n/μ_p which is shown in Fig. 4. As mentioned at

https://doi.org/10.31436/iiumej.v22i2.1559

the end of section 3.1, the argument that we use to rationalize the results shown in Fig. 3 is applicable at any V_a, and therefore, the same argument can be used to rationalize the results shown in Fig. 5.

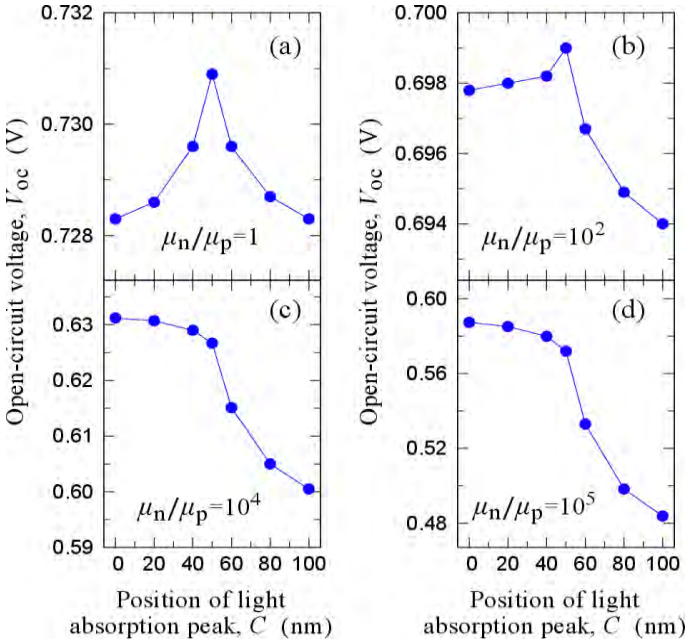


Fig. 4: Open-circuit voltage V_{oc} as a function of the position of the light absorption peak C when (a) $\mu_n/\mu_p = 1$, (b) $\mu_n/\mu_p = 10^2$, (c) $\mu_n/\mu_p = 10^4$, and (d) $\mu_n/\mu_p = 10^5$.

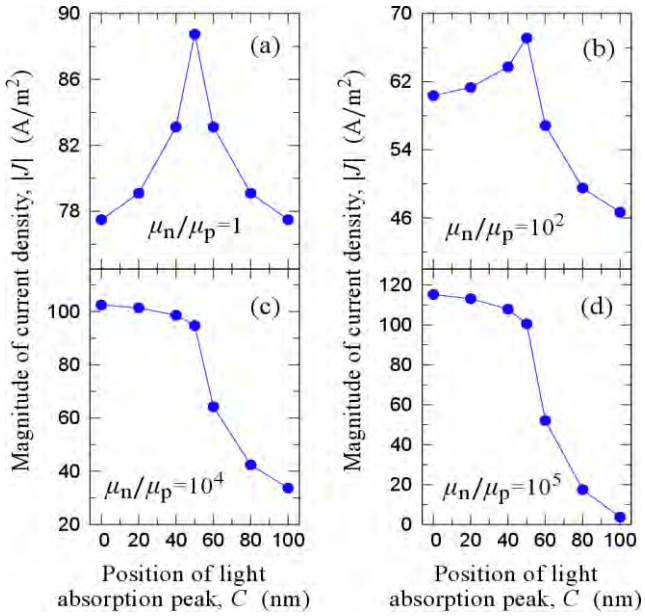


Fig. 5: Magnitude of the current density $|J|$ as a function of the position of the light absorption peak C for (a) $\mu_n/\mu_p = 1$ at $V_a = 0.7$ V, (b) $\mu_n/\mu_p = 10^2$ at $V_a = 0.68$ V, (c) $\mu_n/\mu_p = 10^4$ at $V_a = 0.58$ V, and (d) $\mu_n/\mu_p = 10^5$ at $V_a = 0.48$ V. The applied voltage V_a used for each μ_n/μ_p are chosen randomly but is close to the corresponding V_{oc} (e.g. for $\mu_n/\mu_p = 1$ case, V_{oc} is around 0.73V as shown in Fig. 4(a), so we consider $V_a = 0.7$ V here).

https://doi.org/10.31436/iiumej.v22i2.1559

3.3 Effect on Fill Factor

Figure 6 shows the fill factor (FF) as a function of C for each of the studied μ_n/μ_p. To rationalize the trend of the FF as a function of C, it is important to note that the FF occurs somewhere between the short-circuit and the open-circuit. Hence, the trend of the FF as a function of C at a given μ_n/μ_p is simply somewhere between the trend of $|J_{sc}|$ and the trend of V_{oc} as functions of C for the same μ_n/μ_p. For the device considered in this study, the FF occurs at V_a that is significantly closer to V_{oc} than to the short-circuit, and hence the trends of the FF as a function of C as shown in Fig. 6 is very similar to the trends of V_{oc} as a function of C as shown in Fig. 4.

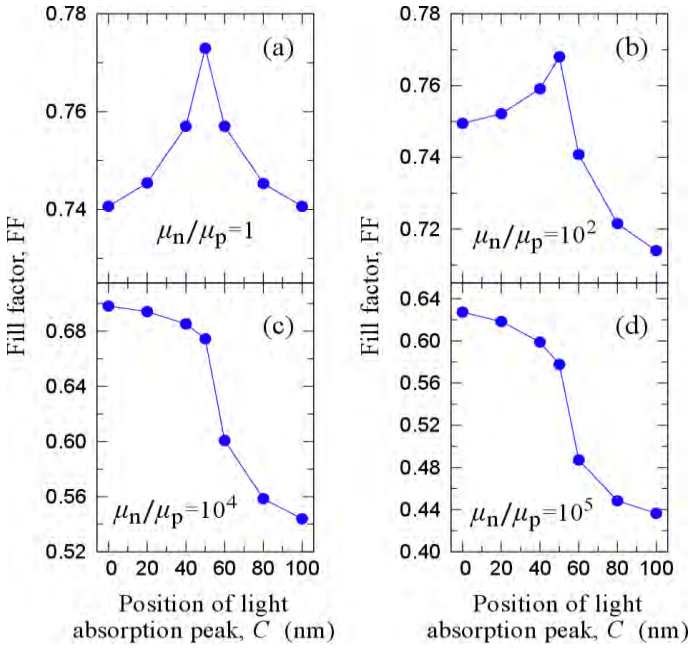


Fig. 6: Fill factor (FF) as a function of the position of the light absorption peak C when (a) $\mu_n/\mu_p = 1$, (b) $\mu_n/\mu_p = 10^2$, (c) $\mu_n/\mu_p = 10^4$, and (d) $\mu_n/\mu_p = 10^5$.

Mescher et al. [14] concluded that enhanced light absorption at the two edges of the active layer (i.e. near the anode and the cathode) lead to a high unfavorable diffusion current which reduces the FF, whereas an enhanced absorption at the center of the active layer leads to a high FF. These conclusions partially agree with our results here. First, we find the drift current is the one that affects the dependence of $|J|$, and thus the dependences of $|J_{sc}|$, V_{oc} and FF on the light absorption profile, whereas the diffusion current basically has no effect (see section 3.1). Second, we find that an enhanced light absorption at the center of the active layer leads to the best FF only if the mobility imbalance (i.e. μ_n/μ_p) is not very high (see Fig. 6).

Tress et al. [15] concluded that an enhanced absorption near the electrode collecting the slower carrier type would give a higher FF even when the mobility imbalance is very low, whereas Islam et al. [16] concluded that when the mobility imbalance is low, an enhanced absorption at the center of the active layer gives the best FF. We find that an enhanced absorption near the electrode collecting the slower carrier type (hole is the slower carrier type in our study here) would give the best FF only when the mobility imbalance is very high (see Fig. 6). When the mobility imbalance is not very high, the best FF is produced when the light absorption is concentrated at the center of the active layer. Therefore, our results here agree with the results from Islam et al. [16], but it is worth

https://doi.org/10.31436/iiumej.v22i2.1559

noting that they did not investigate the role of mobility imbalance on the optimum light absorption profile.

3.4 Effect on Power Conversion Efficiency

Figure 7 shows the power conversion efficiency (PCE) as a function of C for each of the studied μ_n/μ_p. The PCE is given by

$$PCE = FF \times |J_{sc}| \times V_{oc}/P_{in} \tag{19}$$

where P_{in} is the input power per unit area of the incident light (taken to be under the 1 sun condition which is 1000 W/m^2). Hence, the trend of the PCE as a function of C at a given μ_n/μ_p is simply the combination of the trends of $|J_{sc}|$, V_{oc} and FF as functions of C for the same μ_n/μ_p. Therefore, the trends of the PCE shown in Fig. 7 can be understood by combining the trends shown in Fig. 3, Fig. 4, and Fig. 6.

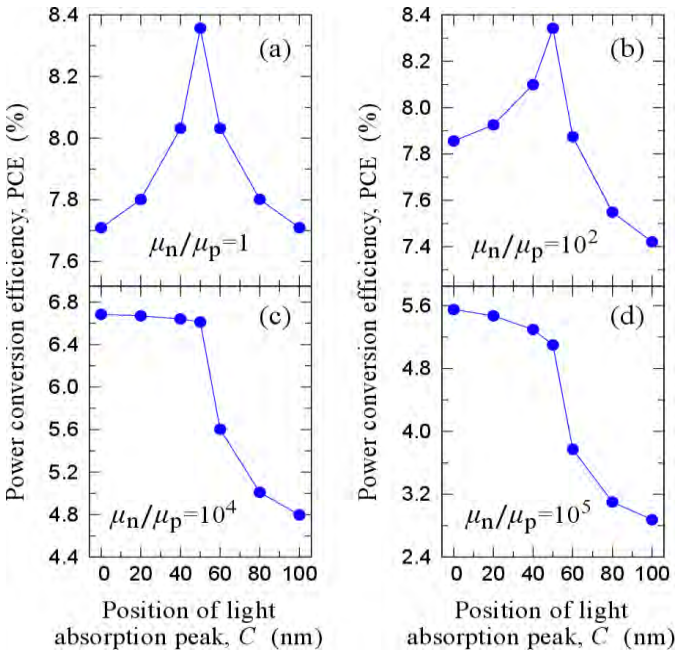


Fig. 7: Power conversion efficiency (PCE) as a function of the position of the light absorption peak C when (a) $\mu_n/\mu_p = 1$, (b) $\mu_n/\mu_p = 10^2$, (c) $\mu_n/\mu_p = 10^4$, and (d) $\mu_n/\mu_p = 10^5$.

According to Mescher et al. [14] and Islam et al. [16], the highest PCE is produced when the light absorption is concentrated at the center of the active layer. However, the conclusion is incomplete according to our results. We find that the highest PCE is produced by the light absorption that concentrates at the center of the active layer only when the mobility imbalance is not very high (see Fig. 7). However, when the mobility imbalance is very high (see Fig. 7), the best PCE is produced when the light absorption is concentrated near the electrode collecting the slower carrier type.

4. CONCLUSION

By using our updated current-voltage model for OPVs, we have investigated how the light absorption profile with an enhanced absorption at a certain position inside the active layer affects the performance of OPVs. It is found that the light absorption profile affects the OPV performance through the drift current. The further the light absorption is

concentrated from the electrode collecting a given charge carrier type, the higher the drift current and the total current for that carrier type, whereas the diffusion current for that carrier type is unaffected. When the carrier mobilities are balanced, the best $|J_{sc}|$, V_{oc}, FF and PCE are produced by the light absorption that concentrates at the center of the active layer. When the mobilities become imbalanced, the $|J_{sc}|$, V_{oc}, FF and PCE produced by the light absorption that concentrates nearer to the electrode collecting the slower carrier type improve relative to the ones produced by the light absorption that concentrates at other positions inside the active layer. When the mobility imbalance is high enough (i.e. reach a threshold value), the best $|J_{sc}|$, V_{oc}, FF and PCE are produced by the light absorption that concentrates near the electrode collecting the slower carrier type. The mobility imbalance threshold values for $|J_{sc}|$, V_{oc}, FF, and PCE are different, where the threshold value for V_{oc} is the lowest, whereas the threshold value for $|J_{sc}|$ is the highest. Therefore, if the mobility imbalance is not very high (when the ratio of the mobility of the faster carrier to the mobility of the slower carrier is less than about 10^3), it is important to ensure that the light absorption is concentrated at the center of the active layer in order to maximize the PCE. However, if the mobility imbalance is very high (when the ratio of the mobility of the faster carrier to the mobility of the slower carrier is about 10^4 or more), the light absorption should be concentrated near the electrode collecting the slower carrier type in order to maximize the PCE.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (grant no. FRGS/1/2017/STG02/UIAM/03/2).

REFERENCES

[1] Scharber MC, Sariciftci NS. (2013) Efficiency of bulk-heterojunction organic solar cells. Progress in Polymer Science, 38(12): 1929-1940. doi: 10.1016/j/progpolymsci.2013.05.001.

[2] Park Sungjun, Heo SW, Lee W, Inoue D, Jiang Z, Yu K, Jinno H, Hashizume D, Sekino M, Yokota T, Fukuda K, Tajima K, Someya T. (2018) Self-powered ultraflexible electronics via nano-grating-patterned organic photovoltaics. Nature, 561(7724): 516-521. doi: 10.1038/s41586-018-0536-x.

[3] Teng N, Yang S, Chen F. (2018) Plasmonic - Enhanced organic photovoltaic devices for low-power light applications. IEEE Journal of Photovoltaic, 8(3): 752-756. doi: 10.1109/JPHOTOV.2018.2797975.

[4] Green MA, Emery K, Hishikawa Y, Warta W, Dunlop ED. (2016) Solar cell efficiency tables. Progress in Photovoltaics, 24(7): 905–913. doi: 10.1002/pip.2788.

[5] Meng L, Zhang Y, Wan X, Li C, Zhang X, Wang Y, Ke X, Xiao Z, Ding L, Xia R, Yip H, Cao Y, Chen Y. (2018) Organic and solution-processed tandem solar cells with 17.3% efficiency. Science, 361(6407): 1094-1098. doi: 10.1126/science.aat2612.

[6] Mandoc MM, Koster LJA, Blom PWM. (2007) Optimum charge carrier mobility in organic solar cells. Applied Physics Letter, 90(13): 133504. doi: 10.1063/1.2711534.

[7] Inche Ibrahim ML, Ahmad Z, Sulaiman K, Muniandy SV. (2014) Combined influence of carrier mobility and dielectric constant on the performance of organic bulk heterojunction solar cells. AIP Advance, 4(5): 057133. doi: 10.1063/1.4881080.

[8] Lee JJ, Lee SH, Kim FS, Choi HH, Kim JH. (2015) Simultaneous enhancement of the efficiency and stability of organic solar cells using PEDOT:PSS grafted with a PEGME buffer layer. Organic Electronics, 26: 191-199. doi: 10.1016/j.orgel.2015.07.022.

[9] Chen A, Zhu K, Shao Q, Ji Z. (2016) Understanding the effects of TCO work function on the performance of organic solar cells by numerical simulation. Semiconductor Science and Technology, 31(6): 065025. doi: 10.1088/0268-1242/31/6/065025.

[10] Koster LJA, Shaheen SE, Hummelen JC. (2012) Pathways to a new efficiency regime for organic solar cells. Advanced Energy Materials, 2(10): 1246-1253. doi: 10.1002/aen.201200103.

[11] Inche Ibrahim ML, Hassan HA. (2019) Influence of active layer thickness on the performance of organic photovoltaics with light trapping. IEEE Transaction Electron Devices, 66(7): 3124-3128. doi: 10.1109/TED.2019.2917594.

[12] Lenes M, Koster LJA., Mihailetchi VD, Blom PWM. (2006) Thickness dependence of the efficiency of polymer:fullerene bulk heterojunction solar cells. Applied Physics Letter, 88(24): 243502. doi: 10.1063/1.2211189.

[13] Fallahpour AH, Gagliardi A, Gentilini D, Zampetti A, Santoni F, Auf Der Maur M, Carlo AD. (2014) Optoelectronic simulation and thickness optimization of energetically disordered organic solar cells. Journal Computational Electronics, 13(4): 933-942. doi: 10.1007/s10825-014-0611-y.

[14] Mescher J, Christ N, Kettlitz S, Colsmann A, Lemmer U. (2012) Influence of the spatial photocarrier generation profile on the performance of organic solar cells. Applied Physics Letters, 101(7): 073301. doi: 10.1063/1.4745510.

[15] Tress W, Merten A, Furno M, Hein M, Leo K, Riede M. (2013) Correlation of absorption profile and fill factor in organic solar cells: The role of mobility imbalance. Advanced Energy Materials, 3(5): 631-638. doi: 10.1002/aenm.201200835.

[16] Islam M, Wahid S, Chowdhury M, Hakim F, Alam M. (2017) Effect of spatial distribution of generation rate on bulk heterojunction organic solar cell performance: A novel semi-analytical approach. Organic Electronics, 46(7): 226-241. doi: 10.1016/j.orgel.2017.04.021.

[17] Gan Q, Bartoli FJ, Kafafi ZH. (2013) Plasmonic-enhanced organic photovoltaics: breaking the 10% efficiency barrier. Advanced Materials, 25(17): 2385-2396. doi: 10.1002/adma.201203323.

[18] Inche Ibrahim ML. (2018) An analytical model for organic solar cells incorporating non-geminate monomolecular and bimolecular recombinations. Semiconductor Science and Technology, 33(12): 125005. doi: 10.1088/1361-6641/aadf72.

[19] Inche Ibrahim ML. (2017) Dissociation of charge-transfer states at donor–acceptor interfaces of organic heterojunctions. Journal of Physics D: Applied Physics, 50(6): 065103. doi: 10.1088/1361-6463/50/6/065103.

[20] Braun CL. (1984) Electric field assisted dissociation of charge transfer states as a mechanism of photocarrier production. The Journal Chemical Physics, 80(9): 4157-4161. doi: 10.1063/1.447243.

[21] Kniepert J, Lange I, van der Kaap NJ, Koster LJA, Neher D. (2014) A conclusive view on charge generation, recombination, and extraction in as-prepared and annealed P3HT:PCBM blends: Combined experimental and simulation work. Advanced Energy Materials, 4(7): 1301401. doi: 10.1002/aenm.201301401.

https://doi.org/10.31436/iiumej.v22i2.1565

STUDY OF EFFECTIVE OMNI-DIRECTIONAL VERTICAL AXIS WIND TURBINE FOR LOW SPEED REGIONS

ALISHER BEKMURDOVICH SAFAROV[1] AND RASUL AKIF-OGLI MAMEDOV[2*]

[1]Department of Energy Audit,
[2]Department of Energy,
Bukhara Engineering Technological Institute, Bukhara, Uzbekistan.

[*]Corresponding author: rasul_91-92@mail.ru

(Received:4th August 2020; Accepted: 16th January 2021; Published on-line: 4th July 2021)

ABSTRACT: This article presents theoretical and experimental studies of an improved vertical axis wind power device that generates electricity in areas with an average wind speed of 3.5-4.5 m/s. An algorithm has been developed for determining the geometrically optimal dimensions of the outer guiding surfaces to improve the efficiency of the device at low wind speeds. The device uses an AFPMG generator with opposite rotation of the stator and rotor. Matlab/Simulink and Solidworks were used to develop a mathematical and physical model of the wind power device. According to the results of the study, it was found that the developed wind power device can reach a rated power of 700 W at a wind speed of 8 m/s. The use of the device in areas with low wind speed is based on the possibility of increasing the efficiency of work by 5-10% at an average wind speed lower than that of other types of wind power devices.

ABSTRAK: Artikel ini memaparkan kajian teori dan eksperimen berkenaan alat kuasa angin paksi menegak yang diperbaharui dan menghasilkan tenaga elektrik di kawasan kelajuan angin berpurata 3.5-4.5 m/s. Algoritma telah dibangunkan bagi menentukan dimensi optimum geometri permukaan berpandu luar dalam meningkatkan kecekapan peranti pada kelajuan angin yang kurang. Peranti ini menggunakan penjana AFPMG dengan putaran stator dan rotor yang berlawanan. Matlab/Simulink dan Solidworks digunakan bagi menghasilkan model matematik dan fizikal peranti tenaga angin. Berdasarkan dapatan kajian, didapati bahawa alat tenaga angin yang dibangunkan ini dapat mencapai daya kuasa sebanyak 700 W pada kecepatan angin 8 m/s. Penggunaan alat ini di kawasan kurang kelajuan angin berkemungkinan meningkatkan efisiensi purata kerja sebanyak 5-10% pada kelajuan angin rendah, iaitu lebih rendah daripada segala jenis peranti tenaga angin lain.

KEYWORDS: *wind energy; low speed generator; external reference surface; vertical axis wind turbine*; *AFPMG generator*

1. INTRODUCTION

The increase in greenhouse gas emissions into the atmosphere, due to the burning of natural fuel sources, leads to such negative consequences as climate change. This, in turn, makes it important to increase the share of humankind that uses clean energy sources. The use of renewable energy sources such as sun, wind, biomass, geothermal energy, and hydropower is rapidly developing [1-5]. Due to the low cost of electricity generated from renewable energy sources, and the fact that it is practically harmless to the environment, such projects are increasingly being implemented, large investments are made in them, and promising developments and research are supported.

Currently, among renewable energy sources, the use of wind energy is one of the most promising projects. Today, more than 100 countries use wind energy to provide people and industries with a reliable and clean energy source. China, India, and Brazil are leading the way in the design and implementation of vertical-axis wind farms. By 2024, the world plans to increase the use of vertical wind power plants by 24% [6].

65-70% of the territory of the Republic of Uzbekistan is steppes and deserts, and 4-5% of the population lives in remote areas. Continuity and reliability of power supply remains the main problem in these areas. To meet the energy needs of the region, the possibility of using renewable energy sources is currently being evaluated. In our region there are opportunities for using solar energy, but in the desert areas, the efficiency of using solar energy is lower. The main reason for this is the high temperature and dustiness of the region.

When assessing the possibility of using the potential of wind energy, it was found that the average wind speed in this area is 3.5–4.5 m/s at a height of 10 m [7,8,9]. We can develop social and economic sectors by providing reliable and uninterrupted power supply to autonomous electricity consumers in remote areas of the region. By developing a new design of a wind turbine that works effectively with variable and weak wind flows, it is possible to provide autonomous consumers with reliable and uninterrupted electricity, thereby developing social and economic sectors in the region.

2. LITERATURE REVIEW

Solomin V. [10] in his candidate's dissertation on "Methodological foundations for the development and creation of vertical-axial wind turbines for the agro-industrial complex of Russia" in 2013 studied a new type of vertical axis wind turbine and performance methodology on the full or partial supply of electricity to small settlements, residential buildings, farms, and small workshops in order to develop infrastructure in remote areas with unstable power supply. Xiongfei P. [11] in a patent he received in 2011 for a two-rotor wind turbine, proposed a vertical-axis wind turbine with an opposite rotating top and bottom wind wheel to make efficient use of low-flow wind speeds.

Zha G. et al. [12] in their patent on a "Vertical axis wind power plant" created a wind power plant with single-rotor outer guide surfaces and inner rotating blades to supply power to urban and remote area power consumers. Pezaris D. [13] in his patent (invention) for a vertical axis wind turbine with external guide surfaces in 2011 achieved high efficiency in obtaining energy at low speed wind currents by applying a new design of external guide surfaces and internal rotating wind wheels in a vertical axis wind turbine. Yadav K. [14] in his research paper "Model and experimental test of a generator applied to a Savonius type wind turbine" simulated in Matlab/Simulink environment a mathematical model of a generator (permanent magnet synchronous generator, (PMSG)) applied to a Savonius type wind turbine. A wind power unit with a capacity of 1 kW was combined with a PMSG generator to study the dynamic performance of the wind turbine system.

Kaliyev V. [15] in his dissertation on "Machine without a core with permanent magnet of the current" in 2015 theoretically and practically studied a small-speed electric generator consisting of vertical axis of coreless permanent magnets. The feasibility of using this generator for low-power wind turbines has been evaluated. Efficiency has been achieved by improving the optimal size of the generator. Taran N. [16] in his research analyzed the possibilities of using axial generators of various designs consisting of permanent magnets with low rotational speeds for wind power and small micro-hydro-electric power station

equipment operating in low-speed wind currents. In his doctoral dissertation written in 2019, the scientist presented and evaluated the methods of mathematical modeling of three types of axial generators. Ansys Fluent and Matlab/Simulink were used to analyze the optimal dimensions and energy performance of these generators for efficient operation. Lim C. et al. [17] conducted research on the possibility of applying the vertical axis of external reference surfaces to wind turbines to increase the efficiency of wind power equipment. In this study, optimal performance indicators for the use of external guide surfaces were obtained and analyzed.

Based on the analysis of the above studies, an optimal design of a wind turbine with a vertical axis, using an AFPMG generator, the stator, and rotor which rotate in opposite directions, operating under variable and weak wind flows is developed.

3. METHODOLOGY

3.1 Mathematical Model of a Wind Power Plant

For the theoretical substantiation of the operation of a wind power plant with a vertical axis, operating in variable and weak wind flows, we will construct mathematical models for the structure and the processes occurring in it.

The rate of rotation of wind turbine blades depends on the wind speed, the radius of the wind turbine wheel, and the steering angle. This ratio is determined by the following expression [18]:

$$\lambda = \frac{\omega_T R}{V} = \frac{2\pi n_T R}{60V} \quad (1)$$

where: ω_T - is the rotation speed of the wind turbine wheel, R - is the radius of the wind wheel, V - is the wind speed, n_T - is the rotational speed of the turbine rotor.

The torque coefficient of a wind turbine is determined by the following expression [19]:

$$C_T = \frac{C_p}{\lambda} \quad (2)$$

The total mechanical moment created in the blades of a wind power device is determined by the following expression [20]:

$$M_T = \sum F_d R \quad (3)$$

where: F_d - is the drag force created by the wind blades; R - is the drag force created by the wind blades.

The expression for determining the drag force created by the wind blades in a stationary position is as follows [21]:

$$F_d = \frac{1}{2} C_d \rho S (V\cos(\alpha))^2 \quad (4)$$

where: ρ – air density; S – frontal area of the blade; V – wind speed; α – the angle between the wind speed direction and the normal direction of the blade surface; C_d –drag coefficient.

When the wind power device rotates about the central axis, the blades are in two positions. They consist of front and back shapes. The drag coefficient value depends on the shape of the blades. Forward torque is generated by the forward blade shape facing the wind direction, while the opposite torque is generated by the rear blade shape facing the wind direction.

The expression for determining the total mechanical moment created by the vanes in a stationary position is as follows [22]:

$$M_{ст} = \frac{1}{2}\rho SR\left[\sum_1^{i=m} C_{d1}\left(V\cos(\alpha_i)\right)^2 - \sum_1^{i=n} C_{d2}\left(V\cos(\alpha_j)\right)^2\right] \quad (5)$$

where: C_{d1} and C_{d2} are the drag coefficients of the front and back surfaces of the blades. m and n - the number of blades facing forward and backward in relation to the direction of the wind. R.D. Blevins gave in his scientific work the drag coefficients of the blades of wind turbines with a vertical axis. It follows from the study that the drag coefficients of the front and rear surfaces are equal to C_{d1} = 2,3 and C_{d2}= 1,1[23].

The total mechanical moment created on the blades when they rotate with a certain angular velocity $\omega_т$ relative to the central axis is determined as follows [24]:

$$M_T = \frac{1}{2}\rho SR\left[\sum_1^{i=m} C_{d1}(V\cos(\alpha_i) - \omega_т R)^2 - \sum_1^{i=n} C_{d2}\left(V\cos(\alpha_j) + \omega_т R\right)^2\right] \quad (6)$$

When using external guide surfaces in a wind turbine, the total mechanical torque generated by the blades is determined as follows:

$$M_T = \frac{1}{2}\rho SR\left[\sum_1^{i=m} C_{d1}(v\cos(\alpha_i) - \omega R)^2\right] \quad (7)$$

The mechanical force generated in a vertical axis wind turbine is determined by the following relationship:

$$P_T = M_T\omega_т = \frac{1}{2}\rho SR\omega_т\left[\sum_1^{i=m} C_{d1}(V\cos(\alpha_i) - \omega_т R)^2\right] \quad (8)$$

To describe the rotational motion of a wind power device, we use the following differential equation [25]:

$$J\frac{d\omega_т}{dt} = M_т - M_g - k\omega_т \quad (9)$$

Here: $M_т$ - is the mechanical torque on the shaft of the wind turbine, M_g is the electromagnetic moment of the generator, k is the friction coefficient, J is the total moment of inertia of the wind turbine, $\omega_т$ is the rotation speed of the wind wheel.

If the blades of the windmill are turned in two directions, the total mechanical speed will be equal to their sum:

$$\omega_T = \omega_1 + \omega_2$$

The total moment of inertia of the rotating parts of the device is equal to the sum of the moments of inertia of the wind turbine rotor and the generator. This relationship is defined by the following expression:

$$J = J_p + J_r \quad (10)$$

Here: J_p - is the moment of inertia of the wind turbine rotor, J_g is the moment of inertia of the generator.

The moment of inertia of a wind turbine rotor is determined by the following expression [26]:

$$J_p = n \cdot J_0 = n \cdot \frac{m_1 \cdot \frac{m_2}{2}}{m_1 + \frac{m_2}{2}} l^2 \quad (11)$$

Here: n - is the number of wind turbine blades, m_1 - is the mass of the turbine blades, m_2 - is the mass of the profile connecting the turbine blades and the rotor shaft, l - is the distance between the centers of the turbine blades and the rotor shaft.

3.2 Mathematical Model of External Guiding Surfaces

As a result of the use of external guiding surfaces, the wind speed is evenly distributed in the inner rotating part, that is, on the blades of the wind power plant, thereby ensuring the stable operation of the structure. The variable movement of the wind flow is explained on the basis of the law of continuity.

The expression for the dependence of the speed of the wind flow entering and leaving the channel of the outer guide surface on the dimensions of the surface channels is as follows: [27]:

$$\frac{V_1}{V_2} = \frac{W_2 \cdot H_2}{W_1 \cdot H_1} = \frac{\left[\frac{D_2 - D_1}{2tg\theta} + m\right] \cdot H_2}{\left[\frac{D_2 - D_1}{2tg\beta} + m\right] \cdot H_1} = \left(\frac{k + tg\theta}{k + tg\beta}\right) \cdot \left(\frac{tg\beta}{tg\theta}\right) \cdot \left(\frac{H_2}{H_1}\right) \quad (12)$$

Here: D_1, D_2- outer and inner diameters of the channel of the guiding surfaces, to which the wind flow is directed; θ, β - angles between the guiding surfaces, to which the wind flow is directed; m is the minimum distance between the guides to which the wind flow is directed; k - is the construction factor; L - is the distance between the outer and inner peripheral part of the guides, to which the wind flow is directed.

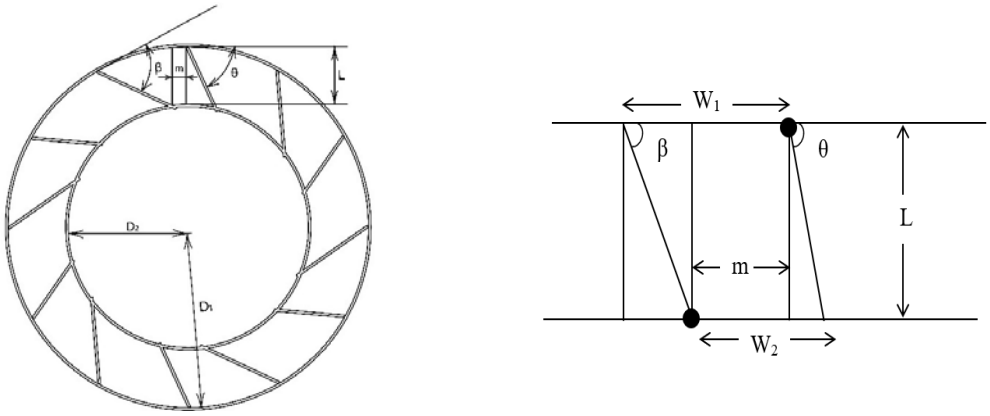


Fig. 1: Designation of the geometric dimensions of the outer guiding surfaces.

Figure 1 shows the determination of the geometric dimensions of the outer guiding surfaces (a, b). Having determined the optimal values of the angles β and θ, it is possible to increase the wind speed, thereby reducing the force that adversely affects the rotating blades of the wind power plant.

Figure 2 shows an algorithm developed to determine the optimal parameters of the outer guiding surfaces, which ensures the stable operation of the device due to the uniform transmission of the wind flow through the channels to the inner part of the wind turbine, that is, on its blades.

3.3 Mathematical Model of the Counter Rotating Generator

To eliminate the disadvantages associated with the use of gearboxes, low-speed generators are used in many low-power wind turbines. The use of low-speed generators in wind power plants intended for use in areas with low wind speeds increases the efficiency and reliability of the devices [28].

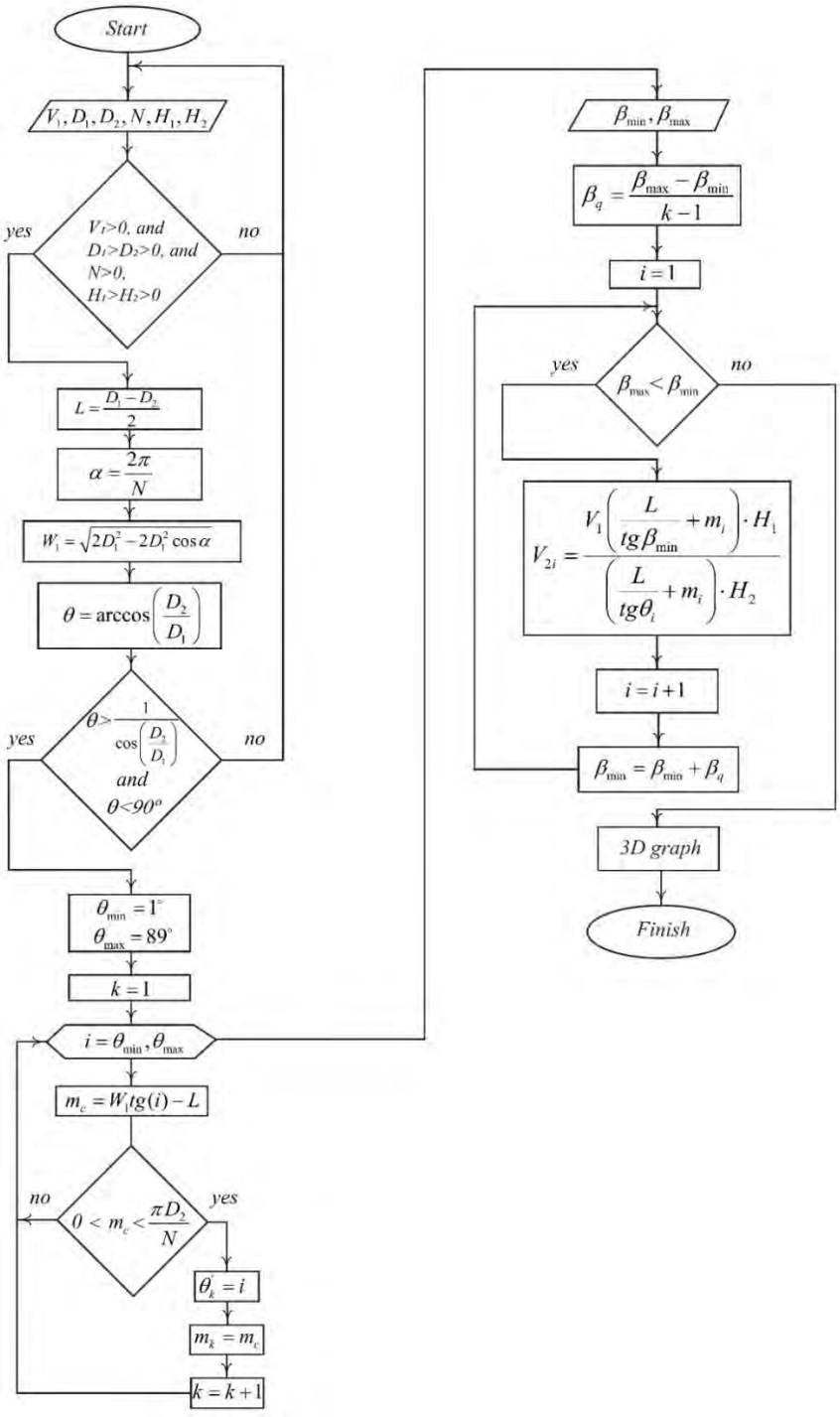


Fig. 2: Algorithm for determining the optimal parameters of external guiding surfaces.

In the developed wind power plant, a stator and rotor generator were used that rotate in opposite directions (AFPMG). For a synchronous machine with noticeable poles without a rotor winding, the voltage equations for the stator circuit [29]:

$$U_d = R_s i_d + L_d \frac{di_d}{dt} - \omega_e L_q i_q \quad (13)$$

$$U_q = R_s i_q + L_q \frac{di_q}{dt} + \omega_e L_d i_d + \omega_e \psi_{PM} \quad (14)$$

The electromagnetic power of a three-phase generator consisting of permanent magnets is determined by the following expression [30]:

$$P_e = \frac{3}{2}\omega_e[\psi_{PM} + (L_d - L_q)i_d]i_q \quad (15)$$

The electromagnetic moment of the generator with the number of pole pairs p is determined by the following expression [31]:

$$T_e = \frac{3}{2}p[\psi_{PM} + (L_d - L_q)i_d]i_q \quad (16)$$

where: ψ_{pm} is the maximum flow per phase coupling for the excitation system

$$\psi_{PM} = \frac{\sqrt{2}E_f}{\omega} \quad (17)$$

By applying this low speed generator to a wind turbine, its efficiency can be greatly improved.

4. RESULTS AND DISCUSSION

Figure 3 shows the appearance of an efficient wind turbine. This wind power plant can be used by electricity consumers located far from the power supply system.

Fig. 3: External view of the wind power plant.

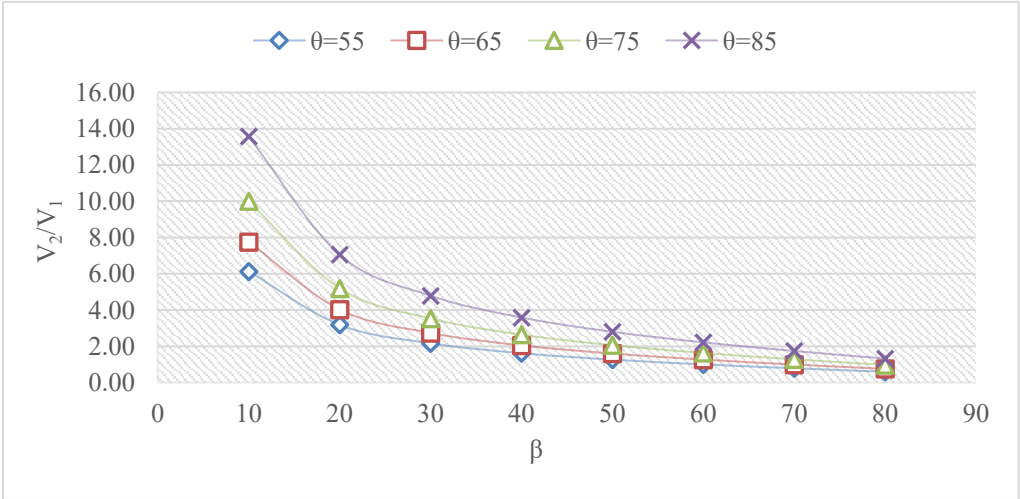


Fig. 4: Graph of the dependence of the angles of the channel surfaces θ and β coefficient of variability of speed (V_2/V_1) from the coefficient of construction k.

Figure 4 shows a graph of the dependence of the angles of the channel surfaces θ and β of the velocity variability coefficient (V_2/V_1) on the design coefficient k. Several values of the angles are taken and analyzed in this graph. In this case, the value of the velocity coefficient will be large for small values of the angle β and large values of the angle θ.

Figure 5 shows the results of the generated physical model of the wind power device in the Solidworks application. Accordingly, according to Bernoulli's law, when a wind flow of 4 m/s is applied to the outer guides, it can be seen that the wind flow speed increases by 30-50%.

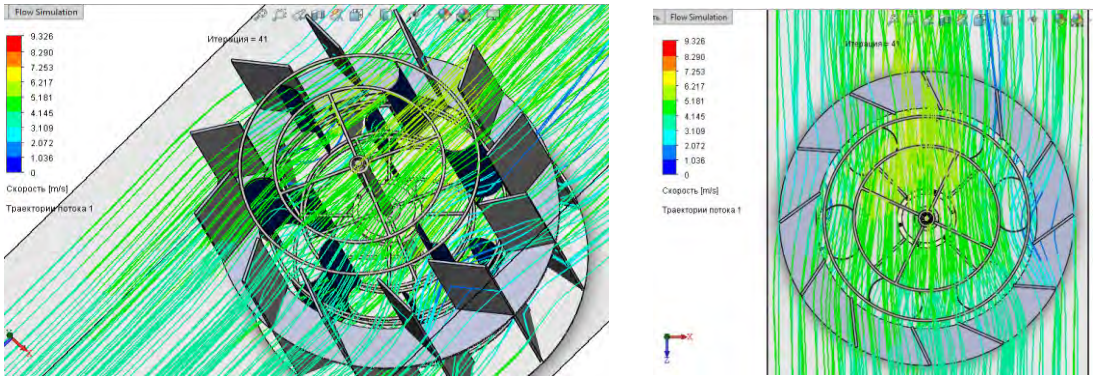


Fig. 5: A physical model of the wind power plant in the Solidworks application.

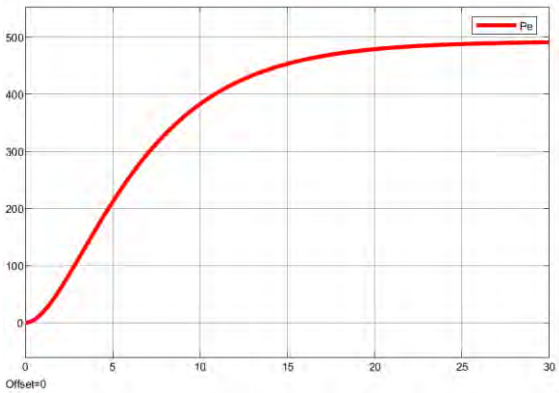


Fig. 6: Dynamic (in time) results of the electromagnetic force of a wind power plant, obtained in the Matlab/Simulink program, for a state in which the stator is stationary and the rotor rotates (W).

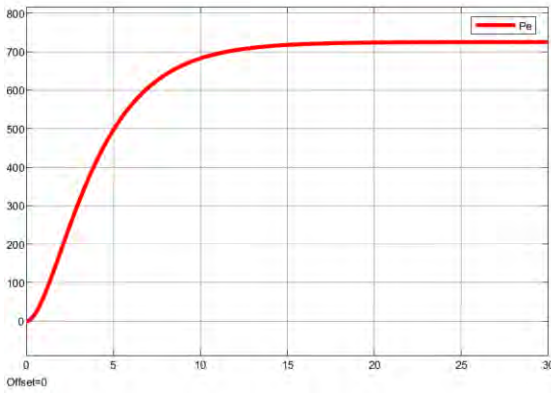
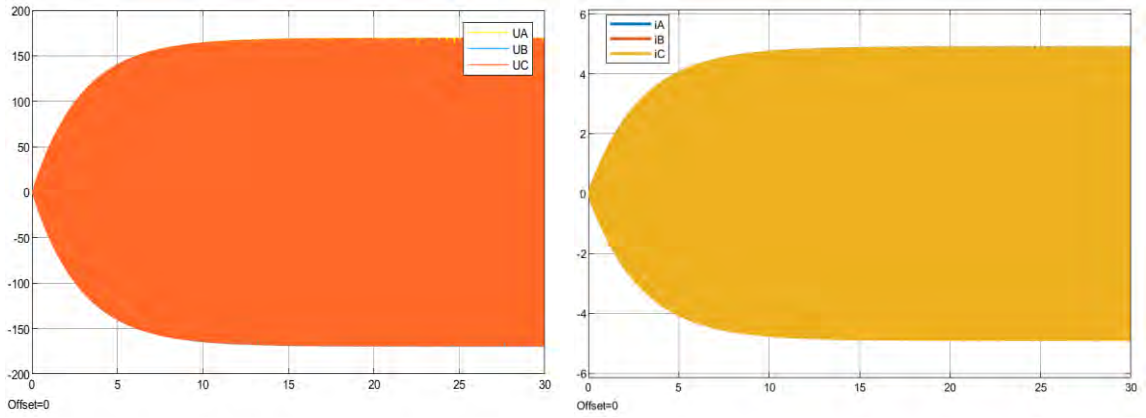


Fig. 7: Dynamic (in time) results of the electromagnetic force of a wind power plant, obtained in the Matlab/Simulink program, for a state in which the stator and rotors rotate in opposite directions (W).

Figures 6 and 7 show graphs of the electromagnetic power (W) determined from the dynamic (in time) expression 15, built in the Matlab/Simulink program of the wind turbine. When plotting the graphs, the nominal wind speed was assumed to be 8 m/s.

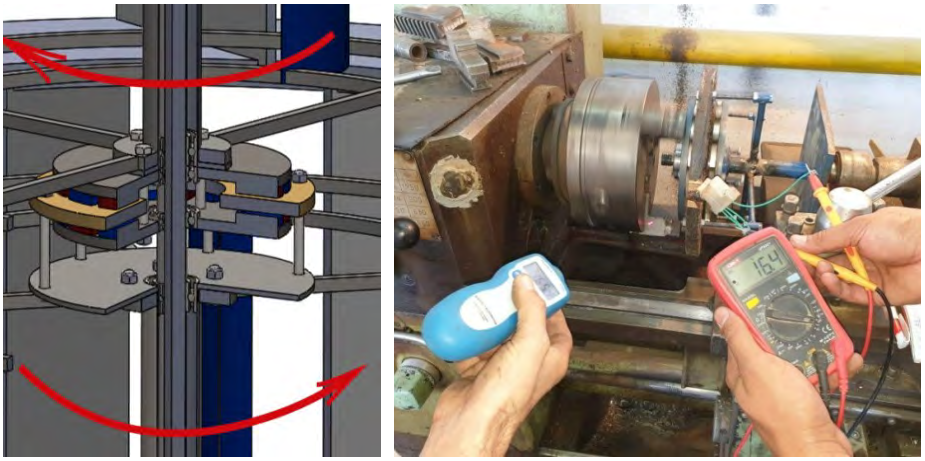
Figure 8 shows the stator voltage (a) and stator current (b) at the generator output of a wind turbine rotating at rated speed. The winding of a three-phase generator is star-connected. To ensure the stable operation of the wind power plant, a low-speed generator was developed, the stator and rotors of which rotate in opposite directions (see Fig. 9).

https://doi.org/10.31436/iiumej.v22i2.1565



(a) (b)

Fig. 8: Results obtained in the Matlab/Simulink program. (a) voltage of the stator winding, (b) current of the stator winding.



(a) (b)

Fig. 9: External view of the developed generator in section (a), the process of experimental tests (b).

Table 1: Main parameters of an electric generator intended for a wind farm

Rated power (W)	700	Magnetic flux, (Wb)	0.00036
Stator winding inductance, (H)	0.000125	Moment of inertia (kg)	0.18
Rated electromagnetic moment, (Nm)	14	Stator thickness (mm)	18
Rated speed of rotation, (rpm)	300	Air gap distance between stator and magnets (mm)	1.5
Number of pole pairs	16	Electromagnetic induction, (T)	0.6
Inductive resistance of the stator winding, (Ohm)	0.7	EMF generated in the stator winding, (V)	150
Active resistance of the stator winding, (Ohm)	12	Current generated in the stator winding, (A)	4.8
Number of windings	12	Number of turns in one winding	350
Number of phases	3	Efficiency, (%)	90

https://doi.org/10.31436/iiumej.v22i2.1565

Table 1 shows the main parameters of an electric generator intended for a wind farm. These types of generators are mainly used in small wind turbines and micro hydropower plants. Figure 10 shows a comparison of the wind power device developed by us with a prototype that does not have external guide surfaces and also has a fixed stator.

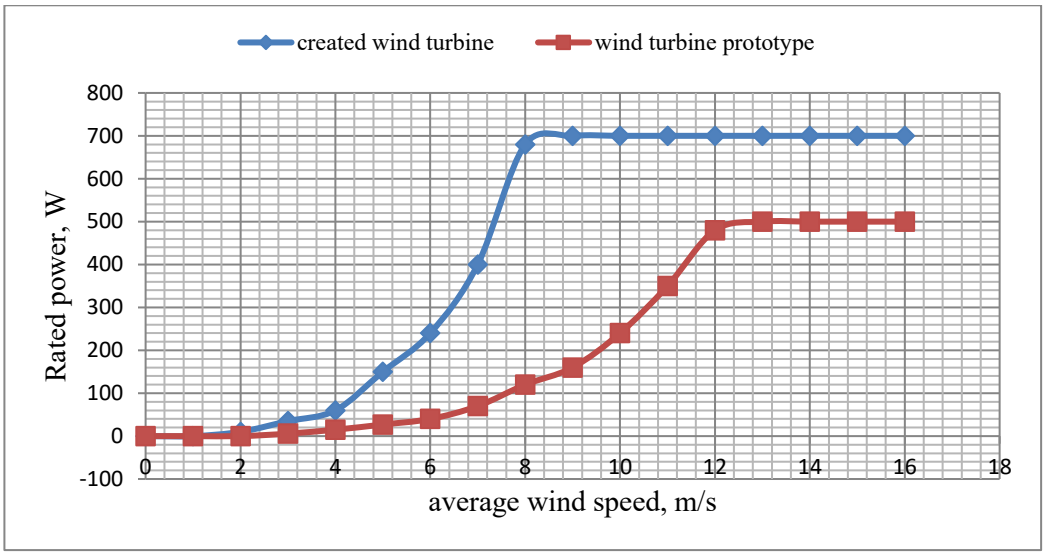


Fig. 10: Graph of the dependence of the wind speed on the rated power of the developed wind power plant from the prototype.

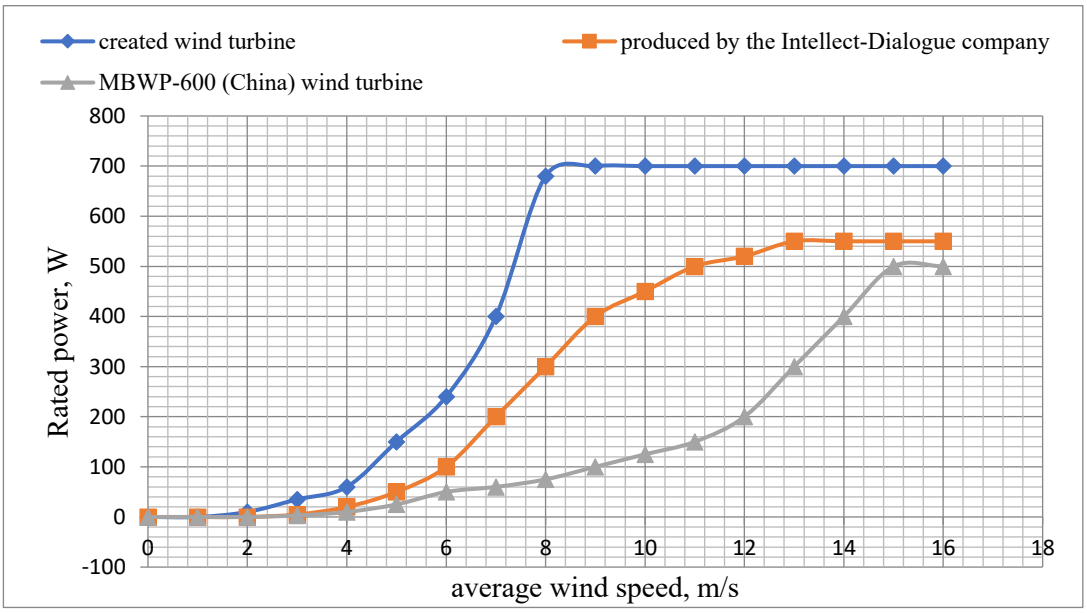


Fig. 11: Graph of the dependence of wind speed on the rated power of the developed wind power plant, as well as wind turbines produced in the Republic of Uzbekistan and a foreign prototype.

Figure 11 shows graphs comparing the wind power plant developed by us, as well as wind turbines produced in the Republic of Uzbekistan by the Intellect-Dialogue company and a prototype made in China. The research showed that the wind power device we developed works more efficiently in areas with weak wind flow in relation to prototypes with a vertical axis of rotation.

5. CONCLUSION

The optimal parameters of the outer guide surfaces applied to the wind turbine are $D_1 = 1.5$ m, $D_2 = 1$ m, $H_1 = H_2$, $\beta = 25^0$, $\theta = 50^0$, which significantly reduces the force that negatively affects the rotation of the blades and creates the opportunity to increase the flow wind up to 30-50%. Using a counter rotating AFPMG generator, it was found that the power rating of the device increased from about 500 to 700 W at a wind speed of 8 m/s. It was found that the developed wind power device is 5-10% higher in performance compared to low-power wind power devices with horizontal and vertical axis, currently used at low wind speeds (3-8 m/s). Due to the use of this wind power device, it is possible to provide autonomous consumers of low-power electricity located far from the central power system, which will lead to the development of social and economic spheres.

REFERENCES

[1] Sadullaev NN, Safarov AB, Nematov ShN, Mamedov RA. (2019) Research on facilities of power supply of small power capability consumers of Bukhara region by using wind and solar energy. International Journal of Innovative Technology and Exploring Engineering, 8(9S2): 229-235.

[2] Norhafana M, Ismail AF, Majid ZAA. (2015) Performance evaluation of solar collectors using a solar simulator. IIUM Engineering Journal, 16(2): 79-90. https://doi.org/10.31436/iiumej.v16i2.606

[3] Karim MA, Amin ZM. (2015) Mathematical modelling and performance analysis of different solar air collectors. IIUM Engineering Journal, 16(2): 43-55. https://doi.org/10.31436/iiumej.v16i2.603

[4] Izanlo A, Gholamian A, Kazemi MV. (2017) Direct power control of a DFIG based wind turbins under unbalanced grid voltage without rotor position sensor. IIUM Engineering Journal, 18(1): 57-71. https://doi.org/10.31436/iiumej.v18i1.676

[5] Jami MS, Mel M, Mohd Ariff AR, Abdulazeez QM. (2018) Investigation of bioflocculant as renewable dewatering aid in sludge treatment. IIUM Engineering Journal, 19(1): 15-23. https://doi.org/10.31436/iiumej.v19i1.735

[6] Wind Turbine Market Growth Statistics [https://www.gminsights.com/industry-analysis/wind-turbine-market]

[7] Sadullaev NN, Safarov AB, Nematov ShN, Mamedov RA. (2019) Statistical analysis of wind energy potential in Uzbekistan's Bukhara region using Weibull distribution. Applied Solar Energy, 55(2): 126-132.

[8] Sadullayev NN, Safarov AB, Nematov ShN, Mamedov RA, Abdujabarov AB. (2020) Opportunities and prospects for the using renewable energy sources in Bukhara region. Applied Solar Energy, 56(4): 291-301.

[9] Sadullaev NN, Safarov AB, Mamedov RA, Kodirov D. (2020) Assessment of wind and hydropower potential of Bukhara region. IOP Conference Series: Earth and Environmental Science: 14-16 October 2020

[10] Solomin EV. (2013) Methodological foundations for the development and creation of vertical-axialwind turbines for the agro-industrial complex of Russia. PhD thesis. South Ural State University, Electrical Engineering and Renewable Energy Department.

[11] Xiongfei P. (2011) Dual power wind turbine. China. CN102287334A.

[12] Zha G, Dano B. (2013) Vertical axis wind turbine. USA. US20130115069A1.

[13] Pezaris CD. (2011) Omnidirectional vertical-axis wind turbine. USA. US20110033288A1.

[14] Yadav YK, (2016) A Savonius wind turbine with electric generator: Model and test. PhD thesis. Clemson University, Mechanical Engineering Department.

[15] Kaliyev V. (2015) Axial Flux Permanent Magnet Coreless Machine. PhD thesis. University of Southern Queensland, Engineering Sciences Department.

[16] Taran N. (2019) Optimum Design of Axial Flux PM Machines based on Electromagnetic 3D FEA. PhD thesis. University of Kentucky, Electrical and Computer Engineering.

[17] Lim YC, Chong WT, Hsiao FB. (2013) Performance investigation and optimization of a vertical axis wind turbine with the omni-direction-guide-vane. Procedia Engineering, 67: 59-69
[18] Bezrukikh PP. (2010) Wind power. (Reference and methodological manual). Moscow, "Energia", 320 p.
[19] Lin SY, Lin YY, Bai CJ, Wang WC. (2016) Performance analysis of vertical-axis-wind-turbine blade with modified trailing edge through computational fluid dynamics. Renewable Energy, 99: 654-662.
[20] Ackermann T, Söder L. (2000) Wind energy technology and current status: a review. Renewable and Sustainable Energy Reviews, 4(4): 315-374.
[21] Bronstein MG. (2011) Harnessing rivers of wind: A technology and policy assessment of high altitude wind power in the U.S. Technological Forecasting and Social Change, 78: 736-746.
[22] Juangsa FB, Budiman BA, Aziz M, Soelaiman TA. (2017) Design of an airborne vertical axis wind turbine for low electrical power demands. International Journal of Energy and Environmental Engineering, 8:293-301.
[23] Blevins RD. (2003) Applied Fluid Dynamic Handbook. Florida, Krieger, 558 p
[24] Ahrens U, Diehl M, Schmehl R. (eds.) (2013) Airborne wind energy. Berlin. Springer, 611 p
[25] Obukhov SG, Sarsikeev EZh. (2012) Low power wind turbine mathematical model in Matlab Simulink. International Scientific Journal for Alternative Energy and Ecology, 2: 42-48
[26] Komass T, Sniders A. (2014) Design and verification of vertical axis wind turbine simulation model. 13th international scientific conference engineering for rural development proceedings; 29-30 May 2014; Engineering for rural development. Edited by L. Malinovska, V. Osadcuks pp. 335-340
[27] Chong WT, Fazlizana A, Poh SC, Pan KC, Hew WP, Hsiao FB. (2013) The design, simulation and testing of an urban vertical axis wind turbine with the omni-direction-guide-vane. Applied Energy, 112: 601-609.
[28] Nikbakhsh A, Izadfar H, Alinejad Beromi,Y. (2017) Design and optimization of permanent magnet synchronous generator for use in hydrodynamic renewable energy by applying ACO and FEA. IIUM Engineering Journal, 18(2): 158-176. https://doi.org/10.31436/iiumej.v18i2.705
[29] Wang W, Wang W, Mi H, Mao L, Zhang G, Liu H, Wen Y. (2018) Study and optimal design of a direct-driven stator coreless axial flux permanent magnet synchronous generator with improved dynamic performance. Energies. Suppl 11:1-22.
[30] Gieras JF, Wang R-J, Kamper MJ. (2004) Axial flux permanent magnet brushless machines. Kluwer Academic Publishers, Dordrecht
[31] Minaz MR, Çelebi M. (2017) Design and analysis of a new axial flux coreless PMSG with three rotors and double stators. Results in Physics, 7: 183-188.

https://doi.org/10.31436/iiumej.v22i2.1597

DOUBLE THRESHOLD SPECTRUM SENSING WITH OPTIMIZATION OF SAMPLES IN COGNITIVE RADIO NETWORKS

CHILAKALA SUDHAMANI

Departmentof Electronics and Communication Engineering,
CMR Technical Campus, Hyderabad, India

Corresponding author: Sudhamani.ece@cmrtc.ac.in

(Received: 26th August 2020; Accepted: 12th November 2020; Published on-line: 5th July 2021)

ABSTRACT:In cognitive radio networks spectrum sensing playsa vital role to identify the presence or absence of the primary user.In conventional spectrum sensing one secondary user will makea final decision regarding the availability of licensed spectrum.But Secondary user fail to make a correct detection about thepresence of the primary user if he is in fading environmentand it causes interference to the licensed users. Therefore toavoid interference to the licensed users and to make correct detection, number of samples is increased, Which increases theprobability of detection. In this paper the optimization of samplesis proposed to reduce the system overhead and also to increase thepropagation time. Simulation results show the optimized valueof samples for a given probability of false alarm and also thevariation of probability of detection with optimized samples andfalse alarm is shown in the results.

ABSTRAK:Dalam rangkaian radio kognitif, penginderaan spektrum memainkan peranan penting untuk mengenal pasti kehadiran atau ketiadaan pengguna utama. Dalam penginderaan spektrum konvensional, seorang pengguna sekunder akan membuat keputusan akhir mengenai ketersediaan spektrum berlesen. Tetapi pengguna Sekunder gagal membuat pengesanan yang betul mengenai kehadiran pengguna utama jika dia berada dalam persekitaran yang pudar dan menyebabkan gangguan kepada pengguna yang berlesen. Oleh itu untuk mengelakkan gangguan kepada pengguna berlesen dan membuat pengesanan yang betul, jumlah sampel meningkat, yang meningkatkan kemungkinan pengesanan. Dalam makalah ini pengoptimuman sampel dicadangkan untuk mengurangi overhead sistem dan juga untuk meningkatkan waktu penyebaran. Hasil simulasi menunjukkan nilai sampel yang dioptimumkan untuk kebarangkalian penggera palsu dan juga variasi kebarangkalian pengesanan dengan sampel yang dioptimumkan dan penggera palsu ditunjukkan dalam hasil.

KEYWORDS: C*ognitive radio; spectrum sensing; detection threshold; optimization*

1. INTRODUCTION

Due to the increased use of 3G and 4G mobile servicesand accelerated use of wireless internet services, there is asoaring demand for radio frequency available for wirelesscommunication. For utilizing particular frequency bandlicenses are required. In every country these licenses areauctioned by the governing agencies like Department of Telecommunication (DoT) in India, Federal CommunicationsCommission (FCC) in USA, etc. and assigns to licensed usersor Primary Users (PUs). However, research conducted by FCC concluded that the actual dispute is not limited bandwidth but improper usage of available spectrum, and also a largeamount of band is underused [1] by the licensed users.

When the spectrum is not occupied by the PU then it is termed as spectrum holes. This problem can be solved with the help of Cognitive Radio (CR) concept [2].

The main focus of cognitive radio is to select wireless spectrum dynamically without disturbing primary user and in accordance to that its transmission and reception parameters changes. This selection of spectrum which is available for communication is done by spotting the presence of PUs and enables Secondary Users (SUs) without disturbing PU communication. This process of selection is known as spectrum sensing [3]. Spectrum sensing techniques are divided into three major categories: Matched filter method, Energy detection method and Future detection. These three methods have its own pros and cons [4-6]. Energy detection method is having more advantages compared to remaining methods because of its simple architecture but the only problem is at low signal to noise ratio (SNR). Double threshold method has been proposed to overcome low SNR [7-10]. In conventional spectrum sensing, the single cognitive radio is used to detect the presence of primary users which may cause problems like fading, shadowing, and uncertainty [11-13]. In cooperative spectrum sensing, the sensing is done with the cooperation data given by multiple SUs.These individual decisions are combined and come to a single conclusion about the presence or absence of primary users at the fusion center with the help of fusion rules. But in CSS, the energy consumption increases with the increased number of SUs.

The accuracy of detection can be analyzed by two probabilities: the probability of detection and the probability of false alarm. Probability of detection represents that the CR is deciding that the primary user is present when he is actually present i.e. spectrum is busy. Where as CR is represented with the probability of false alarm when the report given by CR is that PU is present when PU is absent. Probability of false alarm will affect system efficiency because the user will not use the free band [14-15]. In [16], the authors used multitaper spectrum estimation to give a closed-form expression for the probability of detection and false alarm. By comparing multitaper spectrum estimation with an energy detector, authors concluded that energy detector requires more number of samples to attain the same detection and false alarm probabilities.

The performance analysis is given by the author in [17] for accurate detection of different hard and soft fusion rules in a real scenario. Authors explained about the randomness of signal to noise ratio in the real environment.The author in [18], derived the probability of miss detection in Eigen ratio spectrum sensing using a random matrix theory approach. The mathematical expression derived as a function number of CR, samples, and signal to noise ratio of PU. In [19], the authors proposed a two stage spectrum sensing technique compared with existing models for varying SNR. In this paper we proposed optimization of samples to reduce the system overhead and to increase the overall detection probability.

2. SYSTEM MODEL

In cognitive radio networks, spectrum sensing plays an important role and in this paper we considered energy detection method as suitable one for increasing the detection performance because of its advantages. Block diagram of energy detection method is shown in Fig. 1.

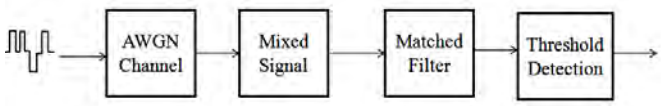


Fig. 1: Block Diagram of Conventional Energy Detection Method

Incoming signal is passed through band pass filter to reduce the noise and then it is squared and integrated to get total signal energy. The received energy is compared with the predefined threshold value to identify the presence or absence of the PU.This is done through hypothesis testing, if PU is present then it is represented by H_1 else it is represented by H_0.

$$x(n) = p(n) + w(n) \ : H_1 \qquad (1)$$

$$x(n) = w(n) \qquad : H_0 \qquad (2)$$

Where $p(n)$ is a PU signal, $w(n)$ is a Gaussian noise and $x(n)$ is a received signal respectively. Energy of a signal is estimated through number of samples and is given by [8,18].

$$Y(n) = \sum_{n=1}^{N} |x(n)|^2 \qquad (3)$$

Where N represents the number of samples used to estimate the total energy of a signal. Decision statistics of spectrum sensing are probability of detection (p_d) and probability offalse alarm (p_f), which are given by [18]

$$p_d = Prob(Y > T|H_1) = Q\left(\frac{T - N(\sigma_p^2 + \sigma_w^2)}{\sqrt{2N(\sigma_p^2 + \sigma_w^2)^2}}\right) \qquad (4)$$

$$p_f = Prob(Y > T|H_0) = Q\left(\frac{T - N(\sigma_w^2)}{\sqrt{2N(\sigma_w^2)^2}}\right) \qquad (5)$$

Where T is detection threshold with which the received energy of SU is compared with to make a final decision about the presence or absence of the PU.

For a fixed false alarm probability, detection threshold is calculated from equation (5) as

$$Y(n) = \sum_{n=1}^{N} |x(n)|^2 \qquad (6)$$

In Eq.(3), the number of samples used to estimate the energy of a given signal determines the system performance. For large number of samples the performance is good compared to the less number of samples which is shown in the Fig. 2. As the number of samples increases detection probability increases because more number of samples gives more accurate values, same can be observed from this figure.

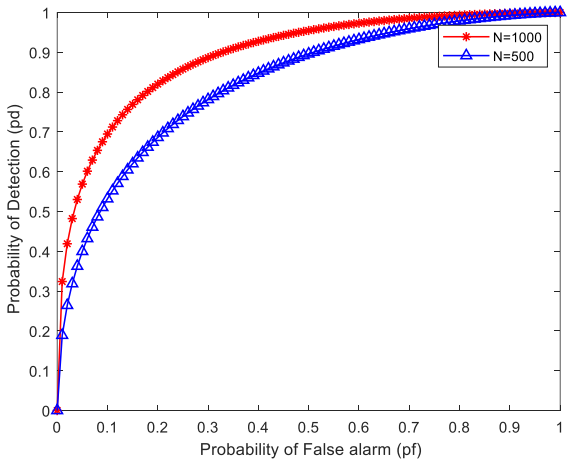


Fig. 2: Probability of detection with probability of false alarm.

Detection threshold, probability of detection and probability of false alarm depends on the number of samples. If the number of samples is go on increasing to increase the performance of a system then the processing time increases which leads to more overhead on this system. The time taken to estimate the detection threshold, probability of detection and false alarm increases, which may cause spectrumun-utilization. If the spectrum is actually free and same is identified by the SU after long processing time then it causes spectrum wastage i.e. no one is going to utilize it and also with in the estimation time of SU, if primary user want to use then SU may cause interference to the PU by accessing the spectrum.

Therefore, to overcome this problem, number of samples used to estimate the detection probability is optimized in this paper. These optimized numbers of samples are estimated by maximizing the detection threshold. Differentiating Eq.(6) and equating to zero gives the optimal number of samples,which is given as

$$\frac{dT}{dN} = \frac{d}{dN}(\sigma_w^2(N + \sqrt{2N}\left(Q^{-1}(p_f)\right)) = 0 \qquad (7)$$

After differentiation it gives the optimal number of samplesas

$$N_{opt} = \frac{1}{\sigma_w^4\left(Q^{-1}(p_f)\right)^2} \qquad (8)$$

Optimized number of samples increases the processing time and reduces the system over head. This increases the system performance and also spectrum utilization. Now the proposed threshold is given as

$$T_{opt} = \sigma_w^2(N_{opt} + \sqrt{2N_{opt}}(Q^{-1}(p_f)) \qquad (9)$$

In Eq. (9) new detection threshold is estimated toincrease the detection probability. To further improve thedetection probability, double threshold method has been proposed. Figures 3 and 4 show the conventional and proposed double threshold spectrum sensing methods.

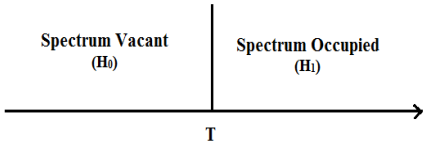


Fig. 3: Conventional spectrum sensing.

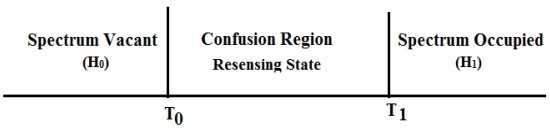


Fig. 4: Double threshold spectrum sensing.

In the double threshold spectrum sensing new thresholds are given as (T_0) and (T_1). If the received signal is less than the detection threshold (T_0) then the PU is declared as absent. If the received is greater than the detection threshold (T_1) then PU is present. If the received signal lies between (T_0) and (T_1) then SU will re-sense the spectrum. This method reduces themiss detection and false alarm probabilities. New detection thresholds are given as

$$T_1 = T_{opt} + KT_{opt} \qquad (10)$$

$$T_0 = T_{opt} - KT_{opt} \qquad (11)$$

Where K represents the improvement factor of detection threshold.

3. SIMULATION RESULTS

In this paper we considered optimization of number of samples and double threshold method. Figure 5 shows the relation between optimal number of samples and probability of false alarm. From this figure it is observed that the optimal *N* value decreases as the *pf* increases from 0 to 0.5 and thenit increases as pf increases from 0:5 to 1. This also showsthat the optimal false alarm is 0.5. At this point we need less number of samples to increase the detection probability.

Figure 6 shows the variation of pd with respective to *pf* for variable improvement factor *K*. The factor *K* is varied from 5% to 50%. From this figure it is observed that the detection probability is high for small values of *K* because the upper threshold is small for small values of *K* and it is easy to identify the PU presence or absence. As *K* increases, upper threshold increases which reduces the detection probability.Therefore we can use small improvement factor for increasing the detection probability. From this figure also it is observed that the optimal p_f is 0.5.

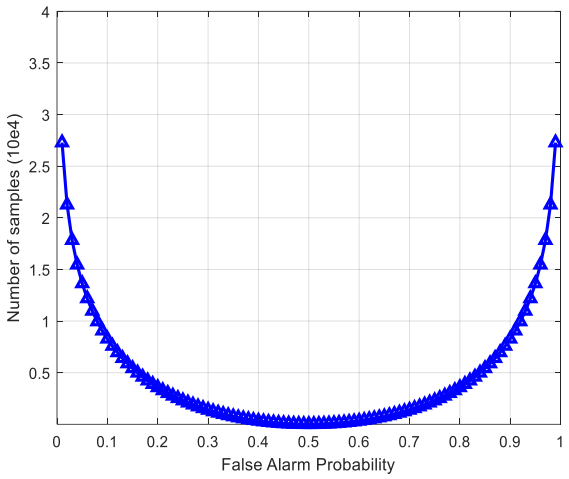


Fig. 5: Optimized samples with probability of false alarm.

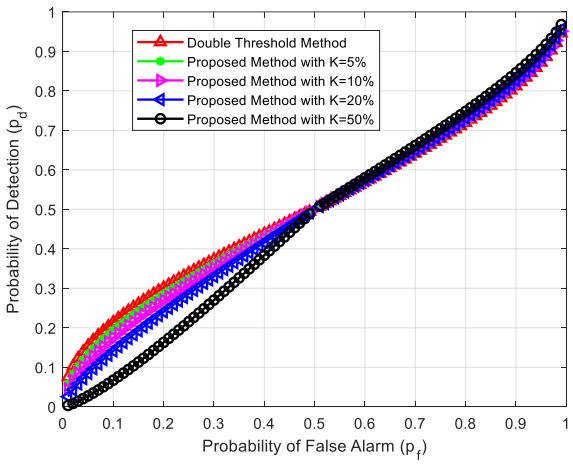


Fig. 6: Probability of detection with probability of false alarm using *K*.

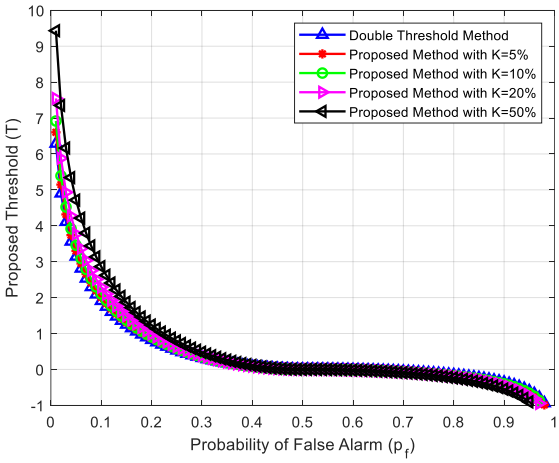


Fig. 7: Detection threshold with probability of false alarm.

Figure 7 shows the variation of threshold with false alarm.Threshold is high at low p_f, because the spectrum sensing depends on detection threshold. At low detection threshold

probability of false alarm is high which is also shown inthe figure. So, we need high detection thresholds to improve the system performance. With the improvement factor thethreshold value decreases for a given p_f.

From the three figures it is observed that the optimal falsealarm is 0:5. Which is used to increase the overall system performance and the optimal number of samples required isalso less at this p_f. This reduces the overhead on the system and time taken to conclude the presence or absence of the PU.

4. CONCLUSION

In this paper, we estimated the probability of detection with variable number of samples and false alarm. From the results it is observed that the detection probability is high for large number of samples. But it increases the sensing and decision making time. So to reduce this optimization of samples has been proposed in this paper. The optimal samples and optimal false alarm are estimated from the results, which improves the detection performance and also reduces the processing time and system overhead. This increases the overall spectrum usage than compared to the conventional method.

ACKNOWLEDGEMENT

This work is supported by grant no.SR/FST/College 017/2017 under DST FIST program-2017, received by the Department of Electronics &Communication Engineering, CMR Technical Campus, Hyderabad.

REFERENCES

[1] Federal Communications Commission. (2002) Spectrum policy task force. ET Docket No. 02-35.

[2] Sharma V, Joshi S. (2018) A literature review on spectrum sensing in cognitive radio applications. Second International Conference on Intelligent Computing and Control Systems: July 2018.

[3] Mitola J, Maguire MG. (1999) Cognitive radio: Making software radios more personal. IEEE Personal Communications, 6:13-18.

[4] Cabric D, Mishra SM, Brodersen RW. (2004) Implementation issues in spectrum sensing for cognitive radios. Proceedings of Asilomar Conference on Signals, Systems, and Computers, 1:772-776.

[5] Yucek T, Arslan D. (2009) A survey of spectrum sensing algorithms for cognitive radio applications. IEEE Communications Surveys and Tutorials, 11:116-130.

[6] Ali SS, Liu C, Jin M. (2014) Minimum eigenvalue detection for spectrum sensing. International Journal of Electrical and Computer Engineering (IJECE), 4:623-630.

[7] Liu SQ, Hu BJ, Wang XY. (2012) Hierarchical cooperative spectrum sensing based on double thresholds energy detection. IEEE Communications Letters, 16:1096-1099.

[8] Xie J, Chen J. (2012) An adaptive double threshold spectrum sensing algorithm under noise uncertainty. Computer and Information Technology, 1:824-827.

[9] Ribas ADOP, Dias US. (2015) On the double threshold energy detection-based spectrum sensing over fading channel. Radio and Wireless Symposium: 2015, San Diego, CA.

[10] Steve Arul U, Salai Chandira Rajan S. (2016) Spectrum management techniques using cognitive radios. Cognitive Radio Technology: International Journal of Data Mining Techniques and Applications, 5: 79-82.

[11] Ganesan G, Li YG. (2007) Cooperative spectrum sensing in cognitive radio–part I: two user networks. IEEE Transaction on Wireless Communications, 6: 2204–2213

[12] Christian I, Sangman M, Chung I, Lee J. (2012) Spectrum mobility in cognitive radio networks. IEEE Communications Magazine, 50(6): 114-121.
[13] Xuping Z, Jianguo P. (2007) Energy-detection based spectrum sensing for cognitive radio. IET Conference on Wireless, Mobile and Sensor Networks: 944-947.
[14] Lee SH, Oh DC, Lee YH. (2009) Hard decision combining based cooperative spectrum sensing in cognitive radio systems. International Conference on Wireless Communications and Mobile Computing: 2009, Leipzig, Germany. 906-910.
[15] Hou F, Chen X, Huang H, Jing X. (2016) Throughput performance improvement in cognitive radio networks based on spectru prediction. 16th International Symposium on Communications and Information Technologies: Qingdao, China.
[16] Wu J, Luo T, Li J, Yue G. (2009) A cooperative double threshold energy detection algorithm in cognitive radio systems. IEEE International Conference on Wireless Communication, Networking & Mobile Computing., 1-4. doi10.1109/WICOM.2009.5303377
[17] Owayed A, Mohammed ZA, Mosa AA. (2010) Probabilities of detection and false alarm in multitaper based spectrum sensing for cognitive radio systems in AWGN. IEEE International Conference on Communication Systems: 2010 Singapore. 579-584.
[18] Verma P, Singh B. (2015) Simulation study of double threshold energy detection method for cognitive radios. 2nd International Conference on Signal Processing and Integrated Networks: Noida. 232- 236.
[19] Sengupta A, Chattopadhyay S, Ghatak SR, Biswas V. (2019) Two-Stage spectrum sensing model for varying SNR conditions in cognitive radio network. International Conference on Electrical, Electronics and Computer Engineering: ALIGARH, India.

https://doi.org/10.31436/iiumej.v22i2.1612

AN INVESTIGATION OF THE SENSITIVITY OF POLYMER-COATED SURFACE ACOUSTIC WAVE-BASED GAS SENSORS IN THE DETECTION OF VOLATILE ORGANIC COMPOUNDS

ALIZA AINI MD RALIB* AND AMIRAH SYAHIRAH SYAMSIL OMAR

Department of Electrical and Computer Engineering, Kulliyyah of Engineering, International Islamic University of Malaysia, Jalan Gombak, 53100, Kuala Lumpur, Malaysia

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

(Received: 3rd August 2020; Accepted: 23rd January 2021; Published on-line: 4th July 2021)

ABSTRACT: Surface acoustic wave sensors (SAWs) are excellent at detecting volatile organic compounds (VOCs) since a sensing layer can be created by spreading a thin film of material across the delay line. This critically enhances performance as it is sensitive to the physical phenomena of interest. This study aims to provide a thorough investigation of the sensitivity of polymer-coated SAW-based gas sensors to VOCs using simulations via the finite element method (FEM). As such, quartz was chosen as the piezoelectric substrate while polymeric materials were chosen as the sensing layers due to their high sensitivity, low energy consumption, short response time, performance at room temperature, and reversibility after exposure to an analyte. The polymeric materials chosen were: (1) polyisobutylene (PIB), (2) polydimethylsiloxane (PDMS), (3) polyisoprene (PIP), (4) polyimide (PI), and (5) phenylmethyldiphenylsilicone (OV25). The VOCs chosen for investigation were: (1) dichloromethane (DCM), (2) trichloroethylene (TCE), (3) 1,2-dichloroethylene (DCE), and (4) carbon tetrachloride (CCl_4). The performance of each polymer-coated SAW sensor was evaluated in terms of frequency shift and sensitivity to each VOC in FEM simulations. Our study found that the PIB-coated sensor had the highest sensitivity (4.0571 kHz/ppm) to DCM vapor and good sensitivity (45.257 kHz/ppm) to TCE vapor. However, the performance of each polymer-coated sensor varied depending on the type of VOC being tested. As an example, while the OV25-coated sensor was more sensitive (52.57 kHz/ppm) than the PIB-coated sensor (53.54 kHz/ppm) to TCE vapor regardless of the concentration, the PIB-coated sensor was more sensitive to DCM vapor at both low (4.06 kHz/ppm) and high (3.54 kHz/ppm) concentrations than the OV25-coated sensor. Therefore, the results of our FEM simulations indicate that polymer-coated SAW-based gas sensors are highly capable of self-powered VOC detection.

ABSTRAK: Sensor gelombang akustik permukaan (SAW) adalah sangat baik dalam mengesan sebatian organik meruap yang tidak stabil (VOCs), kerana lapisan pengesan dapat dihasilkan dengan melapis nipis bahan pada lapisan garis tunda. Cara ini dapat menambah baik prestasi kerana ianya sensitif kepada fenomena fizikal yang dituju. Kajian ini bertujuan bagi menyediakan kajian menyeluruh terhadap kesensitifan sensor gas berasaskan SAW bersalut polimer pada VOC menggunakan simulasi melalui kaedah unsur terhingga (FEM). Oleh itu, kuarza dipilih sebagai substrat piezoelektrik manakala bahan polimer dipilih sebagai lapisan penginderaan berdasarkan kepekaan tinggi, penggunaan tenaga rendah, respon masa singkat, prestasi suhu bilik, dan faktor keboleh-balikan setelah terdedah kepada analit. Bahan polimer yang dipilih adalah: (1) polisobutilena (PIB), (2) polidimethilsiloxana (PDMS), (3) polisoprena (PIP), (4) polimida (PI), dan (5) phenilmethildiphenilsilikon (OV25). VOC terpilih bagi kajian adalah: (1) diklorometana

(DCM), (2) trikloretilena (TCE), (3) 1,2-dikloroetilena (DCE), dan (4) karbon tetraklorida (CCl_4). Prestasi setiap sensor SAW bersalut polimer dinilai berdasarkan peralihan frekuensi dan kesensitifan pada setiap VOC simulasi FEM. Dapatan kajian menunjukkan sensor bersalut-PIB mempunyai kesensitifan paling tinggi (4.0571 kHz/ppm) terhadap wap DCM dan kepekaan yang baik (45.257 kHz / ppm) terhadap wap TCE. Walau bagaimanapun, prestasi setiap sensor bersalut polimer adalah berbeza bergantung kepada jenis VOC yang sedang diuji. Sebagai contoh, sensor bersalut OV25 adalah lebih sensitif (52,57 kHz/ppm) daripada sensor bersalut PIB (53,54 kHz/ppm) pada wap TCE tanpa mengira kepekatan. Manakala sensor bersalut PIB lebih sensitif terhadap wap DCM pada kedua-dua kepekatan rendah (4.06 kHz/ppm) dan tinggi (3.54 kHz/ppm) daripada sensor bersalut-OV25. Oleh itu, hasil simulasi FEM menunjukkan bahawa sensor gas berasaskan SAW bersalut polimer adalah sangat berpotensi sebagai pengesan VOC berkuasa sendiri.

KEYWORDS: *surface acoustic wave; gas sensor; polymer sensing layer; sensitivity; frequency shift*

1. INTRODUCTION

Breath analysis has recently emerged as a new and improved diagnostic tool in health examinations with technical advantages, such as being non-invasive, painless, cost-effective, user-friendly, easily repeatable, and offering real-time testing methods; over existing methods [1]. Current diagnostic methods involve complex, costly, and invasive procedures, such as blood and urine tests, endoscopy, biopsies, imaging etc.; to diagnose diseases. Moreover, some diagnostic imaging studies, such as MRIs, X-rays, and CT scans; cannot be repeated in a short period of time due to exposure to high amounts of radiation [2].

Breath analysis is a systematic review of the volatile organic compounds (VOCs) exhaled in human breath. Each exhaled VOC is a biomarker containing clinical information for disease diagnosis. Several studies have shown that certain exhaled VOCs are clinical biomarkers of metabolic disorders and other diseases [3]. Alkanes, for instance, are biomarkers of lung cancer while acetone is used to diagnose diabetes [4]. As such, various detection devices, such as piezoelectric sensors and optical sensors, have been created specifically for this purpose. A SAW-based gas sensor is a piezoelectric sensor that measures physical changes such as fluctuations in mass [5]. They are also widely used in medicine and research studies to examine human health via breath analysis. Therefore, this study chose to investigate SAW-based gas sensors as they are not only highly sensitive and offer faster disease detection but are environmentally friendly as well. The ability to introduce a sensing layer further enhances its performance in terms of sensitivity and selectivity as it enables mass loading that results in changes in acoustic wave velocity and resonant frequency [6]. This study aims to thoroughly investigate the performance of polymer-based sensing layers in SAW-based gas sensors through simulation via the finite element method (FEM). Previous studies have explored the usage of SAW-based and Quartz Crystal Microbalance gas sensors for electronic nose applications [7]. FEM simulation has also been used to measure the sensitivity of bulk acoustic wave sensor for breath analysis application [8]. However, the sensitivity and selectivity of SAW-based gas sensors to different types of VOCs warrants further investigation. Therefore, this study created simulations to test the ability of five different polymers in sensing four different VOCs in a SAW-based gas sensor. The performance of each polymer-coated sensor was evaluated in terms of frequency shift and sensitivity to each VOC.

2. DESIGN CONCEPT: SURFACE ACOUSTIC WAVE AND PIEZOELECTRICITY

Piezoelectricity is the ability of some solid materials to generate electrical energy when subjected to mechanical stress. A piezoelectric transducer is basically a "crystal" sandwiched between two metal plates and compressed to generate electricity [9]. Direct piezoelectricity occurs when a piezoelectric material is compressed and flows through the metal plates to produce mechanical energy. Therefore, the crystal reacts to the pressure of the applied stress and converts it into electrical energy. Meanwhile, an inverse piezoelectric effect occurs when a crystal is deformed by the electrical energy applied and converts it to mechanical energy [9]. A SAW sensor consists of both input and output interdigital transducers (IDTs) on a piezoelectric substrate. A thin and highly sensitive layer, capable of changing the resonant frequency of the SAW sensor, is then placed between both the IDTs. Figure 1 shows the basic structure of a SAW sensor. The choice of piezoelectric substrate is not only important but varies according to the primary function of the sensor [10]. Popular choices of substrate materials include zinc oxide (ZnO), quartz, and lithium niobate ($LiNbO_3$) [10]. A SAW-based gas sensor is used to detect minute changes in surface mass caused by gas absorption by the sensing layer [9] that then causes the resonant frequency of the sensor to shift i.e., the higher the shift in resonant frequency, the higher the sensitivity of the sensor. Apart from the sensing layer, the type of material used to coat the substrate also affects sensor performance.

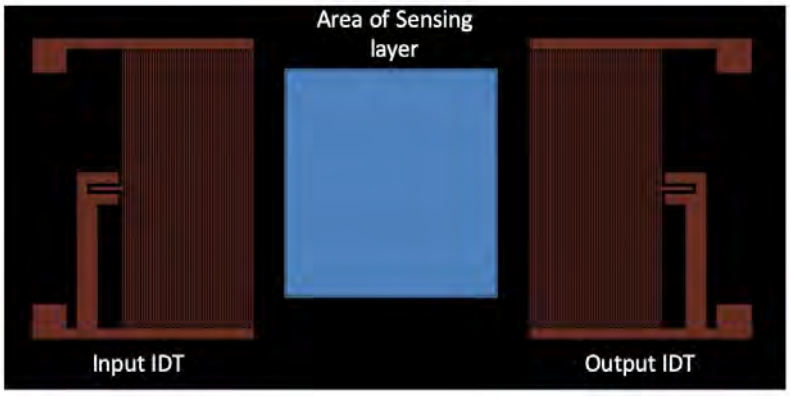


Fig. 1: Schematic diagram of a surface acoustic wave (SAW) sensor.

Resonant frequency, *fr* is defined as the specific frequency of an acoustic wave oscillating at higher amplitudes. It can be calculated using Eq. (1) where, λ is the wavelength and v is the velocity of the acoustic wave. Based on these two parameters, a wavelength is a traveling wave that follows the dimensions of the IDTs [11].

$$f = \frac{v}{\lambda} \tag{1}$$

The absorption of gasses by the sensing layer changes its elemental properties, such as density, thickness, elastic modulus, and acoustic velocity, which causes a frequency shift in the SAW-based gas sensor. The changes in frequency due to the absorption of gases by non-piezoelectric, non-conducting, isotropic thin polymer sensing layers are given below where, Δf is the frequency shift, h is the thickness of the polymer layer, f_0 is the operating frequency, ρ is the density of the polymer layer, k_1 and k_2 are the constants of the substrate, and λ and μ are the Lamé constants [12].

$$\Delta f = (k_1 + k_2) f_0^2 h\rho - k_2 f_0^2 h \left[\frac{4\mu}{v_R^2}\left(\frac{\lambda + \mu}{\lambda + 2\mu}\right)\right] \quad (2)$$

The sensitivity of a SAW sensor is determined by the difference between the input and output measures. It is also can be determined by the amount of shift in resonant frequency and mass sensitivity as shown in Eq. (3) where, c_m is the mass sensitivity coefficient independent of frequency, h' is the thickness of the coating containing gas molecules, and $\Delta\rho_s$ is the change in mass density due to absorption [13].

$$\Delta f = -c_m f_0^2 h' \Delta\rho_s \quad (3)$$

3. SAW SENSOR SIMULATION VIA FINITE ELEMENT METHOD

The SAW sensors were simulated using the FEM. As shown in Fig. 1, each SAW-based gas sensor consisted of an input IDT, a mass loading area, and an output IDT. With an acoustic wave velocity of 3159 m/s [10], quartz was our choice of piezoelectric substrate. Based on Equation (1), the resonant frequency of the SAW sensors was 40 MHz. Aluminum electrodes, with a λ/4-width and 800 nm-thickness, were used as the IDTs. Three reflectors, with a width and spacing of λ/8, were placed on either side of the sensors. Figure 2 shows a 2D model of the simulated SAW-based gas sensor while its physical parameters are detailed in Table 1.

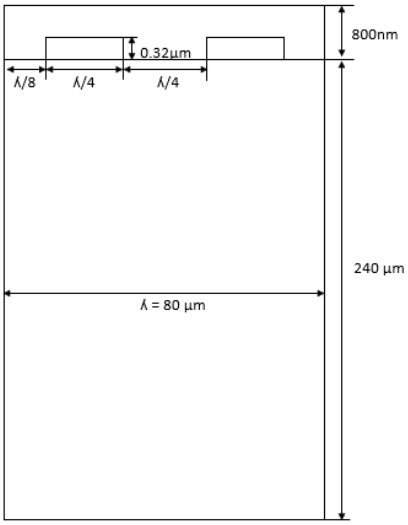


Fig. 2: 2D model of the simulated SAW-based gas sensor.

3.1 Polymeric Sensing Layers

Polymers are highly sensitive, making them an ideal choice of sensing layer. When exposed to VOCs, the sensing layer absorbs and desorbs these vapor analytes resulting in physical changes in mass and dielectrical properties that affect the resonant frequency, which is then converted into an electrical output signal [14]. Polymer-based SAW sensors for gasses are highly sensitive, have low energy consumption, short response times, and perform well at room temperature. Table 1 lists the polymers used in our FEM simulations: (1) polyisobutylene (PIB), (2) polydimethylsiloxane (PDMS), (3) polyisoprene (PIP), (4) polyimide (PI), and (5) phenylmethyldiphenylsilicone (OV25) as well as the physical parameters of the simulated SAW-based gas sensors.

Table 1: Physical parameters of the simulated SAW-based gas sensors.

Parameter	Expression
Substrate material	Quartz
Resonant frequency	47 MHz
IDTs material	Aluminum (Al)
Wavelength	80 µm
Width	20 µm
Sensing layer material	Polyisobutylene (PIB) Polydimethylsiloxane (PDMS) Polyisoprene (PIP) Polyimide (PI) Phenylmethyldiphenylsilicone (OV25)
Thickness	800 nm
VOC vapor material	Dichloromethane (DCM) Trichloroethylene (TCE) Dichloroethylene (DCE) Carbon tetrachloride (CCl_4)

3.2 Volatile Organic Compounds (VOCs)

A sensor's response depends on the interaction between the VOC vapors and the polymeric sensing layer. To better understand these interactions, as well as the transduction mechanism of the device, these responses should be comparable. It is well established that a SAW sensor coated with a permeable material will produce responses proportional to the partition coefficient, K [15]. An increase in sensing layer density indicates vapor absorption. In our FEM simulations, sensors are exposed to VOCs at room temperature and atmospheric pressure. Therefore, as a consequence of mass loading, the simulations will experience changes in the mechanical and electrical boundaries. Table 2 lists the VOCs used in our FEM simulations: (1) dichloromethane (DCM), (2) trichloroethylene (TCE), (3) 1,2-dichloroethylene (DCE), and (4) carbon tetrachloride (CCl_4) as well as the critical parameters used to evaluate SAW-based gas sensor performance; partition coefficient (K) and molar mass (M).

Table 2: Parameters of the simulated VOCs used to calculate the density of vapors absorbed by the polymeric sensing layers

Type of VOC	Partition coefficient, K	Molar Mass, M (g/mol)
Dichloromethane (DCM)	$10^{1.4821}$	84.93
Trichloroethylene (TCE)	$10^{2.3994}$	131.4
Dichloroethylene (DCE)	$10^{1.9215}$	96.95

4. SAW SENSOR SIMULATION VIA FINITE ELEMENT METHOD

The frequency shift is calculated by evaluating the eigenfrequency before and after sensing the absorbed vapors at different concentrations. Subtracting the output eigenfrequency value, obtained after exposure to a specific vapor concentration, from the input eigenfrequency value, obtained before exposure to a VOC, provides the frequency shift. Therefore, the frequency shift of a SAW sensor can be determined by dielectrically repeating this procedure [15]. In our FEM simulations, quartz was used as the piezoelectric substrate. Figure 3 shows the resonant frequency of the SAW-based gas sensor before exposure to any VOCs. This resonance is caused by constructive interference as most of the

acoustic wave propagations occur on the surface and their amplitude decreases according to material depth [16].

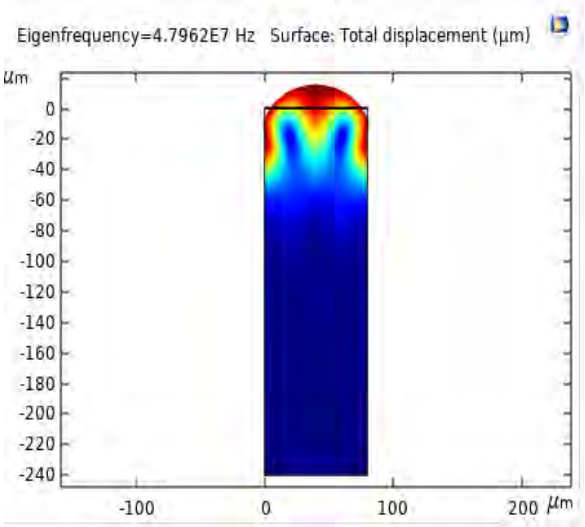


Fig. 3: 2D model of the simulated SAW-based gas sensor.

4.1 SAW Sensor Sensitivity According to Polymer

The five polymer-coated sensors were used to detect DCM vapor. Shifts in frequency were then calculated for analysis. The results of our FEM simulation for each polymer-VOC interaction are shown in Fig. 4. As frequency shift is affected by vapor concentration, the DCM vapor was set at concentrations of 500 ppm to 3000 ppm. The PIB-coated sensor had the highest sensitivity (4.0571 kHz/ppm) followed by the OV25-coated sensor (3.7143 kHz/ppm). PIB is highly sensitive as it is highly permeable, thereby allowing quicker absorption, desorption, and reversibility post-vapor analyte exposure which is ideal for mass-sensitive sensors such as the SAW-based gas sensor [17].

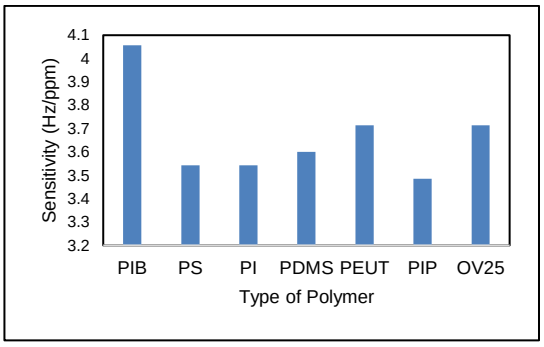
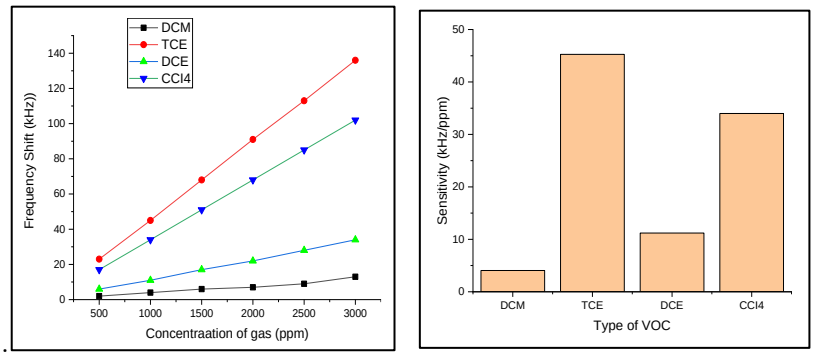


Fig. 4: The sensitivity of polymer-coated sensors to DCM vapor.

4.2 SAW Sensor Sensitivity According to VOCs

As it had the highest sensitivity, the PIB-coated sensor was used to detect each VOC which is measured by a downward shift in resonant frequency [15]. The frequency shifts, as a result of increased surface density caused by vapor absorption, of the PIB-coated sensor after exposure to DCM, TCE, DCE, and CCl_4 vapors were calculated. As shown in Figure 5(a), the higher the vapor concentration, the higher the density of vapor absorbed by the PIB-sensing layer, the higher the shift in resonant frequency. Figure 5(b) shows that the PIB-coated sensor had the highest sensitivity to TCE vapor (45.257 kHz/ppm) and lowest

sensitivity to DCM vapor (4.0571 kHz/ppm). This indicated that the PIB-coated sensor had excellent selectivity towards TCE vapor.



(a) (b)

Fig. 5: PIB-coated sensor sensitivity to DCM, TCE, DCE, and CCl_4 vapors. (a) Frequency shift according to vapor concentrations (b) PIB-coated sensor sensitivity according to type of VOC.

4.3 SAW Sensor Sensitivity According to VOCs

To further demonstrate this approach, the sensitivity of two polymeric sensing materials; PIB and OV25; to TCE, DCE, and DCM vapors was investigated. Quartz was maintained as the substrate material in this simulation. The sensor responses in relation to vapor concentrations were performed to determine the sensitivity of each polymer-vapor combination while individual sensor responses to various pairs of vapors would validate their sensitivity. Figures 6 and 7 show the responses of PIB- and OV25-coated sensors to TCE, DCE, and DCM vapors.

Figure 6 shows the frequency shifts of PIB- and OV25-coated sensors exposed to low concentrations of TCE and DCE vapors. The OV25-coated sensor had the highest sensitivity to both concentration in the range of 50 ppm to 300 ppm (52.57 kHz/ppm) concentrations of TCE vapor in comparison to the PIB-sensing layer.

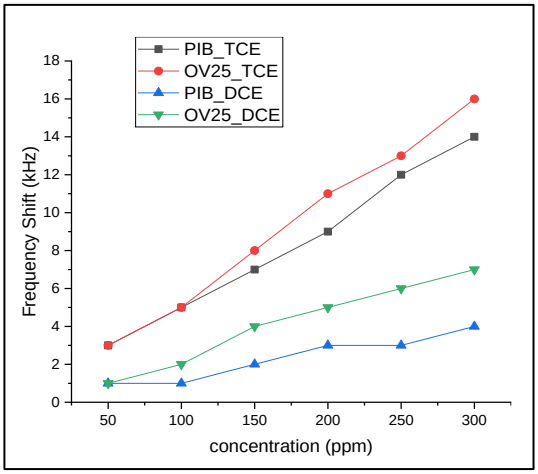
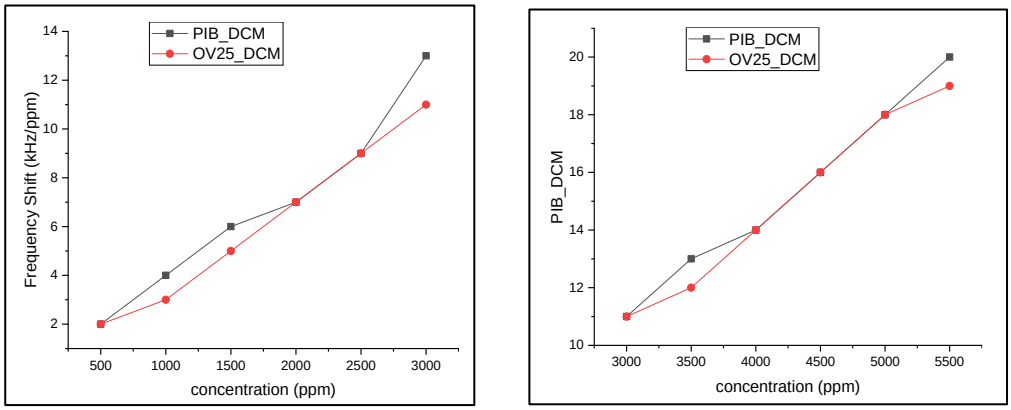


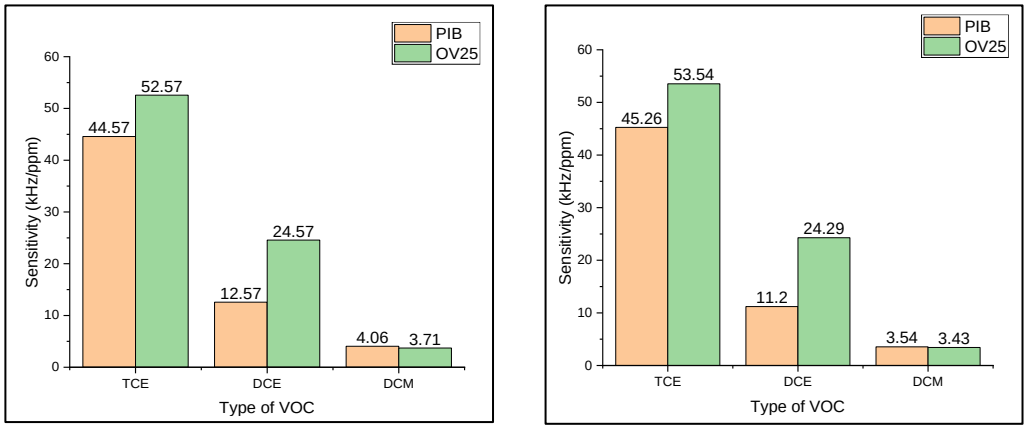
Fig. 6: Frequency shift results of the FEM simulations of PIB- and OV25-coated sensors exposed to concentrations of TCE and DCE vapors in the range of 50 ppm to 300 ppm.

DCM vapor was selected to evaluate the performance of PIB- and OV25-coated sensors as shown in Fig. 7. The PIB-sensing layer had marginally higher frequency shifts in reaction to both low (4.06 kHz/ppm) and high (3.54 kHz/ppm) concentrations of DCM vapor in comparison to the OV25-coated sensor. Figure 8 summarizes the sensitivity of both sensing layers to low (500 ppm to 3000 ppm) and high (3000 ppm to 5500 ppm) concentrations of TCE, DCE, and DCM vapors. The FEM simulations clearly indicate that frequency shifts and sensitivity vary according to the polymer-coating and type of VOC tested. Therefore, the choice of polymer-coating depends entirely on the VOC to be detected.



(a) (b)

Fig. 7: Frequency shift results of the FEM simulations of PIB- and OV25-coated sensors exposed to (a) low concentration and (b) high concentrations of DCM vapor.



(a) (b)

Fig. 8: Sensitivity results of the FEM simulations of PIB- and OV25-coated sensors exposed to (a) low concentration and (b) high concentrations of TCE, DCE, and DCM vapors.

Table 3 shows the comparison of simulation results with other previous work. The results show that the sensor is having higher shift in resonance frequency at high concentrations for both previous works and the current work.

Table 3: Comparison of simulation results with other previous work

Parameter	Lukose (2009) [12]	Zhao (2009) [15]	This work
Substrate material	Quartz	ST-X Quartz	Quartz
Resonant frequency	303.79 MHz	158 MHz	47 MHz
IDTs material	Nickel		Aluminum (Al)
Sensing layer material	Fluoroalco-holpolysiloxane (SXFA) film with varies thickness of h/lambda (0 – 0.22)	PIB, OV25	PIB, PDMS, PI
VOC vapor material with range of concentration	DMMP (10- 110 mg/m^3)	TCE, DCE Low concentration (1– 160 mg/m^3) High concentration (0.2 – 2.0 g/m^3)	DCM, TCE, DCE Low concentration (500 to 3000 ppm) High concentration (3000 to 5500 ppm)
Frequency shift (Hz)	220 Hz	Low concentration (2 – 35 kHz) High concentration (2 – 500 kHz)	Low concentration (2 – 14 kHz) High concentration (1 – 20 kHz)

5. CONCLUSION

The finite element method (FEM) simulations conducted by this study were able to successfully demonstrate the resonant frequency shifts and sensitivity of five polymeric materials: (1) polyisobutylene (PIB), (2) polydimethylsiloxane (PDMS), (3) polyisoprene (PIP), (4) polyimide (PI), and (5) phenylmethyldiphenylsilicone (OV25) to both low and high concentrations of four VOCs: (1) dichloromethane (DCM), (2) trichloroethylene (TCE), (3) 1,2-dichloroethylene (DCE), and (4) carbon tetrachloride (CCl_4). It is evident that resonant frequency shifts vary according to polymer-coated sensor and VOC interactions. The 800 nm thick PIB-sensing layer had the highest sensitivity (4.0571 kHz/ppm) to DCM vapor as well as good sensitivity (45.257 kHz/ppm) to TCE vapor. However, the choice of polymer-coating depends entirely on the VOC to be detected. The OV25-sensing layer had the highest sensitivity to both low (52.57 kHz/ppm) and high (53.54 kHz/ppm) concentrations of TCE vapor. The PIB-sensing layer had the highest sensitivity to both low (4.06 kHz/ppm) and high (3.54 kHz/ppm) concentrations of DCM vapor. It is evident that VOC concentrations also affect the sensitivity of the polymer-coated sensors. Therefore, the findings of this study indicate that polymer-coated SAW-based gas sensor are highly capable of self-powered VOC detection.

ACKNOWLEDGEMENT

This work was supported by the Ministry of Higher Education (MOHE) Malaysia under the Fundamental Research Grant Scheme (FRGS) (Grant No.: FRGS 17-032-0598).

REFERENCES

[1] Kim ID, Choi SJ, Kim SJ, Jang JS. (2015) Exhaled Breath Sensors. in: Kyung CM. (eds) Smart Sensors for Health and Environment Monitoring. KAIST Research Series. Springer, Dordrecht. https://doi.org/10.1007/978-94-017-9981-2_2

[2] Lee BC, Tward DJ, Mitra PP, Miller MI. (2018) On variational solutions for whole brain serial-section histology using a Sobolev prior in the computational anatomy random orbit model. PLoS Comput Biol 14(12): e1006610. https://doi.org/10.1371/journal.pcbi.1006610

[3] Wang C, Sahay P. (2009) Breath analysis using laser spectroscopic techniques: breathbiomarkers, spectral fingerprints, and detection limits. Sensors (Basel, Switzerland), 9(10): 8230-8262. https://doi.org/10.3390/s91008230

[4] Kim KH, Jahan SA, Kabir E. (2012) A review of breath analysis for diagnosis of human health. TrAC Trends in Analytical Chemistry, 33: 1-8.

[5] Afzal A, Iqbal N, Mujahid A, Schirhagl R. (2013) Advanced vapor recognition materials for selective and fast responsive surface acoustic wave sensors: A review. Analytica Chimica Acta, 787: 36-49.

[6] Aslam, M.Z.; Jeoti, V.; Karuppanan, S.; Malik, A.F.; Iqbal, A. FEM Analysis of Sezawa Mode SAW Sensor for VOC Based on CMOS Compatible AlN/SiO_2/Si Multilayer Structure. Sensors 2018, *18*, 1687. https://doi.org/10.3390/s18061687

[7] Ralib AAM, Nordin AN, Salleh H. (2010) A comparative study on MEMS piezoelectric microgenerators. Microsystem Technologies, 16(10): 1673-1681.

[8] Arshak, K., Moore, E., Lyons, G.M., Harris, J. and Clifford, S. (2004), "A review of gas sensors employed in electronic nose applications", Sensor Review, Vol. 24 No. 2, pp. 181-198. https://doi.org/10.1108/02602280410525977

[9] Hazadi AHF, Ralib AAM, Saidin N. (2018) Design and Simulation of Electroacoustic Sensor for exhaled breath analysis. in 2018 7th International Conference on Computer and Communication Engineering (ICCCE) pp. 388-391.

[10] Mujahid, A.; Dickert, F.L. Surface Acoustic Wave (SAW) for Chemical Sensing Applications of Recognition Layers. Sensors 2017, *17*, 2716. https://doi.org/10.3390/s17122716

[11] Choi, K., Kim, H., Ali, K. *et al.* Hybrid Surface Acoustic Wave- Electrohydrodynamic Atomization (SAW-EHDA) For the Development of Functional Thin Films. Sci Rep 5, 15178 (2015). https://doi.org/10.1038/srep15178

[12] Lukose V, Nemade HB. (2019) Finite element simulation of one-port surface acoustic wave resonator with thick interdigital transducer for gas sensing. Microsystem Technologies, 25(2): 441-446.

[13] Haresh M., Pandya, (2010) Design and Modelling of Surface Acoustic Wave (SAW) Devices and Sensors", Ph.D. Thesis, Bharathiar University, Coimbatore: 8.

[14] Liu X, Cheng S, Liu H, Hu S, Zhang D, Ning H. (2012) A survey on gas sensing technology. Sensors, 12(7): 9635-9665.

[15] Zhao YG, Liu,M, Li DM, Li JJ, Niu JB. (2009) FEM modeling of SAW organic vapor sensors. Sensors and Actuators A: Physical, 154(1): 30-34.

[16] Johnson S, Shanmuganantham T. (2014) Design and analysis of SAW based MEMS gas sensor for the detection of volatile organic gases. Carbon, 119(5): 0-041316.

[17] Sayago I, Fernández MJ, Fontecha JL, Horrillo MC, Vera C, Obieta I, Bustero I. (2012) New sensitive layers for surface acoustic wave gas sensors based on polymer and carbon nanotube composites. Sensors and Actuators B: Chemical, 175: 67-72.

https://doi.org/10.31436/iiumej.v22i2.1613

A REVIEW OF FLOW CONFLICTS AND SOLUTIONS IN SOFTWARE DEFINED NETWORKS (SDN)

MUTAZ HAMED HUSSIEN KHAIRI[1,2*], SHARIFAH HAFIZAH SYED ARIFFIN[1], NURUL MUAAZAH. ABDUL LATIFF[1], KAMALUDIN MOHAMED YUSUF[1] AND MOHAMED KHALAFALLA HASSAN[1,2]

[1]School of Electrical Engineering, Faculty of Engineering, University Technology Malaysia, Johor, Malaysia
[2]Faculty of Engineering, Future University, Khartoum, Sudan

[*]Corresponding author: taza1040@gmail.com

(Received: 17th September 2020; Accepted: 16th February 2021; Published on-line: 4th July 2021)

ABSTRACT: Software Defined Networks (SDN) are a modern networking technology introduced to simplify network management via the separation of the data and control planes. Characteristically, flow entries are propagated between the control plane layer and application or data plane layers respectively while following flow table instructions through an OpenFlow protocol. More often than not, conflicts in flows occur as a result of traffic load and priority of instructions in the data plane. Several research works have been conducted on flow conflicts in SDN to reduce their adverse effect. Solutions to flow conflict in SDN have three main limitations. First, the OpenFlow table may still cause a defect in the security module according to the priority and action matching in the OpenFlow of the control plane. Second, flow conflict detection requires more time due to flow tracking and incremental update, whereas in such a case, delay affects the efficiency of SDN. Besides, the SDN algorithm and mechanism have substantially high memory requirement for instruction and proper functioning. Third, most of the available algorithms and detection methods used to avoid flow conflicts have not fully covered the security model policy. This study reviews these limitations and suggest solutions for future research directions.

ABSTRAK: Rangkaian Perisian Tertentu (SDN) adalah teknologi rangkaian moden yang diperkenalkan bagi memudahkan pengurusan rangkaian melalui pecahan data dan kawalan permukaan. Seperti biasa, aliran kemasukan disebar luas antara lapisan permukaan kawalan dan aplikasi atau lapisan permukaan data masing-masing, sambil mengikuti arahan meja melebar melalui protokol aliran terbuka. Kebiasaannya konflik dalam aliran berlaku disebabkan oleh beban trafik dan keutamaan arahan pada permukaan data. Beberapa kajian dibuat terhadap konflik aliran SDN bagi mengurangkan kesan konflik. Solusi konflik aliran dalam SDN mempunyai tiga kekurangan besar. Pertama, jadual Aliran Terbuka mungkin masih menyebabkan kekurangan dalam modul keselamatan berdasarkan keutamaan dan tindakan persamaan dalam aliran terbuka permukaan kawalan. Kedua, pengesanan aliran konflik memerlukan lebih masa bagi pengesanan aliran dan peningkatan kemaskini, kerana setiap penangguhan memberi kesan terhadap kecekapan SDN. Selain itu, algoritma SDN dan mekanisme memerlukan memori yang agak besar bagi memproses arahan dan berfungsi dengan baik. Ketiga, kebanyakan algoritma dan kaedah pengesanan yang digunakan bagi mengelak konflik pengaliran tidak sepenuhnya dilindungi polisi model keselamatan. Oleh itu, kajian ini meneliti kekurangan dan memberi cadangan penambahbaikan bagi arah tuju kajian masa depan yang terbuka.

KEYWORDS: *software defined networking; open flow table; flow conflicts*

1. INTRODUCTION

Software Defined Network (SDN) is a modern network architecture that has been proposed to mitigate the shortcomings related to traffic management, security, virtualization and complexity of traditional networks [1]. The main characteristics of SDN include: 1) separation and abstraction of the control and data plane; 2) logical centralization of network intelligence that enables global view of the entire network and swift response changing needs; 3) ability to develop different applications from the underlying network infrastructure; 4) programmable data plane for networking automation and flexibility; 5) much higher innovation speed to accelerate business innovation and possibility of reprograming of the network by the IT network operators in real time to meet specific business needs. Moreover, through the virtualization process, the network infrastructure can be extracted from individual network services [2].

OpenFlow is a SDN protocol designed to facilitate server communication with network switches; particularly regarding sending and receiving packets. As a flow-based network virtualization architecture, it allows multiple logical networks to share the same physical infrastructure. While the network virtualization layer allows for a set of controllers to manage multiple switches per slice, controllers are responsible for installing flow entries in the assigned domain of switches. In such a design, one physical switch could belong to multiple virtual networks which may be controlled by one or more set of controllers. This design setup leads to flow conflicts [3]. In today's Internet, network traffic is routed based on the destination address prefixes. While this approach allows an efficient implementation of shortest-path (and more complex) routing protocols, it does not provide fine-grained control over network traffic. However, many proposals for future internet architectures require that the network data plane implement routing and forwarding at the level of individual connections or their aggregate e.g., for network virtualization, or for network services [4,5].

2. PROBLEM BACKGROUND AND OBJECTIVES OF THE RESEARCH

Conflict in flow entries manifest in different ways in SDN such as where a physical switch belongs to multiple virtual networks that are often controlled by one or a set of controllers; thus, resulting in the occurrence of flow conflict. Flow conflicts may also occur when vague packets match flow entries. Based on the priority of the flow entry, flow entry conflicts in an open flow table can cause defects in the security module that runs on top of the SDN. An approach used, is to match flow entry packets based on the priority of flow entries. However, because of the similarities between different flow entries, a single packet can be matched to several flow entries, thereby resulting in flow entry conflict [1].

While representative algorithms and detection methods have been proposed to mitigate flow entry conflicts [1-3], they have been shown to have high memory requirements and require a significant amount of time to apply their instructions in the flow table. Specifically, the time taken to add or update flows in case of conflict in the flow entry is high; thus, affecting the performance of the SDN network. In addition, most of these past studies used old versions of flow table in their simulations and experiments.

The present study was initiated,

i. To study and analyze flow entry conflicts in OpenFlow table.

ii. To review the existing detection methods used for flow conflict resolution in OpenFlow table.

3. SDN ARCHITECTURAL COMPONENTS

The three core components of the SDN architecture are: the SDN controller, also known as control layer; the data plane layer, which is made up of a collection of network devices like routers, switches and OpenFlow switches that act as middle boxes for network communication; and the application layer, where the execution of all network applications take place. Figure 1 shows the SDN architecture. It can be clearly observed that it is characterized by the separation of the control and infrastructure layers. This is advantageous in that the application layer can be developed and modified by application developers by using a northbound interface that provides advanced policy applications and services as well as programmable Application Programming Interfaces (API) for this purpose. Moreover, the southbound interface also provides the OpenFlow protocol, which is a standard API [4,5].

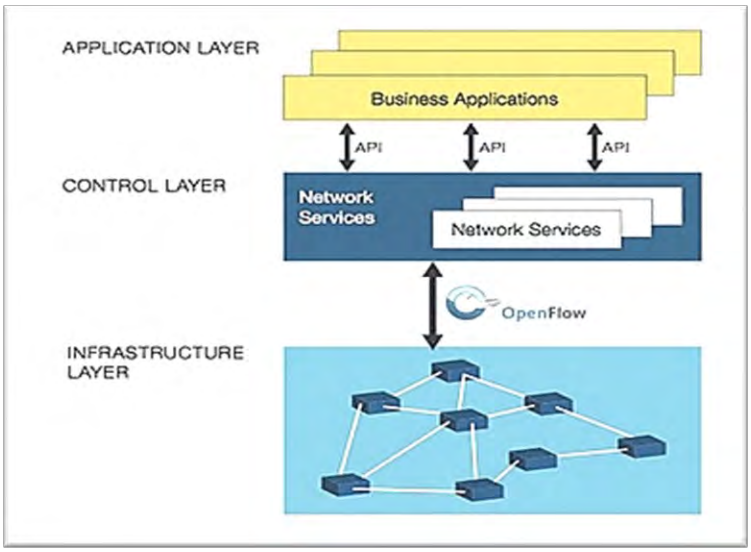


Fig. 1: Software Defined Networking framework [6].

4. OPENFLOW (OF) IN SDN

OpenFlow was one of the first software defined networking (SDN) standards. It was originally a communication and connection protocol in SDN environments that allowed the controller to directly interact with the forwarding plane of network devices for better adaptation to changing business requirements [7].

4.1 OpenFlow Switch and Protocol

OpenFlow was driven by the characteristic separation of the control and forwarding planes in SDN. The OpenFlow switch architecture is illustrated in Fig. 2. The OpenFlow switch, which also contains a number of flow tables, is connected to the controller via an OpenFlow channel. Likewise, OpenFlow switches also have an abstraction layer through which they communicate with the controller via the OpenFlow protocol. The flow table checks which packet belongs to which flow in order for the packets to be processed and delivered [1,4].

https://doi.org/10.31436/iiumej.v22i2.1613

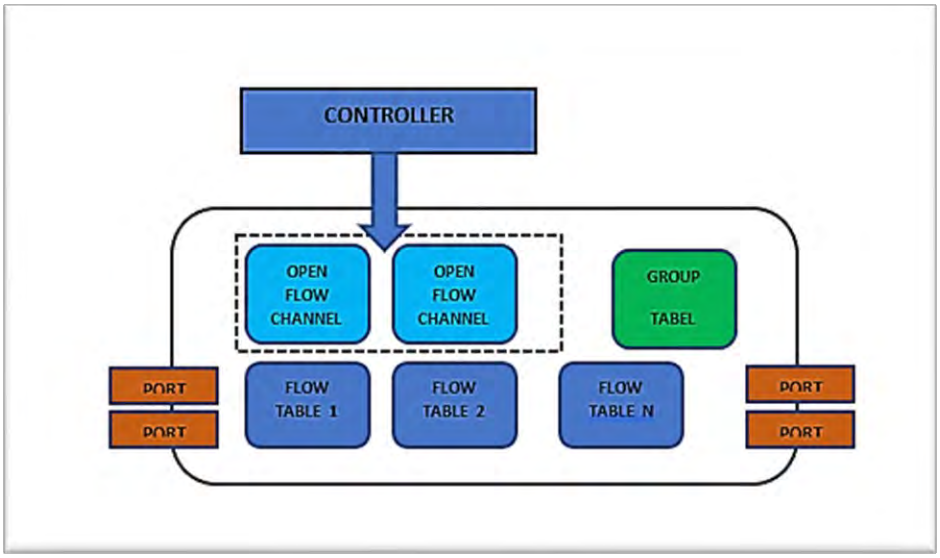


Fig. 2: OpenFlow switch [5].

- Group table: for packet lookup and forwarding.
- OpenFlow channel: responsible for connection between controller and switch.
- Flow table: flow entry – match field – counter and set instruction.

4.2 OpenFlow Ports

OpenFlow ports perform the function of transferring packets between open flow switches and other devices in the network such as routers or hosts. All OpenFlow switches in the SDN network are connected through OpenFlow ports. Some network interfaces may also be disabled in the OpenFlow switch [8].

4.3 OpenFlow Table and Entry

The OpenFlow switch is comprised of several flow tables, each of which has flow entries. The processing pipeline of OpenFlow stipulates the interaction between packets and flow tables. Depending on the application and network structure, an OpenFlow switch may have one or more flow tables [8].

4.4 Flow Table

Flow tables consist of several flow entries as shown in Table 1.

Table 1: Flow entries and their definitions [8].

Match Fields	Priority	Counters	Instructions	Timeouts	Cookie
Used to match between packets	For matching priority of flow entries	To update the proper packets	To adjust the action set of a flow entry	Maximum time or inactive time before switch terminates a flow entry	For reporting a modified or deleted flow entry to controller

5. EXISITING SOLUTIONS TO FLOW CONFLICTS IN SDN

This section discusses the research that has been conducted on detecting flow conflicts in SDN. The studies are explained based on their method, controller type, switch, conflict type, topology, and target application respectively, as in Table 2.

Table 2: List of Existing Solution on Flows Conflict in SDN

Method or Techniques	Reference Research	Controller Model	Switch Model	Conflict type	Network topology	Target application
Mechanism or Algorithm	Lo et al. [1]	-	-	Flow entries	Single topology	Network optimization performance
	Hu et al. [3]	Floodlight	OpenFlow Switch	Flow rule	Tree topology	Security
	Fang and Lu [7]	-	-	Flow rule	Single topology	Network efficiency
	Pisharody et al. [9]	Open Daylight	OpenFlow switch	Flow rule	Tree topology	Security
	Lu et al. [10]	NOX	OpenFlow Switch	Policy rule	Tree topology	Security
	Wang and Youn [11]	Floodlight	Open vSwitch	Flow entries	Single topology	Network optimization performance
	Cui et al. [12]	Floodlight	OpenFlow Switch	Flow rule	Tree topology	Network efficiency
	Hao et al. [13]	NOX	OpenFlow switch	Flow rule	Tree topology	Network optimization performance
	Halder et al. [14]	OF Controller	OpenFlow switch	Flow rule	Tree topology	Security
	Batista et al. [15]	OF Controller	OpenFlow switch	Flow entries	Single topology	Security
	Hong and Wey [16]	Floodlight	Open vSwitch	Flow rule	Tree topology	Network optimization performance
Analytical method	Metter et al. [2]	-	-	Flow rule	Single topology	Network optimization performance
	Lin et al. [17]	Ryu	OpenFlow switch	Flow entries	Tree topology	Network optimization performance
	Tran and Danciu [18]	-	-	Flow rule	Tree topology	Network optimization performance
	Yoshioka et al. [19]	-	-	Flow rule	Single topology	Link utilization

5.1 Purpose of Flow Conflict Detection Techniques

Flow entry conflict detection methods in SDN applications have been used for different purposes such as security, load balance, and firewall amongst others, as shown in Table 2. A number of detection algorithms have attempted to optimize the performance of the

controller by checking and arranging its workload. In [16], the controller was used to control elephant and safety-critical flows using interactive routing. In the study, switch layers were designed to move micro flows through multipath wildcard routing without soliciting or using the controller plane. A firewall policy, named Flow Guard, that is meant to secure the policy contravention in flow table of the OpenFlow switch with controller was proposed in [3]. It tracks flow or path spaces in the network to detect firewall policy contravention within network updates. In a separate work in [20], a method was proposed to check and detect contradiction between the rules installed in a firewall with flow table instructions. The rules applied in the OpenFlow design is to increase the safety and security capabilities for SDN.

5.2 Method and Algorithm Used for Flow Conflict Detection

One of the famous approaches used in the detection of conflicts in flow entries within SDN is the Reduced Bit Vector algorithm. This algorithm uses the concept of bit vectors alongside one of the classification methods to classify flow entries in two main groups. In the first group, flows have the same prefix length while the second group contains a decrease in the redundant bits of vectors [1]. One of the solutions introduced and built for optimizing and securing flow entry in SDN research is a security module that minimizes the amount of controller workload by separating the responsibilities of each controller [16]. A simple analytical model is formulated to enable network performance optimization relative to the signaling rate and table occupancy [2]. The research used a flow guard method for building a firewall to conduct checks and detect violation of firewall policy along the network flow path when updating the status of the network [3]. A routing method is used to ensure optimal link utilization as well as in the minimization of flow entry size. To ensure optimal link utilization, the proposed method assigns the same paths to flows that can be grouped together [19]. A security application, named FRESCO, has been proposed and implemented to address security challenges. The FRESCO framework consists of an application layer and a security enforcement kernel [21] Brew. The proposed policy is used to analyze the controller settings to detect conflicts and create solution modules which avoids conflict in flow rules within a distribution-based SDN system [9]. Source routing has also been used in flow route control. In [22], a Source Routing Based Link Protection Method whose link has paths to dedicated backup was used to direct packets to their backup paths by updating the header of the source router in case of link failure. Performance evaluation of this method has been shown to significantly update the source routing header.

5.3 Experimental Environment Used in Flow Conflict Detection

The experimental environment used for flow entry conflict detection in SDN varies according to the method or algorithm. This section details the experimental environments used in flow conflict detection studies in SDN. [16] developed a simple network with four switches (grid network) connected to one host. Using a floodlight controller with OpenFlow switch version 1.3, all tests and observations were simulated and analyzed with Mininet software. Many studies in SDN use MATLAB software with Mininet and virtual machine software as simulation tools for analyzing and optimizing the traffic and flow entry in SDN. Figure 3 shows an example of a similar research, where the design structure depends on how the OpenFlow switch deals with controller under FRESCO instruction [2].

Additionally, Fig. 4 shows the Brew system model. It uses two virtual switches connected to a controller where all the conflicting flow entries are checked and analyzed in line with the controller plane [9].

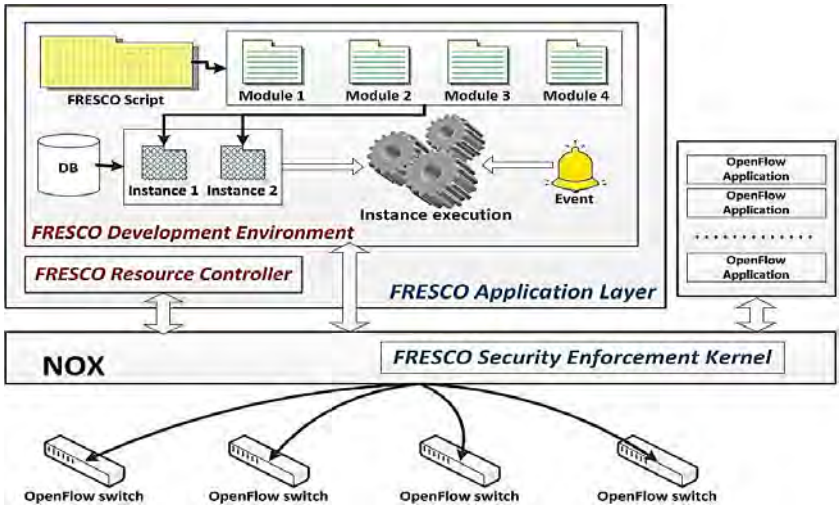


Fig. 3: High level overview of the FRESCO architecture [2].

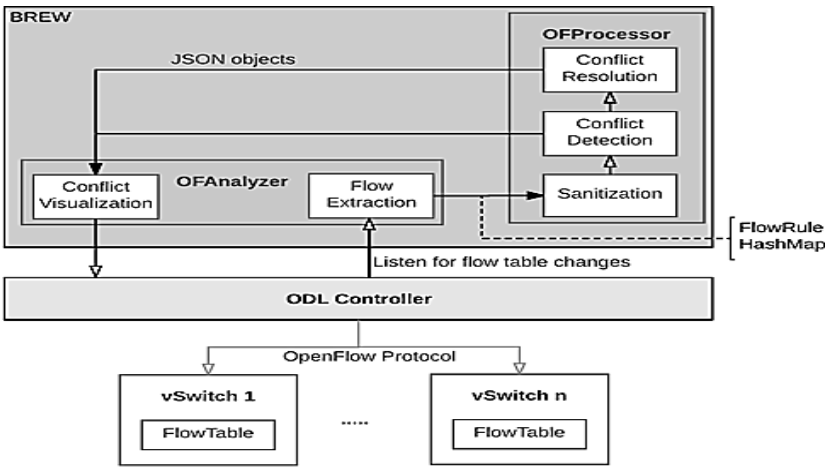


Fig. 4: System architecture showing Brew solution models [9].

5.4 SDN Controller Used in Flows Conflict Detection Techniques

The SDN controller is called the "network brain" [23]. It ensures that different network tasks are performed through effective management of the application modules of the network. The applications typically leverage an API, as in [24], to enable advanced network functionalities while facilitating communication with core controller modules. We discuss both commercial and open-source controllers that are widely used in load balancing modules of the server. While the software load balancing module in the commercial controllers are standalone systems connected to the controller for API utilization purposes, they function as part of the controller system within the application layer in the open-source version. A popular open-source controller is the high-performance java-based OpenFlow controller called Flood Light Controller, which is based on Beacon controller, an experimental OpenFlow controller from Stanford University that now has a large community of developers that supports it. The latest version of Floodlight has support for OpenFlow version 1.4 [25]. Open Daylight (ODL) SDN controller is another type of controller used to implement Brew [9].

5.5 Parameter Measure in Flow Conflict Detection

Flow entry in SDN is generally measured by different types of parameters, we focus only on parameters such as time of search, update time, elephant flows, and Link Bandwidth as measured parameters in the flow entry conflict detection.

6. ANALYSIS AND DISCUSSION

We aim, by this research, to define the types of flow rule conflict that occur in SDN. A number of researchers have used a Static Programming language module by leveraging the OF Analyzer module named Brew; a security policy analysis framework that is developed on an Open Daylight (ODL) based SDN controller [9]. In [10], a method was proposed for accurate detection of universal conflicts. In the proposed mechanism, checks are carried out for policy conflict with machine learning for 500 policies applied in a simulated network. The study also introduced a novel flow management project for MFTs for Open vSwitch in SDN. The memory requirement of the research shows that the Ex table requires a large memory space [11].

The authors of [7] used the MTBDD algorithm to conduct and clarify switch conflict rules in SDN. The test of algorithm offers good functionality on incremental updates of two different rule groups from Stanford's Internet network. In a separate research, TCDR was used to reject illegal flow rules, thus, avoiding substantial conflicts in flow rules while incurring minimal overhead [12].

The detection mechanism for conflicting flows in the rules of a flow table as presented in [13] takes 1ms of time. However, it should be noted that the effect of the firewall rule is not considered for table update. In [14], a graph based technique is applied to produce a guide graph from obtainable flow rules built-up by several controller applications within SDN to determine flow transgression. A classification method is proposed to classify the flow entries in two main groups; with the first group of the flow having same prefix length, while the second group contains a decrease in the redundant bits of vectors . In using RVB, there is still notable delay in time of search and incremental update (over 2 ms) [1]. In [15], the author used an intelligent technique for conflict in flow entries within SDN. While the proposed method effectively detected most of the conflicts that appeared in the flow table, its high memory requirement remained a major shortcoming.

Lin et al. [17] used most of the parameters configured in OpenFlow tables to conduct cross-layer testing between two different flow entries. The priority of flow of the entire inflow table was experimentally observed. The proposed method showed impressive accuracy. The research analytically showed that the conflict of applications in SDN can be detected and analyzed without stopping the application [18].

An analytical model has been proposed for SDN Signaling Traffic and Flow Table Occupancy in [2]. The model was designed to allow network performance optimization relative to signaling rate and table occupancy. The importance of this model could decrease in the near future when switches are upgraded or the size of the flow table is increased. A method of optimizing controller workload with a view to minimizing the amount of controller workload by separating the responsibility of each controller has been demonstrated in [16]. The potential services considered and checked in this research do not include File Transfer Protocol (FTP).

A method, called flow entry aggregation, capable of reducing the maximum number of flow entries while suppressing the maximum link utilization, was proposed in [20].

However, a key drawback of this method is that it cannot by itself prevent any loops from occurring in the network; thus, it is susceptible to flow entry conflicts [19].

7. CONCLUSION

What has been presented in this research is mainly concerned with flow conflict issues in SDN and how they have been mitigated in the flow table. According to the reviewed and discussed articles, flow conflicts can be classified into two major types, flow rules and flow entries. Summarily, researchers have used only two methods to detect flow conflict in SDN; majority of them using an algorithmic approach while a few used analytical methods. Floodlight, Open Daylight, POX and Ryu controllers were the most commonly used controllers in experimental and test studies while only two types of network topologies, tree and single topology, were used in the pertinent literatures. The target application in most studies is network optimization and performance while only a few studies targeted security issues.

From all the reviewed studies, a research involving the use of Artificial Intelligence such as the machine learning method to detect flow conflicts in the OpenFlow table for network security related issues still remain lacking. Thus, this study proposes an anomaly detection algorithm using machine learning to detect and classify flow conflicts in the OpenFlow table as an interesting future research direction. The goals of such an algorithm will be to classify flow rules while concurrently reducing the delay in the update flow table of the controller.

REFERENCES

[1] Lo C, Wu P, Kuo Y. (2015) Flow entry conflict detection scheme for software-defined network. International Telecommunication Networks and Applications Conference (ITNAC), Sydney, Australia, pp. 220-225.

[2] Metter C, Seufert M, Wamser F, Zinner T, Tran-Gia P. (2017) Analytical model for SDN signaling traffic and flow table occupancy and its application for various types of traffic. IEEE Transactions on Network and Service Management, 14(3): 603-615.

[3] Hu H, Han W, Ahn GJ. Zhao Z. (2014) FLOWGUARD: Building robust firewalls for software-defined networks. in Proceedings of the third workshop on Hot topics in software defined networking August 2014. pp 97-102.

[4] Akyildiz, IF, Lee A., Wang P, Luo M, Chou W. (2014) A roadmap for traffic engineering in SDN-OpenFlow networks. Computer Networks, 71: 1-30.

[5] Bozakov Z, Sander V. (2013) OpenFlow: A perspective for building versatile networks. in: Network-Embedded Management and Applications, Clemm A, Wolter R. (eds). Springer, New York, NY. https://doi.org/10.1007/978-1-4419-6769-5_11

[6] sdx central [https://www.sdxcentral.com/networking/sdn/definitions/what-the-definition-of-software-defined-networking-sdn]

[7] Fang Y, Lu Y. (2019) Checking intra-switch conflicts of rules during preprocessing of network verification in SDN. IEEE Communications Letters, 23(9): 1547-1550.

[8] OpenFlow Specification [http://networkstatic.net/wp-content/uploads/2013/02/openflow-spec-v1.3.0.pdf]

[9] Pisharody S, Natarajan J, Chowdhary A, Alshalan A, Huang D. (2019) Brew: A security policy analysis framework for distributed SDN-based cloud environments. IEEE Transactions on Dependable and Secure Computing, 16(6): 1011-1025.

[10] Lu Y, Fu Q, Xi X, Chen Z, Zou E, Fu B. (2019) A policy conflict detection mechanism for multi-controller software-defined networks. International Journal of Distributed Sensor Networks 15 No (5),http://doi.org/10.1177/1550147719844710.

[11] Wang C, Youn HY. (2019) Entry aggregation and early match using hidden Markov model of flow table in SDN. Sensors, 19(10), 2341; https://doi.org/10.3390/s19102341.

[12] Cui J, Zhou S, Zhong H, Xu Y, Sha K. (2018) Transaction-based flow rule conflict detection and resolution in SDN. 27th International Conference on Computer Communication and Networks (ICCCN), Hangzhou, China, pp. 1-9.
[13] Hao W, Jiang Y, J. Gao J. (2017) Detection mechanisms of rule conflicts in SDN based on a path-tree model. 8th IEEE International Conference on Software Engineering and Service Science (ICSESS), Beijing, China, pp. 336-339.
[14] Halder B, Barik MS, Mazumdar C. (2017) A graph based formalism for detecting flow conflicts in software defined network. 2017 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), Bhubaneswar, India, pp. 1-6.
[15] Lopes Alcantara Batista B, Lima de Campos GM, Fernandez MP. (2014) Flow-based conflict detection in OpenFlow networks using first-order logic. 2014 IEEE Symposium on Computers and Communications (ISCC), Funchal, Portugal, pp. 1-6.
[16] Hong ETB, Wey CY. (2017) An optimized flow management mechanism in OpenFlow network. 2017 International Conference on Information Networking (ICOIN), Da Nang, Vietnam, pp. 143-147.
[17] Lin YD, Lai YK, Tsou YL, Lai YC, Liou EC, Chiang Y. (2019) Generic validation criteria and methodologies for SDN applications. IEEE Systems Journal, 13(4): 3909-3920.
[18] Tran CN, Danciu, V. (2020) A general approach to conflict detection in software-defined networks. SN Comput. Sci, 1(1), 9, https://doi.org/10.1007/s42979-019-0009-9
[19] Yoshioka K, Hirata K, Yamamoto M. (2017) Routing method with flow entry aggregation for software-defined networking. 2017 International Conference on Information Networking (ICOIN), Da Nang, Vietnam, pp. 157-162,
[20] Pallavi N., Anisha A.S., Leena V. (2017) Detection of Incongruent Firewall Rules and Flow Rules in SDN. In: Dash S., Vijayakumar K., Panigrahi B., Das S. (eds) Artificial Intelligence and Evolutionary Computations in Engineering Systems. Advances in Intelligent Systems and Computing, vol 517. Springer, Singapore. https://doi.org/10.1007/978-981-10-3174-8_2.
[21] Shin SW, Porras P, Yegneswara V, Fong M, Gu G, Tyson M. (2013) Fresco: Modular composable security services for software-defined networks. in February 2013- 20th Annual Network & Distributed System Security Symposium. http://hdl.handle.net/10203/205914
[22] Huang L, Shen Q, Wenjuan S. (2016) A source routing based link protection method for link failure in SDN. 2nd IEEE International Conference on Computer and Communications (ICCC), Chengdu, China, pp. 2588-2594.
[23] Jarschel M, Zinner T, Hoßfeld T, Tran-Gia P, Kellerer W. (2014) Interfaces, attributes and use cases: A compass for SDN. IEEE Communications Magazine, 52(6): 210-217.
[24] Zhou, J. (2014) Multicatalyst system in asymmetric catalysis. John Wiley & Sons.
[25] OpenFlow Switch Specification OpenFlow Switch Specification, O.S., Version 1.4. 0, October 14, 2013. [https://opennetworking.org/wp-content/uploads/2014/10/openflow-spec-v1.4.0.pdf]

https://doi.org/10.31436/iiumej.v22i2.1618

EXPERIMENTAL VERIFICATION OF THREE MODE CONTROLLER FOR HOME APPLIANCE USING WIRELESS TECHNOLOGY

MASWOODHUR RAHMAN ABDUL WAHIDH AND MURUGANANDAM MASILAMANI

[1]*Engineering Department, University of Technology and Applied Sciences, Ibri, Sultanate of Oman*

**Corresponding author: maswood.abdul@ibrict.edu.om*

(Received: 20th September 2020; Accepted: 12th November 2020; Published on-line: 4th July 2021)

ABSTRACT: An attempt is made to enhance the automation and sophistication of a home appliance. In this work, electric lamp is considered as a home appliance and it is controlled through three different modes. Technology is a never-ending process, the home automation is a well-known field of recent research, which makes the customer to reach the next level of sophistication. One such next level is presented in this article. The proposed paper is about design, experimentation and testing of three-way control for a lamp, using low cost microcontroller, Bluetooth module and android mobile application. The three modes of control are manual switch mode, Light Dependent Resistor (LDR) mode and Bluetooth based Wireless communication mode. The work is initially simulated using Proteus 8 Professional Application and then implemented as a real time working model. Also, the working performances were tested using the Simulation and developed working model. The control in Bluetooth mode is based on an Android Application. A dedicated Android Application is developed for testing the developed working model, using MIT App Inventor.

ABSTRAK: Percubaan dibuat untuk meningkatkan automasi dan kecanggihan perkakas rumah. Dalam karya ini, lampu elektrik dianggap sebagai perkakas rumah dan ia dikendalikan melalui tiga mod yang berbeza. Teknologi adalah proses yang tidak pernah berakhir, automasi rumah adalah bidang penyelidikan baru-baru ini, yang menjadikan pelanggan mencapai tahap kecanggihan yang seterusnya. Satu tahap seterusnya ditunjukkan dalam artikel ini. Makalah yang dicadangkan adalah mengenai reka bentuk, eksperimen dan pengujian kawalan tiga arah untuk lampu, menggunakan mikrokontroler kos rendah, modul Bluetooth dan aplikasi mudah alih android. Tiga mod kawalan tersebut ialah mod suis manual, mod Perintang Bergantung Cahaya (LDR) dan mod komunikasi Tanpa Wayar berasaskan Bluetooth. Karya ini pada mulanya disimulasikan menggunakan Proteus 8 Professional Application dan kemudian dilaksanakan sebagai model kerja masa nyata. Juga, prestasi kerja diuji menggunakan Simulasi dan model kerja yang dikembangkan. Kontrol dalam mod Bluetooth didasarkan pada Aplikasi Android. Aplikasi Android khusus dikembangkan untuk menguji model kerja yang dikembangkan, menggunakan MIT App Inventor.

KEYWORDS: *Proteus 8 Professional; Arduino Uno; Bluetooth module; Android phone; MIT App Inventor*

1. INTRODUCTION

In the modern era of automation and sophistication, there are tremendous changes and impacts in the human life style [1,2]. Undoubtedly, electricity has a major role in the

betterment of human life style. The expectation of sophistication in our day to day life is on the increasing trends. The field of automation, drives the controllability of electrical appliances to an innovative level, for meeting the future requirements. Wireless communications are becoming more popular around the world and the consumers appreciate this wireless life style which gives them relief from the congestion of more wires [3,4].

Usually the control of electric applications is achieved by using switches i.e., when switched ON, the electric supply will be given to the applications and vice versa. If the user wants to control the application, he should be near the switch. In case if he wants to control the application from a faraway place then he should carry a remote always. Bluetooth (wireless) technology finds its usage in the field of home automation [5,6].

The day to day advancements in smart phones and its features, make the users to think about replacement of the recent smart phones as the remote for the home appliances [7,8]. Using a Smart Phone as a remote not only decrease the human intervention, also it reduces the time of control. Adopting the Smart Phones as a remote, to control the appliances will make it easy to do switching operation of an electric appliances, by elderly or handicapped persons. So, the mobile phones were introduced into the monitoring and control of home appliance, by avoiding dedicated remote to control the electrical appliances [9, 10].

Many researches have been made and exhibited related to the control of a Lamp using Arduino UNO microcontroller. But in all the basic controlling strategy we can control the Lamp by only one mode, i.e., the Manual Switch mode. While going for the Sensor based lamp control, an electric lamp can be controlled according to the signal received from the sensor. In this a Light Dependent Resistor (LDR) is used to detect the surrounding light and control (ON / OFF) the lamp accordingly [11-13].

In the recent days, some home automations were also proposed using Bluetooth Communication. The traditional control of home appliance through wall mounted switches were changed into wireless control of appliances. Here home electrical appliances were controlled through Bluetooth communication between the Bluetooth Module (HC-06) and the Bluetooth enabled android smart phone. The Bluetooth mobile phone will work as a transmitter and the Bluetooth Module connected with the microcontroller will act as the receiver. [14-16].

To command the Bluetooth receiver, the Bluetooth enabled Android Phone, i.e., the Bluetooth transmitter, should have the proper Application installed in it, to ensure proper control of electrical appliances. The research work related to the development of an Android Mobile Application development, for the purpose of home automation by Arduino Microcontroller and Bluetooth module has been proposed [17, 18]. These papers gain us ideas about Android Mobile Application development using MIT App Inventor, for lamp control using Bluetooth, through three different modes.

The previous researches were proposed to control the electric lamp in any one mode (or in very few cases two modes) of operation. This will give the user only a certain level of sophistication. The implemented system is designed to integrate all the three individual control techniques into a single control system, which controls the electric lamp in all the above said 3 different modes, say Manual switch mode, LDR mode and Bluetooth mode. This will make the user of lamp to reach the next level of sophistication.

2. THE PROPOSED METHOD

Block diagram of the proposed system is illustrated in Fig. 1. The control of lamp is based on two types of connectivity. A wired channel controls the lamp in Manual mode or

LDR mode and another wireless channel controls the lamp in Bluetooth mode. For the manual switch mode of operation, a two pin SPST switch is used. For the LDR mode, a 10K LDR is used in series with a 10K ohm resistor, as a voltage divider network. The control signal to the microcontroller is taken from the mid-point between the resistor and the LDR.

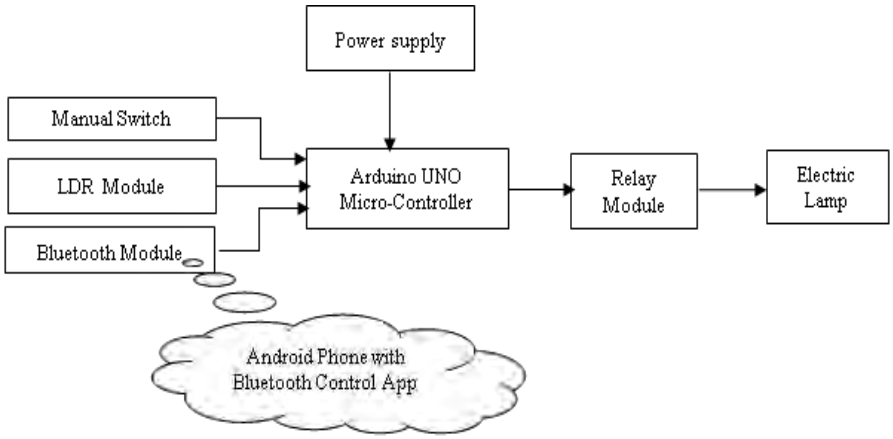


Fig. 1: Block diagram of the proposed system.

The Bluetooth communication is achieved by using HC-06 module. This module will establish a wireless connectivity between the controller and the android mobile phone. Upon receiving the control signal from the android mobile phone, the micro controller will change the lamp status accordingly. The android mobile will have a dedicated application which is developed for controlling the lamp. The lamp is connected to the microcontroller through relay module.

3. HARDWARE DESCRIPTION

3.1 Arduino UNO

Arduino is an open-source electronics platform supported with easy to use hardware and software supports. Arduino UNO is a low-cost microcontroller board, built around ATMEGA328p processor. It has 6 analog pins and 14 digital pins. The board can be powered by USB cable or 12V external power supply. The board can be programmed to control any pins, through Arduino Software (IDE). In recent years the Arduino boards finds it application in all of the electronics projects ranging from beginner level to highest level like scientific solutions etc.

3.2 Relay Module

It is an electrically isolated switch, which isolates the power circuit from control circuit. The control signal (5 V) which will be available in one off the digital pin can be connected to the input side of the relay. Now the output of this relay can even be connected to any 240 V power supply side, if we desire to control a 240 V load. This Switch operates through a relatively small electric current that can turn or off a much larger electric current.

3.3 Bluetooth Module

The Bluetooth module used here is HC-06. It is easy to use Bluetooth Serial Port Protocol Module. This module works on serial communication. HC-06 is a Master/Slave module. By default, every module will be set to Slave. It can be reconfigured to Master, through AT commands. As it is a Slave module, it can only accept connections, whereas Master module can initiate a connection to other devices. The operating range of this module

is around 2.4 GHz, with a distance of about 10 meters at a speed of 1 Mbps. The Smart phone will send serial data to the Bluetooth module when a button is pressed on the app.

3.4 Light Dependent Resistor (LDR)

LDR is a variable resistance whose resistance value will vary according to the intensity of light falling on lit. The LDR will offer less resistance when the Light intensity on the LDR is more and the resistance offered will be more when the light intensity on the LDR is less. Due to this property, it will be commonly used as Day/Night Sensor in Street Lighting Projects. Since on Day time, the resistance of LDR will be less and on Night time resistance of LDR will be high.

4. SOFTWARE DESCRIPTION

4.1 Arduino IDE

The Arduino IDE (Integrated Development Environment) is an open-source software that allows the user to do coding, editing, compiling of program. The written program will be uploaded into the Arduino boards. Arduino IDE was available for all of the major OS. The Arduino IDE has inbuilt functions and libraries that works on Java platform.

4.2 MIT App Inventor

This is also another Integrated Development Environment (IDE) used to develop applications for Android Phone using web browser. This application can be connected to a mobile phone or an emulator. In MIT App Inventor, there will be two main windows, one is the designer window to design the Graphical User Interface (GUI) and the second window is the Block Programming Window. Figures 2 and 3 show the Designer Window and Block programing window respectively.

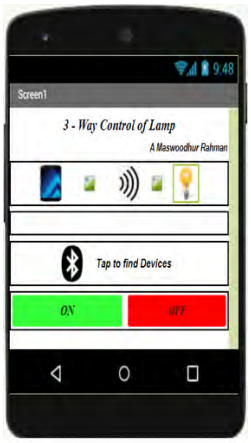


Fig. 2: Android application developed.

In MIT App Inventor, there will be no coding to develop an android app, rather we will be doing a much easier block programming (pick and place programming). This App inventor is available for all major OS and easy to upload the developed app in any android smart phone. The most interesting feature of this paper is the graphical user interface, developed in MIT App inventor. This will be installed in the android smart phone. The smart phone used throughout this paper was Samsung A5-2017.This user interface is developed in such a way that; it will help the user to pair the smart phone with the Arduino Bluetooth module. The communication between the Bluetooth module and the Android Smart phone will be using Bluetooth protocol. The GUI will have a list picker, which will list all the

available Bluetooth devices, around the Android Smart phone. Also, the GUI will have a label, to display the Bluetooth device connectivity status.

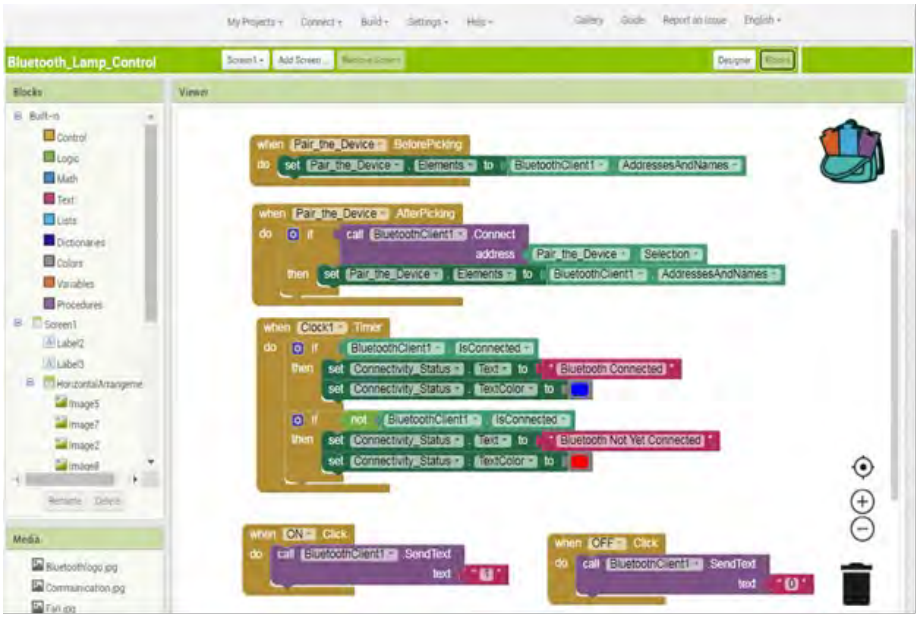


Fig. 3: Android application developed.

In addition, two buttons were provided in the application, one to switch ON the lamp and another button to switch OFF the lamp. To increase the facility for elderly/physically changed people, voice reply is also added in the application to say the status of lamp.

5. SIMULATION OF THE PROPOSED SYSTEM

Proteus is the one of the simulation software to do the all type of electronics project from simple rectifier circuit to advanced level controllers with all the recent sensors, switches, wireless communication devices and controllers. Here the Proteus 8 professional application is used for the proposed system simulation.

Figure 4 depicts the complete circuit diagram of the proposed system with Proteus simulation software. Here the same Arduino code and the android mobile application are used for controlling the lamp for both simulation and for hardware implementation. In simulation the Arduino sketch is initially compiled and the developed hexadecimal code path is identified and it will be entered into the Arduino controller. Similarly, for the Bluetooth, the input COM port number is identified and it will be entered into the Bluetooth device for connectivity between mobile and the computer for simulation.

Figure 5 illustrates the simulation output, in which the lamp is turned ON by manual switch. Out of three mode of control the first mode is simulated and illustrated here. From Fig. 5, it is clear that the LDR torch is near, meaning that more light is fall on the LDR also no signal from Bluetooth device, only the manual switch is turned ON and the lamp is glowing.

https://doi.org/10.31436/iiumej.v22i2.1618

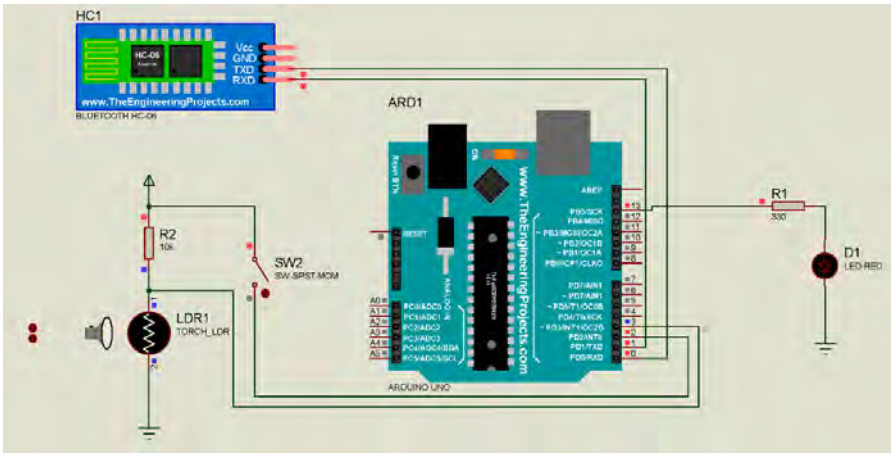


Fig. 4: Proteus simulation – lamp OFF condition.

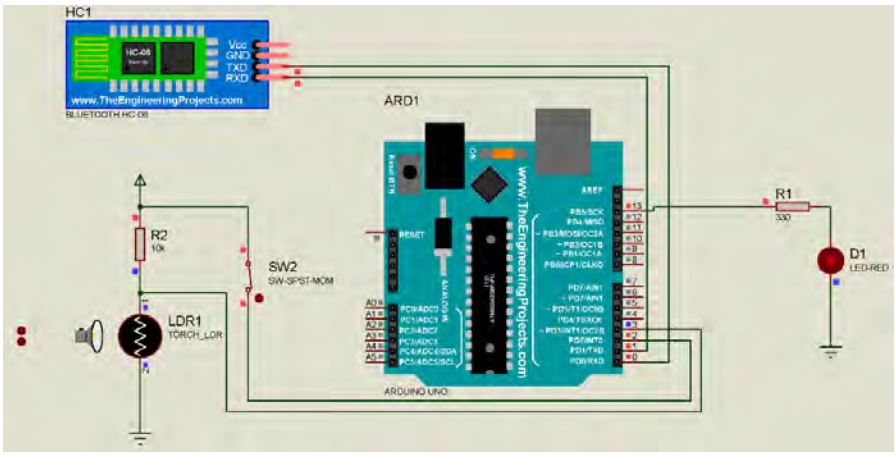


Fig. 5: Proteus simulation – lamp ON through manual switch.

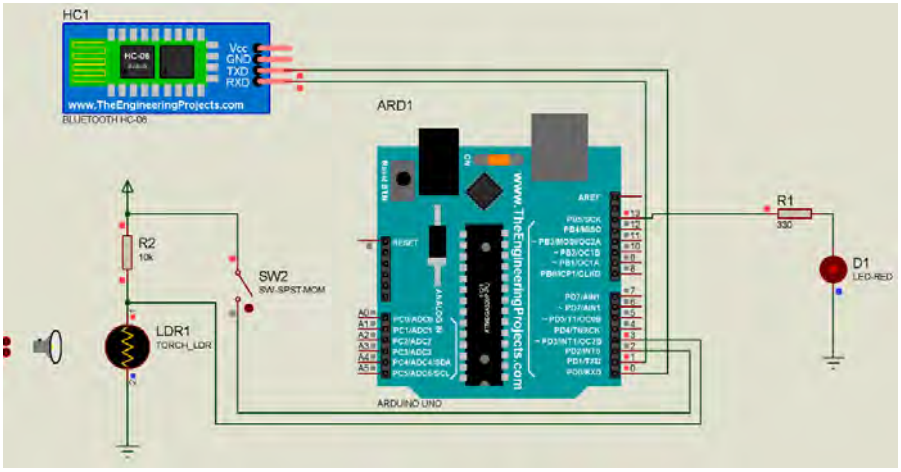


Fig. 6: Proteus simulation – lamp ON through LDR.

Figure 6 demonstrates the simulation output, in which the lamp is turned ON by LDR sensor. Out of three mode of control, the second mode is simulated and demonstrated here. From Fig. 6, it is clear that the manual switch is turned OFF also no signal from Bluetooth device, but the LDR torch is moved away meaning that the light is not falling on the LDR, hence the lamp is glowing. This mode is automation mode whenever the surrounding light intensity is less or dark, the lamp will be turned ON automatically by using the LDR sensor.

https://doi.org/10.31436/iiumej.v22i2.1618

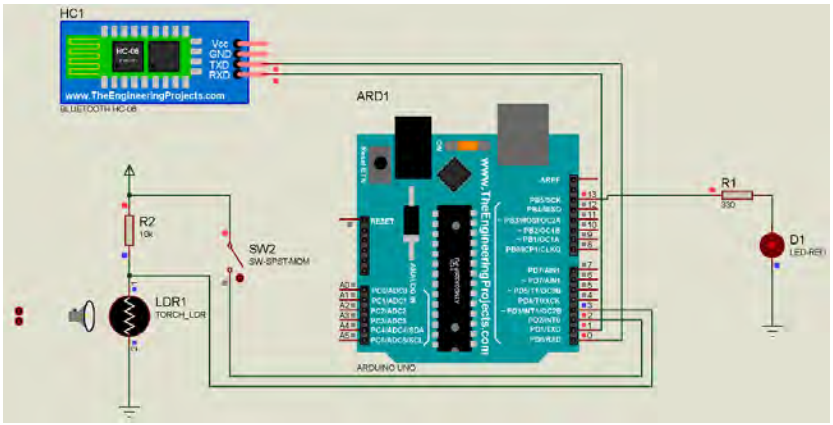


Fig. 7: Proteus simulation – lamp ON through Bluetooth communication.

Figure 7 shows the simulation output, which the lamp is turned ON by Bluetooth communication. Out of three mode of control the third mode is simulated and displayed here. From Fig. 7, it is clear that the manual switch is turned OFF also the LDR torch is near, meaning that more light falls on the LDR, but the Bluetooth device receives signal from the android mobile app which is developed using MIT app inventor, hence the lamp is glowing. This mode is sophistication mode whenever the light illumination is needed then the lamp may be turned ON using the mobile application which was developed and installed on the user's android smart phone.

6. IMPLEMENTATION AND VERIFICATION OF THE PROPOSED SYSTEM

The lamp will be controlled in three different ways. One is the manual mode, by using a SPST switch. Second is the automatic mode. In this mode, the LDR will work as a Day/Night sensor and get the surrounding light intensity. If the surrounding light is dark, the lamp will be ON. If the surrounding light is bright, then the lamp will be made OFF through the relay.

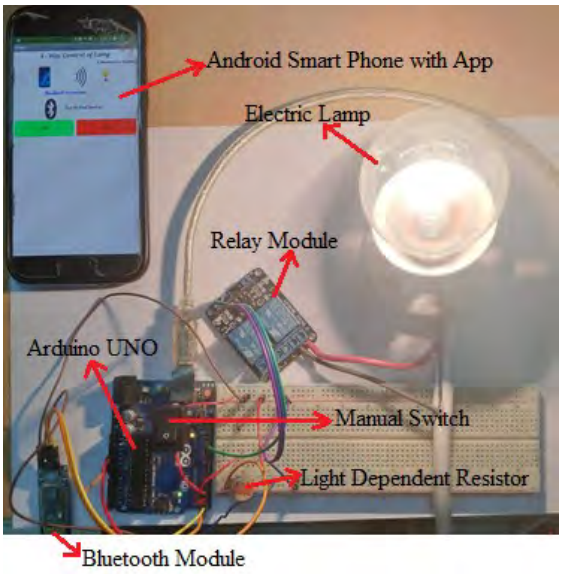


Fig. 8(a): Experimental setup (lamp ON condition).

Fig. 8(b): Experimental setup (lamp OFF condition).

https://doi.org/10.31436/iiumej.v22i2.1618

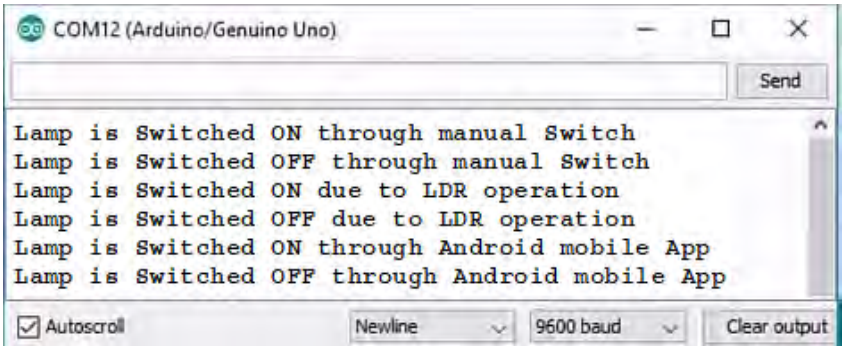


Fig. 8(c): Comment Lines on the Serial Port Monitor of Arduino IDE.

The third mode is through wireless Bluetooth communication. The Bluetooth Module will be connected to the Arduino Controller. Figure 8(a) and (b) shows the lamp status, while Fig. 8(c) show the comment lines which were made printed on the Serial Port Monitor in the Arduino IDE for the corresponding status of the lamp. Depending on the data sent from the Bluetooth module the Arduino controller will make the relay module ON or OFF. The lamp will be turned ON or OFF through the relay driver module. The data sent to Bluetooth module will be varied in the Android Smart Phone. The communication between Android Smart Phone and Bluetooth Module is through serial data communication. All the three modes of operation were represented in Fig. 9 as a flowchart.

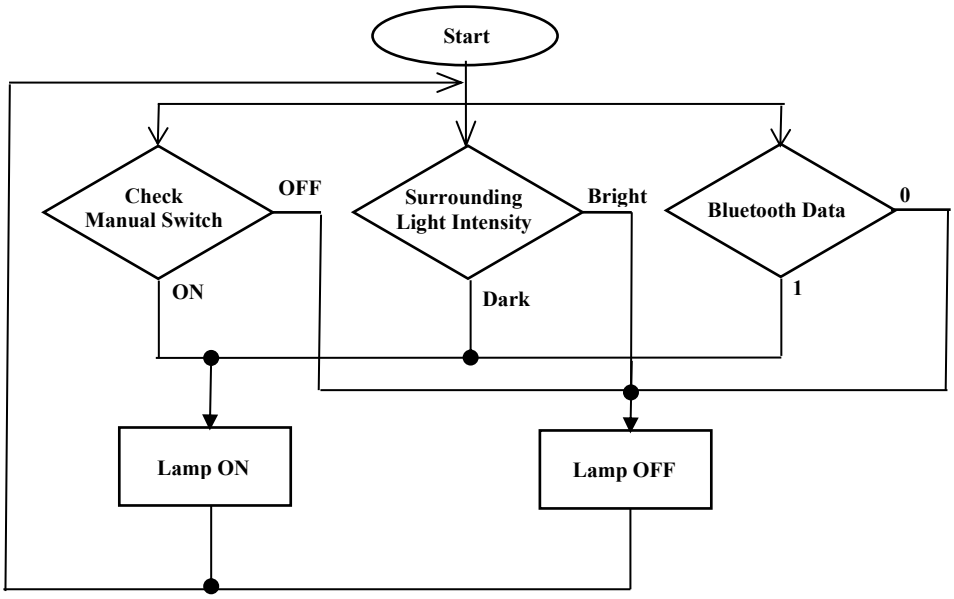


Fig. 9: Flow chart of the operation.

7. RESULTS AND DISCUSSION

The project was tested and the results obtained were represented as in a table format as shown in Table 1. The obtained results were found to be accurate. In manual operation mode, when the manual switch is made ON then the lamp will be made ON through relay. The delay time elapsed by the Lamp for switching form OFF state to ON state was measured to be 2 sec. Similarly, if the manual switch is made OFF, then the lamp will be made OFF. Again, the time elapsed by the lamp was measured to be 2 sec.

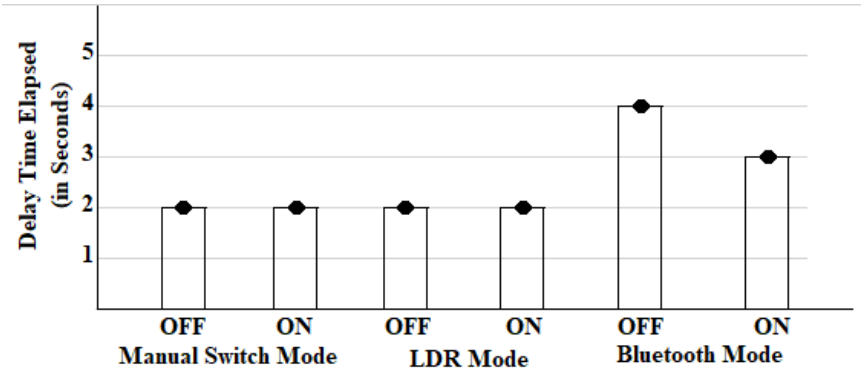


Fig. 10: Control of Lamp in different modes and delay time elapsed in respective mode

In LDR mode of operation, if the surrounding light intensity is Dark, the lamp will be made ON. Similarly, if the surrounding light intensity is Bright, then the lamp will be made OFF. This is made automated that no need of human intervention to change the lamp status, in LDR mode of control. The time taken by the lamp to change from ON state to OFF state or vice versa was measured to be 2 sec.

In wireless mode, the Android Smart phone is paired with the Bluetooth module correctly, either by Media Access Control (MAC) address or by giving a specific name to the Bluetooth Module. The distance between Android Smart phone and HC 06 Bluetooth Module was kept to be 10 meters during the test.

The control word (data) will be sent from the Android Mobile phone through the Android App developed. This data will be received by the Bluetooth module and the same data will be checked by the microcontroller. If the data received is 1 (by pressing ON button in the App), the lamp will be made ON. Similarly, if the data received is 0 (by pressing OFF button in the App), the lamp will be made OFF. In Bluetooth mode the delay time for reaching ON status in the lamp is measured to be 2 sec, and the same for reaching OFF status in the lamp is measured to be 4 sec.

8. CONCLUSION

The endowment of innovation to human kind is to make life more straight forward. In this paper, a three-way control of electric lamp was outlined, exhibited and executed. The lamp is controlled individually without the requirement of Line of Sight. The relay activity and thereafter the lamp can be controlled with an automated mode and also in Bluetooth control mode.

This proposed system not only make it easy to do switching operation of an electric appliances, by elderly or handicapped persons, but also decrease the human intervention into the lamp control cycle and reduce the time delay in manual mode of control. The control over the lamp is secured, to the access from any other intruders, through a pairing password, for the Bluetooth control. Through this Unauthorized Bluetooth device entries were avoided, which give more security to the System. This project can be used for any other electrical appliance that requires on/off switching, without internet connections.

The full functionality of the lamp was tested in all three modes using manual switch, LDR operation and Bluetooth wireless communication, between the android smart phone and the Bluetooth module. The results achieved were 100% accurate.

In the presented paper, the design, verification and GUI were designed and executed for one lamp. The delay time to turn ON the lamp is 3 secs and to turn OFF by Bluetooth mode is 4 sec. The work of this paper can be extended to reduce the delay time in the appliances control, implement "Internet of Things" instead of Bluetooth Communication, increase the distance coverage of the Bluetooth Communication to control the lamp, etc.

REFERENCES

[1] Oditis I, Bicevskis J. (2010) The Concept of Automated Process Control. Scientific Papers, University of Latvia. Computer Science and Information Technologies, 756: 193-203.

[2] Adetiba E, Matthews VO, Awelewa AA, Samuel IA, Badejo JA. (2011) Automatic electrical appliances control panel based on infrared and Wi-Fi. A framework for electrical energy conservation. International Journal of Scientific and Engineering Research, 2(7): 1-7.

[3] Muruganandam M, Maswoodhur Rahman A. (2019) Simulation and Implementation of Loop based Bank Locker Security System using Cost Effective Microcontroller and GSM Module. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 8(9): 2175-2181. https://doi.rog/10.15662/IJAREEIE.2019.0809005

[4] Cotta A, Devidas NT, Ekoskar VKN. (2016) Wireless Communication using HC-05 Bluetooth module Interfaced with Arduino. International Journal of Science, Engineering and Technology Research, 5(4): 869-872.

[5] Piyare R, Tazil M. (2011) Bluetooth based Home Automation System Using Cell Phone. IEEE 15th International Symposium on Consumer Electronics, 837-843. https://doi.rog/10.1109/ISCE.2011.5973811.

[6] Harikrishnan D, Nikhil S, Roy MT, Mathew SV. (2017) Android Based Home Automation Using Bluetooth. International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, 6(4): 2605-2609, https://doi.rog/10.15662/IJAREEIE.2017.0604073

[7] Yoo W-S, Shaik SA. (2016) Development of Home management system using Arduino and App inventor. IEEE 40th Annual Computer software and application Conference, 2: 379-380. https://doi.rog/10.1109/COMPSAC.2016.96.

[8] Gaikwad PV, Kalshetty YR. (2017) Bluetooth Based Smart Automation System Using Android, International Journal of New Innovations in Engineering and Technology, 7(3): 24-29.

[9] Bhadane DS, Wani MD, Shukla SA, Yeole AR. (2015) A Review on Home Control Automation using GSM and Blue-tooth. International Journal of Advanced Research in Computer Science and Software Engineering, 5(2): 419-421.

[10] Badrul Hisham AA, Ishak MHI, Teik CK, Mohamed Z, Idris NH. (2014) Bluetooth-Based Home Automation System Using an Android Phone, Jurnal Teknologi, 70(3): 57-61. https://doi.rog/10.11113/JT.V70.3463

[11] Vancea AP, Orha I. (2018) Smart Home Automation and Monitoring System. Carpathian Journal of Electronic and Computer Engineering, 11(1): 40-43. https://doi.rog/10.1109/10.2478/cjece-2018-0007

[12] David N, Chima A, Ugochukwu A, Obinna E. (2015) Design of a home automation system using Arduino. International Journal of Scientific & Engineering Research, 6(6): 795-801.

[13] Rath DK. (2016) Arduino Based: Smart Light Control System. International Journal of Engineering Research and General Science, 4(2): 784-790.

[14] Das S, Ganguly S, Ghosh S, Sarker R, Sengupta D. (2016) A Bluetooth Based Sophisticated Home Automation System Using Smartphone. IEEE - International Conference on Intelligent Control Power and Instrumentation (ICICPI). 236-240. https://doi.rog/10.1109/ICICPI.2016.7859709

[15] Aung T. (2019) Arduino Based Home Lighting Control by Android Phone. International Journal of Trend in Scientific Research and Development, 3(5): 1634-1638

[16] Preethi M, Dharmalingam R. (2017) Based on the wireless Bluetooth microcontroller controlling home appliances. International Journal of Advanced Research in Electronics and Communication Engineering, 6(5): 440-445.

[17] Adiono T, Anindya SF, Fuada S, Afifah K, Purwanda IG. (2019) Efficient Android Software Development using MIT App Inventor 2 for Bluetooth-Based Smart Home. International Journal of Wireless Personal Communication, 105: 233-256. https://doi.org/10.1007/s11277-018-6110-x

[18] Krishnakanth K, Kavipriya P. (2013) Android Application Development for Environment Monitoring using Smart Phones. International Journal of Mobile Network Communications & Telematics, 3(3): 41-45. https://doi.org/10.5121/ijmnct.2013.3305

https://doi.org/10.31436/iiumej.v22i2.1660

SERIES-SERIES AND SERIES-PARALLEL COMPENSATION TOPOLOGIES FOR DYNAMIC WIRELESS CHARGING

MUHAMMAD AMIRUL ASYRAF ROSLAN, NADIA NAZIEHA NANDA* AND SITI HAJAR YUSOFF

Department of Electrical and Computer Engineering,
Kulliyyah of Engineering, International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia

**Corresponding author: nnazieha.nanda@gmail.com*

(Received: 16th October 2020; Accepted: 2nd February 2021; Published on-line: 4th July 2021)

ABSTRACT: Electric vehicles (EV) have gained worldwide attention since the implementation of a wireless power transfer (WPT) to charge their batteries. With WPT, it can be very convenient for EV to be charged dynamically. Nevertheless, there are some issues in dynamic WPT, such as maintaining the power transfer efficiency. Several factors that lead to these problems include disruption of the alignment and the optimum distance between the transmitter and receiver coils. It is thus contributing to the loss of power efficiency when charging the EV. Not to mention, manufacturers build different specifications of EV charging station for different types of EV models in order to meet customer demands. An incompatible charging device will not utilize EV wireless charging to its maximum potential. Hence, to improve the power output capability as well as stabilizing the maximum power transfer during the charging process, a compensation circuit is added to the system. This article focuses on comparing two available compensation circuits (series-series (SS) topology and series-parallel (SP) topology) under the application of dynamic wireless charging. The simulations are conducted using NI Multisim based on the relationship of power transfer efficiency with resonance frequency, coefficient of coupling, and the load resistance. The WPT efficiency for SP-topology shows that it is sensitive to the change of resonance frequency and coupling coefficient, whereas SS-topology maintains good efficiency during the WPT process. Nonetheless, SS-topology performance suffers efficiency loss when paired with a higher load, while SP-topology acts differently. This article will observe the best conditions on the selected compensation designs for better application in EV charging systems in a moving state.

ABSTRAK: Kenderaan elektrik (EV) telah menarik perhatian dunia sejak pelaksanaan alih kuasa wayarles (WPT) bagi mengecas bateri. Melalui WPT, EV lebih mudah kerana ia boleh dicas secara dinamik. Namun, pengecasan dinamik WPT turut mengalami masalah, seperti mengimbang kecekapan pemindahan kuasa. Beberapa faktor yang membawa kepada masalah ini adalah kerana terdapat gangguan penjajaran dan jarak optimum antara gegelung pemancar dan penerima. Kerana ini, ia menyumbang kepada kehilangan kecekapan kuasa semasa mengecas EV. Pengeluar juga membina spesifikasi stesen pengisian EV berlainan mengikut jenis model EV demi memenuhi permintaan pelanggan. Namun, platform pengecas EV yang berbeza, tidak dapat mengecas EV secara wayarles dengan maksimum. Oleh itu, bagi membaiki keupayaan jana kuasa serta menstabilkan pengeluaran kuasa maksimum semasa proses pengecasan, litar gantian ditambah ke dalam sistem. Artikel ini memberi keutamaan pada dua litar gantian berbeza (topologi bersiri (SS) dan siri-selari (SP)) di bawah aplikasi pengecasan wayarles dinamik. Simulasi dibuat menggunakan NI Multisim mengikut kecekapan pemindahan kuasa dengan frekuensi resonan, pekali gandingan dan rintangan beban. Kecekapan WPT bagi topologi-SP

menunjukkan ianya sensitif pada perubahan frekuensi resonan dan pekali gandingan. Manakala topologi-SS kekal cekap semasa proses WPT. Walau bagaimanapun, prestasi topologi-SS berkurangan ketika diganding dengan beban besar, begitu juga berbeza bagi topologi-SP. Artikel ini akan mengkaji keadaan terbaik pada reka bentuk gantian terpilih bagi aplikasi EV dalam sistem pengecasan bergerak.

KEYWORDS: *electric vehicle (EV); wireless power transfer (WPT); dynamic charging compensation topology; capacitor arrangement*

1. INTRODUCTION

The first electric vehicle (EV) introduction was more than 100 years ago. The invention started from a battery and gradually evolved to an electric motor. In the early 19th century, technology pioneers demonstrated a battery-powered conceptualization vehicle. A British inventor, Robert Anderson built the first-ever a simple electric carriage. In the 1890s, a chemist named William Morrison successfully developed an electric car with a top speed of 14 miles per hour [1]. In the automotive industry, conventional vehicles can be manufactured easily and more affordably for consumers. Still, they present some inconveniences such as being a bit noisy and emitting unpleasant pollutants from their exhaust systems.

In comparison to a conventional vehicle, EV shows significant advantage in terms of noise, and environmental friendliness. Not to mention that EV is a current solution for the concern of environmental pollution, and depletion of energy resources as well as the increased cost for petroleum [1,2]. EV uses electricity as its propulsion, so the main component to store the electricity is the lithium-ion battery. The lithium-ion battery has a shallow discharge rate; its implementation increases the energy stored as well as increasing the travel distances of an EV [3].

The application of wireless charging in EV brings more convenience to its consumer. The ability to charge without any contact reduces the hassle faced by consumers. It increases consumer's safety, is spark-free, has fewer maintenance requirements, provides more reliability, and the charging method is not affected by its surroundings [4]. WPT uses the electromagnetism theory, where energy can be transferred using inductive coupling or capacitive coupling. For the capacitive coupling, the electric fields are used to transfer power between two plates. For inductive coupling, the energy is transmitted between the coils of the receiver and the transmitter via the magnetic field [5]. Ampere's law and Faraday's law are two principles behind the operation of inductive power transfer (IPT).

Figure 1 displays an illustration of the WPT system for EV. The main components that contribute to wireless power transfer are the transmitting coil and the receiving coil [6]. The power is transferred through the air from the coil at the charging pad to the coil installed in the EV. There are also other significant components for EV, which include both the transmitting side and the receiving side. At the transmitter, there is a voltage source, an AC/DC converter, a compensation circuit, and a transmitter coil. While at the receiver side, there is a receiver coil, a compensation circuit, AC/DC converter, and battery storage. The grid will supply power in terms of alternating current (AC). The AC/DC converter will convert it into DC and again convert it to AC as the concept needed to transfer wireless energy needs it to be AC current passing through the coils thus, the WPT phenomenon occurs. Then, at the receiver side, the AC will be converted into DC so it can be stored into the battery bank [7,8].

Through this implementation, EV can be charged in two states, either static or dynamic. Static wireless charging (SWC) is a charging process where the vehicle is idle such as when it is parked in a mall or garage. Dynamic wireless charging (DWC) is a charging process that takes places while the vehicle is moving on the road. However, the combination of both SWC and DWC is known as Quasi-dynamic Wireless Charging (QWC) [9]. As the EV is in state of motion, the power transfer efficiency is varied due to the inconsistency of airgap distances as well as the misalignment between the transmitter coil and the receiver coil. This situation will cause a low magnetic flux density. It is known that flux density is directly proportional to induced current. Low magnetic flux results in a loss of power efficiency due to of low magnetic flux density [10].

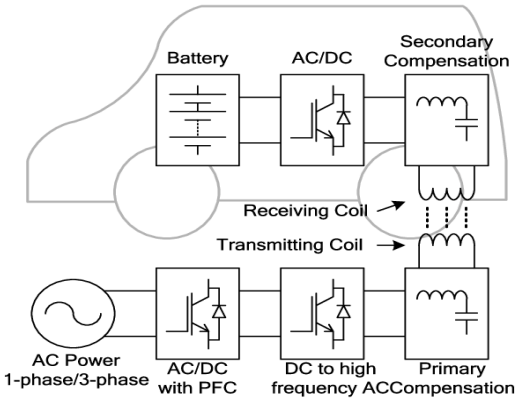


Fig. 1: WPT main components in EV [8].

A compensation circuit can compensate for the power transfer loss. It ensures that the output power for the WPT is maintained at a fixed level across wide variations in coupling [5]. There are four basic compensation designs: Series-Series (SS) topology, Series-Parallel (SP) topology, Parallel-Series (PS) topology, and Parallel-Parallel (PP) topology. Throughout this article, only two compensation circuits are chosen: Series-Series (SS) topology and Series-Parallel (SP) topology. SS-topology and SP-topology are more economically suitable for high power transmission [11]. These two topologies are further studied in finding which compensation designs performed well under DWC in terms of maximum power transfer capability and power transfer efficiency.

The author of [10] used the approach of applying mutual coupling for a different set of compensation designs (SS-topology and SP-topology) to analyse which topology can work well during the charging of an EV battery at constant voltage and constant current. SS-topology resulted in both situations of possible constant voltage or constant current, making it suitable for maintaining the charging process as the position of the primary system is changing with respect to the secondary system, such as an in-motion charging situation. Same value of parameters was used in [13] research, where it focused on transferring the resonant frequency equally to the operating frequency of the power source. Between SS-topology and SP-topology, a larger transmission range can be achieved with the series design in the secondary side. Parallel design, however, is capable of transferring higher output with a drastic load attached. In addition, the author in [15] carried out the test by setting up different compensation designs (SS-topology and SP-topology) using a mutual inductance model at resonance state. Despite that, SS-topology exhibited 70% of maximum efficiency while SP-topology achieved 66%.

2. ANALYSIS OF SS-TOPOLOGY AND SP-TOPOLOGY

With the absence of the compensation circuit in the WPT system, the power transfer efficiency is inefficient compared to a circuit with compensation technology [12]. The inductance leakage in the WPT system causes low power efficiency. Therefore, an additional capacitor is placed either in series or parallel at both the transmitter and receiver side to compensate for the leakage. The capacitive compensation not only solves the inductance leakage but also improves mutual coupling of the primary and secondary coils [13]. Figures 2 and 3 show the schematic circuit diagrams for both SS-topology and SP-topology. V_1, C_1 and L_1 denote the supply voltage, primary capacitance, and primary inductance at the transmitter side of the circuit. C_2, L_2 and R_L denote the secondary capacitance, secondary inductance, and the load resistance at the receiver side of the circuit. The M represents the mutual inductances in this circuit. For the ideal case, the R_1 and R_2 are neglected ($R_1 + R_2 = 0$) so that the circuit operates without the presence of disturbances.

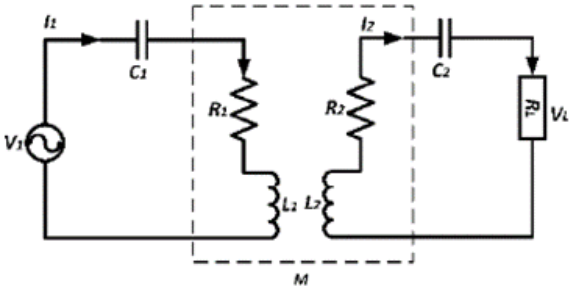


Fig. 2: Series-Series topology [8].

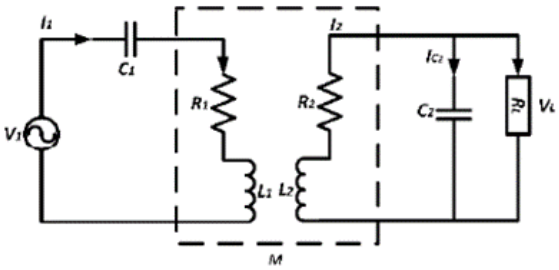


Fig. 3: Series-Parallel topology [8].

The model used for this article is an EV with a power consumption of 3.6 kW with 40 kHz resonant frequency (ω_o). The most common maximized output voltage is 168 V (V_L) for the testing of a rechargeable battery in an EV [23]. Therefore, the input voltage to recharge the EV's battery is 240 V. As the output voltage and output power are already known (P_o equal to P_{in}). The other parameters can be calculated, as shown below:

The load resistance can be found using equation (1). A higher power delivery into the load with maximum power transfer capability and efficiency is improved by adding a compensation circuit at the secondary side, which is where the load is located [13].

$$R_L = \frac{V_L^2}{P_o} \tag{1}$$

The root means square (RMS) value for primary and secondary current can be found using equation (2) and (3). Note that the output power (P_o) is equal to (P_{in}) since the coil resistance is ignored. Meanwhile, the RMS value for the primary voltage is set to be 240 V.

$$I_{1,rms} = \frac{P_{in}}{V_{I,rms}} \tag{2}$$

$$I_{2,rms} = \frac{V_{L,rms}}{R_L} \tag{3}$$

L_2, M and L_1 are calculated using the equation below. The quality factor, QS, and coupling coefficient, k, are set in the calculation as 4 and 0.2. The reason for the coupling coefficient to be 0.2 is that it can avoid the bifurcation issue. Bifurcation is a phenomenon where there exists more than one angle frequency in resonance inductive power transfer. It happens due to an increase in the coupling coefficient to a certain value. Therefore, the value of k should be addressed carefully to avoid a decrease in a voltage gain of the system [10]. It is noted

that the coefficient of coupling is correlated with the distance between the coils. The inductance value will be affected if the value of the coupling coefficient is changed [14].

$$L_2 = \frac{Q_S R_L}{\omega_o} \tag{4}$$

$$M = \frac{I_{2,rms} R_L}{I_{1,rms} \omega_o} \tag{5}$$

$$L_1 = \frac{M^2}{L_2 k^2} \tag{6}$$

ω_o is the resonant frequency, which is equal to $2\pi f$. A compensation circuit placed at the primary side can make the resonant frequency, ω_o be transferred equally to the operating frequency of the power source [13,15]. Using the value of the frequency, L_1 and L_2, primary capacitance, C_1, and secondary capacitance, C_2, can be known. Equation (7) is used to find C_1, whereas Eq. (8) is used for C_2.

$$f_1 = \frac{1}{2\pi\sqrt{L_1 C_1}} \tag{7}$$

$$f_2 = \frac{1}{2\pi\sqrt{L_2 C_2}} \tag{8}$$

Power calculation can be calculated at both the primary and secondary sides using Eq. (9) below. P is the power of the primary/secondary side (W), V is the voltage at the primary/secondary side (V), and I is the current at the primary/secondary side (A). Table 1 lists the required parameters for the simulations.

$$P = (V)(I) \tag{9}$$

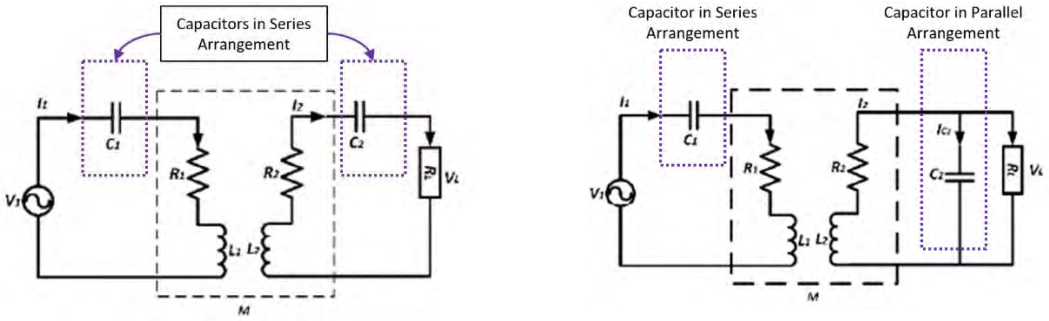
Table 1: Parameters used in wireless power transfer

Component	Parameter
V_P	Primary Voltage
C_P	Primary Capacitance
R_P	Primary Resistance
L_P	Primary Inductance
L_S	Secondary Inductance
C_S	Secondary Capacitance
R_S	Secondary Resistance
R_L	Load Resistance
Q	Quality Factor
k	Coefficient of Coupling

Both compensation designs are usually used in the high-power industry, including WPT. SS-topology and SP-topology provide a constant current and constant voltage during the charging process making them suitable to be applied to EV [10]. The arrangement of the capacitance inside the circuit determines the compensation design used. The series arrangement of the capacitor at the primary side (Fig. 4a) makes it independent of the primary inductance as well as the load, making it suitable for dynamic applications. In contrast, the parallel arrangement of a capacitor is to increase the primary current of the

circuit. At the secondary side, the series arrangement of capacitance (Fig. 4a) is able to stabilize the voltage supply to the load, while the capacitance arrangement in parallel (Fig. 4b) able to supply stable current to the load [10,15].

The compensation arrangement is not limited to only four arrangements, as already mentioned, but there is also a hybrid arrangement of capacitors and inductors [16]. The hybrid compensation circuits are LCC-LCC, LCL-LCL, LCC-P, LCL-P, S-CLC, LCL-S, CCL-S and multi-LCC (Fig. 5 in [17]). However, only SS- and SP-topologies will be further discussed in this article.



(a) (b)

Fig. 4: (a) Series-Series arrangement, (b) Series-Parallel arrangement [8].

Despite that, it also depends on the type of load that is used in the system in order to maximize the capability of the compensation design applied to it. Taking into consideration the EV to work with DWC, SS-topology can be seen as a better application towards charging the EV's battery. The reason is that the coupling coefficient of the compensation design is less affected while the EV is in motion. Also, SS-topology is capable of operating at both constant rates of voltage and current output which makes it more reliable to be implemented for the EV [10]. On the other hand, SP-topology compensation emits higher power output and efficiency with higher loads, making it appropriate to be applied in EV due to its problem in travel distance because of its small built battery capacity [13,15]. Both SS and SP compensation designs display their advantages in implementation under DWC, but further analysis is required to find out which of the two compensations can perform under dynamic situations.

3. RESULTS AND ANALYSIS

SS-topology and SP-topology are constructed as in Fig. 5 and Fig. 6, in NI Multisim software. The letter P or S is the notation for primary or secondary in the circuit diagrams. Both compensation circuits were simulated under three variables. The variables that were manipulated were the resonant frequency, coefficient of coupling, and the resistance load. These variables are chosen based on dynamic situations that the EV may experience in the long run.

Under the ideal case, which there is zero resistance tolerance, SS-topology appeared to be better than SP-topology in terms of maximum output power and power transfer efficiency. The efficiency achieved by SS-topology was 99.94% while SP-topology struggled to transfer power wirelessly with 28.39% efficiency.

https://doi.org/10.31436/iiumej.v22i2.1660

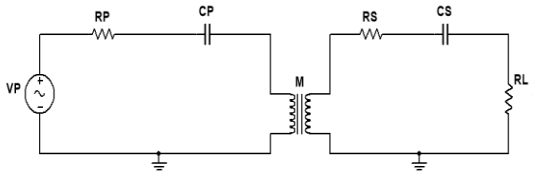


Fig. 5: SS-topology circuit diagram in NI Multisim.

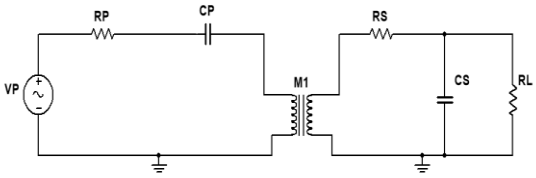


Fig. 6: SP-topology circuit diagram in NI Multisim.

Table 2: Calculated parameters for NI Multisim simulations

Parameter	Value
R_L	7.84 Ω
L_P	398 µH
L_S	124.78 µH
M	44.57 µH
C_P	39.78 nF
C_S	126.87 nF
K	0.2
QS	4

In order to analyse the performance between the two topologies, Fig. 7, Fig. 8, and Fig. 9 show the relationship for each variable with power transfer efficiency. The coupling coefficient changes accordingly, based on the distance as well as the alignment between the coils [20]. The value range tested is between 0.01 and 0.2, in order to avoid the occurrence of the bifurcation issue. It is known that the coupling coefficient decreases with the increase of distance and vice-versa, as is the case with misalignment between the coils [21]. Referring to Fig. 7, from 0.06 (41.3 cm) to 0.2 (25.9 cm), SS-topology is able to transmit better power efficiency from the transmitter coils to the receiver coils. On the other hand, at the distance of 36.8 cm between the coils with 0.08 coupling coefficient, SP-topology reaches its maximum potential with 96.47% of efficiency. When it comes to tolerance to coupling coefficient, SP-topology appears to be less suitable based on its performance compared to SS-topology. SS-topology is capable of maintaining more than 90% efficiency with bigger air-gap fluctuation approximately ≈16 cm.

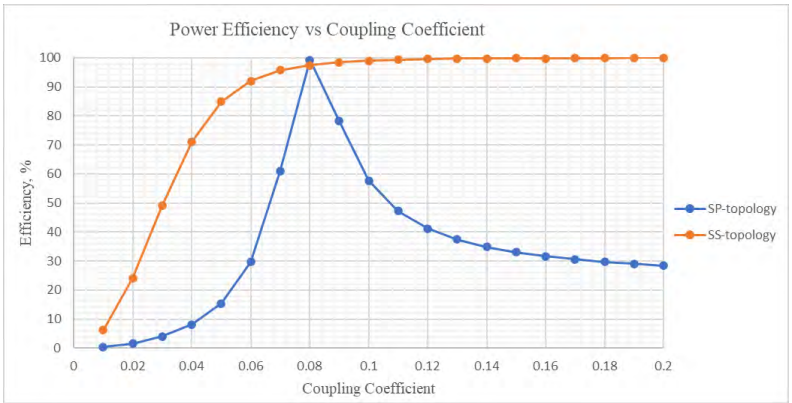


Fig. 7: SS and SP simulation for power efficiency, Ƞ vs coupling coefficient, k.

Another simulation test conducted is the manipulation to the resonant frequency. Based on Fig. 8, both topologies show the fluctuation in power efficiency. The range of the

resonant frequency for the simulation test is chosen between 30 kHz and 50 kHz because it is in the range of EV operational specifications for charging pad manufacturers. The frequency for inductive power transfer can range from 3 kHz to 1 MHz [20]. In terms of maintaining power efficiency, both topologies are unable to consistently maintain power transfer as the frequency is altered. In spite of the inconsistency, SS-topology expresses better efficiency within the frequency range of 38 kHz until 42 kHz while SP-topology has its highest peak with 56.4% of power transfer efficiency when the resonant frequency is 41 kHz. Since the system is calculated and designed with a frequency value of 40 kHz, the outcome for the maximum power efficiency is technically closer to the assigned value of the resonant frequency. SP-topology experienced more significant power loss compared to SS-topology in this study. Hence, it can be observed that for any slight increase or decrease from the theoretical resonant frequency, the power efficiency will drop lower than the maximum power efficiency achieved. This proves that both topologies are sensitive when there is a deviation in resonant frequency [10].

The change in load resistance also affected the efficiency during the power transfer. Load resistances represent the battery capacity in EV. Bigger load signifies more power can be stored, thus requiring less charging time per day. Figure 9 illustrates the results of power efficiency for both topologies under alteration of the load resistance value. In this simulation test, when both topologies were paired to a bigger load (100 Ω), SP-topology operated better than SS-topology with regards to efficiency in WPT. Contrarily, SS-topology attained 99.93% efficiency with a load value at 10 Ω.

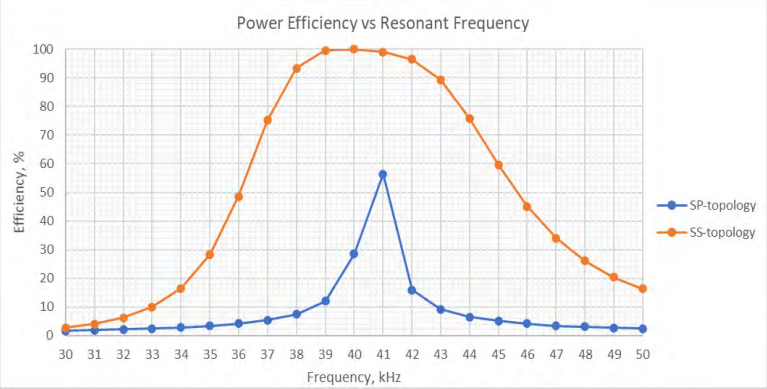


Fig. 8: SS and SP simulation for power efficiency, Ƞ vs resonant frequency, ωo.

Based on the traits analysed for DWC, the relationship between efficiency with resonant frequency, coupling coefficient and load resistance, SS-topology and SP-topology each have their respective strengths in wireless charging. For optimum real application of EV, the coupling coefficient, k was kept at 0.2 to avoid any presence of more than one angle frequency during the resonance inductive power transfer. Because both topologies' efficiency easily affected by the resonant frequency, ω_o=41 kHz was set as ideal operational frequency at the transmitter side. The same parameter values in Table 2 remain constant except for load resistance for each compensation design. Series design at the receiver side resulted in lower voltage but higher current. Parallel compensation at the secondary side made the voltage received to be high while the current received was low. Under the condition of 3.6 kW model, SS-topology had a longer transmission distance with a load of R_L=8.1 Ω. As for SP-topology, the same output power was gained with R_L=130 Ω. Table 3 represents the data simulated under those specifications with the expression of achieving 3.6 kW of power transmission.

https://doi.org/10.31436/iiumej.v22i2.1660

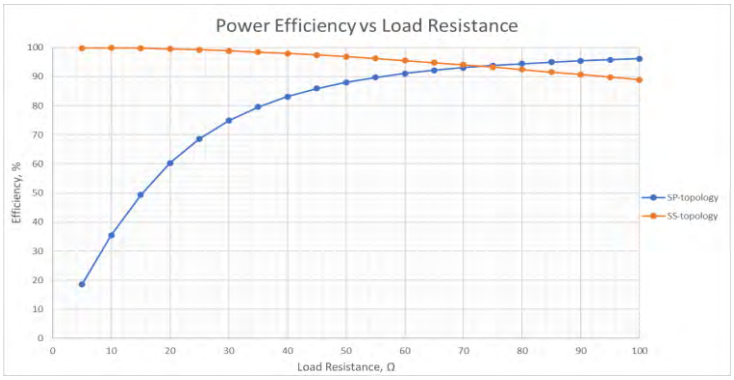


Fig. 9: SS and SP simulation for power efficiency, Ƞ vs load resistance, Ω.

Table 3: Compensation design 3.6 kW simulation paired with respective loads.

	Series-Series (SS) compensation design	**Series-Parallel (SP) compensation design**
Resistance load, R_L	8.1	130
Primary voltage (rms)	240 V	240 V
Secondary voltage (rms)	171 V	683 V
Primary current (rms)	15.20 A	15.16 A
Secondary current (rms)	21.10 A	5.26 A
Power transmitted	3648 W	3640 W
Power received	3608 W	3592 W
Power efficiency, Ƞ	98.9%	98.6%

The main contribution is that both compensation designs have different features for WPT. SS-topology is fitted for small load systems, providing the compensation design of SS tolerance towards the change in frequency and airgap distance between coils. While SP-topology is very relevant for large load systems.

4. CONCLUSION

The two topologies have been simulated based on DWC features when EV is technically in a state of motion. The relationship for power efficiency with resonant frequency, coupling coefficient and load resistances is tested to identify the best fitted compensation design for steady power transfer from the charging pad to the battery. From the simulation results acquired, both designs react differently for each variable. Even using the same parameters, the maximum potential of DWC reflected on the load paired to the system. Great efficiency was achieved as SS-topology used small loads, in contrast to SP-topology which gained great efficiency with bigger loads. Nevertheless, through comparing and analysing the simulation data, SS-topology appeared ideal to implement for DWC. Despite the fact that it can only perform at utmost capability with small loads, it is important to highlight SS-topology's potential to maintain it tolerance towards deviation in resonant frequency and coupling coefficient. For practical applications, the ability to transfer power at a consistent rate without major power loss is an important choice criterion when the vehicle is moving [22]. SP-topology paired greatly with bigger loads. However, it is suitable for longer travel distances but requires more charging time because of the bigger loads. SP-topology seems more suitable for SWC due to its sensitivity towards changes in coupling

coefficient. Not to mention, this article analysed the conditions for DWC which required an EV cruising while charging. For that reason, it touched on different application situations of two chosen compensation designs and elaborated on how to choose the ideal design suitable for practical applications.

ACKNOWLEDGEMENT

We would like to thank partial financial support by the International Islamic University Malaysia under the Kulliyyah of Engineering Postgraduate Tuition Fee Waiver Scheme 2019 (TFW2019) and also this work was partially supported by the Ministry of Higher Education Malaysia under the Fundamental Research Grant Scheme (Grant number FRGS19-067-0675).

REFERENCES

[1] Vepachedu S. (2017) The History of The Electric Car. Andhra Journal of Industrial News. 14-27.

[2] Salah WA, Alsayid B, Albreem MAM, Abu Zneid B, Alkhasawneh M, Al Mofleh A, Abu Sneineh A, Al Aish AA. (2019) Electric vehicle technology impacts on energy. International Journal of Power Electronics and Drive Systems (IJPEDS).10.1.10.11591/ ijpeds.v10.i1.pp1-9.

[3] Ma S, Jiang M, Tao P, Song C, Wu J, Wang J, Deng T, Shang W. (2018) Temperature effect and thermal impact in lithium-ion batteries: A review. Progress in Natural Science: Materials International, 28(6): 653-666. doi: 10.1016/j.pnsc.2018.11.002

[4] Lukic S, Pantic Z. (2013) Cutting the Cord: Static and dynamic inductive wireless charging of electric vehicles. IEEE Electrification Magazine, 1(1): 57-64. doi: 10.1109/MELE.2013.2273228

[5] Afridi K. (2018) Wireless charging of electric vehicles. National Academy of Engineering. in Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2017 Symposium. Washington, DC: The National Academies Press. doi: 10.17226/24906.

[6] Li S, Mi CC. (2015) Wireless power transfer for electric vehicle applications. IEEE Journal of Emerging and Selected Topics in Power Electronics, 3(1): 4-17. doi:10.1109/JESTPE.2014.2319453

[7] Zaini SA, Yusoff SA, Abdullah AA, Khan S, Abd Rahman F, Nanda. NZ. (2020) Investigation of magnetic properties for different coil sizes of dynamic wireless charging pads for electric vehicles (EV). IIUM Engineering Journal, 21(1): 23-32. doi:10.31436/iiumej.v21i1.1108

[8] Mi C, Abdul Masrur M. (2017) Hybrid electric vehicles: Principles and applications with practical perspectives. John Wiley & Sons.

[9] Ahmad A, Alam MS, Chabaan R. (2018) A comprehensive review of wireless charging technologies for electric vehicles. IEEE Transactions on Transportation Electrification, 4(1): 38-63. doi: 10.1109/TTE.2017.2771619

[10] Aditya K, Williamson SS. (2014) Comparative study of series-series and series-parallel compensation topologies for electric vehicle charging. IEEE 23rd International Symposium on Industrial Electronics (ISIE). doi: 10.1109/ISIE.2014.6864651

[11] Shevchenko V, Husev O, Strzelecki R, Pakhaliuk B, Poliakov N, Strzelecka N. (2019) Compensation topologies in IPT systems: Standards, Requirements, Classification, Analysis, Comparison and Application. IEEE Access, 7, 120559-120580. doi: 10.1109/ACCESS.2019.2937891

[12] Azambuja R, Brusamarello VJ, Haffner S, Porto RW. (2017) Full four capacitor circuit compensation for inductive power transfer. Conference Record - IEEE Instrumentation and Measurement Technology Conference, 2-7. doi: 10.1109/I2MTC.2013.6555406.

[13] Chenggang F, Jiancheng S, Lingyan L, Yameng W. (2017) Practical considerations of series-series and series-parallel compensation topologies in wireless power transfer systemapplication. 2017 IEEE PELS Workshop on Emerging Technologies: Wireless Power Transfer (WoW), ID 41321671. doi: 10.1109/WoW.2017.7959404

[14] Rehman, M., Baharudin, Z., Nallagownden, P., & Islam, B. (2017). Modeling and Analysis of Series-Series and Series-Parallel Combined Topology for Wireless Power Transfer using Multiple Coupling Coefficients. IJCSNS International Journal of Computer Science and Network Security, VOL.17 No.11, November 2017.

[15] Rituraj G, Joy ER, Kushwaha BK, Kumar P. (2014) Analysis and comparison of series-series and series-parallel topology of contactless power transfer systems. TENCON 2014-2014 IEEE Region 10 Conference. doi: 10.1109/TENCON.2014.7022440.

[16] Nanda, N. N., Yusoff, S. H., Toha, S. F., Hasbullah, N. F., & Roszaidie, A. S. A Brief Review: Basic Coil Designs for Inductive Power Transfer. Indonesian Journal of Electrical Engineering and Computer Science (IJEECS), Vol. 20, No. 3, December 2020, pp. 1703-1716. doi: 0.11591/ijeecs.v20.i3.pp1703-1716.

[17] Abou Houran M, Yang X, Chen W. (2018) Magnetically coupled resonance WPT: Review of compensation topologies, resonator structures with misalignment, and EMI diagnostics. Electronics, 11: 296. doi:10.3390/electronics7110296

[18] Barman SD, Reza AW, Kumar N. (2014). Efficiency maximization of resonant coupled wireless power transfer system via impedance matching based on coupling tuning. 3rd International Conference on Computer Engineering & Mathematical Sciences, Langkawi, Malaysia, November 2014.

[19] El-Shahat A, Ayisire E, Wu Y, Rahman M, Nelms D. (2019). Electric vehicles wireless power transfer state-of-the-art. Energy Procedia, 162: 24-37. doi: 10.1016/ j.egypro.2019.04.004

[20] Panchal C, Stegen S, Lu J. (2018) Review of static and dynamic wireless electric vehicle charging system. Engineering Science and Technology, an International Journal, 21(5): 922-[a] doi: 10.1016/j.jestch.2018.06.015

[21] Al-Saadi M, Al-Gizi AG, Al-Chlaihawi S, Al-Omari AH. (2018). Inductive power transfer for charging the electric vehicle batteries. Electrotehnică, Electronică, Automatică, 66: 29-39.

[22] Carmeli MS, Castclli-Dezza F, Mauri M, Rossi M, Dolata A, Pedretti M, Simonini I. (2018) Analysis of a quasi-dynamic wireless power transfer system for EV batteries charging. 2018 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 383-388. doi: 10.1109/SPEEDAM.2018.8445238

[23] Aditya K. (2016) Design and implementation of an inductive power transfer system for wireless charging of future electric transportation. PhD thesis. Electrical, Computer and Software Engineering Department, University of Ontario Institute of Technology Oshawa, Ontario, Canada.

https://doi.org/10.31436/iiumej.v22i2.1662

ONLINE NEWS CLASSIFICATION USING MACHINE LEARNING TECHNIQUES

JEELANI AHMED* AND MUQEEM AHMED

Department of Computer Science and Information Technology,
Maulana Azad National Urdu University, Hyderabad, India

**Corresponding author: Jeelani.jk@gmail.com*

(Received: 18th October 2020; Accepted: 31st March 2021; Published on-line: 4th July 2021)

ABSTRACT: A massive rise in web-based online content today pushes businesses to implement new approaches and resources that might support better navigation, processing, and handling of high-dimensional data. Over the Internet, 90% of the data is unstructured, and there are several approaches through which this data can translate into useful, structured data—classification is one such approach. Classification of knowledge into a good collection of groups is significant and necessary. As the number of machine-readable documents proliferates, automatic text classification is badly needed to classify these documents. Unlabeled documents are categorized into predefined classes of labeled documents using text labeling, a supervised learning technique. This paper reviewed some existing approaches for classifying online news articles and discusses a framework for the automatic classification of online news articles. For achieving high accuracy, different classifiers were tried. Our experimental method achieved 93% accuracy using a Bayesian classifier and present in terms of confusion metrics.

ABSTRAK: Peningkatan tinggi pada masa kini pada maklumat dalam talian berasaskan web menyebabkan kaedah baru dalam bisnes telah diguna pakai dan sumber sokongan seperti navigasi, proses, dan pengurusan data berdimensi-tinggi adalah perlu. 90% data di internet adalah data tidak berstruktur, dan terdapat pelbagai kaedah data ini dapat diterjemahkan kepada data berguna, lebih berstruktur — iaitu melalui kaedah klasifikasi. Klasifikasi ilmu kepada koleksi kumpulan baik adalah penting dan perlu. Seperti mana mesin-boleh baca dokumen berkembang pesat, teks klasifikasi automatik juga sangat diperlukan bagi mengklasifikasi dokumen-dokumen ini. Dokumen yang tidak dilabel dikategori sebagai pengelasan pratakrif dokumen berlabel melalui teks label, iaitu teknik pembelajaran berpenyelia. Kajian ini mengkaji semula pendekatan sedia ada bagi artikel berita dalam talian dan membincangkan rangka kerja bagi pengelasan automatik artikel berita dalam talian. Bagi menghasilkan ketepatan yang tinggi, kami menggunakan pelbagai alat klasifikasi. Kaedah eksperimen ini mempunyai ketepatan 93% menggunakan pengelas Bayesian dan data dibentangkan berdasarkan matriks kekeliruan.

KEYWORDS: *text classification; naïve Bayes; support vector machine; news articles*

1. INTRODUCTION

Today, the overwhelming volume of digital content is expanding and growing constantly. Automatic text classification into existing categories is considered a primary method to process and manage this enormous amount of data.

This kind of textual data is generated from various sources and is available in conference materials, editorials, digital/electronic documents, web pages, emails, publications, and journals. Many people use these online sources for day-to-day

https://doi.org/10.31436/iiumej.v22i2.1662

information access instead of being limited to printed material such as newspapers, magazines, and books. The entrance to such information has become more comfortable; however, the organization of knowledge is difficult, making it challenging to manage. Organizing this kind of digital data is considered a critical method to effectively classify this digital information.

Text classification has a vital role in text retrieval, information abstraction and summarization, and question-answering. Usually, the data available on the web used for classification are diverse, from various sources like broadcast or printed news, bulletin boards, newsgroups, advertisements, and movie reviews. Because of their varied nature, like being multi-sourced and having different vocabularies, different formats, and different writing styles for documents, automatic text classification is essential [1]. Due to the enormous amount of text that is stored in the electronic format, it is necessary to understand and examine such data and extract similar details, which may be useful when making decisions [2-4].

When a document is classified under an already defined category this action is termed Text classification (Fig. 1). A document can be single labeled or multi-labeled, depending upon the classes. A document is termed as a single label when it is allocated to a single class and is named multi-label if the document assigns it to multiple classes [5].

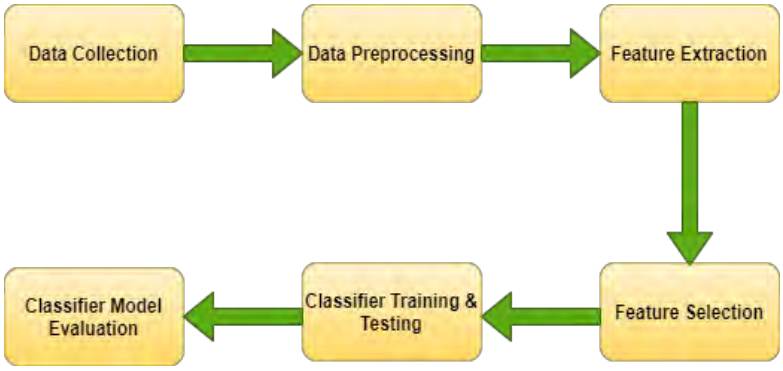


Fig. 1: Text classification process.

Text classification comprises feature selection, document representation, applying the algorithm, and performance assessment. Because of the sudden growth of online transactions and the colossal accessibility of text documents, it is essential to retrieve vital data from documents using classification. Today, organizing and managing text information is considered a crucial practice [6]. Neural Networks [8], Support Vector Machines [7], k-Nearest Neighbor [9], and Naïve Bayesian [10] classification are algorithms that may be used for constructing classification methods.

One area where we encounter a considerable amount of daily text is online news articles. Due to Information Technology developments, persons are worried about their busy lifestyle and thus they desire to only read news articles of their interests [11]. It is a big task to mine pertinent news concerning an individual's interest because much of the news articles are enlightening but could be less significant. An individual's lure can be contingent on many factors such as the news location and the type of news articles [12]. Here, we classify the news articles rendering to the type/category of the news articles. For example, imagine an individual who prefers to read news related to politics. Since the web is a sea full of news articles, it would be challenging for a particular person to read only articles related to politics.

In this work, to customize the news articles for a specific category, we have applied machine-learning techniques. The category may be politics, crime, sports, entertainment, or world news. We trained and tested the classification model using the huff post data set, which contains around 75k news articles. After training and testing, we applied this model for the news articles taken from different live news websites such as the Hindu, the Print, Indian express, and the Quint.

2. BACKGROUND

2.1 Naive Bayes (NB)

This algorithm is used as the classifier for the news articles. It is a prevalent method of classification when dealing with multi-class classification. It works based on the probabilistic technique first proposed by Lewis [13] and drives its roots from the Bayes theorem. This algorithm can run very efficiently with a large dataset. For text classification problems, Naïve Bayes is used as a standard because it has a fast run time as compared to other classifiers. Naïve Bayes is also used to solve problems like spam detection. Due to its simplicity, it outpaces many advanced classification techniques.

2.2 K-Nearest Neighbors (KNN)

The problems of regression and classification domain can be solved using this supervised machine learning algorithm [9]. KNN works by calculating the space between a query and all the instances (K) nearer to the question and then selects the majority recurrent label. It is comprehensible and applied very quickly. It slows down its performance when there is an upsurge in the data sample. It stores all the training data, due to which it is a somewhat costly algorithm. It requires high memory storage as compare to other methods. KNN is used in many areas, such as in politics; it can classify a voter into various classes like 'Will not Vote,' 'Will Vote,' or 'Will Vote to AAP Party'.

2.3 Support Vector Machine (SVM)

SVM is a supervised machine-learning algorithm [7]. SVM works when provided training data along with associated labels. Once the training is over, if supplied a data set, the model assigns a label to it. The support vector machine works well when dealing with linear classification problems. For classification, it makes a hyper-plane by choosing the maximum distance between adjacent data points. It is both a flexible and powerful technique of machine learning that is used for regression and classification. When dealing with high dimensional space, SVM works well and offers excellent accuracy, and it requires very little memory for processing.

2.4 Logistic Regression (LR)

It is the best regression analysis to be conducted where the dependent variable is binary. As for all studies of regression, logistic regression is a statistical process. It describes data and describes the relationship between one conditional dependent variable and one or more interval, ordinal, nominal, or ratio-level independent variables.

3. RELATED WORK

With the rapid increase in demand for managing substantial text databases, text classification is a popular and dynamic research field of data mining. It becomes necessary and essential to competently handle the textual data to search and access any document quickly due to the enormous growth of digital textual information [14,15].

https://doi.org/10.31436/iiumej.v22i2.1662

Text classification uses unique rules to grant classes from a category of already specified classes to unlabeled text documents. Text classification typically operates manually, but because the classification rules are generated manually, such a procedure is time-consuming and costly. Therefore, Machine Learning is another methodology that uses automatic rule creation to classify documents [16].

The numerous classifiers are: Support Vector Machines [7], Neural Networks [8], Bayesian classification [10], Nearest Neighbor (KNN) [9], Association based classification [17-19], Term Graph Model [20,21] and Decision Tree (DT) [22,23], etc. Many authors use the mentioned machine learning techniques; the table below was formulated after a brief study of these techniques.

Table 1: Related work in the field of Text classification by various authors

References & year	Datasets	Algorithm / Technique used	Conclusion/Note
[24] (2014)	Reuter-21578	KNN algorithm	The accuracy of the KNN is maximum compared to Term Graph and Naïve Bayes. Its time complexity is maximum as compared to others is the only drawback.
[25] (2014)	Sohu laboratory corpus	SVM-KNN algorithm	By making recommendations and improving the classifying probability, the SVM-KNN algorithm will improve the classifier's efficiency. This algorithm increases the low amount of processing difficulty. Moreover, SVM is efficient and fast for classification.
[26] (2015)	Single-label news corpus	SVM, TF-IDF	For the classification of Indonesian news posts, the authors used algorithm adoption and problem transformation methods. The authors develop a combination technique for the automatic classifying of the news articles and achieve better results with an f-measure is 85.13%.
[27] (2015)	Reuters-21578	Hyper Rectangular Keyword Extraction	The authors proposed vital extraction based on the document categorization hyper rectangular method. This method generates excellent and accurate results when compared to traditional classification approaches.
[28] (2015)	Czech News Articles	Linear SVC, SGD, PA, NB	The authors proposed a method for enhancing the classification accuracy of multi-document classification. The baseline classifier Naïve Bayes performed better than other classifiers.
[29] (2016)	Various news websites	Naïve Bayes (NB)	Various news webpages were classified using structure attributes and URL content. The naïve Bayes algorithm showrd better results compared to existing approaches when implemented on the same dataset.
[30] (2016)	News articles of emotional topics	NB, C45, DT, SVM vectors, Winnow, Balanced Winnow, and Max Entropy	The authors experimented with hierarchical classification with the purpose of sentiment analysis of news articles. SVM and TF-IDF were used with six and four classification algorithms respectively and achieved better results.
[31] (2016)	Stock new of China	Softmax training algorithm, Weighted sort algorithm, and selection algorithm	For making investment decisions, authors classified stock news by proposing a novel algorithm. The authors achieved better accuracy and availability with this method compared with a general algorithm.

References & year	Datasets	Algorithm / Technique used	Conclusion/Note
[32] (2016)	Online news articles	Neural Network Classifier	The performance of the ANN using feature reduction performed better as compared to the essential ANN and gave better results for the classifications. The training process was prolonged.
[7] (2017)	Various news websites articles	Naïve Bayes, SVM, Random Forest	The performance of the SVM was at the bottom when compared with Naïve Bayes and Random Forest. Also, SVM took much training time.
[33] (2017)	Korean News Articles	KNN, NB, SVM, and Logistic Regression (LR)	The author performed two studies using four classifiers. Study 1, which had the data with less complexity, achieved a higher level of accuracy. Study 2, which had the data with more high complexity, may have harmed the classification results.
[34][2017]	Reuters.com news articles and Gold Standard dataset	SVM, SDG, LR	The authors used various classification algorithms to classify text articles to detect opinion, spam, and fake news. Classifiers achieved good results with an increase in the features.
[35] (2017)	Yahoo News US Edition	Keyword matching, Newsmap	The work developed a semi-supervised classifier for the classification of geographical news. The overall classification accuracy was 0.80, which was relatively low when compared to other traditional classifiers.
[36] (2018)	NLPCC2014, REV1-v2	Attention GRU Bi-RNN	Centered on the process of neural network interest, the authors proposed a bi-directional recurrent neural network algorithm that performed classification on two datasets and achieved a high accuracy of 83.9 compared to other classifiers.
[37] (2018)	BR news dataset US news dataset	ANOVA, SVM	The work performed classification on two news datasets and proved noticeable differences between reliable and unreliable news sources.
[38] (2018)	Indonesian news articles	Latent Dirichlet Allocation (LDA)	Classification results of Indonesian news articles represented using a word cloud, which allowed easy understanding of each category's results and trends; However, more explanation was needed in terms of classification.
[39] (2018)	China News	SVM, CNN, MAXENT	The authors applied three algorithms in two schemes for the classification of Chinese news articles. SVM performed better than the other two, but a detailed explanation of the experiment is needed.
[40] (2019)	News articles and tweets	Word2vec Convolutional Neural Networks	The algorithm used is classified tweets and news articles into related and unrelated words. Word2vec enhanced performance by learning the semantic relations between words.
[41] (2019)	Arabic news articles from websites	SVM, NB, DT, RF, LR	The work classified Arabic news articles collected from various news websites. SVM performed better than all other classifiers and achieved an accuracy of 87%.
[42] (2019)	Indonesian news articles	Convolutional neural networks, recurrent neural networks	The work performed a systematic classification of risk document to get the financial risk information in real-time. Authors achieved better classification results using extensive data. However, manual labeling of new data was a tedious process.

References & year	Datasets	Algorithm / Technique used	Conclusion/Note
[43] (2019)	Thai PBS, Khaosod, and Dailynews	SVM, Decision Tree, Deep Learning	The authors discussed the performance of various classifiers using Thai news as a dataset. Deep Learning outperformed the SVM and Decision Tree.
[44] (2019)	BBC datasets, five groups of 20Newsgroup	SVM, Naïve Bayes, Bi-LSTM, LSTM, CNN	This work used five classifiers for the classification of two Chinese news datasets. They also pointed out the difference between ML and DL classification that ML was a time-consuming classification process that required preprocessing and feature extraction. In contrast, DL did not require cleaning activities. However, using ML, they obtained efficient and reliable results and the best accuracy.
[45] (2020)	Society channel of Sina	C4.5 Decision Tree Algorithm	The authors proposed a classification approach for emotions in news articles. They classified emotions into fear, joy, and sadness. This method achieved a high accuracy of 87.83% and compared it with the SVM classifier.
[46] (2020)	Uzbek Daryo online news	SVM, DT, RF, LR, Multinomial Naive Bayes	The authors used six algorithms for news article classification in Uzbek languages. The approach achieved the highest accuracy of 86.88%.
[47] (2020)	Tigrigna news dataset	SVM, DT, LR, RF, KNN, etc.	The authors constructed the Tigrigna news dataset and have experimented with eight various classifiers. SVM outperformed other classification algorithms. Moreover, it achieved the highest accuracy.

We reviewed several papers; Table 1 contains some recent papers. Several scholars studied and categorized texts in various languages, such as Korean [33] and Chinese [39]. There were still several works available covering Arabic language classification [41], shedding light on academic articles based on the usage of classical classifications of supervised machine learning, like SVM [34,37], KNN [24,25], Decision Tree [45], and Naïve Bayes [7,29]. Although other authors focused on classification through neural networks [32] and deep learning [42], the overall output was higher. Finally, it is evident that the classification algorithm's performance in every language was significantly impaired by the consistency of the data source, as the unnecessary and repetitive characteristics of the data undermined the precision and output of the classifier.

4. RESEARCH DESIGN AND IMPLEMENTATION

Based on the above literature, we proposed a framework to perform text classification of English news articles obtained from various Indian news websites. For ease of understanding, we divide the proposed framework into three main modules: Data Extraction, Data Preprocessing, and Classifier Module as illustrated in Fig. 2.

First, the dataset is cleaned using data preprocessing. Then, we divide the dataset into train and test parts. Train data hold 70%, and test data hold 30% of the dataset. The next module trains the classifier to predict the class labels for the collected news articles and assign a category label in the next step. Thereby performance evaluation of the classifier is carried out using some performance metric. The next section contains a detailed explanation of each module.

https://doi.org/10.31436/iiumej.v22i2.1662

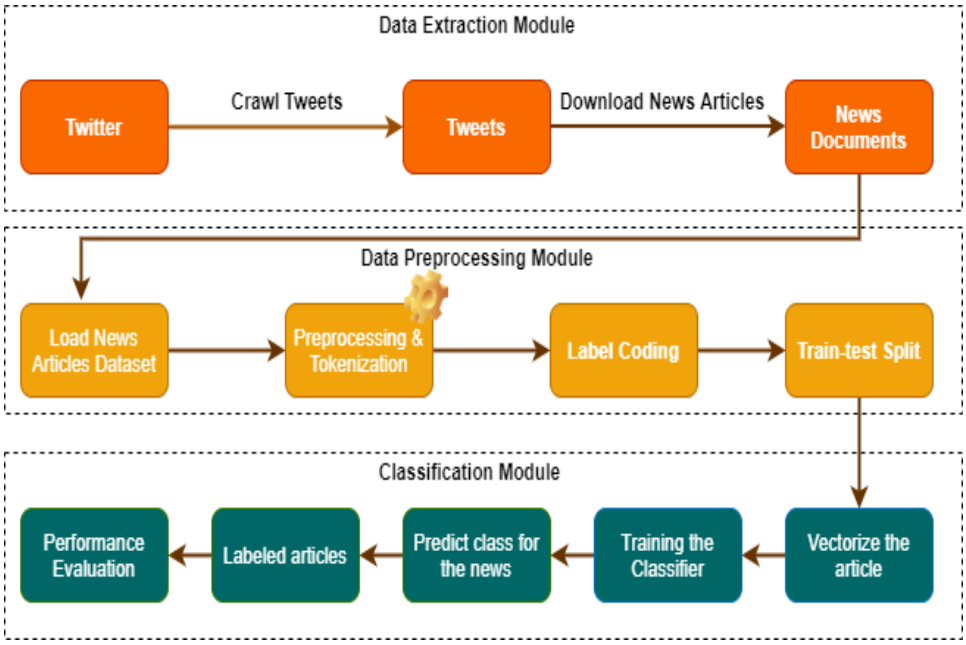


Fig. 2: News article classification process.

4.1 Data Extraction

The process starts with a data collection module, in which news articles from seven different news websites are crawled, as shown in Fig. 2. Firstly, tweets of various news Twitter handles are crawled from Twitter using tweepy. Various news articles were extracted from crawled tweet URLs. Further, news article web pages from each news website are extracted by visiting each URL. In the module, the data extraction module downloads tweets and news articles from seven news websites.

We have also used a dataset that contains around 75k news headlines with its content and category from the year 2012 to 2018 taken from Huff Post. These news articles contain eight categories: entertainment, crime, politics, business, world news, sports, media, and technology. The percentage of each class of news articles is shown in Fig. 3.

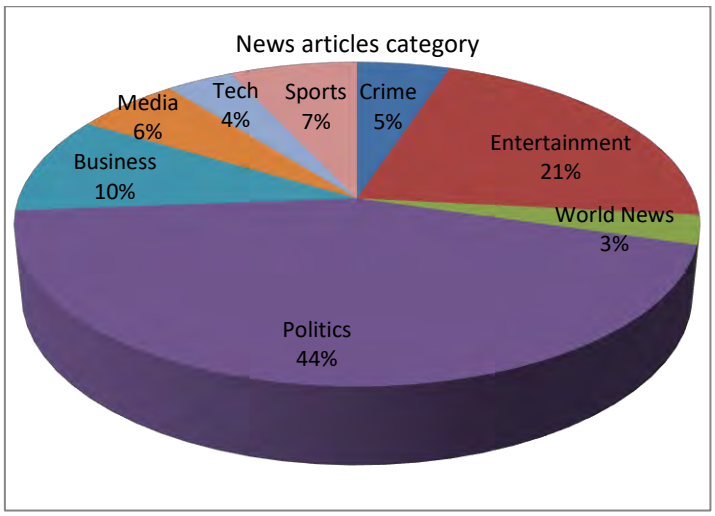


Fig. 3: Percentage of each news article category.

https://doi.org/10.31436/iiumej.v22i2.1662

The dataset that was used is shown in Fig. 4. It contains the first ten lines of the dataset, which have eight categories of news articles in CSV format, as shown below.

category	headline	content
CRIME	there were 2 mass shootings in texas last week but only 1 on tv	she left her husband he killed their children just another day in america
ENTERTAINMENT	will smith joins diplo and nicky jam for the 2018 world cups official song	of course it has a song
WORLD NEWS	south korean president meets north koreas kim jong un to talk trump summit	the two met to pave the way for a summit between north korean and the us
POLITICS	trumps son should be concerned fbi obtained wiretaps of putin ally who met with trump jr	the wiretaps feature conversations between alexander torshin and alexander romanov a convi
CRIME	man faces charges after pulling knife stun gun on muslim students at mcdonalds	â€œwe thought we were going to die one of the students said
BUSINESS	us launches auto import probe china vows to defend its interests	the investigation could lead to new us tariffs similar to those imposed on imported steel and a
SPORTS	trump posthumously pardons boxer jack johnson	the pardoning of the black heavyweight boxer is only the third posthumous pardon in us histor
MEDIA	jake tapper shreds donald trump with a long list of his conspiracy theories	i could go on but this is just an hour show
TECH	selfdriving uber in fatal accident had 6 seconds to react before crash	the ntsb published a preliminary report on the incident thursday
ENTERTAINMENT	brynn cartelli becomes youngestever winner of the voice	the tv singing competitions new champion just turned 15

Fig. 4: Description of the dataset.

4.2 Data Preprocessing

The data preprocessing module performs the cleaning of the dataset, which is considered to be an important task to achieve good results. Firstly, we perform the tokenization of articles; the module changes a group of characters to groups of strings with some recognizable meaning. Next, by using the python nltk package, we remove stop words such as 'what' and 'the' as these are the words that have minimal prominence and occur commonly.

4.2.1 Label Encoding

We used the LabelEncoder class from the Scikit-Learn library in python to transform categorical text data into model comprehensible numerical data. From the Scikit-learn library, we use LabelEncoder class to encode the first column, then fit and transform the data of that column. Then, the new encoded data replaced the existing text data. After running the LabelEncoder piece of code, we get the following table.

Table 2: Label encoding for different news category

Category Name	Category Code
Crime	0
Entertainment	1
World News	2
Politics	3
Sports	4
Business	5
Media	6
Tech	7

4.2.2 Train-test Split

Our work includes a single labeled classification of the news articles collected from seven different news websites. For training and testing purposes, we divided our dataset into 70% for training and 30% for testing. The training data set consisted of 52635 labeled news articles and the testing set consisted 22558 of news articles. In training data, we have used the 10 fold cross-validation, fitted the final model to it, and then evaluated with the unclassified data and obtained evaluation metrics that showed as little bias as possible.

4.3 Training the Classifier

The training set and the set of labels corresponding to it are the input data for the classifier. The labels we used along with training set are: 'Crime', 'Entertainment', 'World News', 'Politics', 'Sports', 'Business', 'Media', and 'Tech'. Based on the pre-decided category tag, processed news articles were numerically labeled. Because the classifier input consists of two vectors, it was essential to vectorize news articles and the labels. They functioned as an input to the classifier after the vectorization of these two entities.

4.4 Testing the Classifier

The classifier model was tested on the testing data once the classifier was trained with the trained data. The classifier predicted the category of the corresponding news articles. An example of the news article's 'business' category from the testing set is presented in Fig. 5. Naive Bayes classifier performed better, with 93% accuracy. Furthermore, our model is sure that the above 'business' article would belong to the same category.

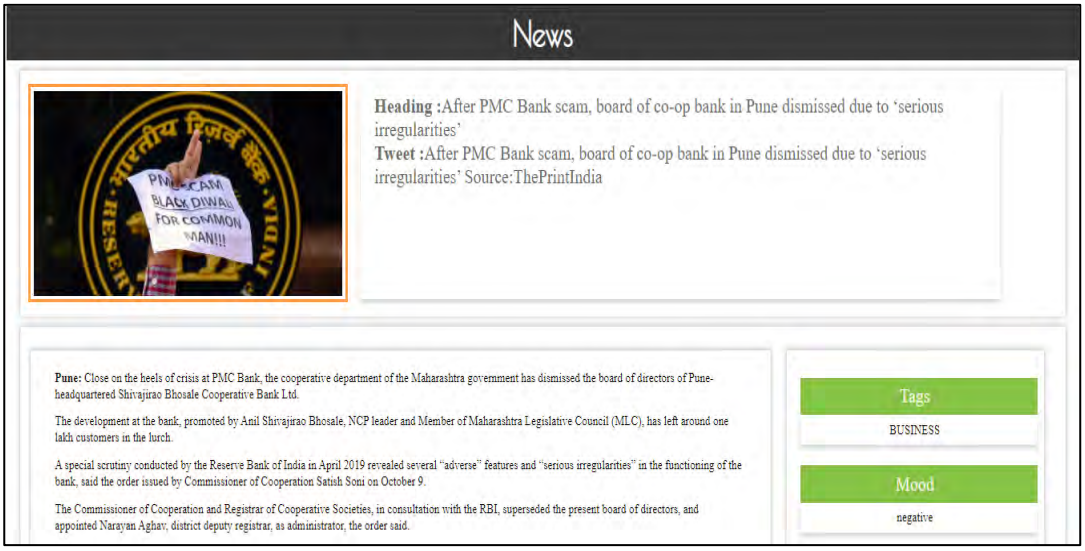


Fig. 5: News article classified as business category.

4.5 Performance Metrics

We made use of Scikit-learn to implement all the classification algorithms. The performance measurements for classifiers were studied to represent the Act imbalances on conventional measuring instruments like accuracy, recall, and precision. These metrics use knowledge about the classes currently and projected classes from classification tasks. The following confusion matrix may reflect all possible scenarios of the findings.

Table 3: Confusion matrix

		PREDICTED	
		Positive	**Negative**
ACTUAL	**Positive**	True Positive (TP)	False Negative (FN)
	Negative	False Positive (FP)	True Negative (TN)

When an instance belongs to a class, it is considered a positive case, and when an instance does not belong to a specific class, it is considered a negative case. An instance is called true positive (TP) when it belongs to the positive case and is appropriately labeled as such. The false negatives (FN) are such cases that are incorrectly classified as negative cases when they belong to the positive cases. The false positives (FP) are such cases that are incorrectly classified as positive when they belong to the negative cases. The instances belonging to the negative case are correctly classified as such. These are called the true negatives (TN). Below we highlight some of the performance metrics based on the confusion matrix.

4.5.1 Accuracy

Accuracy is the number of all right predictions the classifier has made divided by the overall data collection [48]. The accuracy mathematically expressed as

$$Accuracy = \frac{True\ Positive + True\ Negative}{(True\ Positive + False\ Positive + True\ Negative + False\ Negative)} \quad (1)$$

The error rate (ERR) is the sum of all incorrect forecasts separated by the dataset's total number. The highest error rate is 0.0, while 1.0 is the worst. The error mathematically expressed as

$$Error\ Rate = 1 - Accuracy \quad (2)$$

4.5.2. Precision

The number of articles that are properly identified (True Positive) by the number of articles that the classifier estimates corresponds to a specific group (True Positive and False Positive). The precision [48] in mathematical form is

$$Precision = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \quad (3)$$

4.5.3. Recall

The ratio of positive articles correctly predicted to all articles in the actual class is termed a recall [49]. The recall is mathematically expressed as

$$Recall = \frac{True\ Positive}{(True\ Positive + False\ Positive)} \quad (4)$$

A good recall implies that much of the positive cases (TP+FN) are rated as positive (TP). This situation is likely to result in more FP measurements and lower average precision. A poor recall implies we have a significant FN volume (should have been positive but categorized as negative). This ensures that we have greater confidence that this will be a positive case if we find a positive scenario.

On imbalanced classification, when scholars use machine learning techniques, accuracy is not accurate. In imbalanced conditions, higher accuracy can be obtained since all data is predicted as the majority class. Hence, the F1 score and the receiver operating characteristic (ROC) curve are accepted as fair measurements by the machine learning community. The tradeoff between false positive and true positive is indicated by the ROC

curve. When the false positives get overlooked, and the emphasis is distorted, it would probably reflect precision only for the true positives. On the other hand, if the real positives get overlooked, and the emphasis is distorted for false positives, the ratings would most definitely represent the recall. The classifier's efficiency is reflected by the area under the curve (AUC).

4.5.4 F1-Score

The harmonic mean of precision and recall is the F1 score. If we combine Precision and Recall, then it will become the F1 score [50]. The optimal and worst values for F1 score are 1 and 0, respectively. It is efficient to use one value for measurement instead of using two precision and recall values as it combines both. F1 score is presented as

$$F1\ score = \frac{2 * Recall * Precision}{(Recall + Precision)} \tag{5}$$

5. EXPERIMENTAL RESULTS AND PERFORMANCE ANALYSIS

For every classification algorithm checked on our dataset, Table 4 displays the assessment of comparative results of SVM, LR, KNN, and Naive Bayes algorithms using multiple news websites in terms of all mentioned performance metrics. Accuracy and F1 score are nearly similar. Figure 6 illustrates the analysis of all four algorithms in graphical form.

Table 4: Result table

Classifier	Accuracy	Precision	Recall	F1 Score
Naive Bayes	0.93	0.93	0.93	0.9
LR	0.81	0.75	0.79	0.81
SVM	0.76	0.76	0.76	0.75
KNN	0.72	0.64	0.52	0.72

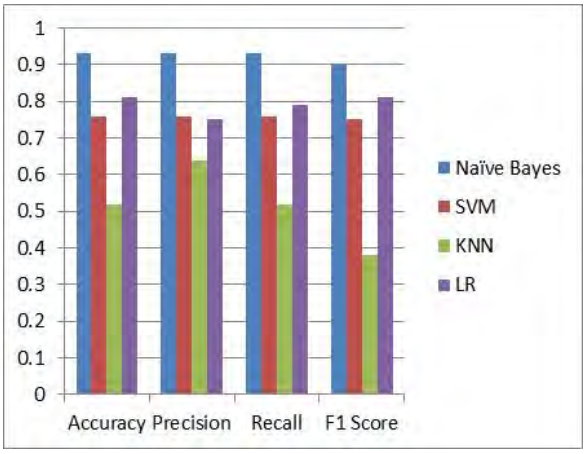


Fig. 6: Comparing all four classifiers.

In all four classifiers, the Naive Bayes classifier performed better than others with 93.0% accuracy. Moreover, the worst result was displayed by the KNN classifier at 72.0%. The best (NB) and worst (KNN) classifier's confusion matrices are shown in Fig. 7 and Fig. 8.

Figure 9 shows classifier output by imbalanced ratios. When training data for minority groups is sparse compared to other algorithms, Naive Bayes worked better than the other

https://doi.org/10.31436/iiumej.v22i2.1662

classifiers. It touched a 0.5 F1 score while the imbalance was 11%, whereas others reached 22% of the imbalance with the similar F1 score. Nonetheless, the efficiency of the classifier will not improve after imbalanced levels dropped to 44%, whereas some algorithms eventually demonstrated better results after more class imbalance training data were produced. Overall, the lowest result was obtained by the KNN classifier. To attain an average F1 score above 0.7, it took an imbalance level of 66 percent, while the majority of the classifiers achieved an average F1 score above 0.7 at a 44 percent disparity rate. In the most imbalanced conditions, Naive Bayes worked the best.

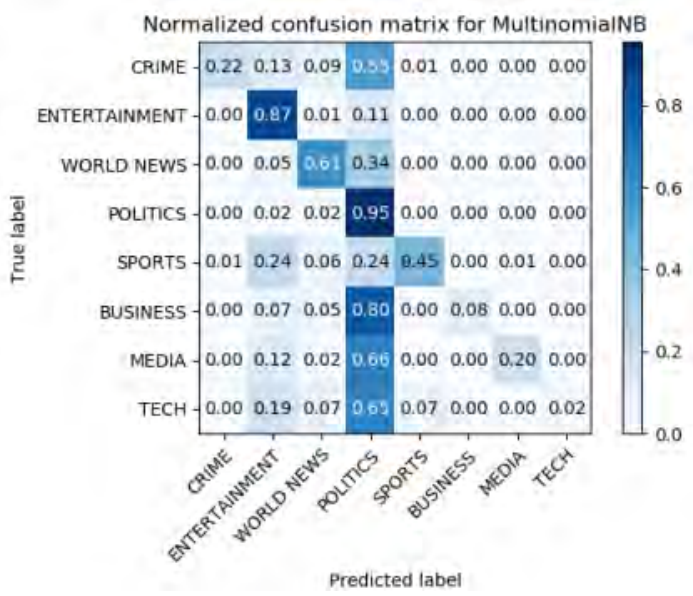


Fig. 7: Confusion matrix for the Naive Bayes classifier.

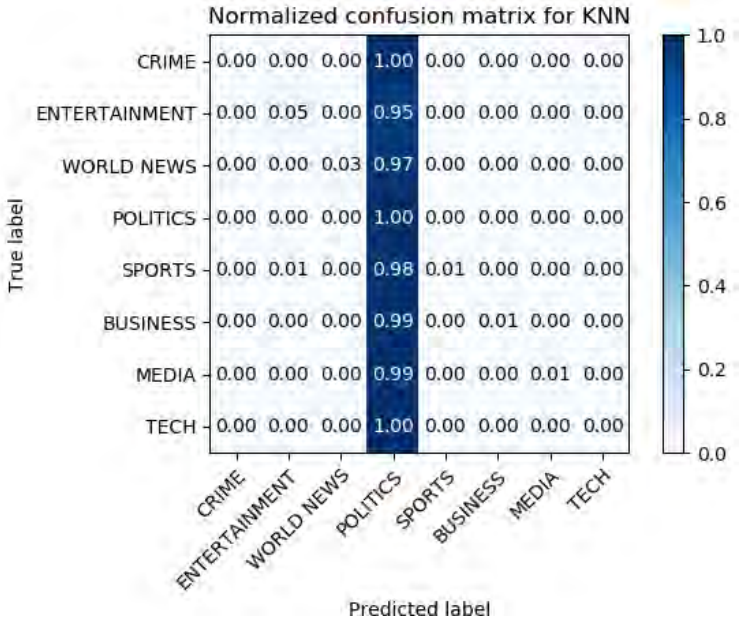


Fig. 8: Confusion matrix for the KNN classifier.

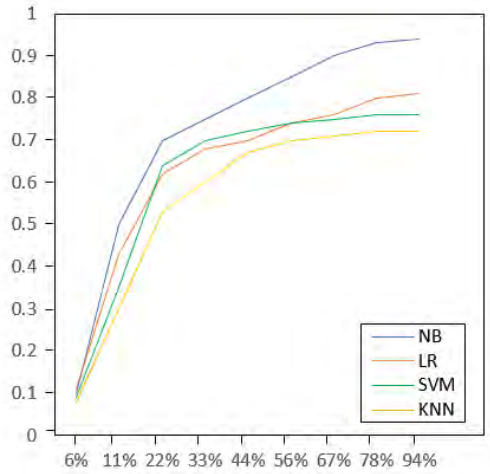


Fig. 9: Performance of classifiers - Naive Bayes performs best while KNN performs worst.

6. CONCLUSION

In this research, for online news posts, we implemented a single-class text classifier system. We presented a sample dataset containing approximately 75k news articles sliced from seven various websites with their tags. We defined the dataset's collection, cleaning, and construction phases. We analyzed our dataset by adding four distinct classifiers. We used Naïve Bayes and several other classification algorithms for classification and performed a comparison of different classifiers' outcomes from seven various sites on the same dataset, and the findings confirm that Naïve Bayes showed better performance than most classifiers; when working with various news datasets, it offers sufficient classification accuracy. This study has many possible extensions. Our future aim is to discover and implement a classification methodology in various regional languages.

ACKNOWLEDGEMENT

The authors send their humble gratitude to the university for providing all the facilities during this research.

REFERENCES

[1] Jindal R, Malhotra R, Jain A. (2015) Techniques for text classification: Literature review and current trends. Webology, 12(2): Article 139. https://www.webology.org/2015/v12n2/a139.pdf

[2] Turney P. (2002) Thumbs Up or Thumbs Down? Semantic Orientation Applied to Unsupervised Classification of Reviews. Computing Research Repository, 417-424. doi:10.3115/1073083.1073153.

[3] Wilson T, Wiebe J, Hoffmann P. (2009) Recognizing Contextual Polarity: An Exploration of Features for Phrase-Level Sentiment Analysis. Computational Linguistics, 35(3): 399-433. doi:10.1162/coli.08-012-r1-06-90

[4] Quan C, Ren F. (2009) Construction of a blog emotion corpus for Chinese emotional expression analysis. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, 3: 1446-1454.

[5] Wang TY, Chiang HM. (2011) Solving multi-label text categorization problem using support vector machine approach with membership function. Neurocomputing, 74(17): 3682-3689. https://doi.org/10.1016/j.neucom.2011.07.001

[6] Harrag F, El-Qawasmah E, Al-Salman AMS. (2010) Comparing dimension reduction techniques for Arabic text classification using BPNN algorithm. In Proceedings of the 2010 First International Conference on Integrated Intelligent Computing, Bangalore, India, 2010, pp 6-11. https://doi.org/10.1109/iciic.2010.23

[7] Sapankevych N, Sankar R. (2009) Time series prediction using support vector machines: A survey. IEEE Computational Intelligence Magazine, 4(2): 24-38. https://10.1109/MCI.2009.932254.

[8] Zhihang Chen, Chengwen Ni, Murphey, YL. (2006) Neural Network Approaches for Text Document Categorization. In Proceedings of the IEEE International Joint Conference on Neural Network Proceedings, pp.1054–1060. https://doi.org/10.1109/ijcnn.2006.246805

[9] Zhang X, Bicheng Li, Xianzhu Sun. (2010) A k-nearest neighbor text classification algorithm based on fuzzy integral. In Proceedings of the Sixth International Conference on Natural Computation, pp 2228–2231. https://doi.org/10.1109/icnc.2010.5584406

[10] Martinez-Arroyo M, & Sucar LE. (2006) Learning an Optimal Naive Bayes Classifier. In Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06), pp 748–752. https://doi.org/10.1109/icpr.2006.748

[11] Pendharkar B, Ambekar P, Godbole P, Joshi S, Abhyankar S. (2007) Topic categorization of RSS news feeds. Group, 4, 1.

[12] Rao V, Sachdev J. (2017) A machine learning approach to classify news articles based on location. In Proceedings of the International Conference on Intelligent Sustainable Systems (ICISS), pp 863-867. https://doi.org/10.1109/iss1.2017.8389300

[13] Lewis DD. (1998) Naive (Bayes) at forty: The independence assumption in information retrieval. In European conference on machine learning, 4-15). Springer, Berlin, Heidelberg

[14] Chen N, Blostein D. (2006) A survey of document image classification: problem statement, classifier architecture and performance evaluation. International Journal of Document Analysis and Recognition (IJDAR), 10(1): 1-16. https://doi.org/10.1007/S10032-006-0020-2

[15] Gupta V, Lehal GS. (2009) A survey of text mining techniques and applications. Journal of Emerging Technologies in Web Intelligence, 1(1): 60-76. https://doi.org/10.4304/jetwi.1.1.60-76

[16] Manning CD, Raghavan P, Schutze H. (2008) Introduction to information retrieval? Cambridge University Press, pp 405-416.

[17] Li W, Han J, Pei J. (2001) CMAR: Accurate and efficient classification based on multiple class-association rules. In Proceedings of the 2001 IEEE International Conference on Data Mining, San Jose, CA, USA, 2001, pp. 369-376. doi: 10.1109/ICDM.2001.989541.

[18] Yin X, Han J. (2003) CPAR: Classification based on predictive association rules. In Proceedings of the 2003 SIAM International Conference on Data Mining, pp 331-335.

[19] Berzal F, Cubero J, Marín N, Sánchez D, Serrano J, Vila A. (2005) Association rule evaluation for classification purposes. Actas del III Taller Nacional de Mineria de Datos y Aprendizaje, pp 135-144.

[20] Jiang C, Coenen F, Sanderson R, Zito M. (2010) Text classification using graph mining-based feature extraction. In Research and Development in Intelligent Systems XXVI, Springer, London, pp 21-34.

[21] Huynh D, Tran D, Ma W, Sharma D. (2011) A new term ranking method based on relation extraction and graph model for text classification. In Proceedings of the Thirty-Fourth Australasian Computer Science Conference, 113: 145-152.

[22] Han J, Kamber M. (2001) Data mining concepts and techniques, Morgan Kaufmann Publishers. San Francisco, CA, pp 335-391.

[23] Chen J, Huang H, Tian S, Qu Y. (2009) Feature selection for text classification with Naïve Bayes. Expert Systems with Applications, 36(3):5432-5435. https://doi.org/10.1016/j.eswa.2008.06.054

[24] Bijalwan V, Kumar, V, Kumari P, Pascual J. (2014) KNN based machine learning approach for text and document mining. International Journal of Database Theory and Application, 7(1):61-70. https://doi.org/10.14257/ijdta.2014.7.1.06

[25] Lin Y, Wang J. (2014) Research on text classification based on SVM-KNN. In Proceedings of the IEEE 5th International Conference on Software Engineering and Service Science, Beijing, China, pp 842-844. https://doi.org/10.1109/ICSESS.2014.6933697.
[26] Rahmawati D, Khodra ML. (2015) Automatic multi-label classification for Indonesian news articles. In Proceedings of the 2nd International Conference on Advanced Informatics: Concepts, Theory and Applications (ICAICTA), Chonburi, Thailand, 2015, pp 1-6. https://doi.org/10.1109/ICAICTA.2015.7335382.
[27] Hassaine A, Mecheter S, Jaoua A. (2015) Text categorization using hyper rectangular keyword extraction: Application to news articles classification. In Proceedings of the International Conference on Relational and Algebraic Methods in Computer Science, pp 312-325. https://doi.org/10.1007/978-3-319-24704-5_19
[28] Lehečka J, Švec J. (2015) Improving multi-label document classification of Czech news articles. In Proceedings of the International Conference on Text, Speech, and Dialogue, pp 307-315.
[29] Arya C, Dwivedi SK. (2016) News web page classification using URL content and structure attributes. In Proceedings of the 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India, 2016, pp. 317-322, https://doi.org/10.1109/NGCT.2016.787743.
[30] Li J, Fong, S, Zhuang, Y, Khoury, R. (2015) Hierarchical classification in text mining for sentiment analysis of online news. Soft Computing, 20(9), 3411-3420. https://doi.org/10.1007/s00500-015-1812-4
[31] Weng, W., Liu, Y., Wang, S., & Lei, K. (2016) A multiclass classification model for stock news based on structured data. In Proceedings of the Sixth International Conference on Information Science and Technology (ICIST), Dalian, China, pp 72-78. https://doi.org/10.1109/ICIST.2016.7483388.
[32] Kaur S, Khiva NK (2016). Online news classification using deep learning technique. International Research Journal of Engineering and Technology, 3(10): 558-563.
[33] Suh Y, Yu J, Mo J, Song L. (2017) A comparison of oversampling methods on imbalanced topic classification of Korean news articles. Journal of Cognitive Science, 18(4): 391-437.
[34] Ahmed H, Traore I, Saad S. (2017) Detecting opinion spams and fake news using text classification. Security and Privacy, 1(1), e9. https://doi.org/10.1002/spy2.9
[35] Watanabe K. (2017) Newsmap. Digital Journalism, 6(3): 294-309. https://doi.org/10.1080/21670811.2017.1293487
[36] Du C, Huang L. (2018) Text classification research with attention-based recurrent neural networks. International Journal of Computers Communications & Control, 13(1): 50-64. https://doi.org/10.15837/ijccc.2018.1.3142
[37] Gruppi M, Horne BD, Adali S. (2018) An exploration of unreliable news classification in Brazil and the US. arXiv preprint arXiv:1806.02875.
[38] Kusumaningrum R, Adhy S. (2018). WCLOUDVIZ: Word cloud visualization of Indonesian News Articles Classification Based on Latent Dirichlet Allocation. Telkomnika, 16(4): 1752-1759.
[39] Cecchini D, Na L. (2018) In Proceedings of the 2018 IEEE International Conference on Big Data and Smart Computing (BigComp), Shanghai, China, pp 681-684. https://doi.org/10.1109/BigComp.2018.00125.
[40] Jang B, Kim I, Kim JW. (2019) Word2vec convolutional neural networks for classification of news articles and tweets. PLOS ONE, 14(8): e0220976. https://doi.org/10.1371/journal.pone.0220976
[41] Qadi LA, Rifai HE, Obaid S, Elnagar A. (2019) Arabic text classification of news articles using classical supervised classifiers. In Proceedings of the 2nd International Conference on new Trends in Computing Sciences (ICTCS), Amman, Jordan, pp 1-6. https://doi.org/10.1109/ICTCS.2019.8923073.
[42] Gumilang M, Purwarianti A, Nurdinasari F. (2019) Systemic risk document classification on Indonesian news articles using deep learning and active learning. In Proceedings of the International Conference on Electrical Engineering and Informatics (ICEEI), Bandung, Indonesia, pp 46-51. https://doi.org/10.1109/iceei47359.2019.8988829

[43] Noppakaow A, Uchida O. (2019) Examinations on the Performance of Classification Models for Thai News Articles. In Proceedings of the 2019 11th International Conference on Information Technology and Electrical Engineering (ICITEE), Pattaya, Thailand, 2019, pp. 1-4 https://doi.org/10.1109/iciteed.2019.8929959

[44] Huang CM, Jiang YJ. (2019) An empirical study on the classification of Chinese news articles by machine learning and deep learning techniques. In Proceedings of the International Conference on Machine Learning and Cybernetics (ICMLC), Kobe, Japan, pp 1-6. https://doi.org/10.1109/icmlc48188.2019.8949309

[45] Winster SG, Kumar MN. (2020) Automatic classification of emotions in news articles through ensemble decision tree classification techniques. Journal of Ambient Intelligence and Humanized Computing, 1–12. https://doi.org/10.1007/s12652-020-02373-5

[46] Rabbimov IM, Kobilov SS. (2020) Multi-class text classification of Uzbek news articles using machine learning. Journal of Physics: Conference Series, 1546(1): 012-097.

[47] Fesseha A, Xiong S, Emiru ED, Dahou A. (2020) Text classification of news articles using machine learning on low-resourced language: Tigrigna. In Proceedings of the 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD), Chengdu, China, pp 34-38. https://doi.org/10.1109/ICAIBD49809.2020.9137443.

[48] Sharma A, Mishra PK. (2020) State-of-the-art in performance metric and future direction for data science algorithm. Journal of Scientific Research, 64(2): 221-238.

[49] Saura JR. (2020) Using Data Sciences in Digital Marketing: Framework, methods, and performance metrics. Journal of Innovation & Knowledge, 6(2): 92-102. https://doi.org/10.1016/j.jik.2020.08.001

[50] Pereira L, Nunes N. (2020) An empirical exploration of performance metrics for event detection algorithms. Non-Intrusive Load Monitoring. Sustainable Cities and Society, 62: 102399. https://doi.org/10.1016/j.scs.2020.102399

https://doi.org/10.31436/iiumej.v22i2.1692

MACHINE-LEARNING-BASED EVALUATION OF CORROSION UNDER INSULATION IN FERROMAGNETIC STRUCTURES

ALI SOPHIAN[1*], FARIS NAFIAH[2,3], TEDDY SURYA GUNAWAN[4], NUR AMALINA MOHD YUSOF[1] AND ALI AL-KELABI[1]

[1]*Smart Structures, Systems and Control Lab, Department of Mechatronics Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur Malaysia*
[2]*School of Engineering, London South Bank University, London, United Kingdom*
[3]*NSIRC, TWI Ltd., Cambridge, United Kingdom*
[4]*Department of Electrical and Computer Engineering, Kulliyyah of Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur Malaysia*

[*]*Corresponding author: ali_sophian@iium.edu.my*

(Received: 10th November 2020; Accepted: 4th April 2021; Published on-line: 4th July 2021)

ABSTRACT: Corrosion under insulation (CUI) is one of the challenging problems in pipelines used in the gas and oil industry as it is hidden and difficult to detect but can cause catastrophic accidents. Pulsed eddy current (PEC) techniques have been identified to be an effective non-destructive testing (NDT) method for both detecting and quantifying CUI. The PEC signal's decay properties are generally used in the detection and quantification of CUI. Unfortunately, the well-known inhomogeneity of the pipe material's properties and the presence of both cladding and insulation lead to signal variation that reduces the effectiveness of the measurement. Current PEC techniques typically use signal averaging in order to improve the signal-to-noise ratio (SNR), with the drawback of significantly-increasing inspection time. In this study, the use of Gaussian process regression (GPR) for predicting the thickness of mild carbon steel plates has been proposed and investigated with no signal averaging used. With mean absolute errors (MAE) of 0.21 mm, results show that the use of GPR provides more accurate predictions compared to the use of the decay coefficient, whose averaged MAE is 0.36 mm. This result suggests that the GPR-based method can potentially be used in PEC NDT applications that require fast scanning.

ABSTRAK: Hakisan di bawah penebat CUI adalah salah satu masalah yang mencabar dalam saluran paip yang digunakan dalam industri gas dan minyak kerana tersembunyi dan sukar dikesan tetapi boleh menyebabkan bencana. Teknik Pulsed eddy current (PEC) telah dikenal pasti sebagai kaedah ujian bukan pemusnah yang berkesan (NDT) untuk mengesan dan mengukur CUI. Sifat kerosakan isyarat PEC umumnya digunakan dalam pengesanan dan pengukuran CUI. Malangnya, sifat tidak tepat yang terkenal dari sifat bahan paip dan kehadiran pelapisan dan penebat menyebabkan variasi isyarat yang mengurangkan keberkesanan pengukuran. Teknik PEC semasa biasanya menggunakan rata-rata isyarat untuk meningkatkan nisbah isyarat-ke-kebisingan (SNR), dengan kelemahan peningkatan masa pemeriksaan dengan ketara. Dalam kajian ini, penggunaan regresi proses Gauss (GPR) untuk meramalkan ketebalan plat keluli karbon ringan telah diusulkan dan diselidiki dan tidak ada rata-rata isyarat yang digunakan. Dengan ralat mutlak (MAE) 0,21 mm, hasil menunjukkan bahawa penggunaan GPR memberikan ramalan yang lebih tepat dibandingkan dengan penggunaan pekali peluruhan, yang rata-

rata MAE adalah 0,36 mm. Hasil ini menunjukkan bahawa kaedah berasaskan GPR berpotensi digunakan dalam aplikasi PEC NDT yang memerlukan pengimbasan pantas.

KEYWORDS: *corrosion under insulation; pulsed eddy current; non-destructive testing; machine learning; fast scanning*

1. INTRODUCTION

Industrial oil and gas infrastructures, including the transmission pipelines that are used in transportation and distribution, require regular inspection as stipulated by governmental safety regulations all around the world to prevent catastrophic accidents. Many of these structures are made of carbon steel materials and covered by insulation for their protection or maintaining the high temperature of the oil or gas flowing inside them. A thin cladding is usually placed outside the insulation to hold the insulation in place and to stop water from seeping into the insulation that may cause or accelerate the growth of corrosion on the outer surface of the pipe. Such structures, whose illustration can be seen in Fig. **1**, will benefit from non-destructive inspection techniques that are able to penetrate through the insulation and cladding, and therefore, detect any presence of the corrosion under insulation (CUI). CUI is considered a crucial problem with these pipes [1,2]. The ability to detect and assess CUI in the pipes without the removal of the insulation is very beneficial [3]. Removal of the insulation takes time, which, in turn, would lengthen the downtime of the plant and would require new costly replacements of the insulation or coating.

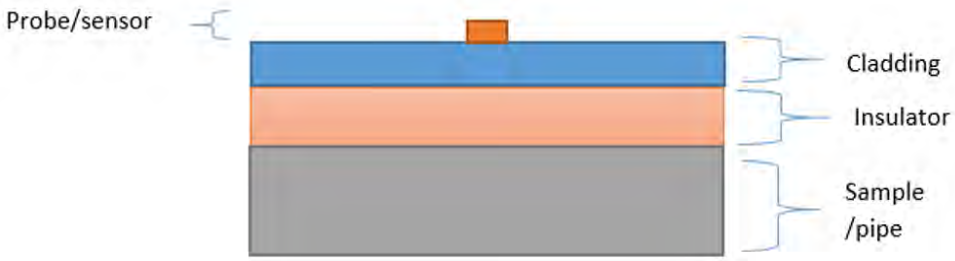


Fig. 1: Typical setting of insulated pipes with a probe sitting on its cladding.

Pulsed eddy current (PEC) non-destructive testing (NDT) has been identified as one of the solutions with the most potential for non-contact evaluation of corrosion in carbon steel structures [4], which are either coated or otherwise. The low frequency components of the excitation field allow deep penetration into the ferromagnetic structure. For such structures, magnetic saturation has been proposed, although many have proposed the use of PEC without magnetically saturating the sample.

There have been several different PEC signal features proposed by researchers [3]. The signal features also depend on the type of the sensing device, whether it is an inductive coil or a magnetic field sensor. Mostly, the features are related to the decay properties of the signals, especially the gradient of the later stage of the decay, thanks to its high correlation to the thickness of the tested structure and its insensitivity to the variation in the distance between the probe and the sample, which is known as the stand-off or lift-off. There have been promising results reported, although most of these have been applied on signals obtained by using inductive coils. A technique based on the signal's inverse time derivative $|\nabla|^{-1}$ is shown to allow in-situ calibration and improved estimation compared to the τ_0 feature, which is the time at which the induced current's

https://doi.org/10.31436/iiumej.v22i2.1692

diffusion phase ceases [1]. An accuracy of 0.67 mm has been achieved with averaging of 16 signals. Time to peak as a feature was proposed and its performance was shown to be good on simulation data, while no experimental data were reported [4]. Other time domain features have also been discussed in other reports, such as [5-7]. The use of a magnetic sensor, instead of an inductive coil, offers benefit in terms of the spatial resolution. However, relatively less studies have been reported where a magnetic sensor was used. Cheng et al. used anisotropic magnet resistance (AMR) sensors and the feature used was the decay coefficient, which is the reciprocal of the decay rate of the magnetic field density in the logarithmic scale [8].

PEC signals are known to be noisy. Some noise reduction techniques have been presented [9,10], and the testing gets even more challenging due to inhomogeneity in ferromagnetic materials [11]. The use of signal averaging is effective to some extent, however it increases the measurement time significantly, depending on the number of signals used. This will prevent fast scanning applications using the PEC. In this paper, machine learning is used to predict the thickness of the insulated ferromagnetic sample without applying signal averaging. Machine learning, including its latest development known as deep learning, has embraced countless applications in various sectors of industry. Non-destructive testing (NDT) is not an exception to this current trend, and eddy NDT current techniques have also seen the applications of machine learning, such as the works reported in [12,13]. In this study, the use of machine learning in processing PEC signals obtained without signal averaging is presented.

2. MODELLING OF PEC AND CUI

An axisymmetric 2D model, as shown in Fig. 2, was built using COMSOL Multiphysics which consists of a few layers, namely the carbon steel sample, the insulation, and the cladding. The inner and outer radii of the excitation coil were 100 mm and 110 mm respectively, while the height of coil was 6 mm. In this work, the model is used to verify the trends in the magnetic field density that will be sensed by the sensor when the lift-off and sample thickness are varied. It is not intended to find the actual magnitudes of the field.

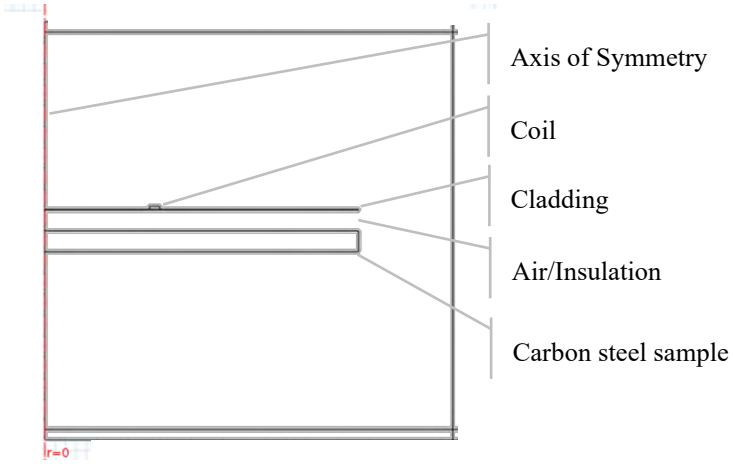
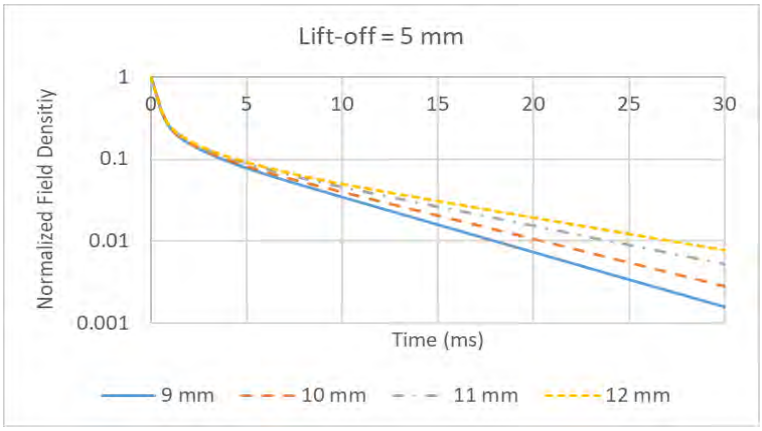


Fig. 2: 2D-axisymmetric finite element model of the PEC system on an insulated sample with a cladding. The model was developed on COMSOL software.

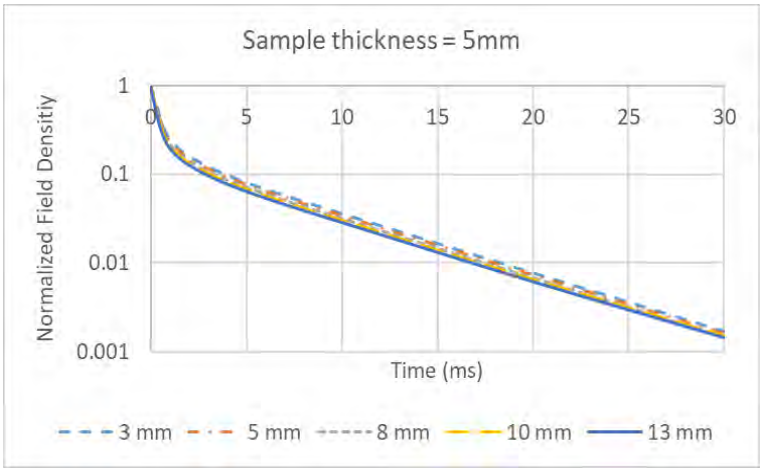
The conductivity and relative permeability of the sample were set to 5 x 10^6 S/m and 100, respectively. The thickness was varied from 9 mm to 12 mm with an increment of 1

mm. For the cladding, its thickness was 0.5 mm and two different sets of permeability and conductivity values were used to simulate stainless steel and aluminium materials.

The model was run for both different sample thicknesses and different insulation thicknesses. The results were plotted in the graphs shown in Fig. 3 and they verify that the slope of the signals is affected by the sample thickness but not by the lift-off, as has been reported by other researchers.



(a)



(b)

Fig. 3: Typical signals obtained by using the developed PEC models, (a) for different thicknesses and (b) for different lift-offs.

3. METHODOLOGY

In this study, the thickness approximation techniques using the gradient coefficient and machine-learning-based regression techniques will be compared on PEC signals that have been obtained without any signal averaging.

3.1 Experimental Setup

A PEC system was built, consisting of a probe, a data acquisition, and a laptop running LabView. The probe consisted of an excitation coil and a Hall-effect device. A ferrite core was used for concentrating the magnetic flux and strengthening the Hall device's output signal. Through the LabView code, the signal sampling was set at 100 kS/s and the pulse width was 25 ms. The LabView code saved the acquired data to be

analysed. The inner and outer diameters of the coil were 100 mm and 110 mm, respectively. Its height was 6 mm and it had 200 turns. A MOSFET was used to switch the high excitation currents on and off with a pulse width of 25 ms driven to the excitation coil. Fig. 4 shows the block diagram of the PEC setup used in this work.

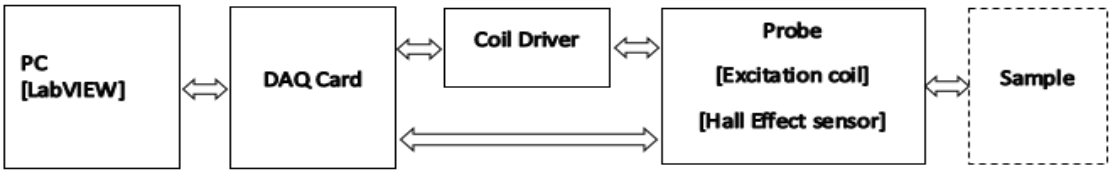


Fig. 4: The block diagram of our PEC system.

Corrosion or wall loss was simulated by thicknesses that were less than the thickness of the sound structure. Square plate samples with surface dimensions of 300 mm × 300 mm were used and they had different thicknesses, which were 9.12, 10.02, 11.06, and 12.08 mm. The stand-off was also varied at 3, 5, 8, 10, and 13 mm. Fifty signals were obtained for each combination of thickness and stand-off. Two different materials were used for the cladding, which were aluminium and stainless steel. The thickness of the cladding was 0.5 mm.

3.2 Signal Processing (Data Pre-processing)

The falling edge of the Hall effect signal and its decay properties were used to predict the thickness. In order to improve the signal-to-noise ratio (SNR), a median filter was applied to reduce the high-frequency noise. The logarithmic values of the signal were used to find the decay coefficient. Normalization to unity was employed to reduce the effects of the variation in the magnitude of the magnetic field that was induced by the excitation coil, which may be caused by the sample's inhomogeneity and the change in the excitation coil's temperature, among others. Fig. **5** shows typical signals that have been averaged in order to increase the clarity of the shape of the signals. For the purpose of determination of the decay coefficient, part of the signal that has been used is 10 ms $\leq t \leq$ 35 ms.

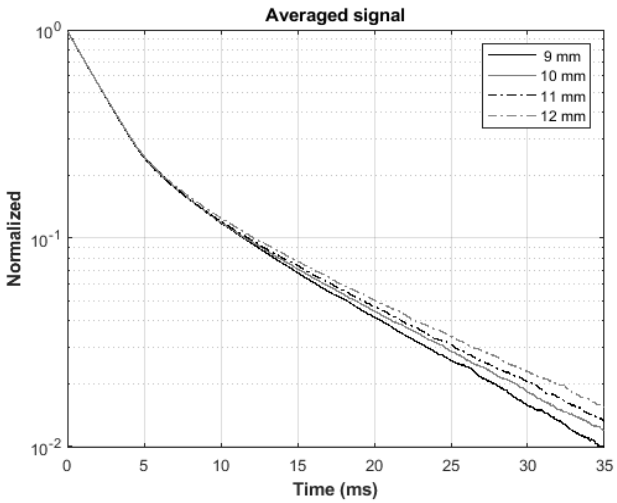


Fig. 5: Typical averaged signals for different sample thicknesses.

3.3 Machine-Learning-based Regression

Based on preliminary observation, the Gaussian process regression (GPR) model consistently provided the best result among other regressions. GPR is a non-parametric and non-linear kernel-based model, and it uses a Bayesian approach to model the uncertainty of the prediction. It has been widely used in various applications, thanks to its simplicity and its good performance when small datasets are available, among others. The kernel function that models the similarity between similar predictors was chosen to be the exponential kernel, which is defined by [14],

$$k(x_i, x_j|\theta) = \sigma_f^2 \exp\left(-\frac{r}{\sigma_l}\right), \quad (1)$$

where σ_l is the characteristic length scale and r is the Euclidean distance between predictors x_i and x_j, and θ is the vector containing the parameters for the kernel.

The decay of the magnetic field that is sensed by the Hall device is set as the predictor variables for the regression model.

3.4 Model Validation and Testing

Validation was performed on machine leaning models in order to test their performance after training, including the generalization of the model. Due to the limited quantity of data and to avoid overfitting the model, a cross validation technique known as k-fold was employed in this work. In this technique, the data were split into k groups and the model was trained and validated k times. In every iteration, one different group was excluded from the training and used only in the validation, therefore each and every group of data was used for validation. In this work, 5-fold was used meaning that the data was split into commonly used 5 groups. Seventy percent of the overall data that were chosen randomly were used for this training and validation exercise. The remaining 30% of the dataset were used to test and compare the model against the performance of the approach that uses the 1/gradient feature.

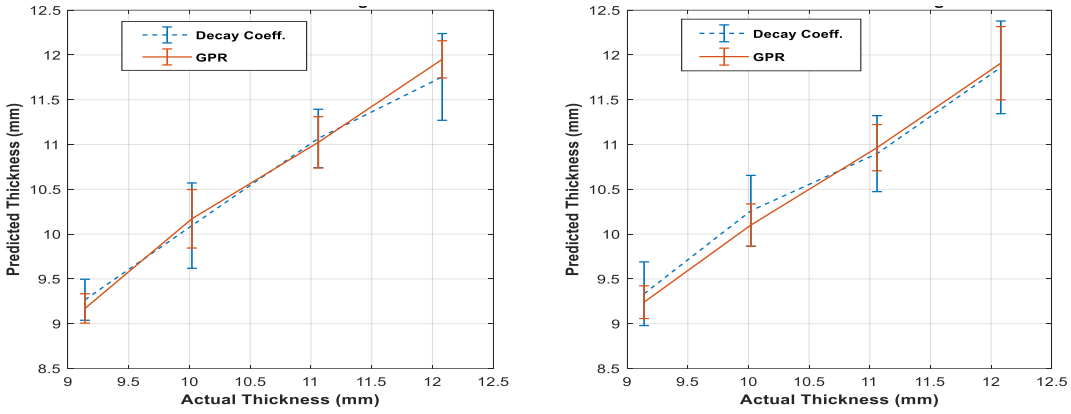
The performances of both methods, namely decay coefficient and GPR, were evaluated and compared using the following metrics: root mean squared error (RMSE), R-squared, mean squared error (MSE), and mean absolute error (MAE). All these metrics are already well known and widely used for evaluating performances of both measurement systems and machine learning models.

4. RESULTS AND DISCUSSION

Fig. **6** exhibits the plots of the predicted thickness generated using both methods with the experimental data obtained from all different lift-offs, sample thicknesses, and cladding materials. The error bars in the plots represent the standard deviation. The performance metrics are shown in Table 1.

Table 1: Performance metrics of both methods

Method	Cladding	RMSE (mm)	R-squared	MSE (mm)	MAE (mm)
Decay coefficient	Stainless steel	0.4611	0.825	0.213	0.373
	Aluminium	0.420	0.856	0.177	0.346
GPR Exponential	Stainless steel	0.298	0.928	0.089	0.215
	Aluminium	0.266	0.942	0.071	0.198



(a) (b)

Fig. 6: The averaged predicted thickness with error-bars representing the standard deviation of the measurement error, (a) aluminum cladding, (b) stainless steel cladding.

Both the plots and the table demonstrate that the GPR exponential model outperformed the approach using the decay coefficient in terms of both the nominal error and the spread of errors. The averaged RMSE, R-squared, MSE and MAE of the GPR model are better by 36.0%, 11.2%, 59.0% and 42.6% respectively than those of predicted thicknesses using the decay coefficient.

5. CONCLUSION

In recent years, machine learning - a subset of artificial intelligence - has developed very rapidly and seen applications in various sectors, with deep learning being the latest development. Insulated industrial pipes develop corrosion or wall thinning that is hidden and challenging to be detected. This work proposed the use of machine learning for processing PEC signals, especially when the SNR was relatively high (as no signal averaging was used in order to improve the scanning speed), and therefore reducing the inspection time. The experimental data analysis shows that the GPR exponential model generated better results with both lower nominal errors and lower error variances, represented in the improvements in the RMSE, R-squared, MSE and MAE, by 36.0%, 11.2%, 59.0%, and 42.6%, respectively. This suggests that the machine learning method can potentially be used in the applications where fast scanning is required.

REFERENCES

[1] Nafiah F, Tokhi MO, Majidnia S, Rudlin J, Zhao Z, Duan F. (2020) Pulsed Eddy current: Feature extraction enabling in-situ calibration and improved estimation for ferromagnetic application. J. Nondestruct. Eval., 39(52). doi: https://doi.org/10.1007/s10921-020-00699-w.

[2] Nicola S, Carreto V, Mentzer RA, Mannan MS. (2013) Corrosion under insulation detection technique. In Proceedings of Corrosion 2013: 17-21 March 2013; Orlando, Florida

[3] Sophian A, Tian G, Fan M. (2017) Pulsed Eddy Current Non-destructive Testing and Evaluation: A Review. Chinese J. Mech. Eng. (English Ed.), 30:500-514

[4] Xu Z, Wu X, Li J, Kang Y. (2012) Assessment of wall thinning in insulated ferromagnetic pipes using the time-to-peak of differential pulsed eddy-current testing signals. NDT E Int., 51:24-29. doi: https://dx.doi.org/10.1016/j.ndteint.2012.07.004.

[5] Ulapane N, Nguyen L. (2019) Review of pulsed-Eddy-current signal feature-extraction methods for conductive ferromagnetic material-thickness quantification. Electron., 8(5):470. doi: https://doi.org/10.3390/electronics8050470.

[6] Ulapane N, Alempijevic A, Calleja TV, Miro JV. (2017) Pulsed Eddy current sensing for critical pipe condition assessment. Sensors, 17(10):2208. doi: https://doi.org/10.3390/s17102208.
[7] Buck JA, Kramer C, Lei J, Lepine BA. (2016) Field implementation of a transient Eddy current system for carbon steel pipe thickness measurements. In Proceedings of 6th International CANDU In-Service Inspection Workshop: 15–17 November 2016; Ontario
[8] Cheng W. (2012) Pulsed Eddy current testing of carbon steel pipes' wall-thinning through insulation and cladding. J. Nondestruct. Eval., 31(3): 215-224. doi: https://doi.org/10.1007/s10921-012-0137-9.
[9] Huang C, Xinjun W, Zhiyuan X, Kang Y. (2010) Pulsed Eddy current signal processing method for signal denoising in ferromagnetic plate testing. NDT E Int., 43(7): 648-653. doi: https://doi.org/10.1016/j.ndteint.2010.06.010.
[10] Huang C, Wu X. (2015) An improved ferromagnetic material pulsed eddy current testing signal processing method based on numerical cumulative integration. NDT E Int., 69: 35-39.
[11] Sophian A, Tian GY, Taylor D, Rudlin J. (2003) Flaw detection and quantification for ferromagnetic steels using pulsed eddy current techniques and magnetization. Laser Metrology and Machine Performance VI, pp 381-390.
[12] Daura LU, Tian GY, Yi Q, Sophian A. (2020) Wireless power transfer-based eddy current non-destructive testing using a flexible printed coil array: WPT based FPC-ECT. Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., 378:20190579 doi: https://doi.org/10.1098/rsta.2019.0579.
[13] Nafiah F, Sophian A, Khan MR, Zainal Abidin IM. (2018) Quantitative evaluation of crack depths and angles for pulsed eddy current non-destructive testing, NDT E Int., 102:180-188. doi: 10.1016/j.ndteint.2018.11.019.
[14] Kernel (Covariance) Function Options [https://www.mathworks.com/help/stats/kernel-covariance-function-options.html]

https://doi.org/10.31436/iiumej.v22i2.1752

CLASSIFICATION OF CHEST RADIOGRAPHS USING NOVEL ANOMALOUS SALIENCY MAP AND DEEP CONVOLUTIONAL NEURAL NETWORK

MOHD ADLI MD ALI[1*], MOHD RADHWAN ABIDIN [2], NIK ARSYAD NIK MUHAMAD AFFENDI[3], HAFIDZUL ABDULLAH[1], DAANIYAL REESHA ROSMAN[1], NU'MAN BADRUD'DIN[1], FAIZ KEMI[1] AND FARID HAYATI[1]

[1]Department. of Physics,
[2]Department of Radiology,
[3]Department of Internal Medicine,
International Islamic University Malaysia, Kuantan, Malaysia

[*]Corresponding author: qunox@iium.edu.my

(Received: 27th August 2020; Accepted: 30th January 2021; Published on-line: 4th July 2021)

ABSTRACT: The rapid advancement in pattern recognition via the deep learning method has made it possible to develop an autonomous medical image classification system. This system has proven robust and accurate in classifying most pathological features found in a medical image, such as airspace opacity, mass, and broken bone. Conventionally, this system takes routine medical images with minimum pre-processing as the model's input; in this research, we investigate if saliency maps can be an alternative model input. Recent research has shown that saliency maps' application increases deep learning model performance in image classification, object localization, and segmentation. However, conventional bottom-up saliency map algorithms regularly failed to localize salient or pathological anomalies in medical images. This failure is because most medical images are homogenous, lacking color, and contrast variant. Therefore, we also introduce the Xenafas algorithm in this paper. The algorithm creates a new kind of anomalous saliency map called the Intensity Probability Mapping and Weighted Intensity Probability Mapping. We tested the proposed saliency maps on five deep learning models based on common convolutional neural network architecture. The result of this experiment showed that using the proposed saliency map over regular radiograph chest images increases the sensitivity of most models in identifying images with air space opacities. Using the Grad-CAM algorithm, we showed how the proposed saliency map shifted the model attention to the relevant region in chest radiograph images. While in the qualitative study, it was found that the proposed saliency map regularly highlights anomalous features, including foreign objects and cardiomegaly. However, it is inconsistent in highlighting masses and nodules.

ABSTRAK: Perkembangan pesat sistem pengecaman corak menggunakan kaedah pembelajaran mendalam membolehkan penghasilan sistem klasifikasi gambar perubatan secara automatik. Sistem ini berupaya menilai secara tepat jika terdapat tanda-tanda patologi di dalam gambar perubatan seperti kelegapan ruang udara, jisim dan tulang patah. Kebiasaannya, sistem ini akan mengambil gambar perubatan dengan pra-pemprosesan minimum sebagai input. Kajian ini adalah tentang potensi peta salien dapat dijadikan sebagai model input alternatif. Ini kerana kajian terkini telah menunjukkan penggunaan peta salien dapat meningkatkan prestasi model pembelajaran mendalam dalam pengklasifikasian gambar, pengesanan objek, dan segmentasi gambar. Walau bagaimanapun, sistem konvensional algoritma peta salien jenis bawah-ke-atas kebiasaannya gagal mengesan salien atau anomali patologi dalam gambar-gambar

perubatan. Kegagalan ini disebabkan oleh sifat gambar perubatan yang homogen, kurang variasi warna dan kontras. Oleh itu, kajian ini memperkenalkan algoritma Xenafas yang menghasilkan dua jenis pemetaan saliensi anomali iaitu Pemetaan Kebarangkalian Keamatan dan Pemetaan Kebarangkalian Keamatan Pemberat. Kajian dibuat pada peta salien yang dicadangkan iaitu pada lima model pembelajaran mendalam berdasarkan seni bina rangkaian neural konvolusi yang sama. Dapatan kajian menunjukkan dengan menggunakan peta salien atas gambar-gambar radiografi dada tetap membantu kesensitifan kebanyakan model dalam mengidentifikasi gambar-gambar dengan kelegapan ruang udara. Dengan menggunakan algoritma Grad-CAM, peta salien yang dicadangkan ini mampu mengalih fokus model kepada kawasan yang relevan kepada gambar radiografi dada. Sementara itu, kajian kualitatif ini juga menunjukkan algoritma yang dicadangkan mampu memberi ciri anomali, termasuk objek asing dan kardiomegali. Walau bagaimanapun, ianya tidak konsisten dalam menjelaskan berat dan nodul.

KEYWORDS: *saliency mapping; chest radiograph; convolutional neural network*

1. INTRODUCTION

The convolutional neural network (CNN) has become the de-facto choice for image classification and object detection. It has shown that the network model can achieve human-level accuracy, including for medical images. Nevertheless, researchers are still finding ways to improve the classification performance with novel ideas. The majority of this research focuses on developing ever more complex and deep architecture. In this paper, we test the idea of changing the input typing rather than the model architecture. Instead of using a regular medical image, the saliency map is proposed to be the alternative input.

1.1 Introduction to Saliency Map

Itti et al. [1] introduced the concept of the saliency map in 1998. A saliency map is a numerical map that localizes an object (or objects) in an image that is deemed interesting (salient). In other words, the map emphasizes relevant features in an image while at the same time suppressing irrelevant features. Saliency maps have been employed in many tasks, including image classification, object detection, and image segmentation [2,3].

Methods for creating a salient map can be divided into the top-down and bottom-up approaches [4]. In the bottom-up approaches, the salient map is constructed based solely on the image's feature. Features such as color mapping, contrast, edges, and objection placement are used to localize the image's salient region. Famous bottom-up algorithms are Binary Normed Gradient for Objectness [5], the Fine-Grained [6], and Spectral Residual [7]. However, [8] stated that medical images produced by conventional modalities such as CXR, computer tomography (CT) scan, and ultrasound are mostly homogenous and possess very few color variants. In situations like this, most conventional bottom-up algorithms will fail to localize any salient object in the image; this is shown in Fig. 4.

Contradicting the previous method, the top-down approach produces a salience map based on the task given. The algorithm takes external cues from a human or model feedback to construct the final salience map. This method is fast becoming the mainstream solution, especially for medical images, as it can produce precise salient region boundaries even in the presence of shades or reflections [9]. However, since the techniques are based on a supervised CNN model, from which it naturally inherits CNN dependencies. First, it requires a large number of annotated samples for training purposes. Secondly, its development and deployment require access to accelerated hardware. These two requirements are an obstacle for the practical deployment of such technology in the medical

field, especially in Malaysia. Currently, Malaysia lacks any open medical image dataset, and very few hospitals are equipped with, or have access to, accelerated hardware.

1.2 Anomalous Saliency Mapping

In this paper, we introduce a new algorithm called the Xenafas algorithm that produces two novel anomalous saliency maps call Intensity Probability Map (IPM) and Weighted Intensity Probability Map (WPM). Different from the bottom-approach which only takes internal image cues, our approach takes cues from the probability mapping of a pixel's intensity relative to a cluster of similar images. However, it also does not require annotated samples or accelerated hardware to create the saliency map, which is practical in the context of Malaysia's clinical settings. Therefore, the algorithm can be considered as a middle ground between the bottom-up and top-down approach. We test the algorithm on a chest radiograph (CXR) dataset to see if the algorithm can create a salience region by highlighting pathological features such as air space opacities, masses, and foreign objects.

2. LITERATURE REVIEW

For readers who want more information on the saliency map, ref [8] provides an extensive review of the subject matter. This paper's literature review will focus on the application of the saliency map in medical image analysis.

The application of saliency in medical images can be separated into two categories, depending on when it is used. The majority of research only applied it post-training and solely for model interpretations; it is not actively involved in model training. For example, in [10], the saliency map produced by the class activation mappings (CAM) [11] is used to validate the feature selected by the CheXNeXt model for its classifications. Similarly, in [12], a saliency map is created via the guided back-propagation method [13], which is then used to provide interpretability for model classification on breast cancer image classifications. Research done by [14-16] also shows similar traits. However, it is vital to mention the finding by [17], in which the author demonstrated that most algorithms used to create this saliency map are inconsistent when repeated. Among all algorithms, the Grad-CAM [18] algorithm shows the most consistency. Thus, the trustworthiness of using a saliency map to validate clinical CNN models is questionable.

The second type of saliency map research actively uses it in the model training. For example, in [19], the saliency map in the form of an attention map reduces a model false-positive rate. Similar to a saliency map, the attention map produced by the Attention Gate (AG) algorithm suppresses irrelevant regions in the image. In [20], localization of pulmonary lesions in CXR images is achieved by extracting a saliency map from a CNN model. Likewise, in [21], a saliency map is generated and used to detect polyps in capsule endoscopy. Various bottom-up saliency algorithms are used for segmenting skin cancer in [22,23].

To the best of the authors' knowledge, there is yet a paper that examines the effect of using a saliency map as input for chest radiograph classification. Therefore, in this paper, we tested the effect of using the proposed saliency map, IPM, and WPM, which will enhance the classification performance of the CNN model. In addition, we also test if the proposed saliency map successfully highlights all pathological features in a CXR image. For the interested reader, a review on the classification of CXR by supervised CNN models can be found from [24].

3. METHODOLOGY

3.1 The Xenafas Algorithm

We proposed the Xenafas method, an algorithm that indicates anomaly regions' location on a CXR image, based on the likelihood of a pixel's intensity (opacities) at a given location. The method starts with creating a control dataset. Images for this dataset must be cherry-picked; avoiding images containing any form of anomalies. Examples of anomalies include but are not limited to; any pathology, foreign body, extreme variation such as dextrocardia, rotated film, and patients in non-standard body positions.

After the control dataset has been created, the images are clustered into several groups using the K-Means algorithm. This step is needed to address the variation in patient body shape, image quality between x-ray machines, and the patient's body's orientation when the x-ray image is taken. The number of clusters, K, depends on the homogeneity of the images in the dataset. A good homogenous dataset will use a K value of 1–3, while a heterogeneous dataset will use a value between 7–10. Next, the 2D pixel intensity distribution or ProbMat is created as shown in Algorithm 1.

Data: Batch of images
Result: 2D Pixel Intensity Distribution (probMat)
initialization;
probMat = [ ][ ];
for *x in image width, y in image height,* **do**
    pxiLs = [pixel intensitity at (x,y) for image in image batch];
    probFunc( ) = non-parametric probability function of pxiLs;
    probMat[x][y] = [probFunc(i) for i in range 0 to 255];
end

Fig. 1: The pseudocode for producing the ProbMat.

There are several ways to create a non-parametric probability function; one of the most popular is to use the Kernel Density Estimation method (KDE). KDE is easy to implement; however, computationally intensive when it is scaled to sample high-resolution images. A dataset with images with 256 by 256 resolution will need 65 536 KDE modeling to create all the necessary ProbMat. This requirement will quickly exhaust the memory resource of a computer. Additionally, there is no clear guideline on determining the appropriate bandwidth value of a KDE model.

We proposed another method as an alternative to KDE. In this method, we first create a histogram of pixel intensity for all CXR images in the subcluster K, at a specific value of x and y. From the histogram, a discrete probability function can be obtained. A continuous function for all possible intensity values is approximate by combining the discrete probability function with cubic spline interpolation. The Savitzky-Golay filter was then applied to smooth the probability distribution function further. Using this continuous probability function, it is now possible to create a matrix (ProbMat) representing the probability of all intensity values at any given location. Pseudocode shown Fig. 2 is used to create the anomalous saliency map. In this part, the ProbMat is used to produce the Intensity

https://doi.org/10.31436/iiumej.v22i2.1752

Probability Map (IPM) and Weighted Intensity Probability Map (WPM) for all CXR images.

```
Data: Test image, ProbMat
Result: IMP or WMP Heatmap
initialization;
heatMap = [ ][ ];
for x in image width, y in image height do
    heatMap[ x ][ y ] = ProbMat[ x ][ y ][ test image pixel intensity at
    (x)(y)];
end
if output type == WMP then
    return wmpFunc( heatMap );
else
    return heatMap as IMP image;
end
```

Fig. 2: Pseudocode to produce IPM and WPM saliency map.

The WPM function is given by Eq. (1),

$$W(x,y) = P(i_{x,y})i_{x,y} \tag{1}$$

The weighted pixel intensity, W at position x, y is equal to the product of its intensity, i, and the intensity likelihood, $P()$, at the same locations. In WPM, the original pixel intensity acts as a weight for the likelihood. Thus, only anomaly regions with high opacities will be shown in WPM; lucent anomalies will be suppressed. In visualizing IPM and WPM images, pixels with lower likelihood will have a higher intensity (appear brighter) than pixels with high likelihood. Thus, a region that is marked brightly (highlighted) is a region that the algorithm considers to have anomaly features.

One of the IPM and WPM images' fundamental weaknesses is that it suppresses anatomical landmarks. Without anatomical landmarks, it is difficult to determine the location of an anomalous region relative to an organ. To solve this issue, we added IPM/WPM heatmap as a layer on top of the corresponding images. Though, only regions that exceed the Otsu threshold [25] are incorporated into the images. An example of IPM, WPM, Infused-IPM (IIPM) and Infused-WPM (IWPM) is shown in Fig. 5. For comparison purpose, Fig. 4 shows the output of conventional bottom-up saliency mapping algorithms called Fine-Grained and Spectral Residual. The implementation of both these algorithms is taken from the OpenCV.

3.2 Classification Method

This paper aims to test whether or not replacing CXR images with IPM and WPM will improve CNN's classification performance. Thus, to ensure any performance changes are due to the input type and not the CNN architectures, only familiar deep CNN models are used. Figure 3 shows the network architecture used, with the base model being MobileNet, DenseNet121, ResNet50, VGG19 and Xception [26-30]. The implementation of base model network architecture is taken from TensorFlow (Ver. 2) library and with the pre-trained weight from ImageNet [31].

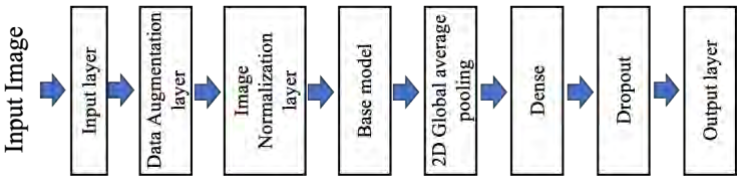


Fig. 3: Classification network architecture.

All models are trained with 100 epochs; however, an early stop is executed if there is no improvement in the loss value after ten epochs. All models were trained and tested on the Google Cloud platforms.

We have chosen the regular classification metric performance for model validation: precision, sensitivity, receiver operating characteristic curve (ROC-AUC), and the area under the precision versus sensitivity curve (PR-AUC).

3.3 Dataset

The dataset that is being used in this research is the Google-NIH dataset [32]. Initially, NIH provides the images while the labels are provided by Google [33]. It is important to note that only labels for the test and validation dataset are provided from the source. To create the training dataset for this study, we split the original validation dataset into a new training and validation dataset with a ratio of 0.3. All datasets are imbalanced datasets.

3.4 Qualitative Validations

To test the clinical relevance of the proposed algorithm, several normal and anomalous CXR images were selected, and the resulting IPM and WPM were examined qualitatively by a certified radiologist. Anomalous CXR that was chosen includes images of a rotated film, foreign body, cardiomegaly, and masses. For model interpretation, the Grad-CAM [18] algorithm is used to visualize which region of the CXR is relevant to the model when making the class classification.

4. RESULTS AND DISCUSSION

4.1 Classification Result

Table 1 shows several classification metrics obtained by various models and input-data types to classify the test dataset for air space opacity. While Table 2 shows similar metrics for the classification of CXR images with masses/nodules. Entries with the highest score for a particular metric is bolded.

The result obtained is not particularly easy to decipher. The highest score in PR-AUC, accuracy, and precision is obtained by ResNet50+Image, ResNet50+IIPM, and ResNet50+IWMP, respectively. Xception+Image and VGG19+Image do have a higher precision score, however, both results were rejected due to their sensitivity score being less than 0.5. This means the models falsely label the majority of positive samples. The model with the highest sensitivity score is VGG199+WPM, with a score of 0.930. However, the model precision is quite low, only 0.672, thus the next model, DenseNet121+IWPM, will be a better choice, having obtained 0.893 in sensitivity and 0.775 in precision. Meanwhile, DenseNet121+Image obtained the highest ROC-AUC score, 0.877.

https://doi.org/10.31436/iiumej.v22i2.1752

Table 1: Classification performance of CNN models with different base models and input data type for the airspace opacity dataset. ROC, PR, ACC, Pre, Sen stand for ROC-AUC, Precision-Recall area under the curve, accuracy, precision, and sensitivity respectively. %Δ represents the percentage difference in score compared to the model with image as input

Input Type	ROC-AUC	%Δ	PR-AUC	%Δ	Accuracy	%Δ	Precision	%Δ	Sensitivity	%Δ
MobileNetV2										
Image	0.844		0.883		0.770		0.814		0.781	
IPM	0.824	-2.4	0.865	-2.1	0.754	-2.1	0.787	-3.3	0.789	1.0
WPM	0.801	-5.2	0.835	-5.4	0.670	-13.0	0.851	4.4	0.522	-33.2
IIPM	0.846	0.2	0.879	-0.5	0.767	-0.5	0.821	0.8	0.763	-2.3
IWPM	0.841	-0.3	0.873	-1.2	0.772	0.3	0.785	-3.6	0.834	6.9
DenseNet121										
Image	**0.877**		0.903		0.776		0.878		0.712	
IPM	0.836	-4.7	0.874	-3.3	0.711	-8.4	0.878	-0.1	0.581	-18.3
WPM	0.820	-6.5	0.860	-4.8	0.736	-5.2	0.795	-9.5	0.733	3.0
IIPM	0.873	-0.4	0.905	0.2	0.763	-1.7	0.877	-0.1	0.686	-3.6
IWPM	0.874	-0.3	0.904	0.1	0.788	1.5	0.775	-11.8	0.893	25.4
Xception										
Image	0.856		0.890		0.683		0.921		0.495	
IPM	0.809	-5.6	0.851	-4.4	0.736	7.8	0.757	-17.9	0.803	62.1
WPM	0.797	-6.9	0.836	-6.1	0.729	6.7	0.797	-13.5	0.714	44.1
IIPM	0.835	-2.5	0.871	-2.1	0.732	7.2	0.856	-7.1	0.646	30.4
IWPM	0.841	-1.8	0.874	-1.8	0.723	5.7	0.876	-4.9	0.606	22.4
ResNet50										
Image	0.873		**0.907**		0.784		0.881		0.725	
IPM	0.838	-4.0	0.879	-3.1	0.757	-3.5	0.831	-5.7	0.728	0.4
WPM	0.829	-5.0	0.874	-3.7	0.759	-3.2	0.802	-9.0	0.775	6.9
IIPM	0.870	-0.4	0.906	-0.2	**0.790**	0.6	0.872	-1.0	0.745	2.8
IWPM	0.867	-0.6	0.905	-0.3	0.743	-5.3	**0.904**	2.6	0.621	-14.3
VGG19										
Image	0.840		0.876		0.654		0.910		0.447	
IPM	0.805	-4.2	0.860	-1.8	0.761	16.4	0.769	-15.5	0.841	88.2
WPM	0.796	-5.3	0.843	-3.8	0.697	6.5	0.672	-26.2	**0.930**	108.3
IIPM	0.858	2.2	0.894	2.1	0.760	16.1	0.875	-3.9	0.683	52.9
IWPM	0.845	0.6	0.887	1.3	0.760	16.1	0.852	-6.3	0.707	58.4

Next, we analyze if using the proposed anomalous saliency mapping as input will result in a better classifier for the airspace opacity dataset. We are particularly interested if such change in input can boost the performance of shallower CNN models (MobileNetV2 and DenseNet121) to comparable performance of deeper CNN models (ResNet50, VGG19 and Xception). What is evident from the result, using the alternative data types as input enhances the model's sensitivity. For example, VGG19+WPM, which obtained the highest sensitivity, obtained a 108.3% improvement compared to VGG19+IMG. This sensitivity improvement is more apparent in deep CNN models (ResNet50, VGG19, and Xception) than shallower CNN models (MobileNetV2 and DenseNet121).

https://doi.org/10.31436/iiumej.v22i2.1752

As one might expect, any improvement in sensitivity tends to reduce model precision. Nevertheless, in most results, the degree of precision reduction is less than the degree of sensitivity gain. For example, the model Xception+IIPM obtained an increase of 30.4% in sensitivity while only reducing its precision by 7.1% compared to Xception+Image.

The answer to which CNN model and input data type perform the best, depends on the purpose of the model. For screening purposes, then DenseNet121+IWPM will be the recommended model as it obtained the second-best sensitivity score while maintaining a reliable precision score. For precise clinical classification, then ResNet50+IWMP is recommended.

Table 2: Classification performance of CNN models with different base models and input data type for the mass/nodule dataset. ROC, PR, ACC, Pre, Sen stand for ROC-AUC, Precision-Recall area under the curve, accuracy, precision, and sensitivity respectively. %Δ represents the percentage difference in score compared to the model with image as input

Input Type	ROC-AUC	%Δ	PR-AUC	%Δ	Accuracy	%Δ	Precision	%Δ	Sensitivity	%Δ
MobileNetV2										
Image	0.588		0.194		0.709		0.209		0.336	
IPM	0.568	-3.4	0.188	-3.2	0.729	2.8	0.201	-4.2	0.268	-20.2
WPM	0.588	0.0	0.204	5.2	0.755	6.5	0.206	-1.4	0.220	-34.3
IIPM	0.577	-1.7	0.182	-6.4	0.805	13.4	0.194	-7.1	0.095	-71.7
IWPM	0.573	-2.4	0.192	-1.0	0.659	-7.2	0.188	-10.2	0.383	14.1
DenseNet121										
Image	0.606		0.199		0.739		0.228		0.308	
IPM	0.587	-3.1	0.198	-0.6	0.716	-3.1	0.205	-10.1	0.308	0.0
WPM	0.607	0.3	0.209	4.9	0.541	-26.8	0.197	-13.8	**0.664**	115.4
IIPM	0.619	2.2	0.215	7.9	0.669	-9.4	0.222	-2.8	0.478	54.9
IWPM	0.617	1.9	0.216	8.2	0.731	-1.0	0.241	5.7	0.366	18.7
Xception										
Image	0.637		0.231		0.702		0.227		0.410	
IPM	0.587	-7.8	0.197	-14.7	0.633	-9.8	0.196	-13.8	0.464	13.2
WPM	0.575	-9.8	0.200	-13.5	0.611	-13.0	0.177	-22.0	0.437	6.6
IIPM	0.602	-5.5	0.206	-10.6	0.737	5.0	0.213	-6.4	0.278	-32.2
IWPM	0.609	-4.5	0.210	-9.2	0.570	-18.7	0.192	-15.5	0.580	41.3
ResNet50										
Image	0.643		0.242		0.689		0.225		0.437	
IPM	**0.644**	0.2	0.268	11.0	0.631	-8.4	0.222	-1.3	0.580	32.6
WPM	0.619	-3.7	0.253	4.5	0.556	-19.3	0.199	-11.4	0.647	48.1
IIPM	0.614	-4.5	0.210	-13.0	0.745	8.1	0.228	1.2	0.292	-33.3
IWPM	0.612	-4.7	0.222	-7.9	**0.820**	19.0	**0.299**	33.2	0.149	-65.9
VGG19										
Image	0.619		0.257		0.741		0.230		0.308	
IPM	0.623	0.6	**0.279**	8.5	0.545	-26.4	0.195	-15.3	0.647	109.9
WPM	0.561	-9.3	0.194	-24.6	0.733	-1.1	0.206	-10.7	0.271	-12.1
IIPM	0.620	0.1	0.245	-4.8	0.743	0.2	0.244	5.9	0.339	9.9
IWPM	0.628	1.4	0.260	1.0	0.752	1.5	0.248	7.7	0.319	3.3

It is worth noting that, since this dataset is imbalanced, the PR-AUC score is more important than the ROC-AUC score. DenseNet121+IWPM also obtains a PR-AUC score of 0.904, a

mere 0.003 less than the highest score, which is 0.907 by ResNet50+Image. In addition to obtaining a reliable classification score, it also has the advantage of requiring fewer computing resources than ResNet50, VGG19, and Xception. Thus, it is more practical to be deployed in Malaysian hospitals.

In Table 2, results show that all models failed to achieve acceptable classification performance for the mass/nodule testing-dataset. No model obtained a precision score of more than 0.5, meaning the majority of positive classifications were actually false. It is worth pointing out that all sensitivity scores for IMG input were lower than 0.5, thus all model missed the majority of the mass/nodule samples. Only DenseNet121+WPM, ResNet50+ WPM/IPM and VGG19+IPM manage to achieve a sensitivity score above 0.5.

4.2 Qualitative Assessment

Figure 4 shows the example of a saliency map produced by the Fine Grained and Spectral Residual algorithm, a conventional bottom-up algorithm. As shown in the figure, the Fine Grained failed to emphasize or suppress any feature in the image. Conversely, the Spectral Residual suppressed almost all featured, making it impossible to extract any meaningful information from its saliency map. Aligned with what was mentioned in [9], the conventional bottom-up saliency map algorithm cannot produce meaningful mapping for CXR images.

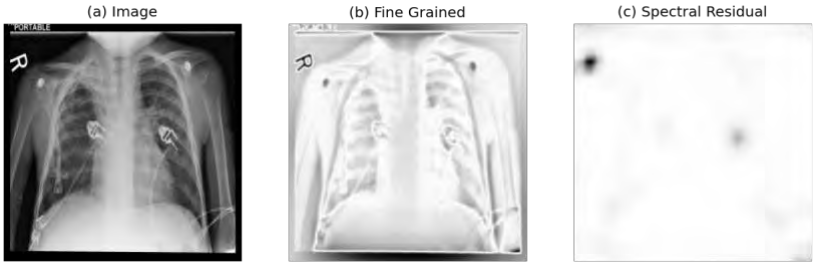


Fig. 4: The saliency map of a CXR image (a) produced by the Fine Grained, (b) and Spectral Residual (c) algorithms.

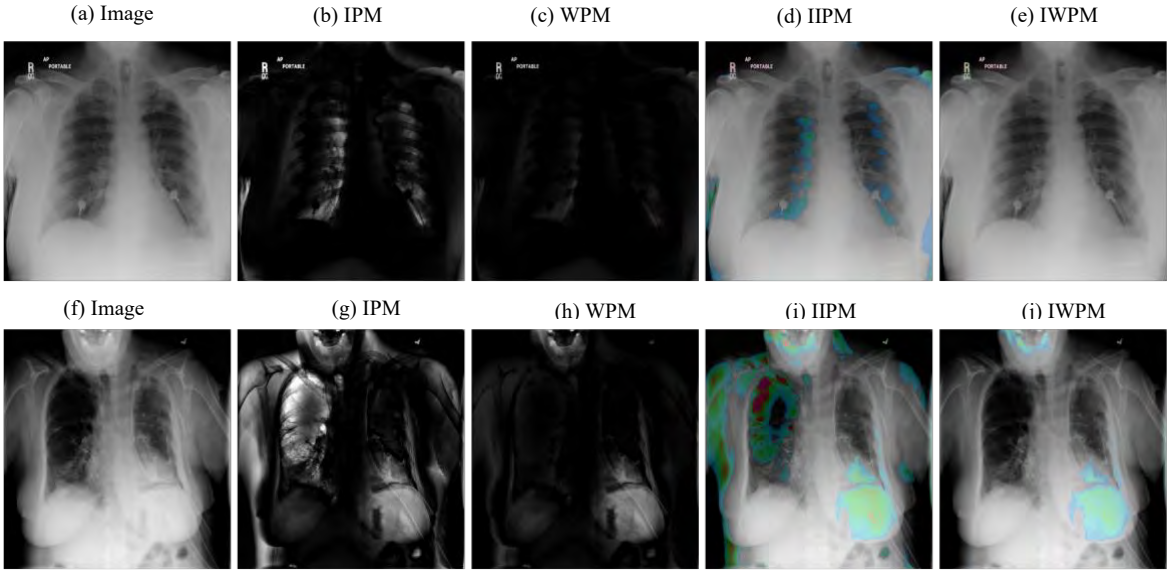


Fig. 5: A comparison between the original image and the corresponding IPM, WPM, IIPM and IWPM for a patient in a normal (a-e) and rotated (f-j) position.

Figure 5(a) shows CXR images with no visible abnormalities; the resulting IPM and WPM are shown the Fig. 5(b) and 5(c), respectively. There is no apparent region highlighted for the WPM image, implying that the algorithm does not identify any anomaly in the original CXR image. However, the perihilar region is incorrectly highlighted in the IPM image; this may suggest that the IPM may be over-sensitive in highlighting anomalies in CXR images.

Next, we examine how the algorithm processed CXR images taken for an incorrectly positioned patient, or for a rotated x-ray film. For example, in Fig. 5(d), the patient's trachea is not located in the midline, suggesting that the patient may be rotated relative to the film. This orientation gives the appearance that the right lung is more lucent than the left. In the produced IPM image, the lucent region is highlighted, whereas WPM does not highlight this feature as WPM suppresses lucent anomalies. Whether or not the cause of the right lung lucency is significant is still an anomaly from the imaging perspective. Thus, the algorithm should highlight this anomaly as in the IPM image and then proceed to validate it by a radiologist.

Another feature that indicates that the patient is in an abnormal position is the presence of teeth in the CXR image. The anomaly is highlighted in the IPM and more evidently in the WPM image. Teeth usually are not presented in a CXR, thus it is a form of anomaly that should be highlighted by a UAS algorithm. However, the algorithm incorrectly highlighted the patient breasts; this error may have been caused by the lack of images containing breasts in the control dataset.

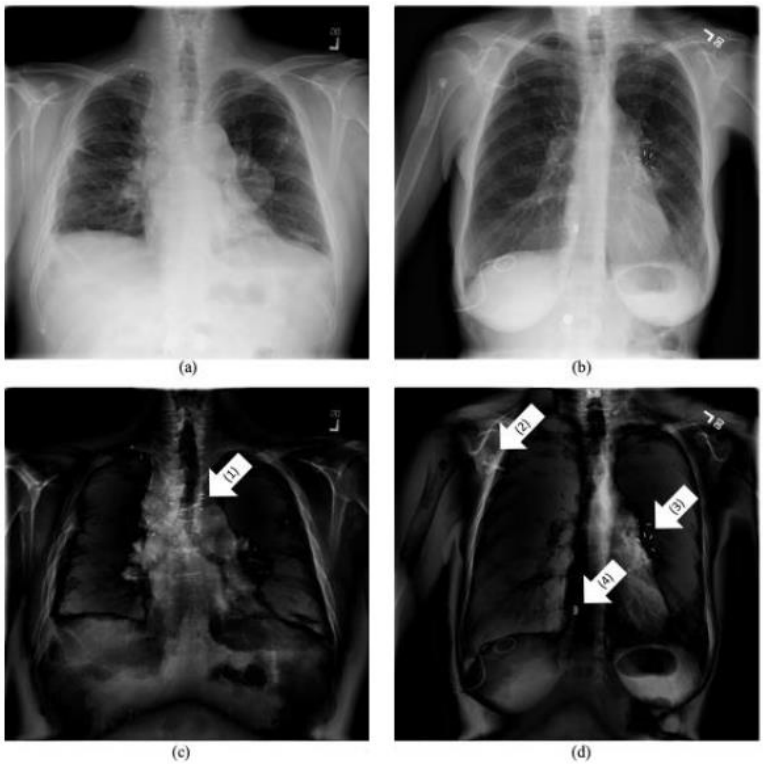


Fig. 6: Chest radiograph with foreign body, (a) and (b).
The resulting WPM images (c) and (d) clearly show the foreign bodies, arrow (1)-(4).

Figure 6(a) and 6(b) show CXR with foreign bodies, and the resulting WPM images are shown in Fig. 6(c) and (d). In both WPM images, the foreign bodies (arrows) are highlighted clearly and are more apparent than in the original CXR images. The opacity of biological

matter around the foreign body is suppressed, making it appear lucent. With this result, it can be assumed that WPM images may help in identifying foreign objects in CXR.

Next, the algorithm capabilities in highlighting pathological changes are demonstrated. Figure 7(a) shows a CXR with cardiomegaly, and it is highlighted clearly in both IPM, Fig. 7(b) and WPM Fig. 7(c) image. On the other hand, Fig. 7(d) shows CXR with a homogenous opacity at the right lower lung zone that does not obscure the cardiac border. This feature is not highlighted in both IPM, Fig. 7(e) and WPM, Fig. 7(f) images.

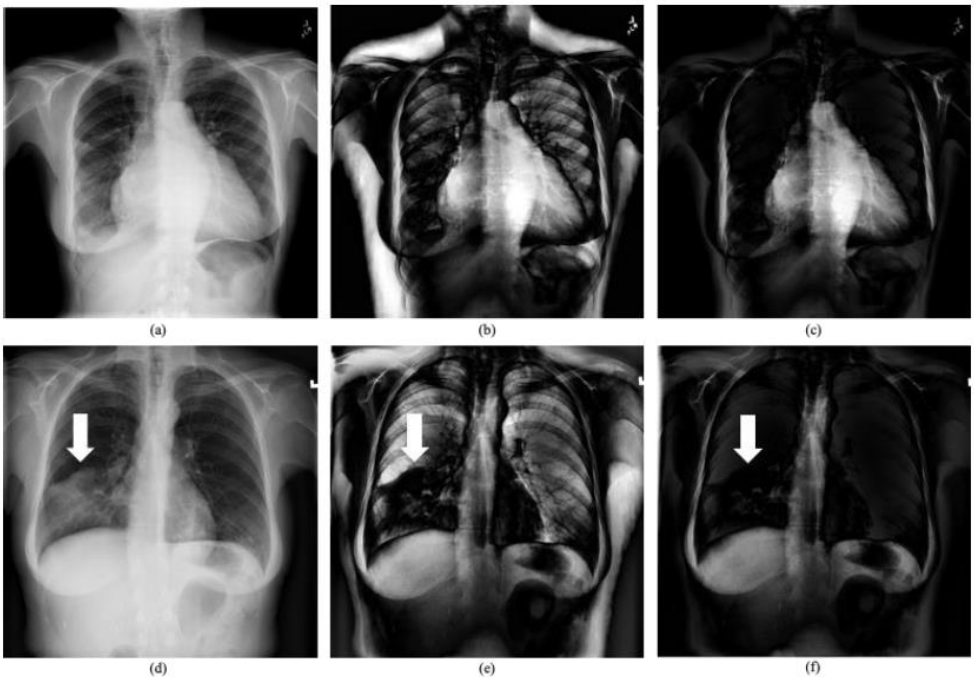


Fig. 7: The cardiomegaly feature in (a) is clearly highlighted in the WPM (b) and IPM (c) image. However the generalized opacity feature in (d) is absent in the IPM (e) and WPM (f) images.

Additionally, the algorithm also frequently failed to highlight opacity due to mass and nodules. The single nodule in Fig. 8(a) was not highlighted in the resulting WPM image, Fig. 8(d). The same can also be said for the multiple nodule-like opacities at bilateral mid and lower lung zones, as shown in Fig. 8(b). Only some of the lung masses are highlighted in resulting WPM images, Fig. 8(e). An example of a correctly highlighted lung mass is shown in Fig. 8(c) and 8(f).

4.3 Grad-CAM Results

To meaningfully deploy a developed model in clinical use, it must show some degree of interpretability and the classification must be validated based on some biological markers. For this reason, we use the Grad-CAM, [18], to visualize which region on the input data is emphasized. Figure 9(a) shows an example of CXR having airspace opacifications. Figure 9(b)-(f) shows the output of the Grad-CAM algorithm for DenseNet121 models that were trained with different input data types. The only models that were trained using the WPM and IIPM as input were correctly labeled the sample.

https://doi.org/10.31436/iiumej.v22i2.1752

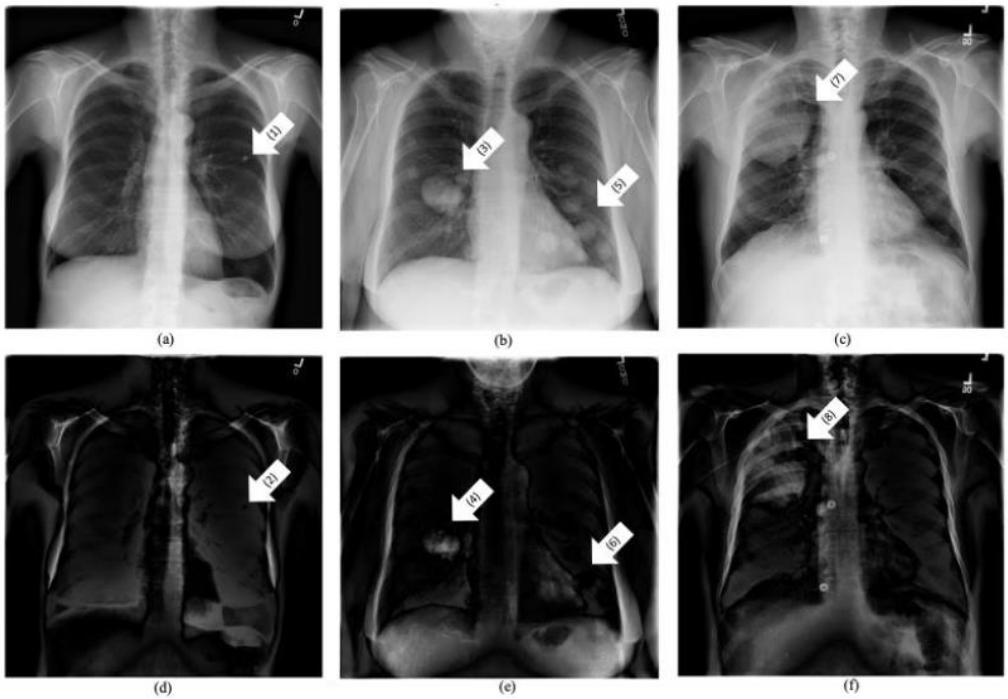


Fig. 8: Most of the lung nodules and masses in (a) and (b) failed to be highlighted in the resulting WPM images, (d) and (e). Only lung masses in (c) were successfully highlighted in (f).

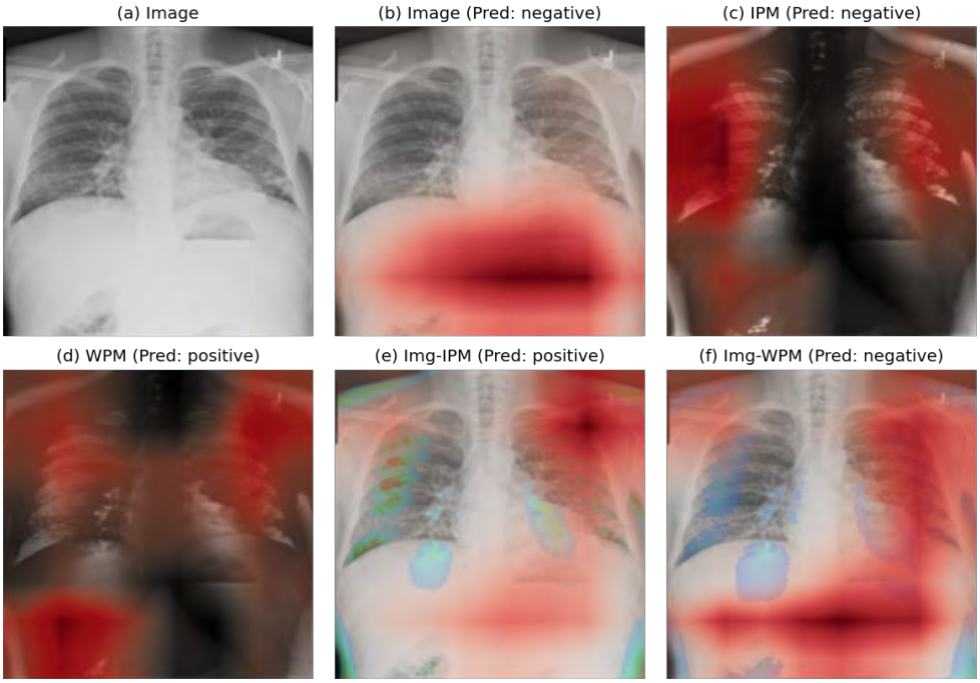


Fig. 9: Output of the Grad-CAM algorithm for air space opacification dataset. Regions that are marked red are regions that the DenseNet-121 deems important.

An obvious pattern that emerges in Fig. 9 is that models that received the original CXR image (image, IIPM, and IWPM) as input tend to mark the lower-left diagram. On the other hand, the model that takes IPM and WPM tends to focus more on the lung and shoulder region. It is not exactly certain why the DenseNet121-IPM model incorrectly labeled the

image even though it correctly emphasized the lung region. One reason that can be attributed is the model emphasized the right lung more than the left lung. Both models that correctly labeled the image, DenseNet121-WPM, and DenseNet121-IIPM, emphasize the left lung region. Even though DenseNet121-WPM also marked the left lung region; it emphasized more on the diagram, hence the wrong labeling.

Figure 10(a) shows a sample of CXR that is positive for mass. The location of the mass is in the left upper and lower lobe. Figure 10(b)-(f) show the output result of the Grad-CAM algorithm for different ResNet50 models that were trained by different input data types. Models trained using images, IIPM and IWPM, show similar marked regions; they extend from the left clavicle bone to the lung right middle lobe. For IIPM and IWPM, the region does not cover any mass. Thus, the model is falsely labeled as negative. ResNet50+IPM correctly marked the left-upper lobe, and the mass contained in it. ResNet50+WPM only weakly marked this region. No model correctly marked the mass at the left lower lobe. From the example of results shown in Fig. 10, it can be concluded that the trained model failed to learn a mass feature. It also emphasizes the need for more effective feature extraction if we want to detect masses in CXR more accurately.

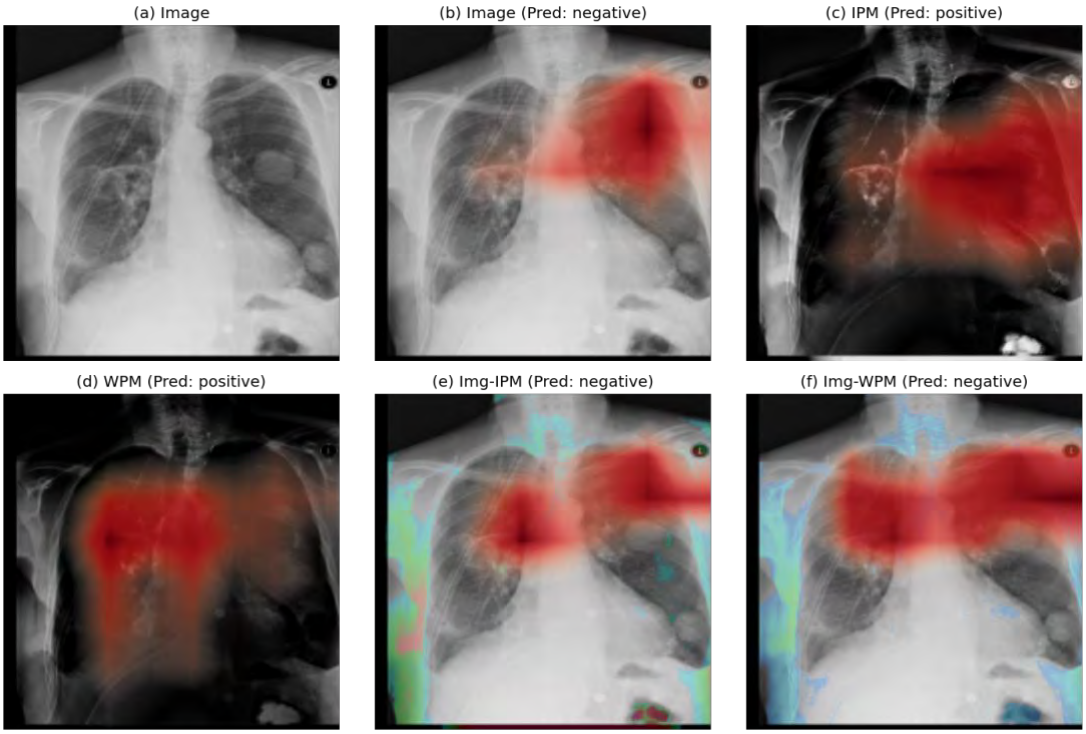


Fig. 10: Output of the Grad-CAM algorithm for the mass/nodule dataset. Regions that are marked red are regions that the ResNet50 deems important.

5. CONCLUSION

In this paper, we introduce the Xenafas algorithm, which creates the IMP and WMP anomalous saliency mapping for CXR images. A qualitative study by a certified radiologist has shown that the algorithm can highlight most foreign objects and cardiomegaly in the CXR samples tested; however, it is inconsistent in highlighting masses and nodules. It has also been shown that using IMP and WMP over regular CXR images increases the sensitivity of most CNN models that were tested. Using the Grad-CAM algorithm, it has been demonstrated that by using the IMP and WMP, the CNN model shifted its focus to a

more relevant CXR image region. The results obtained from the experiment conducted show that the IMP and WMP can be an alternative to regular CXR images for future machine learning development.

ACKNOWLEDGEMENT

This work was supported by the Malaysian Ministry of Higher Education Fundamental Research Grant Scheme [grant no. FRGS17-040-0606].

REFERENCES

[1] Itti L, Koch C, Niebur E. (1998) A model of saliency-based visual attention for rapid scene analysis. IEEE Trans Pattern Anal Mach Intell. 20: 1254-1259.
[2] Liu N. and Han J. (2016) Dhsnet: Deep hierarchical saliency network for salient object detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 678-686.
[3] Rahtu E, Kannala J, Salo M, Heikkilä J. (2010) Segmenting salient objects from images and videos. In European Conference on Computer Vision; pp 366-379.
[4] Itti L, Koch C. (2001). Computational modelling of visual attention. Nature Reviews Neuroscience, 2: 194-203.
[5] Cheng M-M, Zhang Z, Lin W-Y, Torr P. (2014) BING: Binarized normed gradients for objectness estimation at 300fps. In Proceedings of the IEEE conference on computer vision and pattern recognition; pp 3286-3293.
[6] Montabone S, Soto A. (2010) Human detection using a mobile platform and novel features derived from a visual saliency mechanism. Image Vis Comput., 28: 391-402
[7] Hou X, Zhang L. (2007) Saliency detection: A spectral residual approach. In IEEE Conference on Computer vision and Pattern Recognition; pp 1-8.
[8] Borji A, Cheng M-M, Hou Q, Jiang H, Li J. (2019) Salient object detection: A survey. Comput Vis Media; pp 1-34
[9] Castillo JC, Tong Y, Zhao J, Zhu F RSNA Bone-age detection using transfer learning and attention mapping [http://noiselab.ucsd.edu/ECE228_2018/Reports/Report6.pdf]
[10] Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz CP, others. (2018) Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med., 15: e1002686
[11] Zhou B, Khosla A, Lapedriza A, Oliva A, Torralba A. (2016) Learning deep features for discriminative localization. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 2921o2929.
[12] Shen L, Margolies LR, Rothstein JH, Fluder E, McBride R, Sieh W. (2019) Deep learning to improve breast cancer detection on screening mammography. Sci. Rep., 9: 12495.
[13] Springenberg JT, Dosovitskiy A, Brox T, Riedmiller M. (2015) Striving for simplicity: The all convolutional, arXiv preprint arXiv:1412.6806
[14] Ding Y, Sohn JH, Kawczynski MG, Trivedi H, Harnish R, Jenkins NW, Lituiev D, Copeland TP, Aboian MS, Mari Aparici C, others. (2019) A deep learning model to predict a diagnosis of Alzheimer disease by using 18F-FDG PET of the brain. Radiology, 290: 456o464
[15] Norman B, Pedoia V, Noworolski A, Link TM, Majumdar S. (2019) Applying densely connected convolutional neural networks for staging osteoarthritis severity from plain radiographs. J. Digit. Imaging, 32: 471-477
[16] Oh K, Kim W, Shen G, Piao Y, Kang N-I, Oh I-S, Chung YC. (2019) Classification of schizophrenia and normal controls using 3D convolutional neural network and outcome visualization. Schizophr Res., 212: 186-195
[17] Arun NT, Gaw N, Singh P, Chang K, Hoebel KV, Patel J, Gidwani M, Kalpathy-Cramer J. (2020) Assessing the validity of saliency maps for abnormality localization in medical imaging, arXiv preprint arXiv:200600063 Cs

[18] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. (2017) Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision; pp 618-626.
[19] Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, Rueckert D. (2019) Attention gated networks: Learning to leverage salient regions in medical images, arXiv preprint arXiv:180808114 Cs
[20] Pesce E, Withey SJ, Ypsilantis P-P, Bakewell R, Goh V, Montana G. (2019) Learning to detect chest radiographs containing pulmonary lesions using visual attention networks. Med Image Anal, 53: 26-38
[21] Deeba F, Bui FM, Wahid KA. (2020) Computer-aided polyp detection based on image enhancement and saliency-based selection. Biomed Signal Process Control, 55: 101530.
[22] Fan H, Xie F, Li Y, Jiang Z, Liu J. (2017) Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold. Comput Biol Med, 85: 75-85
[23] Khan MA, Akram T, Sharif M, Saba T, Javed K, Lali IU, Tanik UJ, Rehman A. (2019) Construction of saliency map and hybrid set of features for efficient segmentation and classification of skin lesion. Microsc Res Tech., 82: 741-763.
[24] Rahmat T, Ismail A, Aliman S. (2018) Chest x-rays image classification in medical image analysis. Appl Med Inform., 40: 63-73.
[25] Otsu N. (1979) A threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern, 9: 62-66.
[26] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. (2017) Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:170404861
[27] Huang G, Liu Z, Weinberger KQ. (2016) Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 4700-4708
[28] He K, Zhang X, Ren S, Sun J. (2016) Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 770-778
[29] Simonyan K, Zisserman A. (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:14091556
[30] Chollet F. (2016) Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 1251-1258
[31] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. (2009) Imagenet: A large-scale hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition; pp 248-255.
[32] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. (2017) Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; pp 2097-2106
[33] Majkowska A, Mittal S, Steiner DF, Reicher JJ, McKinney SM, Duggan GE, Eswaran K, Cameron Chen P-H, Liu Y, Kalidindi SR, et al. (2020) Chest radiograph interpretation with deep learning models: assessment with radiologist-adjudicated reference standards and population-adjusted evaluation. Radiology, 294: 421-431

https://doi.org/10.31436/iiumej.v22i2.1507

SURFACE ROUGHNESS PREDICTION IN TURNING PROCESS BY APPLYING COMPUTER VISION METHOD

OMER WATHIQ TAHA[1] AND OSAMAH FADHIL ABDULATEEF[2*]

[1]Al-Mustafa University College, Baghdad, Iraq.
[2]Department of Automated Manufacturing Engineering, University of Baghdad, Baghdad, Iraq.

[*]Corresponding author: drosamah@kecbu.uobaghdad.edu.iq

(Received: 12th June 2020; Accepted: 17th January 2021; Published on-line: 4th July 2021)

ABSTRACT: This paper reports the utilization of computer vision and backlight techniques to determine the surface roughness of a workpiece under a variety of process parameters. A CCD (Charge-Coupled Device) camera was used to capture the image of the edge of the workpiece of the turned components using backlight technology to provide an edge roughness profile. The image was processed using SRVISION software developed in MATLAB to extract the profile of the workpiece and calculated the arithmetic average value of roughness (Ra) and root mean square roughness (Rq). The experiments are carried out with AISI 1045 (medium carbon steel), using various feed rates and cutting speeds, comparison is then made of the surface roughness values achieved through the conventional stylus probe method and the image processing technique. The comparison indicates that the vision method provides precise and consistent results with a correlation up to 0.99 with the traditional stylus method. The mean variations in Ra and Rq between the two methods were just 1.65 and 1.433 percent, respectively. As the vision method is a non-contact procedure, it can be significant potential for application without damaging the machined surfaces in the in-process inspection of the components as well as aids monitoring of the components in a shorter period.

ABSTRAK: Kajian ini menggunakan visual komputer dan teknik cahaya belakang bagi memperoleh kekasaran permukaan sesuatu bahan pada pelbagai proses parameter. Kamera jenis CCD (Peranti Terganding-Cas) telah digunakan bagi memperoleh imej tepi bagi komponen yang dipusing menggunakan teknologi cahaya belakang bagi menghasilkan profil imej tepi yang jelas. Imej ini diproses menggunakan perisian SRVISION MATLAB bagi menghasilkan profil bahan dan purata kiraan kekasaran permukaan (Ra) dan punca purata kuasa dua kekasaran permukaan (Rq). Eksperimen dijalankan menggunakan AISI 1045 (besi karbon pertengahan), menggunakan pelbagai kadar suapan dan kelajuan potongan. Perbandingan kemudian dibuat pada nilai kekasaran permukaan yang diperoleh melalui kaedah prob jarum stilus konvensional dan melalui teknik pemprosesan imej. Perbandingan menunjukkan kaedah visual memberikan ketepatan dan dapatan konsisten yang munasabah dengan korelasi sehingga 0.99 dengan kaedah prob jarum stilus tradisi. Purata variasi pada nilai Ra dan Rq antara dua kaedah adalah sebanyak 1.65 dan 1.433 peratus, masing-masing. Adapun kaedah visual adalah prosedur tanpa-sentuh, ianya sesuai dijalankan tanpa merosakkan permukaan mesin dalam proses penilaian komponen, juga membantu mengawasi komponen dalam waktu singkat.

KEYWORDS: *image process; roughness measurement; stylus method; non-contact method*

https://doi.org/10.31436/iiumej.v22i2.1507

1. INTRODUCTION

Turning is a common machining process that removes material from a rotating cylindrical component using a single-point cutting tool. The turned component has a certain surface roughness that acts as a significant parameter in the performance of its work, much like friction, wear, lubrication, electrical and thermal resistance, fluid dynamics, vibration, and noise. Different parameters such as feed rate, cutting speed, cutting depth, cutting tool configuration, machine tool, and material of the component all affect the performance of the required product feature and surface roughness values at an appropriate cost. Roughness may be evaluated using two basic methods: contact and non-contact methods. The contact method utilizes a stylus, which is drawn across the measured surface. The surface waveform is collected through an electronic sensor, commonly a linear differential variable transformer, that calculates parameters of surface roughness, like root mean square roughness Rq, average roughness Ra, maximum peak-to-valley height Rt, etc. The main disadvantages of the stylus device are that: (1) it requires direct physical contact, (2) it limits the measuring speed, (3) it cannot be used as an online measurement because the workpiece needs to be withdrawn from the machine for monitoring, and (4) it has restricted versatility in handling the specific geometric component to be measured [1].

Non-contact methods may be divided into many categories based on the lighting system used and image analysis. Several investigations were performed utilizing non-contact vision methods for the surface roughness assessment. Lee et al. [1] employed computer vision techniques to predict a workpiece's surface roughness under many cutting operations. The workpiece surface image was first acquired with a digital camera, and then the surface image feature was extracted. A polynomial grid was implemented utilizing a self-organizing adaptive modeling method to create relations between the surface image characteristics and real roughness of the surface through various turning operations. Gadelmawla [2] implemented a vision system to capture images for surfaces to be characterized and software was designed to investigate the captured images based on the "Gray Level Co-occurrence Matrix (GLCM)". 3D plots of the GLCMs for different captured images were implemented, compared, and discussed. Also, several statistical parameters were calculated from the GLCMs and compared with the arithmetic average roughness, Ra.

Al-Kindi et al. [3] developed a technique for using computer vision data to achieve accurate measurement of surface roughness parameters. Stylus-based measurements were obtained utilizing standard and non-standard roughness parameters and compared to vision-based measurements. Two light reflection models were adopted and implemented, namely the "Intensity-Topography Compatible (ITC) model" and the "Light-Diffuse model", to explain the obtained vision data and to allow appropriate roughness parameter calculation. Results revealed that the "ITC model" performed better than the "Light-Diffuse model", with notable similar values to those obtained by conventional stylus-based data of roughness parameters. Zhongxiang et al. [4] employed a method for determining the three-dimensional roughness of the surface using profile information. They suggested a three-dimensional measuring technique that was used to investigate surface roughness components on the basis of the digital image processing technology, and set up a three-dimensional surface roughness assessment system containing hardware and software architecture. Fadare et al. [5] developed a computer vision system appropriate for on-line surface roughness measurement of machined components utilizing an "artificial neural network (ANN)" depending on a digital image processing of the machined surface, consisting of a CCD camera, computer, Microsoft Windows Video

Maker, digital image processing software, and two light sources. The machined surface images were captured; analyzed and optical roughness characteristics were assessed using the "2-D fast Fourier transform (FFT) algorithm". They concluded that the optical roughness values predicted by ANN were considered to be in good agreement with the measured values (R^2-value = 0.9529).

Shahabi et al. [6] proposed a different method for measuring roughness using a 2-D contour extracted from an edge image of the workpiece surface. A comparison with a stylus type device indicated a maximum variation of 10% in the measurement of average roughness Ra utilizing the visual method. Sridhar et al. [7] used a machine vision method to determine surface roughness through image processing and backlight technique on the turned components. The comparison was then made of the surface roughness values achieved through the image processing technique and the conventional stylus method, which showed that the suggested method provided close and dependable results similar to the traditional stylus method. Balasundaram et al. [8] calculated the amplitude and spacing, in addition to functional surface roughness parameters through the dry cutting of AISI 1035 carbon steel utilizing machine vision. A "DSLR camera" with high shutter speed was employed to capture a blur-free image of the workpiece surface profile perpendicular to the cutting tool. The edge of the surface profile was identified to sub-pixel precision using the grey level constant moment and the roughness parameters were calculated using the profile. Srivani et al. [9] presented a methodology to characterize the nature of the surface using a computer vision system. For further investigations, a computerized optical microscope was used to collect surface images, and those images were fed into MATLAB software.

Qingqun et al. [10] suggested a different method of on-line turned surface inspection by observing the characteristics of the grey value of the surface digital image. The uniformity of the surface image was evaluated and analyzed by fractal analysis, wavelet transform, and discreteness analysis of the wavelength of the texture profile. The normal texture image was extracted from the average wave profile, which indicated the state of the process and turned surface conditions. The results indicated that the turned surface condition could be effectively checked on-line. Naresh et al. [11] used the technique of machine vision to observe the surface roughness when turning composite MMCs. The machining surfaces were identified during machining operation utilizing machine vision technology and the stylus probe instrument was used to measure the surface finish of the machined surfaces. Patel et al. [12] introduced a computer vision system that captured the surface texture contours of the machined surfaces and extracted images. Using the gray-level co-occurrence matrix, the texture function parameters were extracted and compared to various surface roughness parameters reported from a surface profilometer of a contact form. The image analysis was carried out for the extraction of texture characteristics at various levels of roughness. The variation between the characteristics of each texture and the parameter of surface roughness was examined. Multiple regression models were expanded to estimate individual surface roughness parameter (Ra) estimation and good recognition of surface roughness degree. The linear detection model was found to have better output features compared with a nonlinear recognition model. The findings showed that surface roughness estimation utilizing a linear regression model was a robust method for non-contact measurement. Patel et al. [13] presented a surface roughness prediction approach utilizing "Computer Vision", "Image Processing", and "Machine Learning". Two machine learning algorithms, "Stochastic Gradient Boosting" and "Bagging Tree" were compared and assessed on the basis of statistical parameters. It was found that "Stochastic Gradient Boosting" effectively estimated surface roughness for training as

well as Ten-fold cross-validation. The methods may be utilized for online monitoring of machined components and good evaluation.

In this paper, a computer vision system for tracking and predicting the surface roughness of the turned components with different cutting conditions (cutting speed, feed rate, and cutting depth) utilizing image processing and backlight technique is presented. The surface roughness values that will obtained by the image processing technique and the conventional stylus method will then be compared.

2. METHODOLOGY AND EXPERIMENTATION

The average surface roughness (Ra) and root mean square roughness (Rq) are commonly used as index of measurements to assess a machined surface finish. Estimation of roughness parameters has a significant role in distinguishing difficulties in industrial sectors like contact deformation, friction, and tightness of joint contact precision.

2.1 Stylus Method Description

The machining process was performed on a WILTON lathe (model no. 52TL1440-3) by 18 medium carbon steel AISI 1045 workpieces having a 30 mm diameter and a 300 mm length. The chemical composition and mechanical properties of AISI 1045 material were measured as shown in tables 1 and 2 respectively.

Table 1: chemical composition of AISI 1045 material

Component	c	Si	Mn	P	S	Cr	Fe
Wt %	0.324	0.236	0.578	0.002	0.028	0.103	Residual

Table 2: Mechanical properties of AISI 1045 material

Elastic Modulus (Gpa)	Tensile strength (MPa)	Hardness (HB)	Yield strength (MPa)
216	671	170	353

The experiments were conducted using the Taguchi method by changing working parameters like feed rate, cutting speed, and a fixed cutting depth. The direction of the workpiece rotation was fixed in a counterclockwise direction. No cooling was concerned throughout the turning process. Table 3 indicates the amount of the parameters during the turning cutting process. A stylus device was used as a contact method for measuring the surface roughness of machined components. It contained a diamond stylus probe that was moved perpendicularly to the direction of roughness, and a characteristic of surface roughness was recorded at the other end. Because of its advantages, it is the most widely used technique and generates an object's profile in a clear direction. Surface roughness measurements of 18 turned components were performed on the stylus roughness tester type (SRT-6210).

2.2 Computer Vision System Description

The fundamental components of the vision system designed to capture images of the surfaces to be inspected consist of two parts: hardware and software systems. The hardware system consists of four main items: (1) a Sony DSC-WX100 CCD digital camera with a resolution of 18.2 megapixels, (2) an LED illumination source, (3) a black tube of cardboard to prevent the effect of environmental light, and (4) a personal computer

(PC) with MATLAB program for image processing as software. The camera was fixed using a special frame designed to move horizontally and vertically to ensure that the view of the camera was always perpendicular to the surface of the workpiece and could scan any area that needed to be measured. A software system named "SRVISION" was developed using MATLAB software. It was developed to work on any Windows environment. The image of the surface that needed to be measured was opened by the software, and then the variation of the surface profile was plotted and the surface parameters were calculated. The actual and schematic configuration of the on-machine measurement system of roughness is shown in Figs. 1 and 2, respectively.

Table 3: Cutting condition values utilized in the experimental work [10]

Workpiece material	**Cutting speed (V) [rpm]**	**Feed rate (f) [mm/sec]**	**Cutting depth (d) [mm]**
AISI 1045	140	0.2, 0.29, 0.39,0.77	**0.25**
	250	0.18, 0.36, 0.74	
	650	0.18, 0.26, 0.36, 0.72	
	950	0.16, 0.22, 0.34, 0.54	
	1350	0.15, 0.23, 0.34	

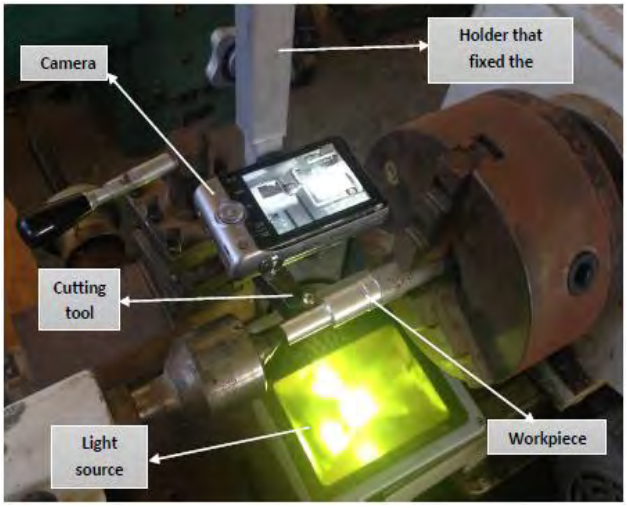


Fig. 1: The actual setup of the on-machine roughness measurement system.

2.3. System Calibration

The horizontal and vertical scaling factors were obtained using a standard block with a length of 2 mm to transform the image dimensions from pixels to real dimensions in microns; the block was located at the same level as the shaft. The block width (in pixels) was calculated using the camera calibration toolbox in MATLAB software and the calibration factors were calculated using the following equation [2]:

$$f = \frac{no.of\ the\ pixels\ for\ gauge\ block\ width\ (length)}{actual\ (standard)\ width\ (length)\ of\ the\ block} \quad (1)$$

https://doi.org/10.31436/iiumej.v22i2.1507

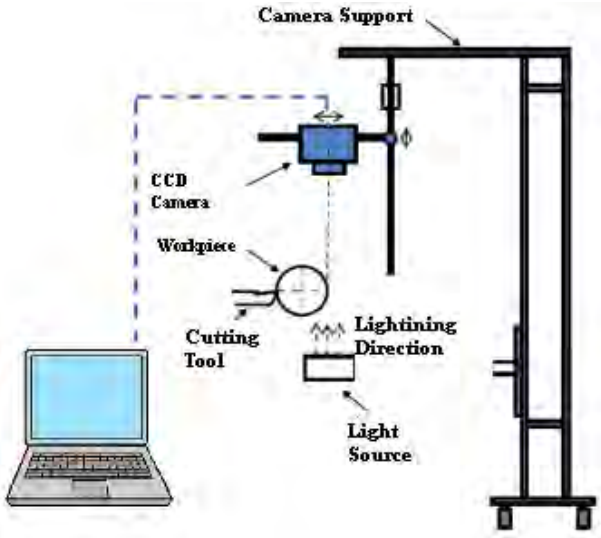


Fig. 2: Schematic diagram of the on-machine roughness measurement system [14].

2.4 Measuring Procedure

A procedure for the assessment of surface roughness using the image processing method is described below:

1- Preparation of specimen: 18 medium carbon steel AISI 1045 components were turned by adjusting the cutting speed and feed rate. The surface roughness values were measured by a stylus type roughness tester.

2- Components were put under the "CCD camera" and modified for appropriate illumination of the LED; the "CCD camera" was focused to get an obvious contour image of the specimen edge. An image of the contour edge of the component being turned was captured and stored in the computer using a USB cable.

3- The captured image was converted to a grayscale version to reduce the operating time of the algorithm.

4- The area to be measured was cropped from the original image, and unnecessary areas around the shaft edge were deleted.

5- The image stored in the computer was recovered and treated using a median filter (mask size 3X3) to remove the noise present in the image.

6- The developed SRVISION software, calculated the image gradient in the Y direction to find the change in the intensity from white to black and found the edge of the workpiece.

7- Converted the grayscale amount of the image into black and white with a binarization technique, the limits were applied on the image so that the component area was black and the rest was white.

8- An algorithm was written for scanning the first row to find the first white pixel in the profile, then scanning the second row to find the second white pixel. This method was repeated to find the whole white pixels lying in the image, these pixels reflected the profile of the workpiece's surface profile.

9- The best fit line was drawn to the contour image to get a mean line of the contour by least-square fitting.

10- Assessment of average surface roughness (Ra) and root mean square roughness (Rq) by image processing technique from the image contour was performed by subtracting each pixel of the counter profile from the calculated mean line and using the following relationships:

$$Ra = \frac{f}{n}\sum_{i=1}^{n} hi \tag{2}$$

$$Rq = f\sqrt{\frac{\sum_{i=1}^{n} hi}{n}} \tag{3}$$

where: n is the number of data points, hi is the absolute distance of the i[th] point on the profiling measure from the mean line, and f is a scaling factor.

Figure 3 indicates the different steps of roughness measurement in the SRVISION. After loading the image into the software, where the software gives the option to choose the area to measure.

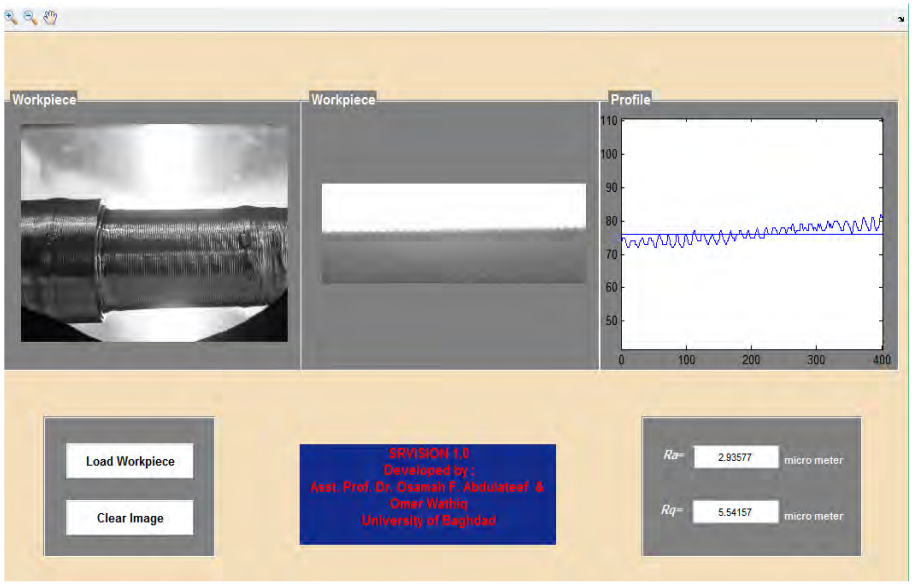


Fig. 3: The main interface of software developed (SRVISION).

3. RESULTS AND DISCUSSION

The results of the stylus and vision methods for measured surface roughness are presented in this section. Additionally, their demand results were compared and discussed.

3.1 Measuring Surface Roughness Using Stylus Method

A stylus instrument was utilized to compare with the values of measured roughness by the vision system. Every surface was measured 5 times at different positions of the workpiece utilizing a cutoff of 0.8 mm. The minimum and maximum values of surface roughness achieved by the stylus method are indicated in Table 4. The variation ΔRa between the minimum and maximum Ra values changed between 0.24 µm and 0.844 µm for the 18 specimens. The maximum variation as a percentage of the minimum Ra value was 13.22% for each workpiece. The difference ΔRq between a minimum and maximum Rq values ranged from 0.13 µm to 1.94 µm. The maximum variation as a percentage of the

minimum Rq value was 15.36 % for each workpiece. The different values of surface roughness at the same workpiece were a result of instability in the machining process performed by the traditional turning machine.

Table 4: Minimum and maximum roughness achieved by stylus device

No.	V [rpm]	f [mm/rev]	Ra(max) [µm]	Ra(min) [µm]	Δ Ra [µm]	Δ Ra [%]	Rq(max) [µm]	Rq(min) [µm]	Δ Rq [µm]	Δ Rq [%]
1	140	0.206	6.474	6.026	0.448	7.43	8.291	7.4	0.891	10.75
2		0.2886	7.887	7.563	0.324	4.28	8.928	8.736	0.192	2.15
3		0.396	9.9	9.362	0.538	5.75	13.34	11.4	1.94	14.54
4		0.77	13.1	12.86	0.24	1.87	15.45	15.04	0.41	2.65
5	250	0.182	6.512	6.107	0.405	6.63	7.766	6.957	0.809	10.42
6		0.364	9.909	9.302	0.607	6.53	12.813	12.193	0.62	4.84
7		0.742	10.67	10.152	0.518	5.10	12.75	11.24	1.51	11.84
8	650	0.179	6.426	5.946	0.48	8.07	8.353	7.365	0.988	11.83
9		0.256	6.821	6.526	0.295	4.52	7.723	7.154	0.569	7.37
10		0.361	8.696	8.129	0.567	6.98	10.03	9.626	0.404	4.03
11		0.732	15.85	15.22	0.63	4.14	18.351	17.95	0.401	2.19
12	950	0.161	4.53	4.125	0.405	9.82	5.594	4.735	0.859	15.36
13		0.22	8.09	7.321	0.769	10.50	10.11	9.1	1.01	9.99
14		0.335	9.334	8.811	0.523	5.94	11.16	10.27	0.89	7.97
15		0.541	10.43	9.586	0.844	8.80	12.42	11.76	0.66	5.31
16	1350	0.147	4.326	3.821	0.505	13.22	5.246	5.116	0.13	2.48
17		0.228	6.633	6.35	0.283	4.46	8.534	7.321	1.213	14.21
18		0.34	10.212	9.659	0.553	5.73	11.752	11.196	0.556	4.73

3.2 Measuring the Surface Roughness Using the Vision Method

Every image of the workpiece was measured 4 times at different positions. Table 5 indicates the minimum and maximum surface roughness values achieved by the machine vision system. The difference ΔRa between the minimum and maximum Ra values ranged from 0.256 µm to 1.184 µm. The maximum variation for each workpiece as a percentage of the minimum Ra value was 12.7%. The difference ΔRq between a minimum and maximum Rq values ranged from 0.377 µm to 0.973 µm. The maximum variation for each workpiece as a percentage of the minimum Rq value was 10.76 %.

3.3 Comparison of Roughness Values Achieved by Stylus and Vision Methods

The results of measurements of average surface roughness (Ra) and root mean square surface roughness (Rq) using the suggested method of a vision system and comparison with the stylus method are shown in Table 6. The results show that the maximum Ra and Rq differences between the two methods were 3,744% and 3,727% respectively. The mean and the standard deviation between the two Ra measurements were 1.65% and 1.0% respectively. Also, the mean and the standard deviation of the difference for Rq were 1.433% and 1.0%, respectively. Figures 4 and 5 show a plot of average roughness and root mean square roughness respectively found by the suggested vision method (Ra (v)) versus the average roughness found by the stylus measurement (Ra(s)). The data were fitted with a linear trend line, and the correlation value was determined in Microsoft Excel using linear regression. A correlation value would specify a perfectly linear relationship between the two data groups. The high correlation of 0.99 indicates that the visual method is

capable of giving dependable roughness values for the measurements obtained in this study.

Table 5: Minimum and maximum roughness achieved by vision method

No.	V [rpm]	f [mm/rev]	Ra(max) [µm]	Ra(min) [µm]	Δ Ra [µm]	Δ Ra [%]	Rq(max) [µm]	Rq(min) [µm]	Δ Rq [µm]	Δ Rq [%]
1	140	0.206	6.326	6.0129	0.3131	5.21	8.368	7.756	0.612	7.31
2		0.2886	7.869	7.613	0.256	3.36	9.146	8.659	0.487	5.32
3		0.396	9.961	9.581	0.38	3.97	12.834	11.861	0.973	7.58
4		0.77	12.956	12.698	0.258	2.03	15.542	14.731	0.811	5.22
5	250	0.182	6.621	6.265	0.356	5.68	7.856	7.17	0.686	8.73
6		0.364	10.09	9.781	0.309	3.16	12.672	11.981	0.691	5.45
7		0.742	10.679	10.293	0.386	3.75	12.896	12.224	0.672	5.21
8	650	0.179	6.816	6.483	0.333	5.14	7.963	7.214	0.749	9.41
9		0.256	6.608	6.174	0.434	7.03	8.125	7.643	0.482	5.93
10		0.361	8.924	8.361	0.563	6.73	10.22	9.654	0.566	5.54
11		0.732	15.624	14.865	0.759	5.11	18.236	17.658	0.578	3.17
12	950	0.161	4.621	4.209	0.412	9.79	5.49	5.113	0.377	6.87
13		0.22	7.953	7.521	0.432	5.74	9.981	9.218	0.763	7.64
14		0.335	9.496	8.432	1.064	12.62	10.861	10.159	0.702	6.46
15		0.541	10.51	9.326	1.184	12.70	12.224	11.476	0.748	6.12
16	1350	0.147	4.286	3.843	0.443	11.53	5.012	4.542	0.47	9.38
17		0.228	6.716	6.283	0.433	6.89	7.962	7.105	0.857	10.76
18		0.34	10.246	9.514	0.732	7.69	11.742	10.954	0.788	6.71

Table 6: Comparison between roughness's achieved by stylus and vision methods

No.	V [rpm]	f [mm/rev]	Ra(v) [µm]	Ra(s) [µm]	Δ Ra [µm]	Δ Ra [%]	Rq(v) [µm]	Rq(s) [µm]	Δ Rq [µm]	Δ Rq [%]
1	140	0.206	6.1277	6.282	0.1543	2.456	8.0184	7.8465	0.1719	2.191
2		0.2886	7.7274	7.684	0.0434	0.565	8.954	8.817	0.137	1.554
3		0.396	9.7024	9.67	0.0324	0.335	12.253	12.183	0.0698	0.573
4		0.77	12.727	13.02	0.293	2.250	15.024	15.245	0.2209	1.449
5	250	0.182	6.3845	6.31	0.0745	1.181	7.4028	7.3705	0.0323	0.438
6		0.364	9.9862	9.7	0.2862	2.951	12.278	12.45	0.1713	1.376
7		0.742	10.506	10.351	0.155	1.497	12.651	12.23	0.4214	3.446
8	650	0.179	6.397	6.337	0.06	0.947	7.475	7.549	0.074	0.980
9		0.256	6.6552	6.795	0.1398	2.057	7.957	7.875	0.082	1.041
10		0.361	8.6246	8.45	0.1746	2.066	9.9618	9.882	0.0798	0.808
11		0.732	15.157	15.343	0.186	1.212	18.08	18.12	0.04	0.221
12	950	0.161	4.487	4.3375	0.1495	3.447	5.2212	5.177	0.0442	0.854
13		0.22	7.7202	7.8	0.0798	1.023	9.5548	9.68	0.1252	1.293
14		0.335	8.74	9.08	0.34	3.744	10.623	10.715	0.0921	0.860
15		0.541	10.1	10.055	0.045	0.448	11.983	11.89	0.0928	0.780
16	1350	0.147	3.9539	4.004	0.0501	1.251	4.8807	5.026	0.1453	2.891
17		0.228	6.4611	6.4	0.0611	0.955	7.4762	7.7656	0.2894	3.727
18		0.34	9.8284	9.964	0.1356	1.361	11.25	11.4	0.15	1.316

https://doi.org/10.31436/iiumej.v22i2.1507

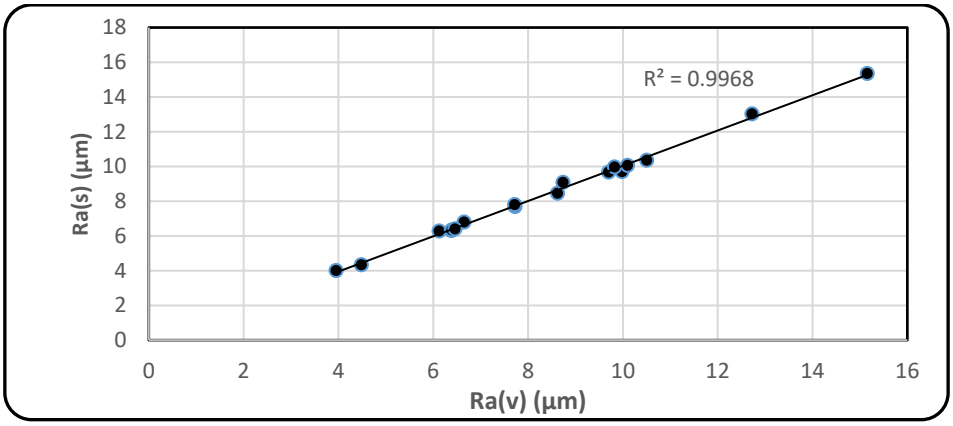


Fig. 4: Comparison between Ra values achieved by stylus and vision methods.

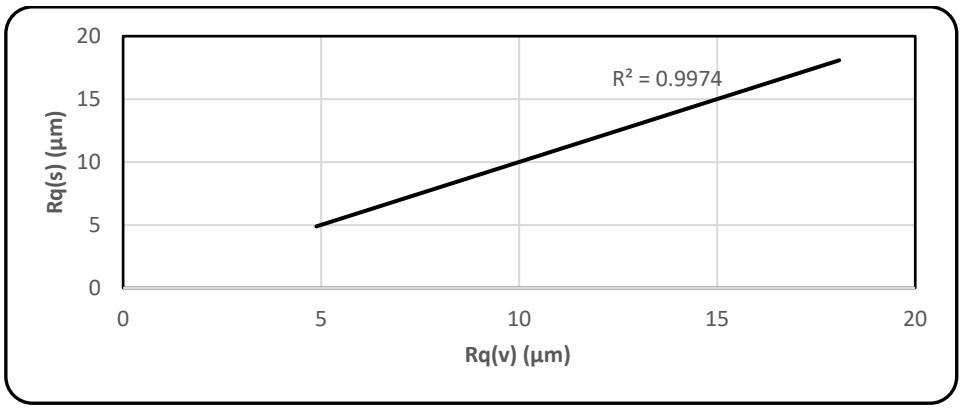


Fig. 5: Comparison between Rq values achieved by stylus and vision methods.

Also, the comparison plot of estimated average surface roughness values (Ra) and root mean square surface roughness values (Rq) using the stylus approach and vision approach are shown in Figs. 6 and 7. It's quite obvious that the estimated values of roughness by the two approaches are accurate with an R-squared of 0.997.

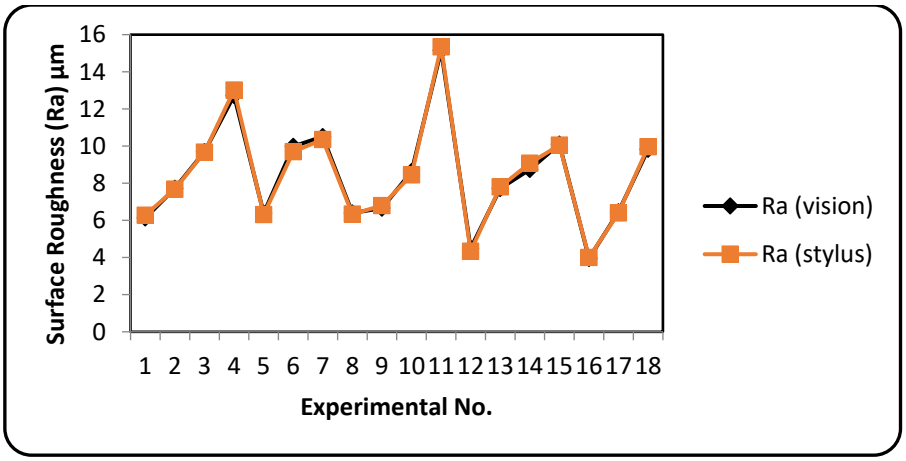


Fig. 6: The analogy of stylus and vision values for average surface roughness (Ra).

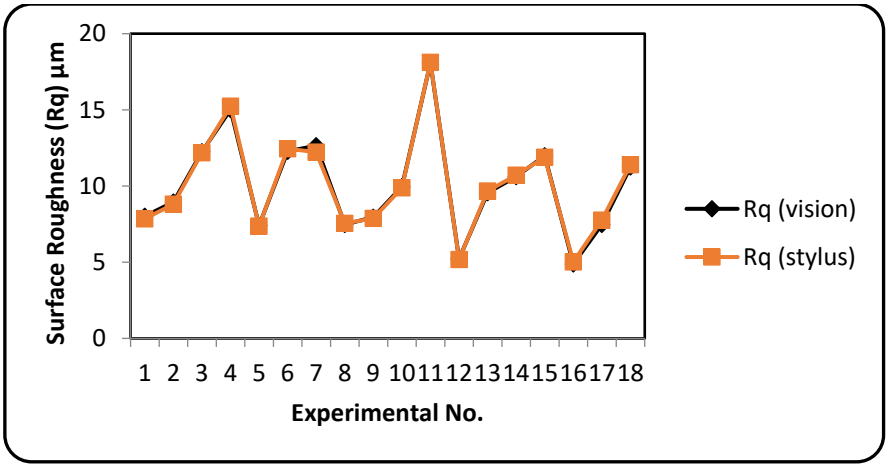


Fig. 7: The analogy of stylus and vision values for root mean square surface roughness (Rq).

4. CONCLUSION

A computer vision system and backlight method for assessing the surface roughness of turned medium carbon steel AISI 1045 specimens under different machining conditions were proposed in this study. The computer vision system captured and stored the enlarged contour edge images of the specimens as they were being turned. SRVISION software was developed for calculating the surface roughness immediately from the specimen's contour image. The advantage of using a backlighting device is that it is not influenced by industrial environment lighting conditions. The precision of the vision method was compared with the stylus method for many experiments. Comparison graphs drawn between the vision and stylus methods demonstrated the percentage error obtained a maximum variation of 3.75 % and the coefficient of correlation (R^2) values were close to one. Hence the vision method is reliable and appropriate for on-line, non-contact surface roughness measurement of machined components.

REFERENCES

[1] Lee BY, Tarng YS. (2001) Surface roughness inspection by computer vision in turning operations. International Journal of Machine Tools and Manufacture, 41(9): 1251-1263. https://www.sciencedirect.com/science/article/abs/pii/S0890695501000232

[2] Gadelmawla ES. (2004) A vision system for surface roughness characterization using the gray level co-occurrence matrix. NDT & E International, 37(7): 577-588. https://doi.org/10.1016/j.ndteint.2004.03.004

[3] Ghassan A.Al-Kindi, Bijan Shirinzadeh. (2007) An evaluation of surface roughness parameters measurement using vision-based data. International Journal of Machine Tools and Manufacture, 47(3-4): 697-708. https://doi.org/10.1016/j.ijmachtools.2006.04.013

[4] Hu Zhongxiang, Zhu Lei, Teng Jiaxu, Ma Xuehong, Shi Xiaojun. (2009) Evaluation of three-dimensional surface roughness parameters based on digital image processing. International Journal of Advanced Manufacturing Technology, 40(3): 342-348. https://link.springer.com/article/10.1007/s00170-007-1357-5

[5] Fadare DA, Oni AO. (2009) Development and application of a machine vision system for measurement of surface roughness. ARPN Journal of Engineering and Applied Sciences, 4(5): 30-37.

[6] Shahabi HH, Ratnam MM. (2010) Noncontact roughness measurement of turned parts using machine vision. The International Journal of Advanced Manufacturing Technology, 46: 275-284. https://link.springer.co/artmicle/10.1007/s00170-009-2101-0

[7] Sridhar VG, Adithan M. (2012) An in-process approach for monitoring and evaluating the surface roughness of turned components. European Journal of Scientific Research, 68(4): 534-543.

[8] Mohan Kumar Balasundaram, Mani Maran Ratnam. (2014) In-process measurement of surface roughness using machine vision with sub-pixel edge detection in finish turning. International Journal of Precision Engineering and Manufacturing, 15(11): 2239-2249. https://link.springer.com/article/10.1007/s12541-014-0587-3

[9] Srivani A, Anthony Xavior M. (2014) Investigation of surface texture using image processing techniques. 12th Global Congress on Manufacturing and Management, GCMM 2014, Procedia Engineering, 97 (2014): 1943-1947. https://doi.org/10.1016/j.proeng.2014.12.348

[10] Qingqun Mai, Yanming Quan, Peijie Liu, Guo Ding. (2016) A new method of on-line turned surface monitoring by digital image processing. MATEC Web of Conferences, MMME 2016, 63. https://doi.org/10.1051/matecconf/20166304030

[11] Naresh P, Syed Altaf Hussain, Durga Prasad B. (2019) Surface Roughness Measurement of Machined Surfaces by Machine Vision Technique. International Journal of Recent Technology and Engineering (IJRTE), 7(ICETESM): 129-134. https://www.ijrte.org/wp-content/uploads/papers/v7iicetesm18/

[12] Dhiren R Patel, Mysore BK, Vinay Vakharia. (2020) Modeling and prediction of surface roughness using multiple regressions: A noncontact approach. Wiley, Engineering Reports. 2(e12119): 1-15. https://doi.org/10.1002/eng2.12119

[13] Dhiren R Patel, Harshit Thakker, Kiran MB, Vinay Vakharia. (2020) Surface Roughness Prediction of Machined Components Using Gray Level Co-occurrence Matrix and Bagging Tree. FME Transactions, 48(2): 468-475.

[14] Kumar BM, Ratnam MM. (2015) Machine vision method for non-contact measurement of surface roughness of a rotating workpiece. Sensor Review, 35(1): 10-19. https://doi.org/10.1108/SR-01-2014-609

https://doi.org/10.31436/iiumej.v22i2.1547

THE EFFECT OF LATERAL LIFTING TASKS ON HAND GRIP AND PINCH STRENGTH MEASUREMENTS

JOE YEE TAN[1], AZRUL AZWAN BIN ABDUL RAHMAN[2*], NADIAH BINTI AHMAD[3] AND ARFAUZ BIN A. RAHMAN[3]

[1]Faculty of Manufacturing Engineering
[2]Advanced Manufacturing Centre
[3]Centre of Smart System and Innovative Design,
Faculty of Manufacturing Engineering,
Universiti Teknikal Malaysia Melaka, Malaysia.

**Corresponding author: azrulazwan@utem.edu.my*

(Received: 7th July 2020; Accepted: 16th November 2020; Published on-line: 4th July 2021)

ABSTRACT: In lateral material handling tasks, which is very common in industries, warehouse systems, and other sectors, the workers can lead to work-related musculoskeletal disorders (WMSDs) because of the task and work designs, especially in upper extremities. WMSDs in hand, wrist, and fingers cause workers to have health problems, in pain and uncomfortable, hence decrease their working productivity and efficiency. The workstations distances are one of the factors that might affect workers when they perform manual material handling laterally, which can lead to WMSDs in hands. Hence, there is a need to study the relationship between transfer distances with the hand grip and pinch strengths. An experiment was carried out with 30 male participants to identify the relationship of lateral transfer distances with the hand grip and pinch strengths. The results obtained from the experiment were further investigated and analysed by using repeated measure one-way MANOVA and graphs. The results had proved that in the distances of 1.0 m, 1.25 m and 1.5 m did not affect one's hand grip and pinch strengths. But, the postures and movements were varied based on distances.

ABSTRAK: Pengangkutan barang atau bahan secara lateral amat biasa dalam industri, gudang dan sektor-sektor lain, dan kerja ini akan menyebabkan pekerja menghadapi penyakit gangguan muskuloskeletal berkaitan kerja (WMSDs) terutamanya di bahagian tangan. WMSDs yang melibatkan tangan dan jari menyebabkan pekerja mempunyai masalah kesihatan secara kekal dan mengalami kesakitan serta tidak-keselesaan. Secara tidak langsung, masalah ini telah mengurangkan prestasi mereka ketika bekerja. Syarikat terpaksa memberi bayaran yang tinggi kepada pekerja untuk kos perubatan and mengalami kerugian besar kerana pekerja yang tidak datang bekerja disebabkan penyakit tersebut. Jarak antara dua stesen kerja adalah faktor yang menyebabkan penyakit ini dihadapi oleh pekerja. Oleh itu, kajian diperlukan untuk mengkaji hubungan antara jarak dan kekuatan genggaman tangan serta jari. Satu esperimen yang melibatkan 30 orang lelaki responden telah dijalankan untuk mengenal pasti hubungan antara jarak dan ganggaman tangan serta jari. Hasil daripada esperimen telah dikaji dan dianalisiskan dengan menggunakan MANOVA dan grafs. Hasil kajian telah membuktikan bahawa jarak dalam 1.0m, 1.25m dan 1.5m tidak membawa apa-apa kesan terhadap genggaman tangan dan jari. Cara dan pergerakan responden adalah berbeza dan disebabkan oleh jarak, walaubagaimanapun, cara dan pergerakan responden tidak membawa sebarang kesan terhadap kekuatan genggaman tangan dan jari mereka.

KEYWORDS: *lateral lifting tasks; transfer distance; hand grip strength; pinch strength; postures; movements*

https://doi.org/10.31436/iiumej.v22i2.1547

1. INTRODUCTION

Work-related musculoskeletal disorders (WMSDs) are a group of painful disorders of muscles, tendons, and nerves, for example, carpal tunnel syndrome, tendonitis, thoracic outlet syndrome, and tension neck syndrome (Canadian Centre of Occupational Health and Safety, 2018). WMSDs are also commonly called repetitive motion injury, repetitive stress injury and overuse injury, which had indicated that WMSDs are caused by a singular cause for damage to the musculoskeletal system, which is repetition and stress. Almost all work requires the use of the arms and hands. Therefore, most of the WMSDs affect the hands, wrists, elbows, neck, and shoulders. However, work using the legs can lead to WMSD of the legs, hips, ankles, and feet. Some back problems also result from repetitive activities [1].

Work tasks that are high in frequency and involve repetitive movements or activities with awkward postures that cause WMSDs bringing the effects on human muscles which may be painful during work or at rest. However, the traumatic injuries of the muscles, tendons, and nerves due to accidents are not considered to be WMSDs [2]. Manual material handling is a task that required in almost all working environments, for examples workers in construction, agriculture, hotels, factories, warehouses, building sites, farms, hospitals, offices and restaurants, where the tasks are most likely to be exposed to heavy loads for a long period and repetitive works which lead to cumulative disorders due to gradual and cumulative deterioration of the musculoskeletal system[3].

Lateral lifting tasks are considered one of the manual material handling tasks, which had indicated highly repetitive lifting movements. In lateral manual lifting tasks, the repetition of discomfort body postures and the overexertion of the force of loads cause the contracted muscles to squeeze the blood vessels and limit the flow of blood down to the working hand muscles. The reduced blood supply causes muscle fatigue, making hands and fingers more prone to injury [4]. The challenge faced by ergonomists are repetitive lifting tasks cannot be avoided, hence the range of the postural deviations and moments in forwarding flexion can be reduced by raising lift origins and destinations positions [5]. However, the lateral transfer distances and the load weights are hard to control and determined.

The workplace design plays a crucial role in the development of a WMSD. Certain workplace conditions, for example, the layout of the workstation, the speed of work especially in conveyor-driven jobs, and the weight of the objects being handled are important because these factors highly influence the risk factors and the conditions of the workers [2].

A workstation is a place a worker occupies when performing a job. The workstation may be occupied all the time, or it may be one of several places where work is done. The distances between workstations are very important because it is the factor that might cause WMSDs in workers [6]. If the workstation is properly designed, the worker should be able to maintain a correct and comfortable body posture [7]. Therefore, the distances between the workstations needed to investigate and find out how distances can affect human hand activities and strengths.

The bending of fingers and force exerts onto the hand by the load when carry in certain distances repetitively will also bring the effect of WMSDs of hands and fingers. A longer time is needed to recover when extra forces are exerted onto the muscles. There is insufficient time for recovery during repetitive work, hence increasing in forceful movements increases muscle fatigue speed [2]. Repetitive movements together with extra forces are dangerous and can lead to permanent WMSDs when involved in the same joints

and muscle groups over and over with the same motion for too long. The survey carried out by CPWR (2018) had been proved that the rate of overexertion injuries resulting in days away from work in 2015 in transportation sectors is the highest [8].

The costs of MSDs needed for employees for treatments are the highest among the others in 2013, in which the annual costs are 80 billion USD for only workers in United State [9]. Based on the 2016 Liberty Mutual Workplace Safety Index, workplace injuries and accidents that cause employees to miss six or more days of work cost U.S. employers nearly 62 billion USD in 2013. The injuries caused by overexertion involving outside sources are the highest with 24.4% of the total cost of the most disabling workplace injuries or 15.08 billion USD among others. This high cost paid by companies for overexertion injuries is considered too high and hence WMSDs that caused by overexertion should be decreased by preventions and improvements.

WMSDs can be prevented and decrease if appropriate methods and preventions are done. Certain improvements and modifications on lateral manual handling tasks in industries can be done to decrease the effect on the hand grip and pinch strengths and improve efficiency. Therefore, studies and experiments can be carried out to identify the effect of transfer distances and load weights in lateral lifting task.

2. METHODS

2.1 Participants

A total of thirty healthy right-handed male participants (mean age was 21±3 years, mean body mass was 75±15 kg and mean height 175±10 cm) were recruited as subjects for the experiments. All the subjects were non-smokers and free of any history of upper-extremities musculoskeletal disorders or any injuries that might affect the way they performed the tasks. All the subjects were given an informed consent form and participation form for the details of the experiments. All participants were volunteers.

2.2 Equipment and Tools

2.2.1 Jamar Hand Dynamometer and Pinch Meter Gauge

Jamar hand dynamometer was used as presented in Fig. 1 to measure the hand grip strength of participants before and after the experiment in this study. Ven-Stevens et al. [10] mentioned that grip strength is defined as the measurable ability to exert pressure onto an object or the force applied by the hand and fingers. During manual tasks, muscles are contracting much harder with excessive force, thus causes stress on the muscles, tendons, and joints. The amount of force exerted onto the objects depends on the type of grip, the weight of an object, body posture, the type of activity and the duration of the task [11].

Fig. 1: Jamar hand dynamometer.

There are three types of forearm positions when hand grip strength is measured, which is pronation, neutral and supination. In this study, the forearm neutral position (Fig. 2) was chosen because this is the posture of the hands when carried out lateral lifting tasks.

Fig. 2: The position of participant's hand grip strength forearm neutral position.

Jamar pinch meter gauge was used as presented in Fig. 3 to measure the pinch strength of participants before and after the experiment. There are three types of pinch prehension which are lateral pinch (key pinch), three-point pinch (Palmer, three-jaw chuck pinch), and two-point pinch (tip to tip pinch) [12].

Fig. 3: Jamar pinch meter.

The pinch strength was measured in the lateral pinch position (Fig. 4) in this study due to the finger posture of carried out lateral lifting tasks. Petersen et al. [13] previously had tested the "10% rule" which stated that the dominant hand possesses 10% greater grip strength than the non-dominant hand. However, the study had concluded that the 10% rule is only valid for right-handed persons and grip strengths are equivalent in both hands for left-handed persons. The concept is explained by the bilateral difference in the right-handed will be increased by the differential stress due to hand dominance and reduced in the left-handed persons [13].

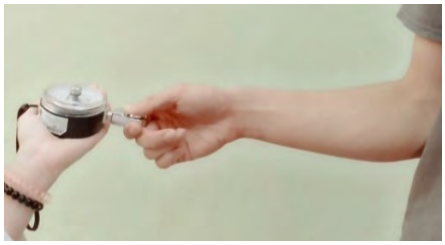


Fig. 4. The position of participant's pinch strengths lateral pinch position was taken.

2.2.2 Loads and Workstations

A total of six units of A4 paper boxes with handle were prepared where each box was weighed 10kg. Two tables were placed as the workstations for the participants to carry out the lateral lifting tasks. The marking of the lateral distances on the floor for experiment purposes was using the clothes tape. Measuring tape and weighing scale were used to measure the height of the participants, distances between two workstations (tables) and weight of the participants.

2.3 Participant Counterbalance

https://doi.org/10.31436/iiumej.v22i2.1547

According to Zeelenberg and Pecher [14], sequential effects happen when performance in the current condition is affected by the conditions preceding it. Hence, the counterbalance condition of the trials is important in order to minimize the sequential effects of the trials and unwanted order effects from causing differences between conditions. There are three different lateral distances to be studied, hence there is three arrangement of the trials (3! = 3 x 2 x 1 = 6) for both neutral forearm position of hand grip strength measurements and lateral pinch strength measurements. Figure 5 showed the tree diagram of counterbalancing of all 3 lateral distances. Counterbalancing is performed by creating Latin squares [14]. The measurement for the first participant would follow the first arrangement, the second participant would follow the second arrangement and so on. The arrangement process was repeated every 6 participants.

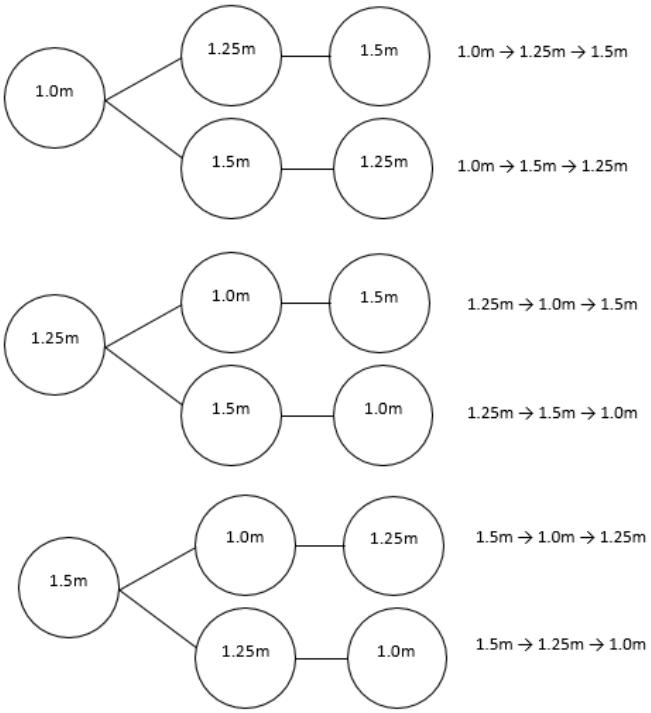


Fig. 5: Tree diagram for 6 trials counterbalances arrangements in the experiment.

2.4 Experimental Procedures

2.4.1 Preparations for Experiments

Six boxes of 10 kg were prepared with proper handles. Two tables with a height of 82.66 cm were prepared and put side to side with specific distances in between (Fig. 6). The distances of 1.0 m, 1.25 m, 1.5 m were marked by using a cloth tape. The box was placed align with the bench horizontal edge. Before the participants arrived, it was necessary to ensure that the tools needed in the experiments are in good condition and precise scales. The Jamar Hand Dynamometer and Pinch Meter Gauge were checked and make sure that the equipment was working properly. Every time before running an experiment on a participant, the checking had been done once. The gauge needle needed to be in the box that contained the number zero but should not touch the stopper. Before the experiment begins, the participant's hand grip and pinch strength were measured. The measurements were taken twice for each hand grip and pinch strength for both hands and the larger value was calculated and recorded.

https://doi.org/10.31436/iiumej.v22i2.1547

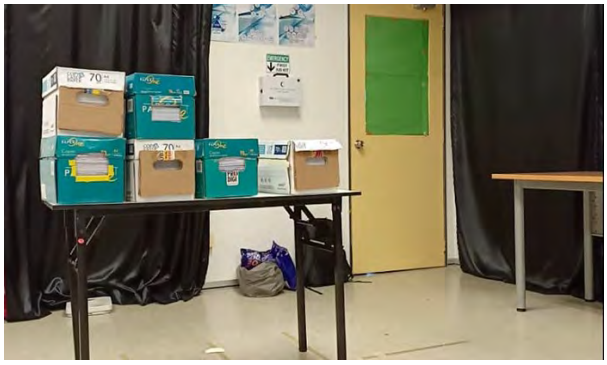


Fig. 6: Six boxes were prepared, and two tables were placed 1.25 m apart.

2.4.2 Preparations for Experiments

The distances that the participants would be carried out first were following the counterbalance sequences that had been distributed to the participants. The participant was required to stand still in the middle of the two benches and facing the front (Fig. 7). The participant required to hold the handle of the box and transfer the box by using both hands from right to left in a single direction only. The boxes were picked up from the first table to the second, which the first table located on the right side was the origin position. The second table that acted as the destination position was continuously empty for participants to place the boxes (Fig. 8). After the three types of distances had been finished by the participants, the survey forms were distributed to the participants and the participants were required to answer the survey immediately.

Fig. 7: A participant standing in the middle of two tables.

Fig. 8: A participant transferring the boxes from right to left and the second table was continuously emptied.

2.5 Measurements

2.5.1 Hand Grip and Pinch Strengths Measurements

The arrangement of the hand grip and pinch strength data to be measured would follow the counterbalance sequencing design of the experiment. After each set of lateral distance's lifting task was carried out, the hand grip and pinch strength were measured again for twice each for both hands based on the same postures of hand and finger fixed during the measurements of hand grip and pinch strengths before the experiment started. The larger value between two readings taken for the hand grip and pinch strength measurement each was chosen for the analysis phase. The participants were given to rest one minute after each reading. The hand grip and pinch strength were read in kilograms and recorded in the

participant form for further analysis. The postures and movements of participants were also been observed and recorded.

2.5.2 Qualitative Data Collections

During the survey session, the participants were given questionnaires and required to answer some questions that can provide useful information and confirmation on the quantitative data that had collected. The survey was carried out in Likert scale format and the questionnaires were collected and transferred into Microsoft Word and Excel for further analyzation. The postures and movements of the participants were observed and recorded in the participation form. The postures and movements of participants were focussed on the way the participants transfer load, especially the legs.

2.6 Statistical Analysis

The statistical analysis involved both hand grip and pinch strength measurements (quantitative data) and qualitative data. The comparison of forces of hand grip and pinch strength before and after the experiments had done by constructing the statistical analysis on the data gathered and collected during the experiment sessions. The multivariate analysis of variance (MANOVA) method was used to investigate the inter-relationship between one factor with the hand grip strength data and pinch strength data. The repeated measure one-way MANOVA test had been chosen to analyze how the transfer distance can affect the hand grip and pinch strengths. The repeated measure was chosen because the same group of participants was tested all along with the experiment. The MANOVA test was conducted by using SPSS (IBM SPSS Statistics v26.0.0).

There were some assumptions in order to perform the one-way MAVOVA test, which was the data must be a normal distribution. A normality test is used to test and ensure that the data were normally distributed. MANOVA test was also used to examine that there was a reduction in the hand grip and pinch strength data before and after the experiments. The percentages of reduction of forces were calculated and tested to identify the differences between the percentages for the distances.

The information gathered from the survey forms (qualitative data) were interpreted by using Microsoft Excel. The results were analyzed and generated into bar charts, pie charts, and line graphs that represented the participants' percentages of perceptions on the hand grip and pinch strengths after the experiment. The results obtained from the qualitative data analysis were being compared with the results obtained from the quantitative statistical analysis to find out the relationships. Besides, the observations data were also transferred into Microsoft Excel for analysis purposes and generated using bar charts and line charts. The results were compared to identify the relationship between the distances and the postures with hand grip and pinch strengths.

3. RESULTS

3.1 Shapiro-Wilk Test

The statistical analysis of MANOVA has certain assumptions to be fulfilled before running the analysis, which included normality testing to ensure that the raw data collected is in normal distribution. The raw data were transferred into an Excel file format in terms of ages, races, body weights, body heights, hand grip readings and pinch strength readings as shown in Table 1, Table 2 and Table 3.

https://doi.org/10.31436/iiumej.v22i2.1547

Table 1: Participants data collection for hand grip strength of both hands before and after experiments

No.	Age	Weight	Height	Hand Grip												
	[y, m]	[kg]	[m]	Initial R	Initial L	1.0m R	% 1.0m R	1.0m L	% 1.0m L	1.25m R	% 1.25m R	1.25m L	% 1.25m L	1.5m R	% 1.5m R	1.5m L	% 1.5m L
1	21, 10	54	1.69	38.00	32.50	37.00	2.63	29.00	10.77	36.00	5.26	32.00	1.54	35.00	7.89	30.00	7.69
2	22, 0	65	1.70	45.50	40.00	42.00	7.69	36.50	8.75	39.50	13.19	35.50	11.25	30.00	34.07	35.00	12.50
3	21, 6	63	1.76	39.00	30.00	32.00	17.95	31.00	-3.33	31.00	20.51	30.00	0.00	36.00	7.69	30.00	0.00
4	22, 11	72	1.80	38.50	39.00	35.50	7.79	32.50	16.67	31.00	19.48	39.00	0.00	37.00	3.90	39.00	0.00
5	21, 10	57	1.74	35.50	31.50	34.50	2.82	30.00	4.76	33.00	7.04	30.00	4.76	33.00	7.04	30.00	4.76
6	21, 1	65	1.65	44.50	35.00	44.00	1.12	36.00	-2.86	39.50	11.24	35.00	0.00	39.00	12.36	33.50	4.29
7	21, 10	71	1.76	33.00	35.00	27.00	18.18	35.50	-1.43	32.00	3.03	36.00	-2.86	28.00	15.15	30.00	14.29
8	21, 3	55	1.65	39.50	30.00	34.00	13.92	26.00	13.33	35.00	11.39	30.00	0.00	37.00	6.33	30.00	0.00
9	23, 11	75	1.74	33.50	29.00	32.00	4.48	27.50	5.17	30.50	8.96	26.00	10.34	31.50	5.97	27.00	6.90
10	23, 3	70	1.80	40.00	39.50	36.00	10.00	36.00	8.86	40.00	0.00	36.50	7.59	36.50	8.75	37.00	6.33
11	23, 7	60	1.68	34.50	37.00	29.50	14.49	32.00	13.51	33.00	4.35	33.00	10.81	33.50	2.90	34.00	8.11
12	23, 5	63	1.75	32.00	26.50	29.00	9.38	26.50	0.00	31.00	3.13	26.00	1.89	28.00	12.50	24.00	9.43
13	24, 0	70	1.72	34.00	35.00	32.00	5.88	34.00	2.86	32.00	5.88	33.50	4.29	31.00	8.82	29.00	17.14
14	23, 1	68	1.75	45.50	41.50	45.00	1.10	39.00	6.02	43.00	5.49	40.00	3.61	41.50	8.79	40.00	3.61
15	23, 8	65	1.62	40.00	34.50	36.00	10.00	30.00	13.04	35.50	11.25	32.00	7.25	37.50	6.25	30.00	13.04
16	22, 5	99	1.70	37.50	35.00	33.00	12.00	28.50	18.57	34.50	8.00	32.00	8.57	35.00	6.67	33.00	5.71
17	22, 11	66	1.65	32.50	30.50	27.00	16.92	24.00	21.31	31.50	3.08	30.00	1.64	29.50	9.23	29.50	3.28
18	23, 11	84	1.80	36.00	36.50	36.00	0.00	34.00	6.85	34.00	5.56	34.50	5.48	36.00	0.00	33.00	9.59
19	22, 4	45	1.65	34.00	27.00	32.00	5.88	26.50	1.85	31.00	8.82	25.00	7.41	30.50	10.29	24.00	11.11
20	23, 4	80	1.84	39.00	45.00	37.00	5.13	42.00	6.67	38.00	2.56	37.00	17.78	31.00	20.51	32.00	28.89
21	22, 6	80.5	1.65	52.50	44.00	45.50	13.33	40.50	7.95	50.00	4.76	44.50	-1.14	44.00	16.19	39.00	11.36
22	23, 11	78.1	1.74	44.00	47.00	44.00	0.00	39.00	17.02	44.00	0.00	45.00	4.26	43.00	2.27	44.00	6.38
23	23, 3	102	1.65	42.00	44.00	41.00	2.38	37.00	15.91	40.00	4.76	35.50	19.32	41.50	1.19	44.00	0.00
24	23,11	67.5	1.72	36.00	33.00	33.00	8.33	33.00	0.00	35.50	1.39	33.00	0.00	34.00	5.56	36.00	-9.09
25	22, 5	84	1.72	26.00	24.00	21.00	19.23	23.00	4.17	22.00	15.38	20.00	16.67	20.00	23.08	18.00	25.00
26	22, 8	86.23	1.63	41.50	32.50	36.00	13.25	31.50	3.08	35.00	15.66	29.00	10.77	39.00	6.02	29.50	9.23
27	23, 9	64	1.70	37.00	34.00	34.50	6.76	32.00	5.88	35.00	5.41	34.00	0.00	36.00	2.70	28.00	17.65
28	23, 7	69	1.68	39.50	36.00	38.00	3.80	33.00	8.33	37.00	6.33	35.00	2.78	33.00	16.46	32.00	11.11
29	24, 1	65	1.73	39.50	33.00	30.00	24.05	29.00	12.12	30.50	22.78	26.50	19.70	38.50	2.53	31.00	6.06
30	23, 5	65	1.70	46.50	43.50	45.00	3.23	39.50	9.20	43.50	6.45	41.00	5.75	46.00	1.08	42.50	2.30

https://doi.org/10.31436/iiumej.v22i2.1547

Table 2: Participants data collection for pinch strength of both hands before and after experiments

No.	Age	Weight	Height	Pinch													
	[y, m]	[kg]	[m]	Initial R	Initial L	1.0m R	% 1.0m R	1.0m L	% 1.0m L	1.25m R	% 1.25m R	1.25m L	% 1.25m L	1.5m R	% 1.5m R	1.5m L	% 1.5m L
1	21, 10	54	1.69	9.25	6.75	8.50	8.11	6.75	0.00	8.75	5.41	6.00	11.11	9.25	0.00	6.50	3.70
2	22, 0	65	1.70	9.50	9.25	9.25	2.63	8.60	7.03	9.00	5.26	8.00	13.51	9.25	2.63	7.75	16.22
3	21, 6	63	1.76	8.00	7.00	6.25	21.88	6.00	14.29	7.00	12.50	6.50	7.14	7.25	9.38	6.25	10.71
4	22, 11	72	1.80	9.00	8.75	8.75	2.78	8.50	2.86	8.50	5.56	7.75	11.43	7.00	22.22	8.50	2.86
5	21, 10	57	1.74	9.00	7.25	8.75	2.78	6.75	6.90	8.50	5.56	7.00	3.45	8.75	2.78	6.50	10.34
6	21, 1	65	1.65	10.00	6.85	9.00	10.00	6.75	1.46	9.50	5.00	6.75	1.46	10.50	-5.00	6.00	12.41
7	21, 10	71	1.76	8.50	6.75	8.25	2.94	6.25	7.41	7.25	14.71	6.50	3.70	7.50	11.76	6.75	0.00
8	21, 3	55	1.65	9.50	8.25	9.00	5.26	7.25	12.12	9.25	2.63	7.75	6.06	8.75	7.89	8.00	3.03
9	23, 11	75	1.74	9.00	8.85	7.50	16.67	7.25	18.08	9.00	0.00	7.50	15.25	7.25	19.44	7.50	15.25
10	23,3	70	1.80	7.75	7.25	6.00	22.58	6.75	6.90	7.50	3.23	7.00	3.45	7.00	9.68	6.25	13.79
11	23, 7	60	1.68	7.50	6.25	7.25	3.33	5.75	8.00	6.25	16.67	5.50	12.00	6.75	10.00	5.00	20.00
12	23, 5	63	1.75	8.00	7.50	8.00	0.00	7.25	3.33	7.75	3.13	7.25	3.33	7.25	9.38	7.00	6.67
13	24, 0	70	1.72	8.50	6.75	7.50	11.76	6.50	3.70	7.75	8.82	6.75	0.00	8.25	2.94	6.25	7.41
14	23, 1	68	1.75	6.75	6.75	6.25	7.41	6.00	11.11	6.50	3.70	6.50	3.70	6.50	3.70	6.50	3.70
15	23, 8	65	1.62	9.50	8.50	9.50	0.00	7.50	11.76	9.50	0.00	8.00	5.88	9.50	0.00	7.75	8.82
16	22, 5	99	1.70	9.50	8.50	9.00	5.26	8.00	5.88	8.75	7.89	8.25	2.94	9.00	5.26	8.00	5.88
17	22, 11	66	1.65	9.50	8.00	7.25	23.68	6.00	25.00	7.50	21.05	6.25	21.88	8.00	15.79	8.25	-3.13
18	23, 11	84	1.80	9.50	8.50	8.75	7.89	7.00	17.65	9.25	2.63	7.75	8.82	9.25	2.63	7.50	11.76
19	22, 4	45	1.65	6.00	4.75	6.00	0.00	4.25	10.53	5.75	4.17	4.15	12.63	5.25	12.50	4.50	5.26
20	23, 4	80	1.84	6.50	7.25	6.00	7.69	6.50	10.34	5.00	23.08	5.25	27.59	5.00	23.08	5.50	24.14
21	22, 6	80.5	1.65	9.25	7.25	8.25	10.81	6.75	6.90	8.25	10.81	6.75	6.90	8.00	13.51	6.50	10.34
22	23, 11	78.1	1.74	9.50	8.00	8.50	10.53	8.25	-3.13	9.25	2.63	6.00	25.00	8.50	10.53	7.00	12.50
23	23, 3	102	1.65	11.75	10.75	11.00	6.38	9.00	16.28	10.25	12.77	9.75	9.30	11.50	2.13	10.00	6.98
24	23,11	67.5	1.72	7.00	7.50	6.00	14.29	6.75	10.00	6.25	10.71	6.50	13.33	5.25	25.00	7.00	6.67
25	22, 5	84	1.72	4.75	4.75	4.00	15.79	4.00	15.79	4.25	10.53	3.75	21.05	3.75	21.05	4.00	15.79
26	22, 8	86.23	1.63	9.00	8.50	9.00	0.00	8.50	0.00	9.00	0.00	8.00	5.88	8.00	11.11	8.50	0.00
27	23, 9	64	1.70	7.50	6.75	7.25	3.33	6.50	3.70	7.50	0.00	6.50	3.70	7.25	3.33	6.25	7.41
28	23, 7	69	1.68	8.00	7.25	7.50	6.25	7.00	3.45	7.50	6.25	6.75	6.90	7.50	6.25	7.00	3.45
29	24, 1	65	1.73	9.25	7.50	9.00	2.70	7.00	6.67	9.00	2.70	6.50	13.33	9.25	0.00	7.00	6.67
30	23, 5	65	1.70	8.50	8.75	7.75	8.82	8.75	0.00	8.50	0.00	8.00	8.57	8.25	2.94	8.25	5.71

There are two tests used in SPSS software to test the normality, which is Shapiro-Wilk and Kolmogorov-Smirnov. Since the sample size of this experiment is only 30, hence the test that chosen to test for normality in SPSS is the Shapiro-Wilk test, where the Shapiro-Wilk test is restricted for the sample size that less than 50 [15]. The normality test was run for two separate sections where the first section would be the hand grip data (Table 4) and the second section would be the pinch strength data (Table 5). In normality testing, the hypothesis was fixed as below:

H_0: The data are normally distributed.

H_1: The data are not normally distributed.

This hypothesis was set for both hand grip and pinch strength measurements normality tests. The p-value, which is the significance value generated by SPSS is used to determine whether the data is normally distributed. Since the confidence level of 95% is chosen, the p-value needed to be more than 0.05 to indicate that the data is normally distributed. If the p-value is less than 0.05, the data is not normally distributed. The Shapiro-Wilk test result was run at confidence level of 95%.

Table 3: Participants data collection for stepping during experiments

No.	Age [y, m]	Race	Weight [kg]	Height [m]	Stepping		
					1.0m	1.25m	1.5m
1	21, 10	C	54	1.69	1	2	2
2	22, 0	C	65	1.7	1	2	2
3	21, 6	C	63	1.76	2	2	2
4	22, 11	C	72	1.8	0	2	2
5	21, 10	C	57	1.74	1	1	2
6	21, 1	C	65	1.65	2	2	2
7	21, 10	C	71	1.76	1	1	2
8	21, 3	C	55	1.65	1	1	2
9	23, 11	C	75	1.74	1	1	1
10	23,3	C	70	1.8	0	0	0
11	23, 7	C	60	1.68	1	2	2
12	23, 5	M	63	1.75	1	1	1
13	24, 0	C	70	1.72	1	1	1
14	23, 1	M	68	1.75	2	2	2
15	23, 8	M	65	1.62	1	2	2
16	22, 5	M	99	1.7	1	1	2
17	22, 11	M	66	1.65	1	1	1
18	23, 11	I	84	1.8	0	2	2
19	22, 4	M	45	1.65	1	1	2
20	23, 4	M	80	1.84	2	2	2
21	22, 6	M	80.5	1.65	1	2	2
22	23, 11	C	78.1	1.74	1	1	2
23	23, 3	M	102	1.65	1	1	2
24	23,11	C	67.5	1.72	1	1	2
25	22, 5	M	84	1.72	1	1	1
26	22, 8	M	86.23	1.63	1	1	1
27	23, 9	C	64	1.7	2	2	2
28	23, 7	C	69	1.68	1	1	1
29	24, 1	C	65	1.73	1	1	1
30	23, 5	C	65	1.7	2	2	2

Table 4: Normality test results for hand grip strength measurements before and after the experiments for both hands of 30 participants.

	Kolmogorve-Smirnov[a]			Shapiro-Wilk		
	Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Initial R	0.126	30	0.200*	0.977	30	0.742
1.0m R	0.120	30	0.200*	0.960	30	0.313
1.25m R	0.148	30	0.094	0.946	30	0.133
1.5m R	0.069	30	0.200*	0.979	30	0.795
Initial L	0.125	30	0.200*	0.975	30	0.687
1.0m L	0.072	30	0.200*	0.983	30	0.898
1.25m L	0.083	30	0.200*	0.982	30	0.879
1.5m L	0.128	30	0.200*	0.959	30	0.289

Table 5: Normality test results for pinch strength measurements before and after the experiments for both hands of 30 participants.

	Kolmogorve-Smirnov[a]			Shapiro-Wilk		
	Statistic	**df**	**Sig.**	**Statistic**	**df**	**Sig.**
Initial R	0.172	30	0.023	0.933	30	0.058
1.0m R	0.114	30	0.200*	0.948	30	0.154
1.25m R	0.154	30	0.067	0.937	30	0.075
1.5m R	0.111	30	0.200*	0.972	30	0.608
Initial L	0.156	30	0.059	0.947	30	0.143
1.0m L	0.125	30	0.200*	0.94	30	0.092
1.25m L	0.155	30	0.064	0.936	30	0.071
1.5m L	0.129	30	0.200*	0.974	30	0.656

Based on the Shapiro-Wilk test in Table 4, the results of the hand grip strength measurements were summarized in a table where all the hand grip strength data were having the p-value or significance value of more than 0.05 ($p > 0.05$). This indicated that the H_0 is accepted and the data were all normally distributed. The normal Q-Q plot graphs for each hand grip strengths data were generated and shown in Fig. 9 and Fig. 10.

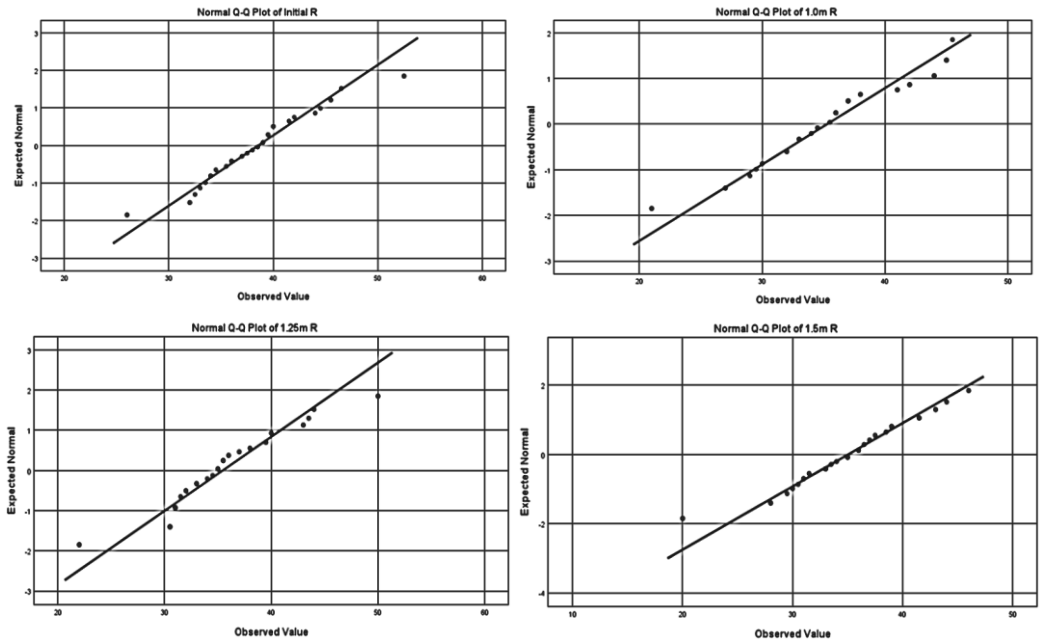


Fig. 9: Normal Q-Q plot graphs for right hand grip strengths data.

https://doi.org/10.31436/iiumej.v22i2.1547

The results of pinch strength measurements were summarized in a table (Table 5) where all the pinch strength data were having the p-value or significance value of more than 0.05 ($p > 0.05$). This indicated that the H_1 is rejected and the data are all normally distributed. The normal Q-Q plot graphs for each pinch strengths data were generated and shown in Fig. 11 and Fig. 12.

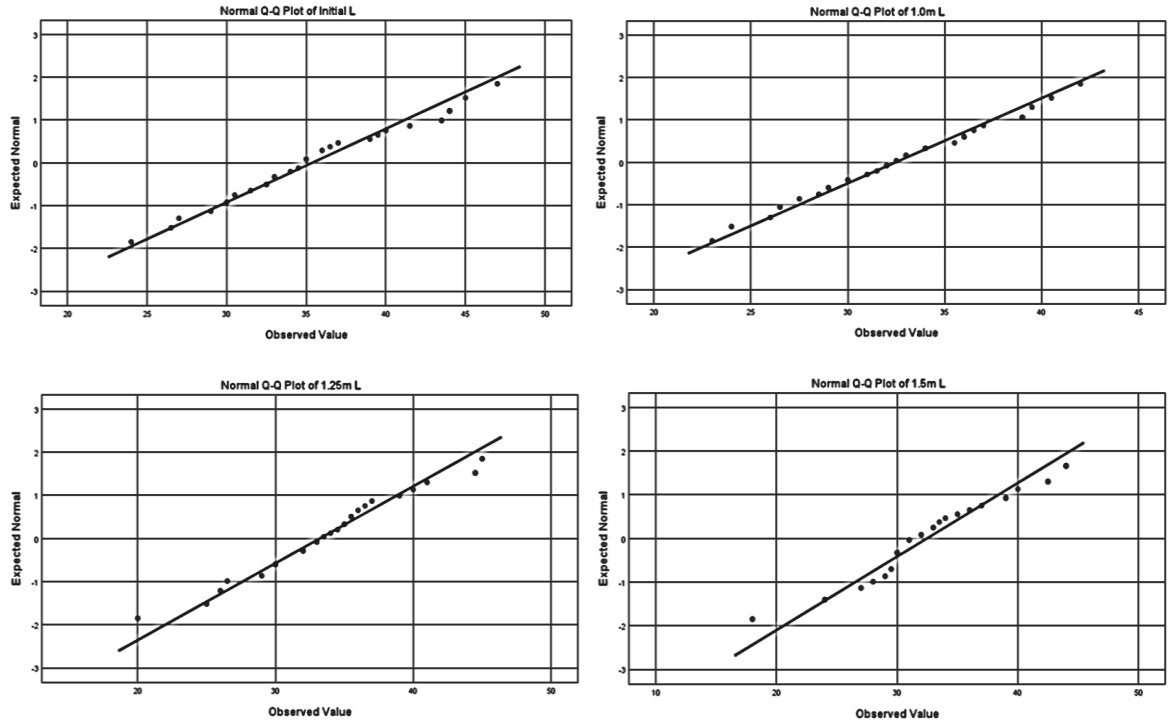


Fig. 10: Normal Q-Q plot graphs for left hand grip strengths data.

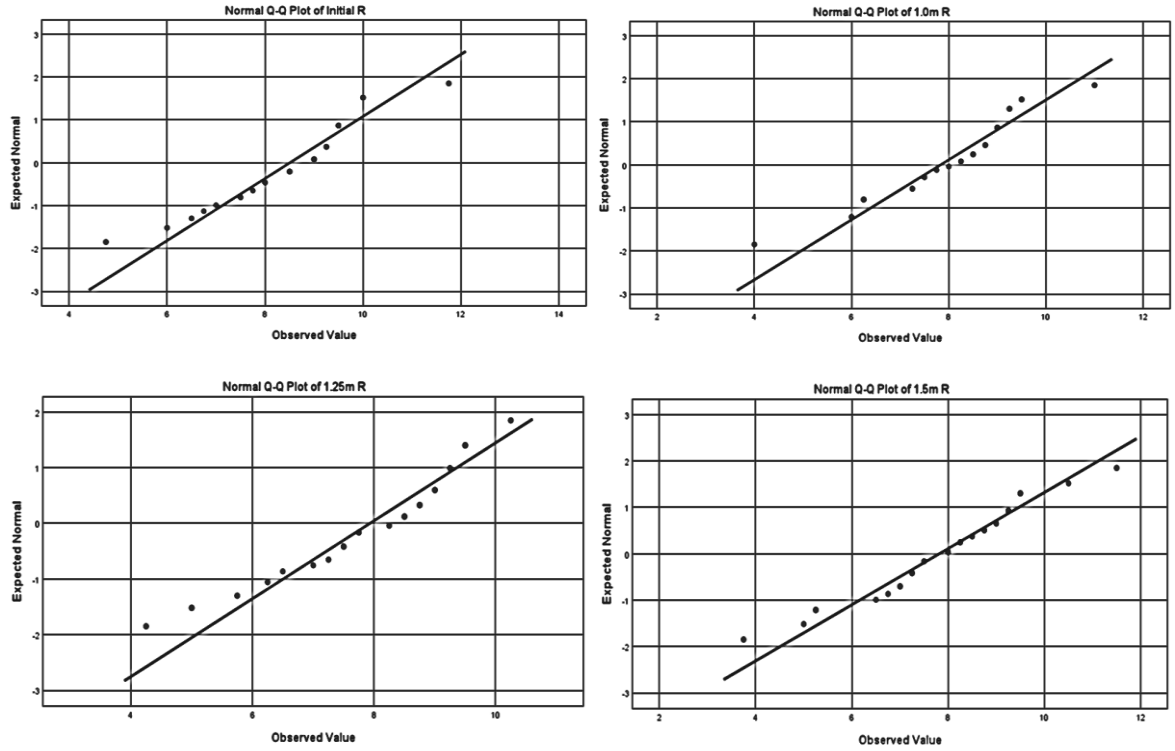


Fig. 11: Normal Q-Q plot graphs for right hand pinch strengths data.

https://doi.org/10.31436/iiumej.v22i2.1547

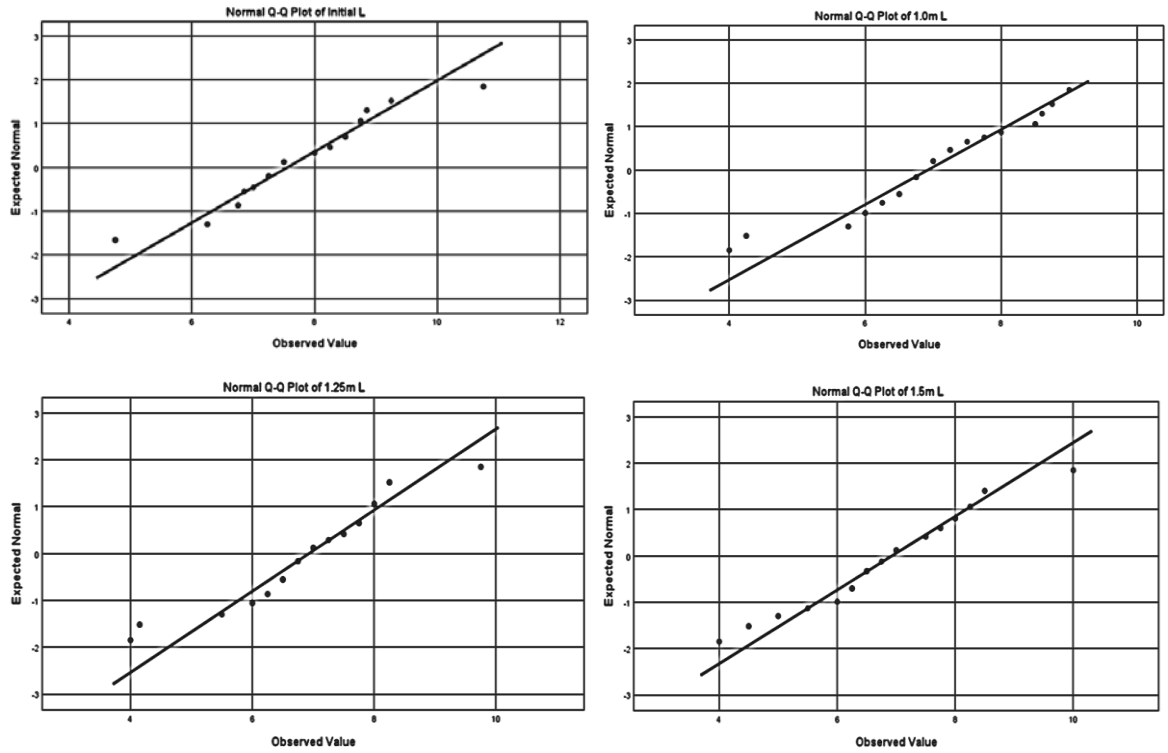


Fig. 12: Normal Q-Q plot graphs for left hand pinch strengths data.

3.2 Wilks' Lambda Test

Before comparing in terms of lateral distances, the reduction of forces between the initial hand grip strengths and pinch strengths with the hand grip and pinch strengths after the experiments had to be identified. The Wilks' Lambda test is a test statistic used in MANOVA to test whether there are differences between the means of identified groups of subjects on a combination of dependent variables [16]. Hence, Wilks' Lambda score generated in SPSS had been chosen to identify whether there are any differences between the initial readings and the results readings after the experiments. Both hand grip and pinch strengths data had been investigated separately for two sections which are the right hands (Table 6) and the left hands (Table 7). In this reduction of forces MANOVA test, the hypothesis was fixed as below:

H_0: There are no significance differences in the hand grip and pinch strengths data between the initial measurements and after-task measurements.

H_1: There are significance differences in the hand grip and pinch strengths data between the initial measurements and after-task measurements.

This hypothesis was set for both right-hand and left-hand hand grip and pinch strength measurements reduction of forces MANOVA tests.

The result displayed by Wilks' Lambda test had shown that the significance value which was the p-value of the data was 0 for both right hand and left hand, which was smaller than 0.05 ($p < 0$), indicated that there were statistically significance differences in the hand grip and pinch strength data between the initial measurements and after the lifting task experiment measurements for both hands, therefore reject H_0.

Table 6: Identification of differences between the initial readings and the results readings in right-hand grip and pinch strengths

		Multivariate[a,b]				
Within Subjects Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**
Distance	**Pillai's Trace**	0.571	11.583	6.000	174.000	0.000
	Wilks' Lambda	0.430	15.036[c]	6.000	172.000	0.000
	Hotelling's Trace	1.322	18.722	6.000	170.000	0.000
	Roy's Largest Root	1.320	38.267[d]	3.000	87.000	0.000

Table 7: Identification of differences between the initial readings and the results readings in left hand grip and pinch strengths

		Multivariate[a,b]				
Within Subjects Effect		**Value**	**F**	**Hypothesis df**	**Error df**	**Sig.**
Distance	**Pillai's Trace**	0.522	10.246	6.000	174.000	0.000
	Wilks' Lambda	0.489	12.308[c]	6.000	172.000	0.000
	Hotelling's Trace	1.019	14.440	6.000	170.000	0.000
	Roy's Largest Root	0.995	28.868[d]	3.000	87.000	0.000

3.3 Percentage Reductions

The calculation of reductions of forces between the initial hand grip and pinch strength readings with the respective distances hand grip and pinch strength readings of 1.0 m, 1.25 m, and 1.5 m was generated in an Excel file spreadsheet in terms of percentages as shown in Table 8. The calculation of percentage reductions of forces was using the formula as Eq. (1) below.

$$\frac{R_L - R_I}{R_I} \times 100\% = \%\ of\ reduction \quad (1)$$

Where:

R_L = Larger readings calculated from data measurements,

R_I = Initial readings of the data.

3.4 Removal of Outliers

Besides normality testing, the other assumption needed to be ensured is the exception of outliers for the data before the MANOVA test starts running. The outliers in the data set were categorized as the negative values of the percentage of forces reductions. The data were analyzed manually by excepting all the cases and data set that involved the negative value of the percentage of forces reduction in both hand grip and pinch strengths for both hands by using the filter command in SPSS. There was a total of 7 cases that showed the filter number of "0" which had negatives values in the percentage of reduction in hand grip and pinch strength measurements. These cases with filter number "0" as shown in Table 8, were considered outliers and did not include in the next stages of the analysis.

Table 8: Percentages of reduction of forces of hand grip and pinch strength tabulation

Right Hand Grip (%)			Right Pinch (%)			Left Hand Grip (%)			Left Pinch (%)			Filter
1.0m	1.25m	1.5m	1.0m	1.25m	1.5m	1.0m	1.25m	1.5m	1.0m	1.25m	1.5m	
2.63	5.26	7.89	8.11	5.41	0.00	10.77	1.54	7.69	0.00	11.11	3.70	1
7.69	13.19	34.07	2.63	5.26	2.63	8.75	11.25	12.50	7.03	13.51	16.22	1
17.95	20.51	7.69	21.88	12.50	9.38	-3.33	0.00	0.00	14.29	7.14	10.71	0
7.79	19.48	3.90	2.78	5.56	22.22	16.67	0.00	0.00	2.86	11.43	2.86	1
2.82	7.04	7.04	2.78	5.56	2.78	4.76	4.76	4.76	6.90	3.45	10.34	1
1.12	11.24	12.36	10.00	5.00	-5.00	-2.86	0.00	4.29	1.46	1.46	12.41	0
18.18	3.03	15.15	2.94	14.71	11.76	-1.43	-2.86	14.29	7.41	3.70	0.00	0
13.92	11.39	6.33	5.26	2.63	7.89	13.33	0.00	0.00	12.12	6.06	3.03	1
4.48	8.96	5.97	16.67	0.00	19.44	5.17	10.34	6.90	18.08	15.25	15.25	1
10.00	0.00	8.75	22.58	3.23	9.68	8.86	7.59	6.33	6.90	3.45	13.79	1
14.49	4.35	2.90	3.33	16.67	10.00	13.51	10.81	8.11	8.00	12.00	20.00	1
9.38	3.13	12.50	0.00	3.13	9.38	0.00	1.89	9.43	3.33	3.33	6.67	1
5.88	5.88	8.82	11.76	8.82	2.94	2.86	4.29	17.14	3.70	0.00	7.41	1
1.10	5.49	8.79	7.41	3.70	3.70	6.02	3.61	3.61	11.11	3.70	3.70	1
10.00	11.25	6.25	0.00	0.00	0.00	13.04	7.25	13.04	11.76	5.88	8.82	1
12.00	8.00	6.67	5.26	7.89	5.26	18.57	8.57	5.71	5.88	2.94	5.88	1
16.92	3.08	9.23	23.68	21.05	15.79	21.31	1.64	3.28	25.00	21.88	-3.13	0
0.00	5.56	0.00	7.89	2.63	2.63	6.85	5.48	9.59	17.65	8.82	11.76	1
5.88	8.82	10.29	0.00	4.17	12.50	1.85	7.41	11.11	10.53	12.63	5.26	1
5.13	2.56	20.51	7.69	23.08	23.08	6.67	17.78	28.89	10.34	27.59	24.14	1
13.33	4.76	16.19	10.81	10.81	13.51	7.95	-1.14	11.36	6.90	6.90	10.34	0
0.00	0.00	2.27	10.53	2.63	10.53	17.02	4.26	6.38	-3.13	25.00	12.50	0
2.38	4.76	1.19	6.38	12.77	2.13	15.91	19.32	0.00	16.28	9.30	6.98	1
8.33	1.39	5.56	14.29	10.71	25.00	0.00	0.00	-9.09	10.00	13.33	6.67	0
19.23	15.38	23.08	15.79	10.53	21.05	4.17	16.67	25.00	15.79	21.05	15.79	1
13.25	15.66	6.02	0.00	0.00	11.11	3.08	10.77	9.23	0.00	5.88	0.00	1
6.76	5.41	2.70	3.33	0.00	3.33	5.88	0.00	17.65	3.70	3.70	7.41	1
3.80	6.33	16.46	6.25	6.25	6.25	8.33	2.78	11.11	3.45	6.90	3.45	1
24.05	22.78	2.53	2.70	2.70	0.00	12.12	19.70	6.06	6.67	13.33	6.67	1
3.23	6.45	1.08	8.82	0.00	2.94	9.20	5.75	2.30	0.00	8.57	5.71	1

3.5 One-way Repeated Measure Multivariate Analysis of Variance (MANOVA)

This stage of analysis used repeated measure one-way MANOVA for analysis of data in order to achieve the third objective of this project which is to investigate and compare the differences between the hand grip and pinch strength after carrying out lateral lifting tasks in terms of lateral transfer distances. The analysis was carried out using SPSS with the factor of 3 which is % reduction in 1.0 m; % reduction in 1.25 m; and % reduction in 1.5 m to investigate the relationships between these data. The MANOVA analysis was carried out separately for two sections which the first set was the right-hand grip and pinch strengths data and the second set was the left-hand grip and pinch strengths data with the confidence level of 95%. In this MANOVA test, the hypothesis was fixed as below:

H_0: There are no significance differences in the percent reduction of hand grip and pinch strengths between lateral distances of 1.0 m, 1.25 m, and 1.5 m.

H_1: There are significance differences in the percent reduction of hand grip and pinch strengths between the lateral distances of 1.0 m, 1.25 m, and 1.5 m.

This hypothesis was set for both right-hand and left-hand hand grip and pinch strength measurements MANOVA tests.

Table 9: Percentage reduction of right-hand grip and pinch strength measurements in three different distances

Within Subjects Effect		Value	F	Hypothesis df	Error df	Sig.
Distance	**Pillai's Trace**	0.053	0.764	4.000	112.000	0.551
	Wilks' Lambda	0.947	0.759c	4.000	110.000	0.554
	Hotelling's Trace	0.056	0.753	4.000	108.000	0.558
	Roy's Largest Root	0.053	1.482d	2.000	56.000	0.236

Table 10: Percentage reduction of left-hand grip and pinch strength measurements in three different distances

Within Subjects Effect		Value	F	Hypothesis df	Error df	Sig.
Distance	**Pillai's Trace**	0.045	0.505	4.000	88.000	0.732
	Wilks' Lambda	0.956	0.495c	4.000	86.000	0.740
	Hotelling's Trace	0.046	0.484	4.000	84.000	0.747
	Roy's Largest Root	0.032	0.712d	2.000	44.000	0.496

The Wilks' Lambda test results analyzed and generated by SPSS had shown that the significance value or the p-value is 0.554, which was more than 0.05 ($p > 0.05$), means that there were no significance differences in reduction percentages for both right-hand grip and pinch strengths (Table 9) between the three distances of 1.0 m, 1.25 m, and 1.5 m, hence accept H_0.

The Wilks' Lambda test results analyzed and generated by SPSS had shown that the significance value or the p-value was 0.74, which was more than 0.05 ($p > 0.05$), concluded that there were no significance differences in the percentages of reductions in both left-hand grip and pinch strengths (Table 10) between the three distances of 1.0 m, 1.25 m, and 1.5 m, hence accept H_0.

4. DISCUSSION

4.1 Reduction of Forces

Line charts had been plotted from the data collected from the experiment by calculated the mean value of each category of right-hand and left-hand grip readings to identify the pattern for further confirmation. From the line chart (Fig. 13), the right-hand data for both hand grip and pinch strengths were having an obvious decrease trend from the initial readings compared to the others. This can be concluded that there is a reduction of forces between the initial and after the experiment's right-hand grip and pinch strengths readings. From the line chart (Fig. 14), the left-hand data for both hand grip and pinch strengths were having an obvious decrease trend from the initial readings compared to the others. This can be concluded that there is a reduction of forces between initial and after experiments' left-hand grip and pinch strengths readings.

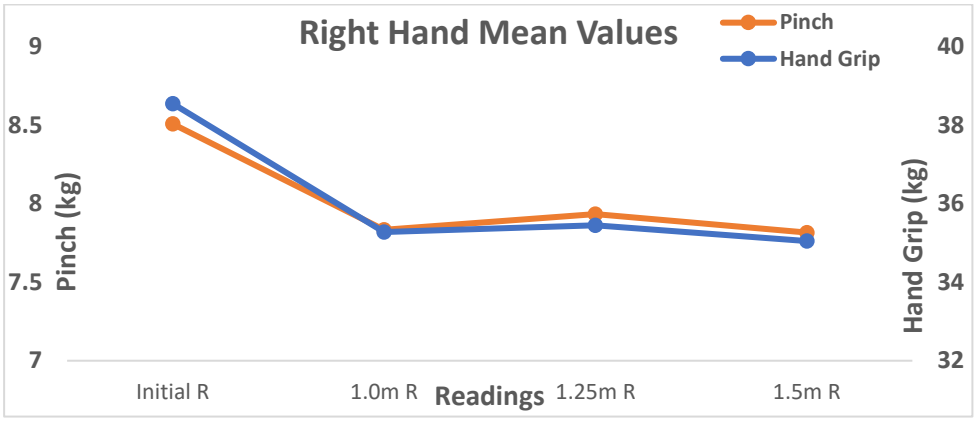


Fig. 13: Line chart for right-hand grip and pinch strengths based on the mean value for each category.

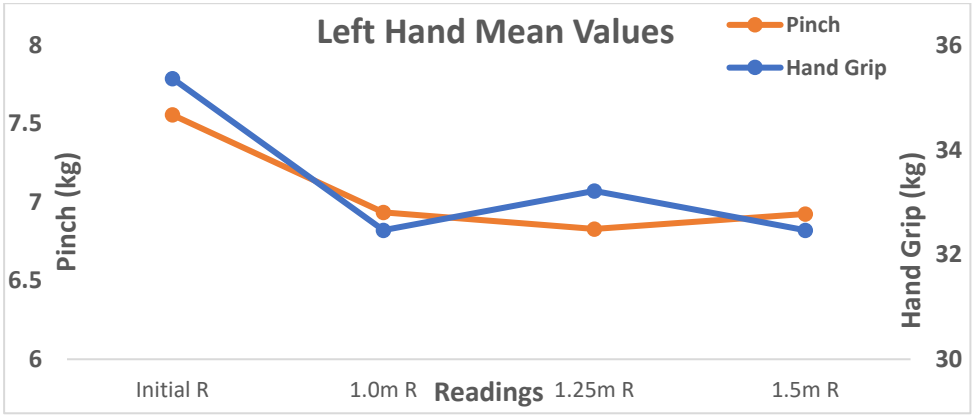


Fig. 14: Line chart for left-hand grip and pinch strengths based on the mean value for each category.

4.2 Transfer Distances

The overall results generated by using the one-way repeated measure MANOVA test and graphical charts had indicated that there was a high significance reduction of forces between the initial and the after experiments data for both hand in hand grip and pinch strengths. However, when the relationship between the three types of distances with two dependent variables of hand grip and pinch strengths were tested, the results showed that there were no significance differences in the relationship. This statistical result indicated that distances (1.0 m, 1.25 m, and 1.5 m) will not affect the hand grip and pinch strengths during the lateral lifting tasks.

The surveys carried out in each participant had shown the perceptions of participants on the distance that caused pain in hands (Fig. 15) and fingers (Fig. 16) during lateral lifting tasks during the experiment. The result showed in Figure 11 indicated that there were 56.7% of people disagreed that at 1.0m distance would cause pain in hand, while 43.3% of people agreed that 1.5 m distance would cause pain in hand. The high percentage for disagreed and neutral score at 1.25m of 43.3% respectively showed that the majority of the people did not felt pain when transferring load at 1.25 m. The percentage of disagreed and neutral at a distance of 1.5 m was in a total of 56.6%, which is higher than the percentage of agreed, which is 43.3%, indicated that more than half of the participants did not felt pain in hand at 1.5 m

https://doi.org/10.31436/iiumej.v22i2.1547

distance. The result had clearly concluded that the majority of participants did not feel any pain in either of these three distances during the lateral lifting tasks experiment.

The result in Fig. 16 had shown that the highest percentage for all type of distances were from the disagreed score, which was 63.3% at 1.0 m, 50% at 1.25 m and 43.3% at 1.5 m. The neutral score for three distances had a consistent pattern which laid around 20% to 30%. The agreed score had an inclined trend pattern followed the sequence of 1.0 m, 1.25 m, and 1.5 m which had indicated that there were less than 40% of participants agreed that they felt pain in fingers during lifting tasks in these distances. The result had clearly shown that majority of the participants did not felt pain in fingers in either these three distances during lateral lifting tasks experiments.

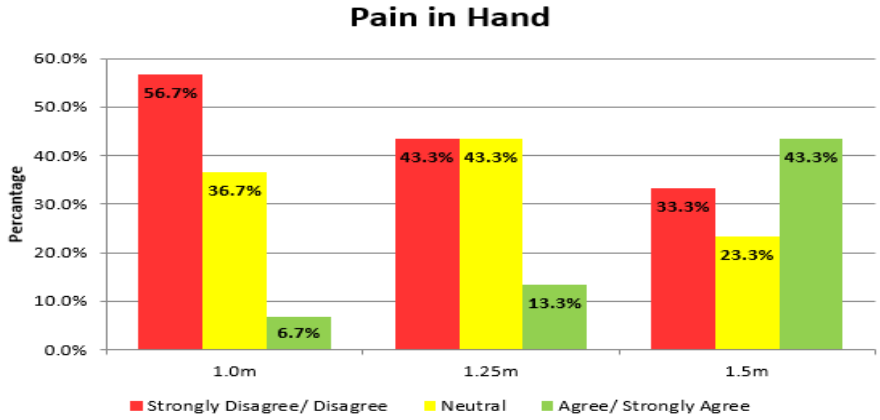


Fig. 15: Perceptions of participants on the distance that cause pain in hands during lateral lifting tasks.

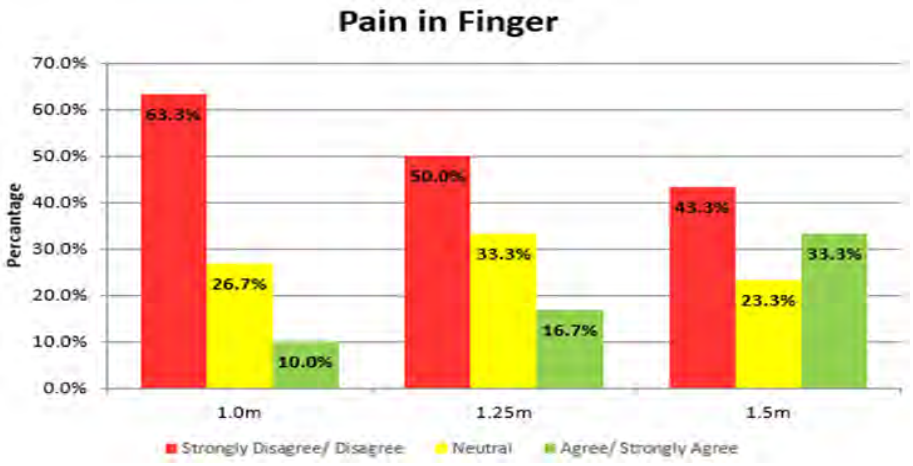


Fig. 16: Perceptions of participants on the distance that cause pain in fingers during lateral lifting tasks.

4.3 Transfer Postures

During the experiment, the transfer process of every participant was observed to identify the way and the postures of participants carried out lateral lifting tasks. Based on Vengata et al. [17] stated that the hand grip strengths and its evaluations are affected by several factors such as age, gender, posture, fatigue, wrist, and forearm position. For example, the difference in grip strength in different posture caused by the change in length of the muscle. Lee and

Sechachalam also indicated that the wrist position can affect one's hand grip strength due to flexion [18].

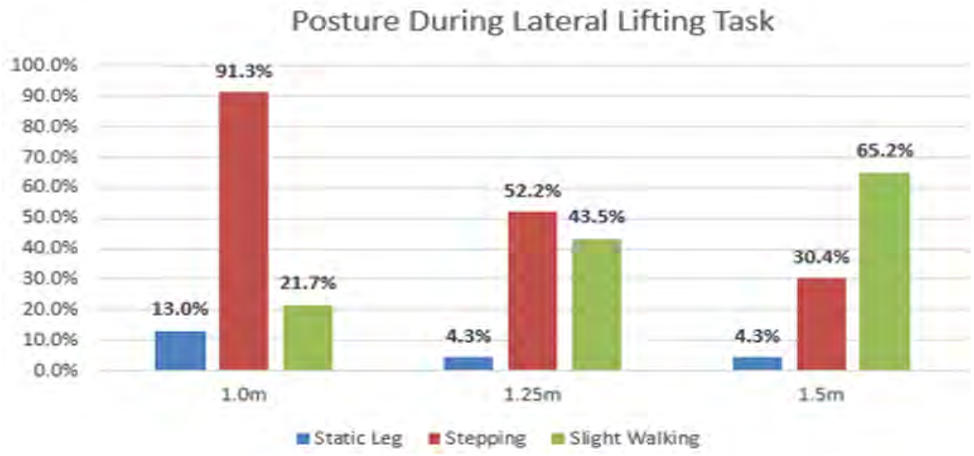


Fig. 17: The posture of participants during lateral lifting tasks.

The posture of the participants during experiments transferring load in 1.0 m, 1.25 m, and 1.5 m was observed and tabulated in an Excel file to generate a graph in Fig. 17. The result showed that the majority of the participants were stepping and slight walking during the transferring process. For every distance, the percentages of participants that were remained static when transferring the load were not more than 13%. For the distance of 1.0m, 91.3% of participants were stepping and 21.7% of participants were slight walking during the transfer process. For distance 1.25 m, about half of the participants were stepping while 43.5% of participants were slight walking during the transfer process. For 1.5 m, the percentage of participants that stepping was continuously decreased to 30.4%, while the percentage of participants that slight walking was increased to 65.2%.

The posture of stepping and slight walking would lower the participants tiring and fatigue level during lateral lifting tasks, since participants were majority stepping and walking during the lifting tasks, hence the result of reduction of hand grip and pinch strength in between three distance did not have significance differences. The high percentage of participants were stepping and slight walking in the lateral transfer tasks was influenced by the knowledge level of the participants themselves towards lateral lifting tasks.

The line graphs for the mean values of hand grip and pinch strength measurements for each distance in static leg, stepping, and slight walking postures of the participants were also generated to investigate the pattern and the trend of the graphs (Fig. 18). The results showed that the patterns were not consistence, and this had indicated that the leg and body posture did not contribute any large effect on the hand grip and pinch strengths measurements in 1.0 m, 1.25 m, and 1.5 m transfer distances.

4.4 Limitation

This study had a few limitations, such as the experiments only involved single gender which is the majority gender that performing material handling tasks in every sector, and only right-handed males can participate in the experiment. The experiment can be expanded further which involved two genders and left-handed people can also participate in the experiments. The relationship between the dominant hands and genders with lateral lifting tasks can be analyzed.

https://doi.org/10.31436/iiumej.v22i2.1547

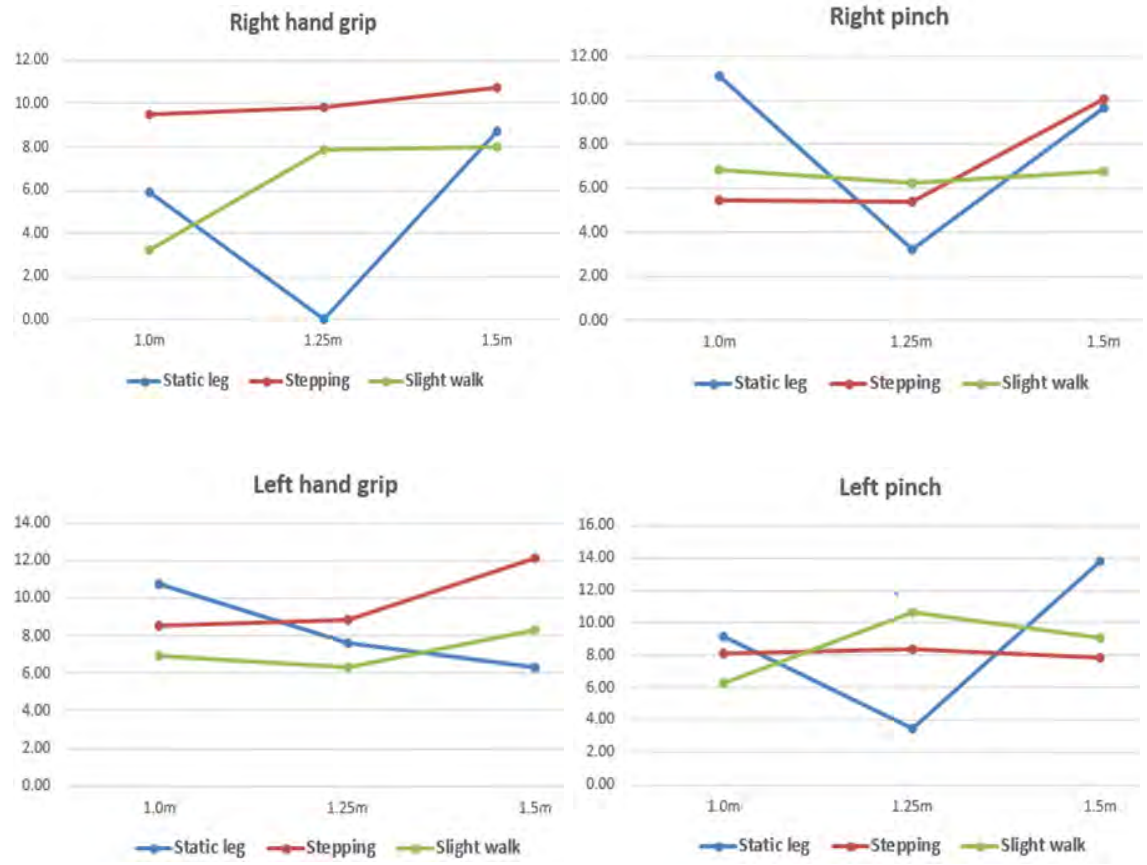


Fig. 18: The mean value of hand grip and pinch measurements for each distance in static leg, stepping and slight walking postures.

Besides, the experiments only tested on a group of participants, which could not give a better chance to make comparisons between different groups of people with specific characteristics. Increased in the number of groups to be testes based on different characteristics such as one group is educated, and the others are not can be further research to identify whether the postures and the way they carried out lateral lifting tasks are different.

Moreover, the experiment involved in this study was only tested one independent variable which is the transfer distance due to the time issues. The weight of the load can also be analyzed to see the relationship between the load weights and the transfer distances with the hand grip and pinch strengths. The factor of presence of proper grip of the load can also be investigated in two groups of participants to identify the differences in hand grip and pinch strength measurements. The postures or the way participants carry the loads can also be observed or identified to investigate the relationship between the transfer distances.

Lastly, the frequency of lateral transfer tasks in this study for the experiment only involved 6 cycles per distance. For the next research, the frequency of transferring loads was suggested to increase the frequency to identify the relationship of the repetitive factor with the hand grip and pinch strengths.

5. CONCLUSION

The experimental study involved the participants to carry out six cycle lateral lifting tasks each in three different distances of 1.0m, 1.25m, and 1.5m following the counterbalancing

sequences. The experiment was carried out in the condition of participants did not fatigue; hence participants were given one minute to rest before proceeding to the next tasks. The observations on the posture of the participants during the experiments also indicated that most of the participants were stepping and slight walking during the transfer process, hence the chances of exhausted occurred in hand were highly reduced. The survey form results also indicated that the participants did not fatigue and did not felt pain in hand during the experiments. The analysis of the results in this study had concluded that the transfer distance of 1.0 m, 1.25 m, and 1.5 m during lateral lifting tasks does not affect the hand grip and pinch strengths. However, further research was needed to identify whether the other distances especially longer distance will affect the hand grip and pinch strengths.

ACKNOWLEDGEMENTS

The authors would like to thank Universiti Teknikal Malaysia Melaka for providing the Jamar Hand Dynamometer and Pinch Meter for the experiments. Many thanks as well to all the participants for the cooperation and contributions in this study. This work was supported by the Faculty of Manufacturing Engineering Universiti Teknikal Malaysia Melaka. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

REFERENCES

[1] Shriniket M. (2018) 992 Ergonomic risk evaluation in an automobile industry for preventing musculoskeletal problems. In 32nd Triennial Congress of the International Commission on Occupational Health (ICOH): 29 April-4 May 2018; Dublin. Edited by BMJ Publishing Group Ltd; pp A267. http://dx.doi.org/10.1136/oemed-2018-ICOHabstracts.756.

[2] Advisory Committee from Hospital Pharmacy. (2018) Summaries of safety labeling changes approved by FDA- Boxed warnings highlights July-September 2017. 53(1): 18-24. doi: 10.1177/0018578717743812.

[3] Phil H, Ferrett Ed. (2011) Workplace hazards and risk control. In introduction to health and safety at work. 5th Ed. London, Routledge; pp 193–227. doi: 10.4324/9780080970714.

[4] Veselina N, Iskra SM. (2018) Work-related musculoskeletal disorders (WMSDs): risk factors, diagnosis and prevention. Scripta Scientifica Salutis Publicae, 4: 15. doi: 10.14748/sssp.v4i0.5076.

[5] Tae HK, Jay PM, Monica RW, Steven AL. (2014) The effects of transfer distance on spine kinematics when placing boxes at different heights. Applied Ergonomics, 45(4): 936-942. https://doi.org/10.1016/j.apergo.2013.11.010.

[6] Charles DR. (2008) Office ergonomics. Industrial safety and health for administrative services. CRC Press, 1st edition; pp 297-312.

[7] Copsey SM, Elke S. (2018) 1167 Mainstreaming gender into occupational safety and health (osh) practice. Women, Health and Work. BMJ Publishing Group Ltd, doi: 10.1136/ oemed-2018-icohabstracts.1494.

[8] Advisory Committee from The Center For Construction (CPWR). (2018) Fatal and nonfatal injuries - Musculoskeletal disorders in construction and other industries. CPWR, 6th Ed.

[9] Snooks H, Foster T, Nicholl J. (2004) Results of an evaluation of the effectiveness of triage and direct transportation to minor injuries units by ambulance crews. Edited by Emerg MJ, 21(1): 105-111. doi: 10.1186/isrctn21222541.

[10] Ven-Stevens LAWVD, Maud JLG, Ruud WS, Ton ARS. (2015) Instruments for assessment of impairments and activity limitations in patients with hand conditions: A European Delphi study. Journal of Rehabilitation Medicine, 47(10): 1-9. doi: 10.2340/16501977-2015.

[11] Jaffar N, Abdul-Tharim AH, Mohd-Kamar IF, Lop NS. (2011) A literature review of ergonomics risk factors in construction industry. Procedia Engineering, 20: 89-97. doi:10.1016/j.proeng.2011.11.142.

[12] Mathiowetz V, Keren WBS, Gloria VBS, Nancy KBS. (1984) Reliability and validity of grip and pinch strength evaluations. The Journal of Hand Surgery, 9(2): 222-226, doi: 10.1016/S0363-5023(84)80146-X.

[13] Petersen P, Petrick M, Connor H, Conklin D. (1989) Grip strength and hand dominance: Challenging the 10% rule. American Journal of Occupational Therapy, 43(7): 444-447. doi: 10.5014/ajot.43.7.444.

[14] Zeelenberg R, Pecher D. (2015) A method for simultaneously counterbalancing condition order and assignment of stimulus materials to condition. Behavior Research Methods, 47(1): 127-133, doi: 10.3758/s13428-014-0476-9.

[15] Nornadiah MR, Bee WY. (2011) Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. Journal of Statistical Modeling and Analytics, 2(1): 21-33.

[16] Helen P, Vanessa S, Elizabeth W, Heidi R. (2000) A comparison of nursing competence of graduates and diplomates from UK nursing programmes. Journal of Clinical Nursing, 9(3): 369-381. https://doi.org/10.1046/j.1365-2702.2000.00331.x.

[17] Vengata SM, Subramanian GS, Isaac J. (2015) Factors affecting hand grip strengths and its evaluation: A systemic review. International Journal of Physiotherapy and Research, 3(6): 1288–1293.

[18] Lee JA, Sechachalam S. (2016) The effect of wrist position on grip endurance and grip strength. The Journal of Hand Surgery, 41(10): 367-373. doi: 10.1016/J.JHSA.2016.07.100.

https://doi.org/10.31436/iiumej.v22i2.1640

EFFECT OF MACHINING PARAMETERS ON SURFACE ROUGHNESS, POWER CONSUMPTION, AND MATERIAL REMOVAL RATE OF ALUMINIUM 6065-SI-MWCNT METAL MATRIX COMPOSITE IN TURNING OPERATIONS

Savina Jaddinagadhe Puttaswamy* and
Raghavendra Bommanahalli Venkatagiriyappa
Department of Mechanical Engineering, JSS Academy of Technical Education, Bangalore-560060, India

**Corresponding author: mailtosavin@gmail.com*

(Received: 4th October 2020; Accepted: 25th March 2021; Published on-line: 4th July 2021)

ABSTRACT: Nanocomposites were prepared with Al-6065-Si and multi walled carbon nanotubes of 1 wt.% as reinforcement through the stir-casting method. Fabricated nanocomposites were machined on a lathe machine using a tungsten carbide tool. The study investigated the multi-objective optimization of the turning operation. Cutting velocity, feed, and depth of cut were considered for providing minimum Surface Roughness of the workpiece. Also, the power consumed by the lathe machine with maximum metal removal rate was examined by surface response methodology. The design of experiments was developed based on rotational central composite design. Analysis of variance was executed to investigate the adequacy and the suitable fit of the developed mathematical models. Multiple regression models were used to represent the relationship between the input and the desired output variables. The analysis indicates that the feed is the most influential factor that effects the surface roughness of the workpiece. Cutting speed and the depth of cut are two other important factors that proportionally influence the power consumed by the lathe tool as compared to the feed rate.

ABSTRAK: Komposit nano disediakan bersama Al-6065-Si dan karbon nanotiub berbilang dinding sebanyak 1 wt.% sebagai bahan penguat melalui kaedah kacauan-tuangan. Komposit nano yang terhasil melalui mesin pelarik ini menggunakan alat tungsten karbida. Kajian ini merupakan pengoptimuman pelbagai objektif operasi pusingan. Kelajuan potongan, suapan dan kedalaman potongan diambil kira sebagai pemberian minimum pada kekasaran permukaan bahan kerja. Tenaga yang digunakan bagi mesin pelarik dengan kadar maksimum penyingkiran logam diteliti melalui kaedah tindak balas permukaan. Rekaan eksperimen yang dibangunkan ini adalah berdasarkan rekaan komposit pusingan tengah. Analisis varian telah dijalankan bagi mengkaji kecukupan dan penyesuaian lengkap bagi model matematik yang dibangunkan. Model regresi berganda digunakan bagi menunjukkan hubungan antara input dan pembolehubah output yang dikehendaki. Analisis menunjukkan pemberian suapan merupakan faktor mempengaruhi keberkesanan kekasaran permukaan bahan kerja. Kelajuan pemotongan dan kedalaman potongan adalah dua faktor penting lain yang mempengaruhi kadar langsung ke atas tenaga yang digunakan oleh mesin pelarik dibandingkan kadar pemberian suapan.

KEYWORDS: *MWCNT (multi-walled carbon nanotube); power consumption; machining; cutting force; central composite design; response surface methodology*

1. INTRODUCTION

Carbon-nanotubes reinforced with AMMC's in the recent technological era are gaining high importance among various classes of composites. Researchers have made attempts to produce MMCs with CNT as reinforcement material. Core research on metals such as aluminum and aluminum alloy materials [1,2], due to their light weight, high strength composites, are sought for dynamic mechanical systems like aerospace industries, robotics and automobiles [3]. Materials with high stiffness, modulus, strength, low density, and high specific surface are associated in carbon nanotubes as an ideal reinforcement [4].

The presence of hard MWCNT reinforcement and its abrasive nature affects the quality of the workpiece resulting in difficulty of machining and wear of the cutting tool [5]. The traditional machining process, if adopted in composite materials with appropriate tool design and the optimal operating conditions can resolve this difficulty. These raw materials are very expensive. In order to reduce their wastage during the final conversion of the composites into engineering products and to obtain the required geometrical parts, it becomes necessary to investigate and study the basic manufacturing and machinability of the material at the time of design of material and selection [6,7]. The machining of composites depends mainly on their various reinforcements and diverse matrix properties. The cutting tool will be alternatively in contact with the base metal and the reinforcing materials and hence the response obtained by the machining of composite materials can be completely different. Thus, the cutting tool's material, geometry, and wear resistance play critical roles for the machining of composite materials. Correspondingly, the different machine operating parameters like the feed, cutting velocity, depth of cut, and other factors such as tool geometry, machining system stability, lubrication, and proper cutting tool selection, all play crucial roles [8]. Considering these many factors, it becomes very difficult to attain fine surface finish and high metal removal rate (MRR). The feed, cutting speed and the depth of cut are those parameters that can be controlled, and proper selection of these parameters yields proper surface finish to the MMCs [9].

The machining operations are classified under two important categories: cutting and grinding process. Process flexibility, yield time, high material removal rate and good surface finish are found to be salient features in the turning process. The proper prediction of cutting forces in the turning process is the primary task to achieve along with high dimensional accuracy and suitable machining system stability [10]. In the industry, metal removing processes are used to get the desired shape and dimension with precise quality. The process that removes metal at a higher rate and power consumed is considered to obtain an economical process [11]. Measuring the power consumption in the metal-removal-by-cutting-tool operation helps for designing machine components, increasing the life of the tool for high productivity, and managing the capacity required by the motor for machine. The objective of the research is to analyze the percentage contribution of machining parameters like feed, cutting velocity, and the depth of cut of the developed MMC on surface roughness, material removal rate, and power consumed using RSM.

2. EXPERIMENTAL DETAILS

2.1 Fabrication of MWCNT-Si-Al Matrix of 1 wt.% by Stir Casting Method

The Al-6065 ingot-castings were placed in the electric furnace and the temperature of the crucible in the furnace was raised and maintained at a temperature of 750 ^{o}C for about 20 minutes. This process was carried out to convert the material into a molten state. The

molten metal was stirred continuously for about 15 minutes. The CNT powder was slowly added into the molten metal.

The stirring created a uniform mixture of the reinforced particle of 1 wt.% of MWCNT and 4% of silicon into molten mix. The molten metal was poured into the metal die and left for solidification. The sample of the specimen was considered for further machining processes [12].

2.2 Turning Machine and the Cutting Conditions

The experiments were conducted on a heavy-duty precision lathe machine as shown in Fig. 1, with a tungsten carbide tool. The cutter being the single point cutting tool, fed right into the rotating workpiece and cut the material as chips. This process was carried out to create the desired shape.

A full bridge strain gauge dynamometer was used to measure the cutting forces. This analog device is highly capable of measuring the cutting force while the turning operations are executed. This data of measured cutting force was utilized for further analysis [13]. The specimen of Ø25mm × Ø75mm size were used for the experimentation.

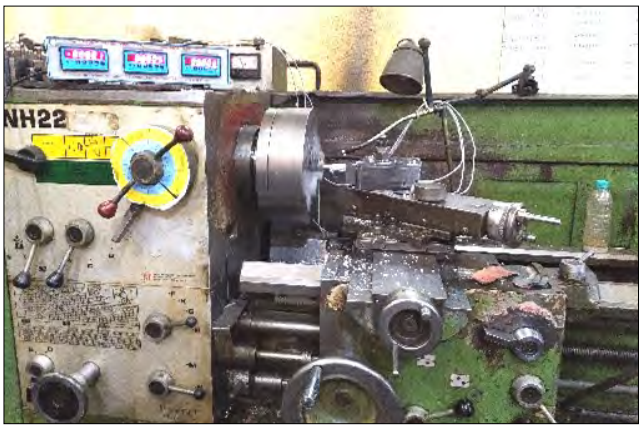


Fig. 1: Heavy-duty precision lathe machine.

2.3 Design of Experiment (DoE) by Response Surface Methodology (RSM)

A group of statistical and mathematical techniques that helps in modelling and analysis of a problem is termed as RSM. The output or the response of it can be controlled and is affected by several input variables. The objective here is to determine the correlation between the responses and variables determined.

RSM is one of several DoE methods, which is efficient for analyzing and planning the problem when certain independent variables will influence the dependent variables or when several obtained responses should yield valid and objective conclusions. Rotatable central composite design (CCD) and Box-Behnken design are the widely used and suitable types of RSM methods that are readily available for investigation purposes [14]. The embedded factorial design is not found in the Box-Behnken design and is an independent quadratic model. This design is normally incorporated while executing the non-sequential investigations. The design uses certain combinations for the treatment at the corners of the face center, the process space and at the center of the design body. The design requires three levels of each factor that are near to rotatable.

As compared to the CCD methods, the Box-Behnken model has less capability for the orthogonal blocking [15]. Thus, CCD methodology is the most commonly used technique

in case of the second-order response models. The CCD provides very few numbers of experiment and, with a rotatable feature, the optimal response can be obtained [16]. Experimental independent variables along with their coded levels for the central composite design (CCD) are represented in Table 1.

Table 1: Experimental independent variables and their coded levels

Independent variable	Levels of variables		
	Low	Medium	High
Cutting velocity (m/min)	30	50	70
Feed (mm/rev)	0.2	0.3	0.4
Depth of cut (mm)	0.5	0.85	1.2

2.4 Surface Roughness (R_a) and Its Measurement

R_a is a subjective property that indicates surface roughness. It is measured in micrometers, and has a crucial characteristic that quantifies high frequency deviations from that of an ideal surface. R_a is the arithmetic mean value. It is based on the mean of normal deviations from a nominal surface. It is generally specified over the "cut-off" length and is represented in Eq. (1).

$$R_a = \frac{1}{n}\sum_{i=1}^{n} y_i \qquad (1)$$

where, R_a is surface roughness,
n is number of measurement points and
y_i is surface deviation at measurement point of 'i'.

From the observation, it is noted that the values obtained are affected by few factors and machining parameters. The commercially available surface profilometers feature a diamond stylus that travels over the workpiece surface to measure the surface roughness [17].

2.5 Power Consumption (P_c)

A dynamometer is a measuring device used to measure tangential force from which power consumed by the machine can be calculated. The product of tangential force and the cutting velocity results in cutting power consumption (P_c). It is represented by Eq. (2). The experiments were conducted using CCD by RSM [18].

$$P = F_z \times v \qquad (2)$$

where, P is power in kilowatts,
F_z is force in newton and
v is cutting speed in meter per second.

2.6 Metal Removal Rate (MRR)

The metal that is removed per unit time is the MRR. The SI unit of MRR is mm^3sec^{-1}. For every revolution associated with the material, a ring-shaped chip of the material is taken out. The MRR can be obtained using Eq. (3).

$$MRR = v \times f \times d \qquad (3)$$

where, v is the cutting velocity in m/min,
f is the feed in mm/rev, and
d is the depth of cut in mm.

3. RESULTS AND DISCUSSION

3.1 Validation for Surface Roughness by Response Surface Methodology

The experiments were executed according to CCD, that was generated by response surface methodology (RSM) with three input variables: cutting speed (v), feed (f) and depth of cut (d). Twenty experimental runs were executed with 95% confidence level. The readings of trial runs were measured for the surface roughness using the surface roughness tester measuring machine and the cutting force using the lathe tool dynamometer. The response output obtained for MRR, experimental R_a, model R_a, and % of error is tabulated in Table 2. Regression coefficients obtained from the experimental results are shown in Table 3 for polynomial regression equation of surface roughness (R_a). Table 4 shows the corresponding analysis of variance and is obtained using Minitab-14 software.

Table 2: Experimental runs and surface roughness responses

Sl. No	Input Variables			Output Responses				
	Cutting velocity (v)	Feed (f)	Depth of cut (d)	MRR ($mm^3\ min^{-1}$)	Exp. (R_a)	Model (R_a)	R_a (Diff.)	% Error
1	30	0.2	1.20	120.000	1.34	1.360	0.020	1.559
2	70	0.2	0.50	116.667	1.68	1.865	0.185	11.058
3	30	0.4	0.50	100.000	2.8	3.023	0.223	7.977
4	70	0.4	0.50	233.333	3.12	3.324	0.204	6.543
5	30	0.4	1.20	240.000	3.1	3.138	0.038	1.254
6	30	0.2	0.50	50.000	1.93	1.935	0.005	0.278
7	70	0.2	1.20	280.000	0.96	0.960	0.0008	0.093
8	50	0.3	0.85	212.500	3.86	3.918	0.058	1.527
9	50	0.3	0.85	212.500	3.91	3.918	0.008	0.229
10	50	0.3	0.85	212.500	3.87	3.918	0.048	1.265
11	70	0.4	1.20	560.000	2.89	3.109	0.219	7.587
12	50	0.3	0.85	212.500	3.84	3.918	0.078	2.056
13	17.34	0.3	0.85	73.695	2.86	2.786	0.073	2.583
14	50	0.4633	0.85	328.171	3.64	3.391	0.248	6.830
15	50	0.3	0.85	212.500	3.91	3.918	0.008	0.229
16	82.66	0.3	0.85	351.305	2.82	2.705	0.114	4.073
17	50	0.3	1.42155	355.388	2.41	2.407	0.002	0.088
18	50	0.3	0.27845	69.613	3.26	3.052	0.207	6.368
19	50	0.3	0.85	212.500	3.9	3.918	0.018	0.486
20	50	0.1367	0.85	96.829	0.71	0.748	0.038	5.472

Table 3: Estimated regression coefficients for R_a in μm

Term	Coeff.	SE Coeff.	T	P
Constant	-8.1365	0.95560	-8.514	0.000
v	0.1049	0.01682	6.237	0.000
f	43.1855	3.57327	12.086	0.000
d	4.7344	0.95350	4.965	0.001
v^2	-0.0011	0.00012	-9.384	0.000
f^2	-69.3314	4.73106	-14.655	0.000
d^2	-3.6393	0.38621	-9.423	0.000
v x f	0.0463	0.03039	1.522	0.159
v x d	-0.0118	0.00868	-1.357	0.205
f x d	4.9286	1.73676	2.838	0.018
S = 0.171930	R^2 = 98.53%	R^2 (pred.) = 88.95%	R^2 (adj.) = 97.21%	

Table 4: Analysis of variance for R_a in µm

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Regression	9	19.8663	19.8663	2.20737	74.67	0.000
Linear	3	9.2669	4.6496	1.54985	52.43	0.000
Interaction	3	10.2385	10.2385	3.41282	115.45	0.000
Square	3	0.3610	0.3610	0.12032	4.07	0.040
Residual Error	10	0.956	0.2956	0.02956		
Lack-of-Fit	5	0.2913	0.2913	0.05826	68.01	0.000
Pure Error	5	0.0043	0.0043	0.00086		
Total	19	20.1619				

The model developed was investigated using the analysis of variance technique. The value of the determination coefficient (R^2 = 95.00%) specified that only the values that are less than the 5% of total variations were not explained by the developed model and this indicated the quality of the model's fit. The value of the adjusted determination coefficient (adj. R^2 = 0.9500) was high and specifies that the obtained model was highly significant. If P-value is less than F-value (P<F), it indicates that the model is highly significant. The results of R_a as obtained using ANOVA is represented in Table 4. It describes the predictability of the obtained model for its surface roughness (R_a) to be at a 99% confidence level. This is due to the average value of 'P' which is less than 0.0001. Figure 2 shows that the experimental output, surface roughness (R_a) is closer to the regression model obtained using RSM method. The R^2 was found to be 98.53% and shows that the regression equation is highly reliable [21]. The lower value of 'P = 0.01' indicates the statistic significance of the model and concludes that the proposed model is a correct one.

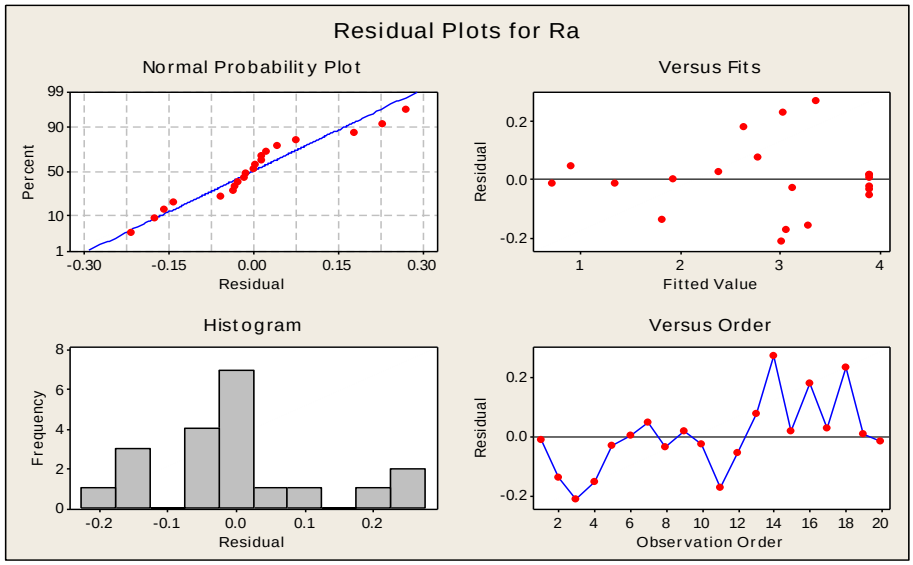


Fig. 2: Residual plots for R_a in µm.

Residual plots for R_a shown in Fig. 2 have been analyzed and are discussed. The probability plot was used to check for the data normality. The distribution of the data points all along the normal line with very few outliers indicated that the data is normally distributed. The residual versus fitted values, as shown in the second plot, do not show any subsequent trends. This indicates the chosen RSM model fits well within the given dataset. The third plot of frequency verses residual illustrates the normal distribution of residuals that are produced by the model and assures that the assumptions made are reasonable and the choice of the model is appropriate. The residue verses observation order as depicted in the last plot

highlights the presence of random data points. The randomness in the plot shows the insignificant experimental order in comparison to first surface response (R_a). A confidence level of 95% has been applied to RSM for R_a. The statistical significance of the surface roughness has been measured and concluded based on all the factors and respective interactions which has a probability value or P-Value to be less than 0.05.

3.2 Optimized Solution for Surface Roughness

The lower values, the upper values and the final target values of R_a responses are fed to the response optimizer tool of Minitab by giving equal weightages. The description of the same is given in Table 5.

Table 5: Response optimization

Goal	Lower	Target	Upper	Weight	Import
Optimization of R_a	0.71	1.1	3.91	1	1

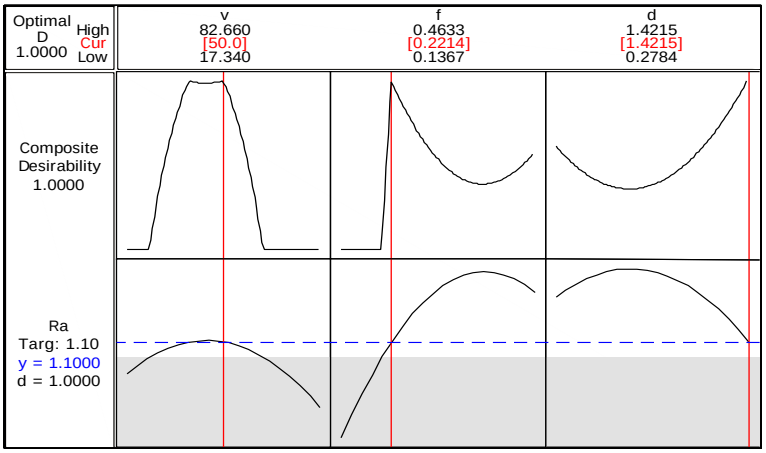


Fig. 3: Predicted response for R_a in μm.

A significant solution of about 100% composite desirability is provided by the software. This is concluded from the analysis of data of experiments. About 50 m/min cutting velocity, 0.2214 mm/rev of feed rate, and 1.4215 mm of DOC along with 1.1 μm of minimum surface roughness is reached. Lower half of Fig. 3 shows the of variation of R_a along with focused input variables. The blue dotted lines indicate the pin-pointing desired optimal points [19].

3.3 Validation for Power consumption by RSM

The experiments were conducted using three input variables such as cutting speed (v), feed rate (f), depth of cut (d) for 20 experimental run and obtain the output responses as cutting forces (F_z) in z direction, power consumption (P_c), difference of power consumption and percentage error are shown in the Table 6. The output of the experiments was measured using tool dynamometer for cutting forces and calculated the power consumption. Regression coefficients are as shown in Table 7 for polynomial regression equation of power consumption (P_c) and Table 8 shows the corresponding analysis of the variance.

The ANOVA results of the P_c is shown in Table 8. It indicates the obtained model for the surface roughness having probability of 99% confidence level. This is due to the average value of P-value, which is less than 0.0001. The R^2 is 99.40% and provides a highly reliable regression equation [21].

Table 6: Experimental run and power consumption results

Sl. No	Input Parameters			Output Responses				
	Cutting velocity (v)	Feed (f)	Depth of cut (d)	F_z (in N)	P_c (in W)	Y (P_c)	P_c (Difference)	% Error
1	30	0.2	1.2	107.910	53.955	51.12824	2.82676	5.239107
2	70	0.2	0.5	49.050	57.225	55.38798	1.83702	3.21017
3	30	0.4	0.5	78.480	39.240	38.20172	1.03828	2.645973
4	70	0.4	0.5	88.290	103.005	107.0857	4.08072	3.961672
5	30	0.4	1.2	166.770	83.385	86.47904	3.09404	3.710547
6	30	0.2	0.5	49.050	24.525	27.37598	2.85098	11.62479
7	70	0.2	1.2	117.720	137.340	139.6482	2.30824	1.680676
8	50	0.3	0.85	127.530	106.275	105.5303	0.74475	0.700776
9	50	0.3	0.85	137.340	114.450	105.5303	8.91975	7.793578
10	50	0.3	0.85	127.530	106.275	105.5303	0.74475	0.700776
11	70	0.4	1.2	186.390	217.455	215.871	1.58396	0.728408
12	50	0.3	0.85	127.530	106.275	105.5303	0.74475	0.700776
13	17.34	0.3	0.85	137.340	39.691	39.13672	0.5545442	1.397144
14	50	0.4633	0.85	147.150	122.625	117.5362	5.088793954	4.149883
15	50	0.3	0.85	117.720	98.100	105.5303	7.43025	7.574159
16	82.66	0.3	0.85	127.530	175.694	167.6571	8.0367482	4.574292
17	50	0.3	1.42155	176.580	147.150	144.282	2.868011577	1.949039
18	50	0.3	0.27845	49.050	40.875	36.06505	4.809945647	11.76745
19	50	0.3	0.85	127.530	106.275	105.5303	0.74475	0.700776
20	50	0.1367	0.85	58.860	49.050	46.46107	2.588926864	5.278138

Residual Plots for P_c shown in Fig. 4 has been analyzed and is discussed. The probability plot is used to check for the data normality. The distribution of the data points all along the normal line with very few outliers indicates that the data is normally distributed. The residual versus fitted values, as shown in the second plot, do not show any subsequent trends. This indicates the chosen RSM model fits well within the given dataset. The third plot of frequency verses residual illustrates the normal distribution of residuals that are produced by the model and assures that the assumptions made are reasonable and the choice of the model is appropriate. The residue versus observation order, as depicted in the last plot, highlights the presence of random data points. The randomness in the plot shows an insignificant experimental order in comparison to first power consumption (P_c). A confidence level of 95% has been applied to RSM for P_c.

Table 7: Estimated regression coefficients for P_c in watts

Term	Coeff.	SE Coeff.	T	P
Constant	-43.838	29.576	-1.482	0.169
v	-1.202	0.521	-2.308	0.044
f	342.726	110.593	3.099	0.011
d	13.983	29.511	0.474	0.646
v^2	-0.002	0.004	-0.429	0.677
f^2	-882.428	146.426	-6.026	0.000
d^2	-47.010	11.953	-3.933	0.003
v x f	5.109	0.941	5.432	0.000
v x d	2.161	0.269	8.039	0.000
f x d	175.179	53.753	3.259	0.009
S = 5.32124	R^2 = 99.40%	R^2 (pred.) = 97.00%	R^2 (adj.) = 98.86%	

Table 8: Analysis of variance for P_c in watts

Source	DF	Seq. SS	Adj. SS	Adj. MS	F	P
Regression	9	46840.3	46840.3	5204.47	183.80	0.000
Linear	3	42495.3	575.7	191.90	6.78	0.009
Interaction	3	1379.0	1379.0	459.68	16.23	0.000
Square	3	2965.9	2965.9	988.65	34.92	0.000
Residual Error	10	283.2	283.2	28.32		
Lack-of-Fit	5	149.5	149.5	29.90	1.12	0.453
Pure Error	5	133.7	133.7	26.73		
Total	19	47123.4				

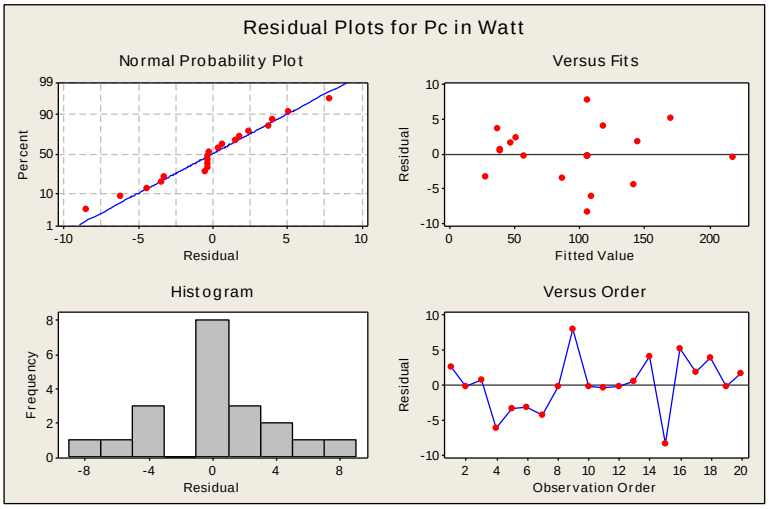


Fig. 4: Residual plots for P_c in watts.

3.4 Optimized Solution for Power Consumption

The lower values, the upper values and the final target values of P_c responses are fed to the response optimizer tool of Minitab by giving equal weightages. The description of the same is given in Table 9.

Table 9: Response optimization for power consumption

Goal	Lower	Target	Upper	Weight	Import
Optimization of P_c	24.525	40	217.455	1	1

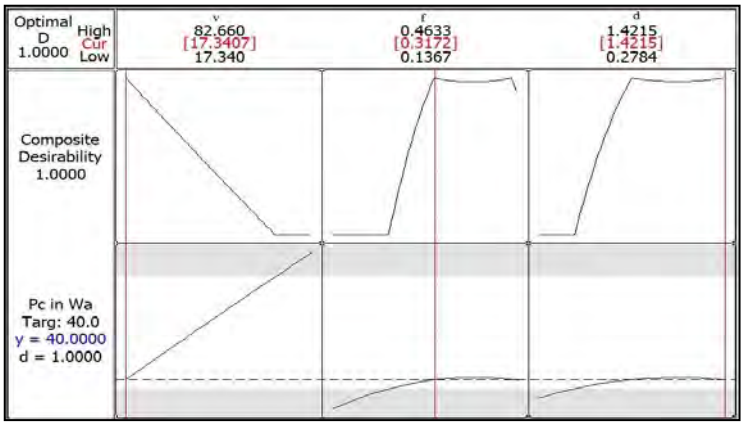


Fig. 5: Predicted response for R_a in μm.

Minimum power consumption (P_c) of 40 watts is obtained for a cutting velocity of 17.3407 m/min, feed rate of 0.3172 mm/rev and depth of cut of 1.4215 mm, as shown in the Fig. 5. The lower half of Fig. 5 shows the of variation of P_c along with focused input variables [19].

3.5 Development of Empirical Models

The relation between the input machining parameters and the output responses is developed by the mathematical models and thereby optimization of the machining process. RSM is an optimization procedure that produces a multi objective function model which can solve the optimization problem. The present study concentrates on the development of Empirical models for the output responses, Surface roughness (R_a) and power consumption (P_c) with regards to the input machining parameters in actual factors by using RSM. [20]. Any further optimization of the machining process can be carried out using the developed models. The second order regression equations of the model that was developed are estimated by regression analysis. The following equations for R_a and P_c were obtained in terms of actual factors.

The polynomial optimization equation of surface roughness (R_a):

$$\boldsymbol{R_a} = -8.1365 + (0.1049 \times v) + (43.1855 \times f) + (4.7344 \times d) - (0.0011 \times v^2) - (69.3314 \times f^2) - (3.6393 d^2) + (0.0463 \times v \times f) - (0.0118 \times v \times d) + (4.9286 \times f \times d)$$

The polynomial optimization equation of power consumption (P_c):

$$\boldsymbol{P_c} = (-11.189) - (0.052 \times v) + (70.789 \times f) + (14 \times d) - (0.000 \times v^2) - (115.384 \times f^2) - (6.358 \times d^2) + (0.125 \times v \times f) + (0.036 \times v \times d) + (21.429 \times f \times d)$$

4. CONCLUSION

This study presents an effective method for estimating the optimal turning operation parameters for surface roughness and power consumption using response surface methodology. The investigation shows that the control factors had varying effects on the response variable. The use of central composite design of response surface methodology was considered successful to obtain optimal responses. The results of power consumption (P_c) as obtained using ANOVA is represented in Table 8. The average value of 'P' was less than 0.001 and hence the model obtained for power consumption to be at a 99% confidence level. Fig. 4 shows that the output response is closer to the regression model obtained using RSM method. The surface roughness was mainly affected by the feed and the cutting speed. Also, an increase in the feed rate resulted in higher surface roughness. It was also observed that the cutting speed varies inversely along with surface roughness, thereby increasing the surface finish of the workpiece. The observed ANOVA response of the cutting velocity and depth of cut majorly influenced the power consumed by the machine. Minimum surface roughness along with minimum power consumption was estimated using the optimized solution parameters. This demonstrates that the RSM can be successfully applied to obtain a maximum amount of information with the least number of experimental runs. This can be successfully implemented to obtain an effective second order polynomial and for optimizing the machining process.

ACKNOWLEDGEMENTS

The authors wish sincere thanks to Basheer Ahmed Mohamed, the Managing Director of Sidhi Vinayaka Fab Engineering Pvt. Ltd., Peenya Industrial Area, Bangalore, India, for providing us the research facilities and supporting us in handling the research work.

REFERENCES

[1] Esawai A, Morsi K. (2007) Dispersion of carbon nanotubes (CNTs) in aluminium powder Composites part-A Applied Science and Manufacturing, 38: 646-650.
[2] Iijima S. (1991) Helical microtubules of graphic carbon. Nature, 354: 56-58.
[3] Bakshi SR, Lahiri D, Agarwal A. (2010) Carbon nanotube reinforced metal matrix composites - A review. International Materials Reviews, 55: 41-64.
[4] Liu ZY, Yu SJ, Xiao BL. (2012) Effect of ball milling time on mechanical properties of carbon nanotubes reinforced aluminium matrix composites. Composites part-A applied Zcience and Manufacturing, 43: 2161-2168.
[5] Baburaj E, Senthil P, Dhanapal P. (2018) Investigation on influence of machining parameters on high speed phase milling of hybrid (LM0-SiC-Fly Ash). Composite Int J Pure and Appl Math, 118(20): 901-906.
[6] Stratford KN, Audy J. (1997) Indirect monitoring of machinability in carbon steels by measurement of cutting forces. J. Mater Processing Tech, 67(1-3): 150-156.
[7] Teti R. (2002) Machining of composite materials. CIRP Annals, 51(2): 611-634.
[8] Ranganath MS, Vipin RSM. (2014) Effect of cutting parameters on MRR and surface roughness in turning of Aluminium-6061. Int J Adv Res and Innov, 2(1): 32-39.
[9] Parhad P, Likhite A, Bhatt J. (2015) The effect of cutting speed and depth of cut on surface roughness during machining of austempered ductile iron. Transactions of the Indian Institute of Metals, 68: 99-108. https://doi.org/10.1007/s12666-014-0439-y.
[10] Thangarasu SK, Shankar S, Thomas AT, Sridhar G. (2018) Prediction of cutting force in turning Process-an experimental approach, IConAMMA-2017, IOP Conf. Series: Materials Science and Engineering, 310: 012119.
[11] Patel MT, Deshpande VA. (2014) Optimization of machining parameters for turning different alloy steel using CNC-Review. Int J Innov Res Sci Eng Tech, 3(2): 2319-8753.
[12] Puttaswamy SJ, Venkatagiriyappa RB. (2020) Investigation on influence of anionic surfactant for homogenisation of MWCNT in Aluminium 6065 matrix. IIUM Engineering Journal, 21(2). https://doi.org/10.31436/iiumej.v21i2.1317
[13] Thangarasu SK. (2018) Prediction of cutting force in turning process-an experimental approach. IOP Conference Series: Materials Science and Engineering, 310: 012119.
[14] Samuel MK, Chandrasekaran C, Paarthasarathy B, Kamalakkannan SD. (2013) Prediction of surface roughness response in turning on Cu graphite MMC fabricated through stir-casting method. IJETAE, IESSN 3(3): 2250-2459.
[15] Geo R, D'Cotha JS. (2014) Effect of turning parameters on power consumption in EN 24 alloy steel suing different cutting tools. Int J Eng Res Gen Sci, 2(6): 2091-2730.
[16] Akcay H, Anagun AS. (2013) Multi response optimization application on a manufacturing factory. Mathematical and Computational Application, 18(3): 531-538.
[17] Kuttolamadom M, Hamzehlouia S, Mears L. (2010) Effect of machining feed on surface roughness in cutting 6061 aluminum. SAE Int J Mater Manuf, 3(1): 108-119. https://doi.org/10.4271/2010-01-0218.
[18] Richard Geo, Jose Sheril D'cotha (2014) Effect of Turning Parameters on Power Consumption in EN 24 Alloy Steel using Different Cutting Tools, International Journal of Engineering Research and General Science, Vol. 2, Issue 6, October-November, ISSN 2091-2730.
[19] Singh BJ, Sodhi HS. (2014) Parametric optimisation of CNC turning for Al-7020 with RSM. Int J Operational Research, 20(2): 180-206.
[20] Ramudu C, Sastry MNP. (2012) Analysis and optimization of turning process parameters using design of experiments. Int J Eng Res Appl, 2(6): 20-27.
[21] Lakshminarayanan AK, Balasubramanian V. (2009) Comparison of RSM with ANN in predicting tensile strength of friction stir welded AA7039 aluminium alloy joints. Transactions of Nonferrous Metals Society of China, 19(1): 9-18.

https://doi.org/10.31436/iiumej.v22i2.1673

POTENTIAL OF FABRICATION OF DURIAN SKIN FIBER BIOCOMPOSITES FOR FOOD PACKAGING APPLICATION THROUGH THE ELECTRICITY IMPACT ANALYSIS

HAZLEEN ANUAR[1*], SITI MUNIRAH SALIMAH ABD RASHID[1], NURFARAHIN MOHD NORDIN[1], FATHILAH ALI[2], YOSE FACHMI BUYS[3], SABU THOMAS[4], NUR AIMI MOHD NASIR[1] AND SYAZEVEN EFFATIN AZMA MOHD ASRI[1]

[1]Department of Manufacturing and Materials Engineering,
[2]Department of Biotechnology Engineering,
Kulliyyah of Engineering, International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia
[3]Department of Mechanical Engineering, Faculty of Engineering,
University of Malaya, 50603 Kuala Lumpur, Malaysia
[4]School of Energy Materials and International and Inter University Centre for Nanoscience and Nanotechnology, Mahatma Gandhi University, Priyadarshini Hills,
Kottayam 686560, Kerala, India

[*]Corresponding author: hazleen@iium.edu.my

(Received: 2nd November 2020; Accepted: 17th March 2021; Published on-line: 4th July 2021)

ABSTRACT: As an effort to replace the petroleum-based polymers and reduce waste-related environmental problems, biopolymers are the best candidate due to their renewable, biodegradable and commercially viable. Initiative have been taken by developing durian skin fibre (DSF) reinforced polylactic acid (PLA) biocomposites with the addition of epoxidized palm oil (EPO). PLA/DSF biocomposites were fabricated via extrusion and then injection moulded. The biocomposites were assessed for its life cycle by developing a system boundary related to its fabrication processes using GaBi software. The life cycle assessment (LCA) of PLA/DSF biocomposites show that global warming potential (GWP) and acidification potential (AP) were the major impacts from PLA/DSF biocomposite. For PLA/DSF biocomposite, the results were 199.37 kg CO_2 equiv. GWP and 0.58 kg SO_2 equiv. AP. Meanwhile, for PLA/DSF/EPO biocomposite, the results obtained were 195.89 kg CO_2 equiv. GWP and 0.57 kg SO_2 equiv. AP. The GWP and AP were contributed by the electricity used in the fabrication of biocomposites. These impacts were due to the usage of electricity, which contributed to the emission of CO_2. However, the PLA/DSF/EPO biocomposite had lower negative impacts because EPO improved the workability and processability of the biocomposite, and hence, reduced the amount of energy required for production. It can be concluded that the plasticized PLA/DSF biocomposite can be a potential biodegradable food packaging material as it has favourable properties and produces no waste.

ABSTRAK: Biopolimer adalah terbaik dalam usaha mengganti polimer berasaskan-petroleum dalam mengurang masalah pencemaran-sisa. Ini kerana biopolimer boleh diperbaharui, biodegradasi dan sangat maju secara komersial. Inisiatif telah diambil dengan menghasilkan sabut kulit durian (DSF) bersama biokomposit asid polilaktik (PLA) dengan penambahan minyak kelapa sawit terepoksi (EPO). Biokomposit PLA/DSF direka melalui kaedah pemyemperitan dan acuan suntikan. Biokomposit ini

dipantau kitar hidupnya dengan membina sistem sempadan berkaitan proses rekaan menggunakan perisian GaBi. Pengawasan kitar hidup (LCA) biokomposit PLA/DSF menunjukkan potensi pemanasan global (GWP) dan potensi pengasidan (AP) menyebabkan impak terbesar komposit PLA/DSF. Dapatan kajian menunjukkan 199.37 kg CO_2 bagi GWP dan 0.58 kg SO_2 bagi AP bagi biokomposit PLA/DSF. Sementara itu, dapatan kajian bagi biokomposit PLA/DSF/EPO adalah 195.89 kg CO_2 bagi GWP dan 0.57 kg SO_2 bagi AP. Kedua-dua GWP dan AP adalah disebabkan oleh penggunaan elektrik dalam proses pembuatan biokomposit. Ini adalah kesan daripada penggunaan elektrik, dan menyumbang kepada pembebasan CO_2. Walau bagaimanapun, biokomposit PLA/DSF/EPO mempunyai kurang kesan negatif, kerana EPO telah menambah baik kebolehkerjaan dan kebolehprosesan biokomposit, menyebabkan kurang tenaga yang diperlukan dalam proses pembuatan. Kesimpulannya plastik biokomposit PLA/DSF berpotensi sebagai bahan biodegradasi bagi pembungkus makanan kerana ianya mempunyai ciri-ciri yang diperlukan dan tidak menghasilkan sisa buangan.

KEYWORDS: *life cycle assessment; cradle-to-grave; durian skin fibre; plasticizer; food packaging; polylactic acid*

1. INTRODUCTION

Earth has ecosystem that consists of humans, plant life, ocean and natural resources. However, the development of engineering has resulted in resource depletion and environmental destruction. Almost all food packaging including single use packaging are petroleum-based polymers that are undegradable in nature and when they are being dumped on the land and in the ocean; it leads to the production of smelly garbage or polluted water. These problems are crucial environmental issues that should not be taken lightly.

Food packaging provide protection to the food, to maintain the quality and safety of the food from the environment and other physical harm. Food packaging is important when transporting the food from one place to another [1]. Most food packaging is made from glass, cloth, metal, paper and polymers. Polymers are the most common materials used in the food packaging industry due to their superior properties.

The most commonly used plastics in packaging include polyethylene terephthalate (PET), polyvinyl chloride (PVC), polyethylene (PE), polypropylene (PP), polystyrene (PS) and polyamide [2,3]. The utilization of these polymers harms the environment and ecosystems via global warming and ozone layer depletion. This is due to the unwanted gas like carbon dioxide is released to the atmosphere. The degradation process of polymers takes many years to be completed. PP for example shows resistance to biodegradation since it is highly hydrophobic, high molecular weight, lacks an active functional group and has a continuous chain of repetitive methylene units [4].

2. METHODOLOGY

The framework was conducted according to ISO 14040 [5] and ISO 14044 [6] which involve four phases; goal and scope definition, inventory analysis, impact assessment and interpretation using GaBi software.

2.1 Goal and Scope Definition

The goal of this study is to assess the environmental impacts of food packaging made of PLA/DSF in comparison with PLA/DSF with the addition of epoxidized palm oil (EPO) as a plasticizer. This comparison was to investigate the effect of EPO on PLA/DSF

on the environmental impact. The functional unit of this study was 50-unit biodegradable food packaging and the reference flow was 850 g for each PLA/DSF and PLA/DSF/EPO biocomposite. The scope of this study was from the preparation of durian skin fibre, preparation of biocomposite and fabrication of biocomposite into food packaging before end-of-life disposal. The scope of impact assessment covered was global warming, acidification, eutrophication and ozone depletion. These impacts were chosen based on the impact on electricity usage as electricity was the main contributor and this could be the main cause for the climate change. The study covered the entire life cycle of PLA/DSF and PLA/DSF/EPO biocomposite from cradle-to-grave.

2.2 System Boundary

The system boundary in this study includes the preparation of DSF, manufacturing of biocomposites and end-of-life stage. The preparation of DSF involves chopping and grinding, washing, drying and alkali treatment process while manufacturing of biocomposite covers extrusion and injection moulding. The inputs and outputs of each biocomposite food packaging manufacturing process was identified based on the inventory analysis elaborated later. From the information gathered in the inventory analysis, system boundaries for both biocomposites were determined. Figures 1 and 2 show the schematic diagrams of input and output flow for PLA/DSF and PLA/DSF/EPO biocomposites, respectively.

2.3 Data Sources and Limitations

In this study, the input and output data for preparation of DSF and fabrication of biocomposites were primary data which was collected during the process. The covered data consists of raw material (durian skin waste) and energy consumption for the fabrication of biocomposite. Data on production of PLA resin and EPO was excluded as these two materials were bought from the supplier and delivered directly to the fabrication lab. The use phase also was neglected as this biocomposite still in prototype food packaging.

2.4 Inventory Analysis

The inventory analysis was gathered which include the material and energy inputs, air emissions, solid waste emissions and waterbone emissions in the life cycle of the PLA/DSF biocomposite with and without EPO. It involved collecting quantitative data for every unit process in the system based on a functional unit of product. Details of each phase starting from durian skin fibre preparation until end-of-life stage are explained next.

2.5 Raw Materials

Durian skins waste were collected from local markets and transported to IIUM. PLA (grade 3052) was obtained from NatureWorks, China; EPO was obtained from Malaysian Palm Oil Board; and databases for raw PLA and EPO were obtained from GaBi software (which were compiled from a literature review by [7-10] for EPO. The database for PLA was obtained from a study by [11]. The composition of the biocomposite was observed based on the weight of the food packaging. The biocomposite was initially prepared for 1000 g including waste after injection moulded. From such preparation, the weight of a single food packaging was determined at 17 g whereas the total weight for 50 units of biocomposite food packaging was 850 g. Compositions for each PLA are shown in Tables 1 and 2.

Table 1: Composition of PLA/DSF biocomposite

Material	Composition (%)	Weight for 1 unit (g)	Weight for 50 units (g)
PLA	70	11.9	595
DSF	30	5.1	255
Total	100	17	850

Table 2: Composition of PLA/DSF/EPO biocomposite

Material	Composition (%)	Weight for 1 unit (g)	Weight for 50 units (g)
PLA	65	11.05	552.5
DSF	30	5.1	255
EPO	5	0.85	42.5
Total	100	17	850

2.6 Preparation of Durian Skin Fibre (DSF)

The durian skin was cut into smaller pieces and washed thoroughly with tap water to remove any dust and adhering particles. The skin was then dried at 70 °C for 24 h. The dried skin was ground to obtain fibre with 100 to 150 µm in length. The weight of the raw durian skin fibres used is approximately 300 g. Sodium hydroxide (NaOH) was used for alkali treatment of DSF.

2.7 Preparation of Biocomposites

Next, fabrication process of the biocomposite consists of extrusion using extruder (Brabender, Germany) and injection moulding processes using Battenfeld HM 600/850 injection moulding machine. The temperature profile from barrel to die was 160 °C to 170 °C at 100 rpm screw speed. For injection moulding, the temperature was 165 °C to 180 °C for all four zones of the injection moulding machine and 45 s cooling time. Finally, the biocomposite food packaging waste was exposed to the environment for end-of-life stage.

2.8 Transportation

Transportation is referring to picking up the waste durian skins from Petaling Jaya, to laboratory at Gombak, Selangor, Malaysia. Transportation data required for conveying the product to the end user and to the transfer station was not available, because the process is assumed to have only evolved within the area of fabrication, which is IIUM. The total journey from IIUM Gombak to Petaling Jaya and back to IIUM Gombak is approximately 92 km. The transportation for obtaining the PLA and EPO were not included in this study since it was delivered to IIUM. Both PLA/DSF biocomposites were still in prototype, thus the distribution of the biocomposites to the consumer was not included for transportation.

2.9 Electricity

Table 3 describes the total power and electricity utilized for each particular process in the production of PLA/DSF and PLA/DSF/EPO biocomposites.

Table 3: Power calculation of each machine used.

Machine	Total time utilised for 50 units (hours)	Total power (kWh)	Energy (MJ)
Oven	a) 24 b) 48	a) 36 b) 72	a) 129.6 b) 259.2
Extruder	10	50	180
Grinder	0.85	3.4	12.24
Injection moulding	1.25	66.94	240.98

2.10 End-of-Life Stage

The PLA/DSF and PLA/DSF/EPO biocomposite were exposed to the environment to observe their degradation time. Both biocomposites took about five to six months to fully biodegrade.

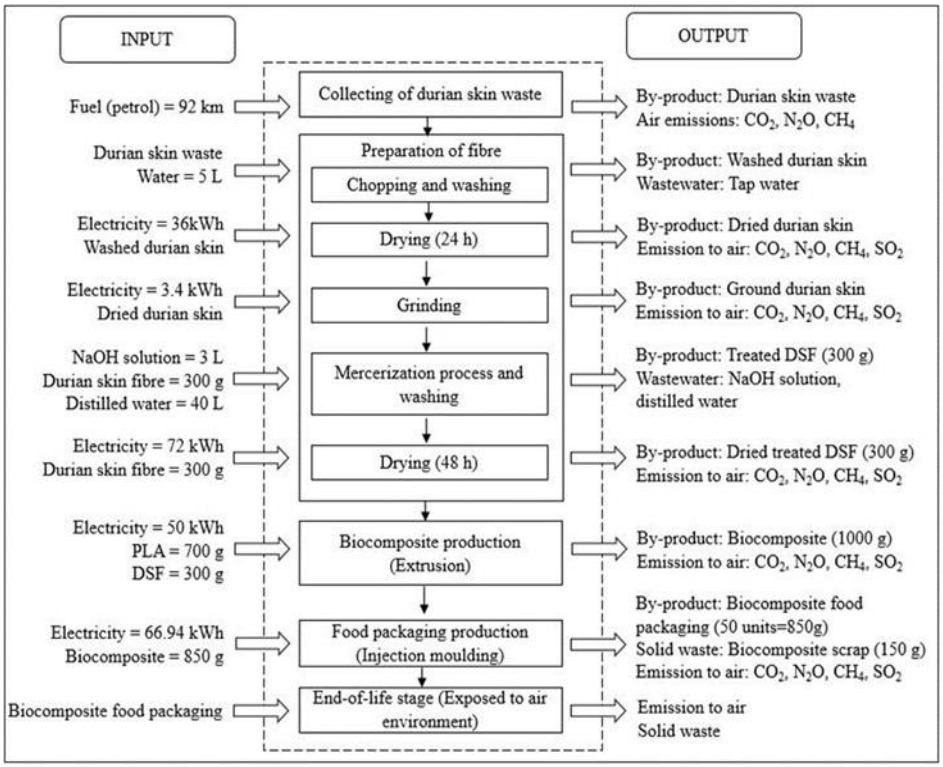


Fig. 1: System boundary of PLA/DSF biocomposite.

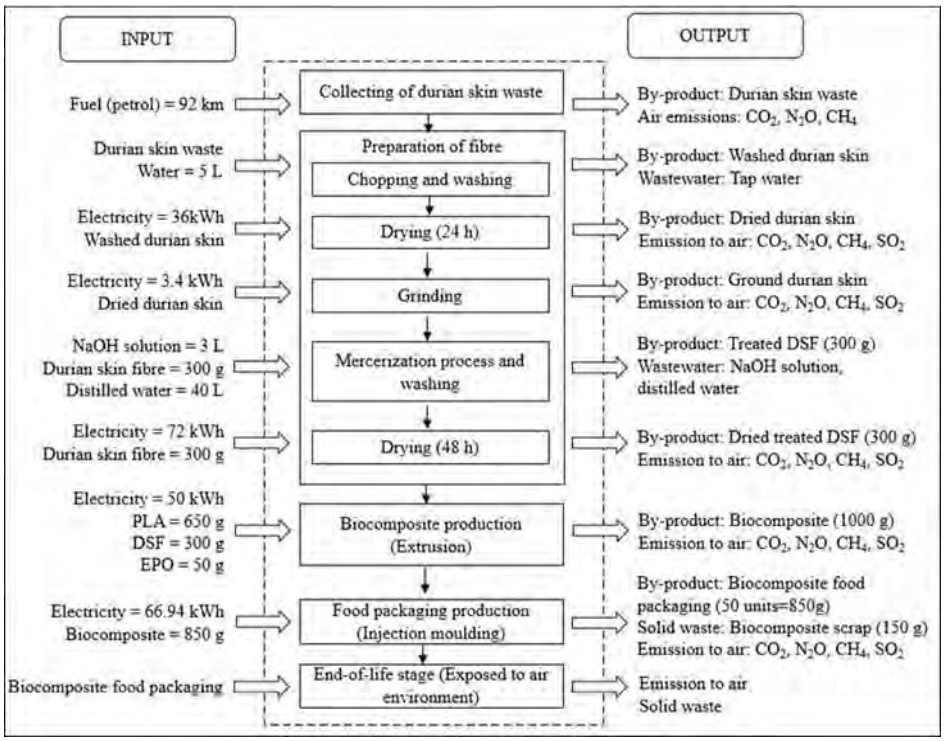


Fig. 2: System boundary of PLA/DSF/EPO biocomposite.

2.11 Impact Assessment

The environment impact assessment was conducted to evaluate and analyse the possibility of activities in the inventory analysis and production of biocomposite food packaging whether to leave negative environmental impacts on the earth or not. For this examination, effects were measured up to the midpoint level. The midpoint level expects to cover ecological issues that stand close to the inventory and endpoint results. Ecological impact categories, under the midpoint level, include carbon footprint, photochemical ozone formation, human toxicity, water footprint, ozone layer depletion, global warming, acidification, eutrophication, etc., depend on the Intergovernmental Panel on Climate Change (IPCC). Among the impact listed, the last four contributed to the highest impact score and play a vital role in the lifecycle of PLA/DSF biocomposite food containers. This study uses the ReCiPe (1.08) Life Cycle Impact Assessment (LCIA) method, run through GaBi software, and based on ISO 14040 framework [5] and ISO 14044-guidelines and requirements [6].

3. RESULTS AND DISCUSSION

The results are analysed for four major environmental impacts comprised of global warming potential, ozone layer depletion potential, acidification potential and eutrophication potential based on the impact in electricity usage as this could be the main cause for climate change. The environmental impacts evaluated was cradle-to-grave analysis started from raw materials, processing, manufacturing and end life of the product which was landfill. Based on the impact score, global warming potential (GWP) possesses the highest impact score with 199.38 kg CO_2 equiv. for PLA/DSF over PLA/DSF/EPO with 195.89 kg CO_2 equiv. These are followed by acidification potential (AP), with an impact score of 0.58 kg SO_2 equiv. for PLA/DSF and 0.57 kg SO_2 equiv. for PLA/DSF/EPO. The remaining two impacts considered in this study produced very low values of impact score; which are deemed insignificant. For all four impacts considered in this study, PLA/DSF showed a higher potential of environmental impacts than PLA/DSF/EPO biocomposites. This was as expected as biocomposite consists of EPO that is bio-based and environmentally friendly plasticizer. According to [12], plasticizers made from natural sources incorporated into a bio-based polymer provide lower negative impacts to the environment.

3.1 Global Warming Potential (GWP)

The findings for global warming potential (GWP) are presented in Table 4 and Fig. 3 as percentage values. From Table 4 and Fig. 3, it can be seen that the injection moulding process had the highest contribution towards the global warming potential for both biocomposites. The process of PLA/DSF showed a slightly higher percentage, at 33.08% over PLA/DSF/EPO at 31.70%. Plasticizer, added to the composites, improved their workability, processability and flexibility, making them easier to be processed with reduced energy needed [13]. Therefore, the percentage contribution of each biocomposites process towards GWP was different as EPO was added at the beginning of the mixing process for the plasticized one.

Figure 3 clearly shows that the major contributors towards GWP are the drying, mixing and injection moulding processes of both biocomposites. These processes contribute up to 95.0% of the total GWP for both PLA/DSF and PLA/DSF/EPO biocomposites. Specifically, these processes require electricity to generate output (i.e., biocomposite food packaging). Consumption of electricity highly contributed towards the

release of CO_2 gas. Energy industries are a major source of greenhouse gases in Malaysia [14,15]. Global warming is mainly caused by the release of greenhouse gases into the environment; primarily CO_2. CO_2 emissions take up to 93.0% and the rest are other greenhouse gases, such as nitrous oxide and methane. [16] stated that the three crucial sources of fossil fuel, that highly affected the power generation operations in Malaysia are natural gas, coal fuel and hydropower. However, hydropower that come from renewable energy sources used only small percentage as compared to other sources. Emissions of greenhouse gases, mainly CO_2 from the burning of fossil fuels, cause environmental degradation and global warming. Therefore, the more electricity utilised in the production of biocomposites raises the emission level of CO_2 gas to the environment; eventually contributing towards global warming. The amount of electricity used are high in the fabrication process due to the utilisation of various machines.

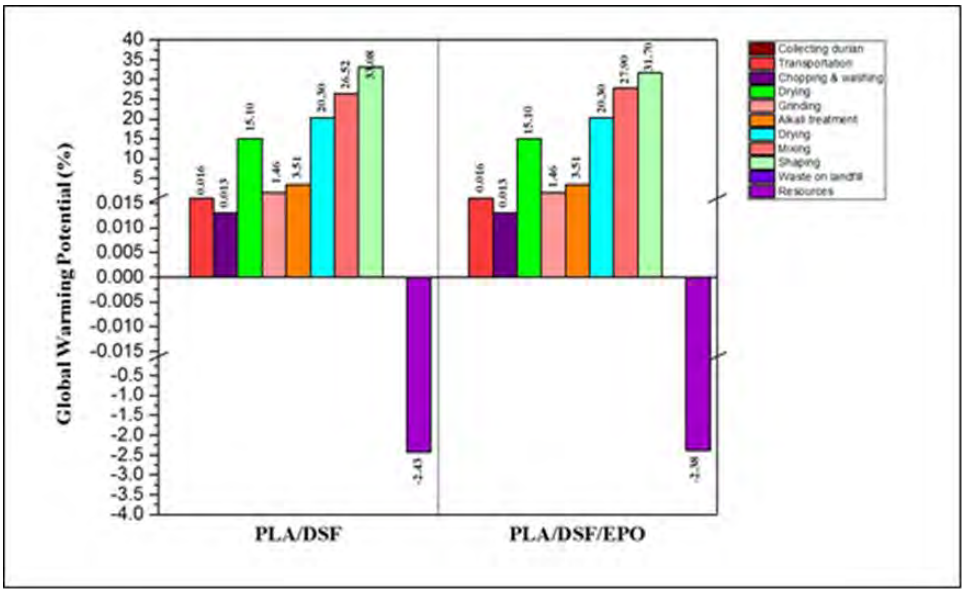


Fig. 3: Percentage contribution of each process to GWP.

Table 4: Percentage contribution of each process to GWP

Process	% of GWP for each process (kg CO_2 equiv.)	
	PLA/DSF	PLA/DSF/EPO
Collecting durian skin waste	0	0
Transportation	0.016	0.016
Chopping and washing	0.013	0.013
Drying (24 h)	15.10	15.10
Grinding	1.46	1.46
Alkali treatment	3.51	3.51
Drying (48 h)	20.30	20.30
Mixing (extrusion)	26.52	27.90
Shaping (injection moulding)	33.08	31.70
Waste on landfill	0	0
Recovery	-2.43	-2.38
	Total GWP value of each type of biocomposite (Including material recovery) (kg CO_2 equiv.)	
	199.38	195.89

Meanwhile, the process of chopping and washing of durian skins contributed less to the potential global warming, about 0.013%, for both biocomposites. This is because only tap water is used in the process and no greenhouse gases are emitted during fibre preparation. A previous study on environmental impact was done by [17] on tap water and compared with bottled mineral water. The study found that tap water imparted a lower environmental impact than bottled mineral water. In Fig. 3, the resource column indicates that resource recovery is involved in the production process. The biocomposites are produced from bio-based materials that originated from plants, such as DSF and PLA. Carbon dioxide can be sequestered, as plants consume it, even though it is released during the production process [18]. In summary, GWP for PLA/DSF/EPO contributed less to global warming than PLA/DSF biocomposite; with the potential of 195.89 kg CO2 equiv. compared to 199.38 kg CO_2 equiv.

3.2 Acidification Potential (AP)

Table 5 and Fig. 4 present the percentage of acidification potential of each process involved in the production process of both biocomposites. From Table 5 and Fig. 4, the injection moulding process shows significantly high potential to acidification as compared to other process, which is responsible for 30.51% of AP from PLA/DSF biocomposite and 31.59% of AP from PLA/DSF/EPO biocomposite. Similarly, the major contributors to AP of both biocomposites are from the processes that need electricity, namely injection moulding, extrusion and drying. The acidification potential of these processes includes electricity generated 97.15% of total AP for PLA/DSF and 98.60% of total AP for PLA/DSF/EPO biocomposites including the grinding process. From the total AP for both biocomposites, more than 57% comes from the emission of nitrogen oxides, followed by sulfur dioxides at 41%. A small portion is made up of ammonia and nitrogen dioxides. This is due to the utilization of electricity during biocomposites fabrication. The ocean and the atmosphere absorb the toxic gases generated and hence acidity is increased. As a conclusion, even though the contribution of electricity to AP for PLA/DSF/EPO is higher than PLA/DSF, the total AP of PLA/DSF/EPO is lower than that of PLA/DSF. Therefore, PLA/DSF/EPO causes less effect to the acidification as compared to PLA/DSF biocomposite.

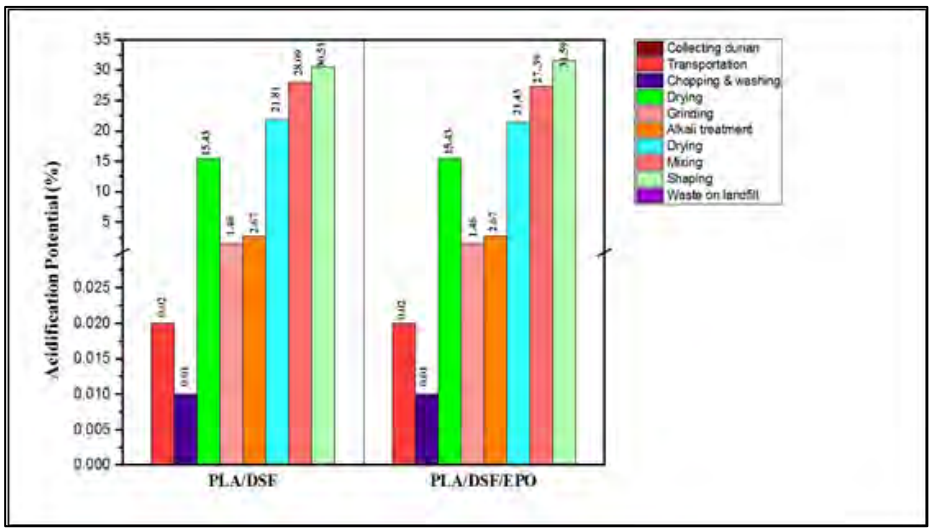


Fig. 4: Percentage contribution of each process to AP.

Table 5: Percentage contribution of each process to AP

Process	% of AP for each process (kg SO_2 equiv.)	
	PLA/DSF	PLA/DSF/EPO
Collecting durian skin waste	0	0
Transportation	0.02	0.02
Chopping and washing	0.01	0.01
Drying (24 h)	15.43	15.43
Grinding	1.46	1.46
Alkali treatment	2.67	2.67
Drying (48 h)	21.81	21.43
Mixing (extrusion)	28.09	27.39
Shaping (injection moulding)	30.51	31.59
Waste on landfill	0	0
	Total AP value of each type of biocomposite (kg SO_2 equiv.)	
	0.58	0.57

3.3 Eutrophication Potential (EP)

Eutrophication is one of several environmental impacts considered in this study; and showed a very low impact score. Eutrophication is measured in terms of kg P equiv. as it is mainly caused by phosphorus and phosphate [19]. The LCA leads to 8.29 x 10^{-5} kg P equiv. for PLA/DSF and 6.95 x 10^{-5} kg P equiv. for PLA/DSF/EPO. Table 6 and Fig. 5 indicate the results generated from the study for eutrophication potential. Figure 5 shows that the alkali treatment process for DSF displayed the highest potential to eutrophication; with 37.27% for both biocomposites. This process consumes sodium hydroxide, tap water and distilled water. The waste is then discharged through the sewage system before going directly into soil and water bodies, stimulates, and eventually causes eutrophication.

When the waste yielded from the alkali treatment process is discharged, it unlocks the phosphate and phosphorus contained within the soil; thus, increasing the nutrients level of the soil. When the concentration levels of the nutrients are too high, they are carried by rain water into rivers and ground waters that subsequently flow into lakes and seas. [20] assured that wastewater is a contributor for eutrophication to occur; where it increases the amount of nitrogen and phosphorus transferred to water bodies.

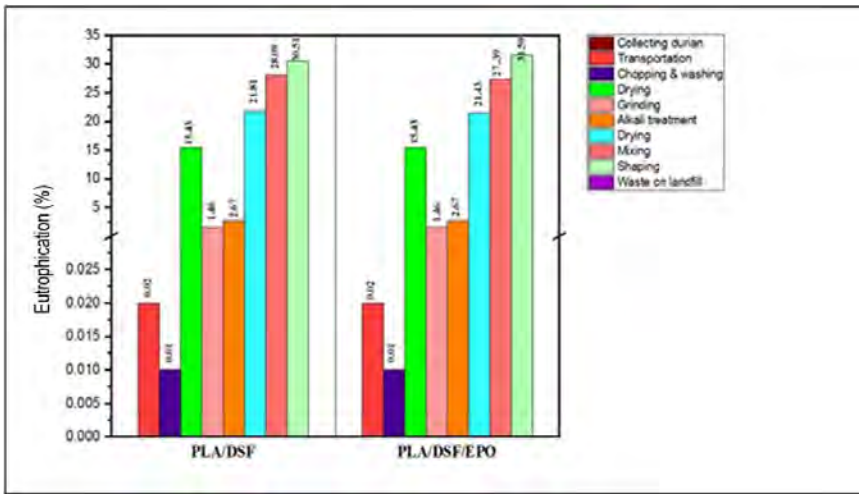


Fig. 5: Percentage contribution of each process to EP.

Table 6: Percentage contribution of each process to EP

Process	% of EP for each process (kg P equiv.)	
	PLA/DSF	PLA/DSF/EPO
Collecting durian skin waste	0	0
Transportation	0.05	0.05
Chopping and washing	9.67	9.67
Drying (24 h)	2.53	2.53
Grinding	0.91	0.91
Alkali treatment	37.27	37.27
Drying (48 h)	16.16	16.16
Mixing (extrusion)	18.51	17.43
Shaping (injection moulding)	14.96	15.98
Waste on landfill	0	0
	Total EP value of each type of biocomposite (kg P equiv.)	
	8.29×10^{-5}	6.95×10^{-5}

The transportation process contributes the lowest impact to eutrophication. This is because transportation does not produce waste that is discharged to soil and water bodies, but contributes towards other environmental impacts through emissions of gases into the air. In conclusion, eutrophication is not affected by the life cycle of the biocomposites based on the low impact score of EP which is nearly zero.

3.4 Ozone Depletion Potential (ODP)

Ozone layer depletion, or ozone depletion potential (ODP), for both biocomposites is further analysed and presented in Table 7 and Fig. 6. ODP is considered as the negligible impact on the biocomposites in this study. This is due to the lowest impact score obtained; which was nearly zero. ODP is the most insignificant impact from the food packaging industry. This is in agreement with a previous finding by [21] which showed that ODP yields a very low impact score compared to other environmental impacts.

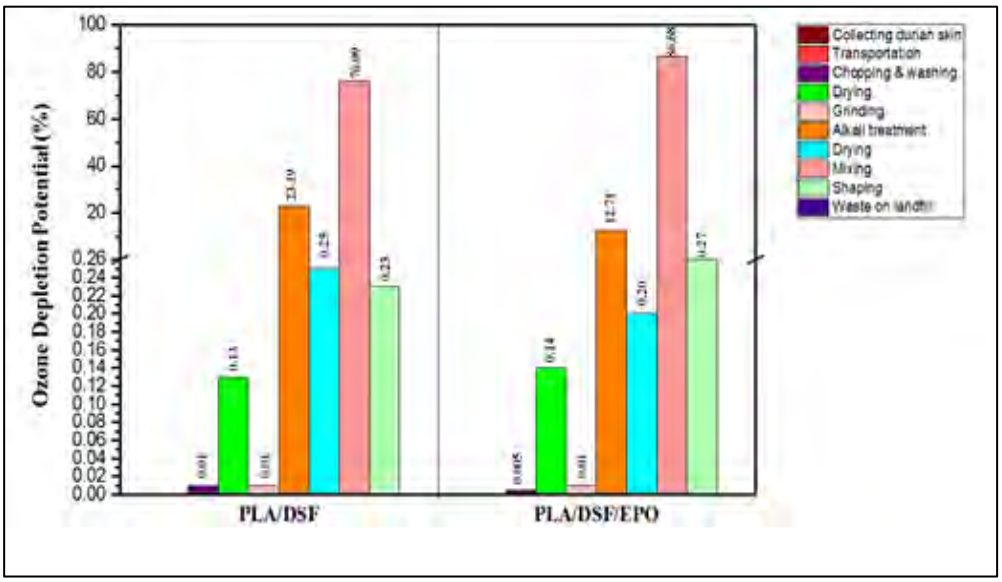


Fig. 6: Percentage contribution of each process to ODP.

Table 7: Percentage contribution of each process to ODP

Process	% of ODP for each process (kg CFC-11 equiv.)	
	PLA/DSF	**PLA/DSF/EPO**
Collecting durian skin waste	0	0
Transportation	0	0
Chopping and washing	0.01	0.005
Drying (24 h)	0.13	0.14
Grinding	23.19	12.71
Alkali treatment	0.25	0.20
Drying (48 h)	76.09	86.68
Mixing (extrusion)	0.23	0.27
Shaping (injection moulding)	0	0
Waste on landfill	0	0
	Total ODP value of each type of biocomposite (kg CFC-11 equiv.)	
	8.28×10^{-10}	**1.28** 10^{-10}

4. CONCLUSION

The fabrication of biocomposites food packaging causes significant global warming and acidification to occur. The impact scores are higher with almost 200 kg CO_2 equiv and 0.5 kg SO_2 equiv. PLA/DSF produced higher impact compared to PLA/DSF/EPO for all categories. GWP and AP were not significantly difference but for EP and ODP the impact difference between PLA/DSF and PLA/DSF/EPO was almost 19%. This impact is due to generation of electricity throughout the production process. The highest emission was achieved by carbon dioxide with the highest percentage value. Ozone layer depletion and eutrophication potential values were very low, almost zero. Meanwhile, eutrophication only affected by the alkali process in the fabrication of biocomposites, which produces waste discharged into the soil and other water bodies.

ACKNOWLEDGEMENT

The authors wish to thank the Ministry of Higher Education (MOHE) Malaysia for the funding provided through the Fundamental Research Grant Scheme (FRGS/1/2019/TK05/UIAM/02/2), Research Initiative Grant Scheme–Post-Doctoral Fellow (RPDF18-010-0010) from the International Islamic University Malaysia (IIUM), Department of Science & Technology (DST)-ASEAN-India Research Training Fellowship (RTF/2019/000083) for the equipment and facilities provided in making these studies a success.

REFERENCES

[1] Duncan TV. (2011) Applications of nanotechnol. in food packaging and food safety: barrier materials, antimicrobials and sensors. J. of Colloid and Interface Science, 363: 1-24. https://doi.org/10.1016/j.jcis.2011.07.017

[2] Siracusa V. (2012) Food packaging permeability behaviour: A report. Int. J. of Polym. Sci., 1-11. https://doi.org/10.1155/2012/302029

[3] Badmus AA, Gauri S, Ali NI, Gomes C. (2015) Mechanical stabilty of biobased food packaging materials. Food Sci. and Quality Management, 39: 41-48.

[4] Xu X, Jayaraman K, Morin C, PecqueuxN. (2008) Life cycle assess of wood-fibre-reinforced polypropylene composites. J of Mater. Processing Tech., 198: 168-177. https://doi.org/10.1016/j.jmatprotec.2007.06.087
[5] ISO, 2006a. ISO 14040: Environmental management - life cycle asses. – princ. and framew. Int Organization for Standardization, Genève, Switzerland.
[6] ISO, 2006b. ISO 14044: Environmental management - life cycle assess. – princ. and framew. Int. Organization for Standardization, Genève, Switzerland.
[7] Subramaniam V, May CY, Muhammad H, Hashim Z, Tan YA, Wei PC. (2010) Life cycle assess.of the production of crude palm oil (part 3). J. of Oil Palm Research, 22: 895-903.
[8] Schmidt JH. (2010) Comparative life cycle assess. of rapeseed oil and palm oil. The Int J. of Life Cycle Assess., 15: 183-197.
[9] Flynn HC, Canals LMI, Keller E, King H, Sim S, Hastings A, Smith P. (2012) Quantifying global greenhouse gas emissions from land-use change for crop production. Global Change Biology, 18: 1622-1635. https://doi.org/10.1111/j.1365-2486.2011.02618.x
[10] Choo YM, Muhamad H, Hashim Z, Subramaniam V, Puah CW, Tan Y. (2011) Determination of GHG contributions by subsystems in the oil palm supply chain using the LCA approach. The Int. J. of Life Cycle Assess., 16: 669-681.
[11] Vink ETH, Davies S. (2015). Life cycle inventory and impact assessment data for 2014 Ingeo™ polylactide production. Industrial Biotechnology, 11: 167-180. https://doi.org/10.1089/ind.2015.0003
[12] Jusoh ER, Ismail MH, Abdullah LC, Robiah Y, Rahman WAWA. (2012) Crude palm oil as a bioadditive in polypropylene blown films. Bioresources, 7: 859-867.
[13] Sanyang ML, Sapuan SM, Jawaid M, Ishak MR, Sahari J. (2016) Effect of plasticizer type and concentration on physical properties of biodegradable films based on sugar palm (Arenga pinnata) starch for food packaging. J. of Food Sci. and Tech., 53: 326–336.
[14] Zaid SM, Myeda NE, Mahyuddin N, Sulaiman R. (2015) Malaysia's rising GHG emissions and carbon 'lock-in' risk : A review of Malaysian building sector legislation and policy. J. of Surv., Construction and Property, 6: 1-1.
[15] Abdullah WSW, Osman M, Ab Kadir MZA, Verayiah R. (2019) The Potential and Status of Renewable Energy Development in Malaysia. Energies, 12(2): 2437. https://doi.org/10.3390/en12122437
[16] Samsudin MSN, Rahman MM, Wahid MA. (2016) Power generation sources in Malaysia : Status and prospects for sustainable development. J. of Adv. Review on Sci. Research, 25: 11-28.
[17] Jungbluth N. (2006) Comparison of the environmental impact of tap water vs. bottled mineral water. Uster, Switzerland: Swiss Gas and Water Association Bull.
[18] Hervy M, Evangelisti S, Lettieri P, Lee KY. (2015) Life cycle assess. of nanocellulose-reinforced advanced fibre composites. Composites Sci. and Tech., 118: 154-162. https://doi.org/10.1016/j.compscitech.2015.08.024
[19] Ortiz-Reyes E, Anex RP. (2018) A life cycle impact assess. method for freshw. eutrophication due to the transp. of phosphorus from agricultural production. J. of Cleaner Production, 177: 474-482.
[20] de Jonge VN, Elliott M. (2001) Eutrophication. Encyclopedia of Ocean Sciences, 2: 852-870.
[21] La Rosa AD, Recca G, Summerscales J, Latteri A, Cozzo G, Cicala G. (2014) Bio-based versus traditional polymer composites. A life cycle assessment perspective. J. of Cleaner Production, 74: 135-144. https://doi.org/10.1016/j.jclepro.2014.03.017

https://doi.org/10.31436/iiumej.v22i2.1715

MODEL ANALYZING FOR REUSING GOLD WIRE CAPILLARY IN THE GOLD WIRE BONDING PROCESS

CHATPON PHIMPHA*AND SOMBAT SINDHUCHAO

Department of Industrial Engineering, Faculty of Engineering,
Ubon Ratchathani University,
Ubon Ratchathani 34190, Thailand

**Corresponding author: Perpetually09@gmail.com.*

(Received: 3rd December 2020; Accepted: 4th February 2021; Published on-line: 4th July 2021)

ABSTRACT: Manufacturing process improvement is necessary for manufacturers to gain business advantages. Re-using or increasing the useful lives of machine parts is considered to be a process of performance improvement. To re-use parts, the manufacturers must know the effects of the factors related to workpieces' qualities to prevent defects. This research study aims at presenting the results of analysing the effects of the factors and mathematical models for bond shear strength when reusing gold wire bonding capillary in the gold wire bonding process of integrated circuit (IC) products using design experiment. The operation factors in the reference experiment, including bond force, bond time, USG current, EFO current and EFO gap, are investigated. The Fractional Factorial Design was used to determine five factors that affect the bond shear strength. The analysis of the results show that the bond force is a significant factor where increasing bond force factors leads to increasing bond shear strength. In the end, a Regression model of bond shear strength is obtained to show the result between the bond shear strength and effect of factors.

ABSTRAK: Proses pembaharuan pengilangan adalah penting untuk para pengilang bagi memperoleh keuntungan bisnes. Guna-semula atau menambah jangka hayat pada bahagian-bahagian tertentu pada mesin adalah dianggar sebahagian proses penambahbaikan prestasi mesin. Bagi mengguna semula bahagian-bahagian ini, pengilang mesti mengetahui akibat sesuatu faktor berkaitan kualiti bahan bagi mengelak kecacatan. Kajian ini bertujuan menyampaikan dapatan kajian melalui kesan faktor dan model matematik pada kekuatan ricihan ikatan apabila mengguna semula wayar emas melalui proses kapilari ikatan wayar emas pada produk litar bersepadu melalui rekaan eksperimen. Faktor operasi melalui rujukan eksperimen dari daya ikatan, masa ikatan, arus USG, arus EFO dan jarak EFO dikaji. Rekaan Faktorial Pecahan digunakan bagi mendapatkan lima faktor yang mempengaruhi kekuatan ricihan ikatan. Dapatan kajian menunjukkan daya ikatan merupakan faktor penting di mana, pertambahan faktor daya ikatan menguatkan ricihan ikatan. Akhirnya, model Regression kekuatan ricihan ikatan diperoleh bagi menjelaskan dapatan kajian antara kekuatan ricihan ikatan dan kesan faktor.

KEYWORDS: *reusing capillary; gold wire bonding; bond shear; design of experiment*

1. INTRODUCTION

Improving the production process to increase efficiency in the working process is an essential part of business. Factories are required to understand and develop their production processes to produce products or services that can meet customer needs and enhance business competitiveness with other factories producing the same product [1,2].

Machine parts that are expired from the manufacturing process are used. The parts are studied, and the effects of the factors are identified in order to re-use the parts or to increase the parts' useful lives. This is considered a strategy for reducing manufacturing costs and improving processes for business competitiveness. To re-use expired parts, the parts must be carefully considered, and the effects of the process factors must be identified in order to prevent defects [3].

Most studies focus on the factors of the best practices for new parts. However, this study aims at studying the effects of the process factors for expired parts. In order to make the manufacturers understand the effects of each factor of the qualities of the work pieces to efficiently use the parts without affecting the work pieces.

The gold wire bonding process is a process of integrated circuit (IC) products, which are the main parts in electronic devices such as mobile phones, televisions, etc., The IC packaging product consists of two main processes, namely, Front of Line (FOL) production, which consists of five sub processes, and End of Line (EOL) [4], which consists of four sub processes. The most important process of IC Packaging production is the gold wire bonding process in the FOL because it connects the main circuit of IC packaging using small gold wire [5,6]. In this process, high-precision machining and a large number of adjustment factors are required to process [7,8]. A machine part that is called a gold wire bonding capillary is an important part of the gold wire bonding process because this part presses on work pieces being wired. The gold wire bonding capillary is a small tube for bonding gold wires through the mentioned tube. By checking the manufacturing process, it was found that the gold wire bonding capillary can manufacture 150,000 work pieces. Then, the gold wire bonding capillary must be changed. Accordingly, this study emphasizes studying the effects of the quality factors of work pieces if the parts are expired in order to provide data for manufacturers who want to improve the efficiency of their processes by re-using parts and to help them know the data and relationships of the factors for manufacturing work pieces according to requirements [3].

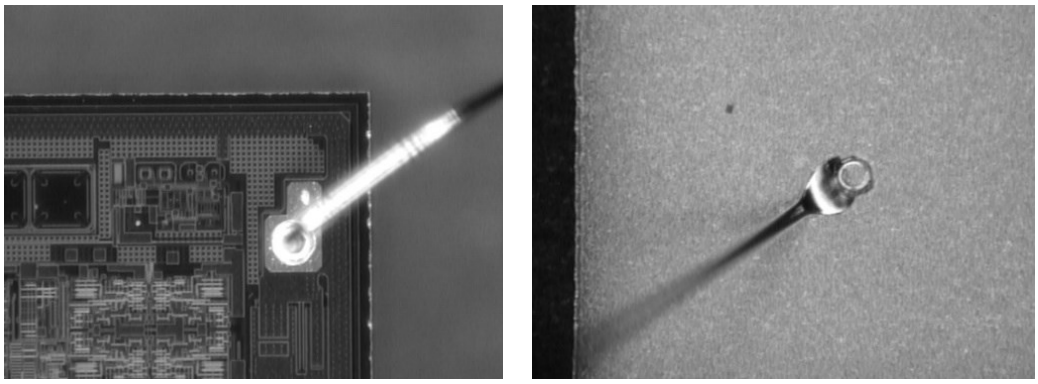
Statistical methods are used to find the effect of factors to adjust the parameters of the gold wire bonding process to maximize bond shear strength. An experimental approach that defines parameters in a process is the design of experiment method [9]. Each of the researched studies demonstrated the application of experimental methods, showing various factors that affected the gold wire bonding process, including showing results of different research. The research work showed that an increase in bonding temperature affected the shear strength of gold bonding. The one-factor experiments were not suitable for the study of most factors in the same experiment [10]. The study also found research related to the use of analysis of variance (ANOVA) to analyse factors that affect the strength of wire bonding. ANOVA is an experiment that can be studied using multiple factors. However, the large number of factors are limitations to the number of the experiments [11]. According to relevant research reviews, the suitable method for determining the optimum conditions of the gold wire bonding process was factorial design (FD) [12]. A full factorial consists of all possible combinations of levels for all factors. The total number of experiments for studying are included in [13]. Consider the factors in which five factors were researched with two levels of factors by factorial, the design of a full factorial experiment should involve a total of 32 experiments. The Fractional Factorial Design (FFD) is one type of factorial experiment. It is an experiment that uses half the number of experiments from the design of a full number of factorial experiments. The objective of this research was to use fractional factorial experiment to determine the effect of

https://doi.org/10.31436/iiumej.v22i2.1715

parameters and show the regression model that is used for bond shear strength in the gold wire bonding process of IC packaging process.

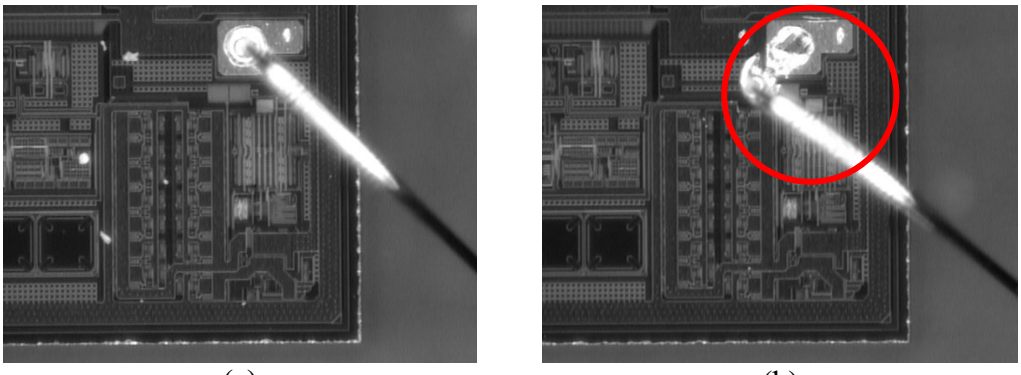
2. MATERIALS AND METHODOLOGY

The process of bonding gold wire is summarized as follows: gold wire bonding starts with an electric current through a device called an EFO wand, which is responsible for distributing electricity to heat the gold wire in the capillary to form a spherical shape. The molten gold wire is then pressed onto the circuit pad, where the first connection is a spherical shape called the first bond, as shown in Figure 1(a). The machine then raises the gold wire and drags the gold wire to connect it to the second connection at the lead frame. The lead frame is the point used to connect the circuit of the IC packaging with the integrated circuits of various devices. At the second point of circuit connection, the gold wire is touched onto the lead frame and pressed to break the gold wire, producing a flat shape. This point connection is called the second bond, as shown in Fig. 1(b). The quality of the gold wire bonding is checked by bond shear strength. The bond shear test uses a machine to push the first bond area with a constant straight force. This process continues until the connected gold wire is damaged and the maximum strength obtained from the test is recorded. Figure 2(a) shows the picture of the product before the bond shear test. Figure 2(b) shows the picture of the product after the bond shear test, where the bonded gold wire was broken by the bond shear test as denoted by the red circle.



(a) (b)

Fig. 1: (a) Appearance of the first bond, (b) appearance of the second bond.

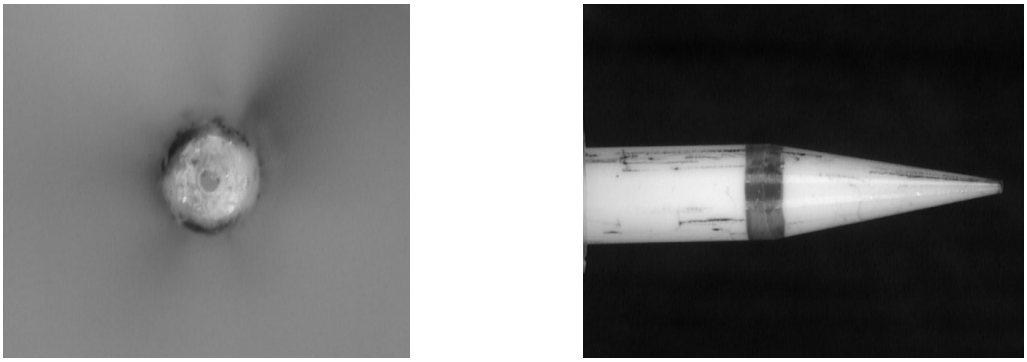


(a) (b)

Fig. 2: (a)The first bond before the bond shear test, (b) the first bond after the bond shear test.

2.1 Machine and Materials

In the experiments, the gold wire bonding capillary that is expired after manufacturing 150,000 work pieces is studied in order to investigate the effects of the factors of the shear forces in the gold wires. The reused capillary is shown in Fig. 3. The machines and materials used in this research include the gold wire bonding machine Model Kns IConn PLUS as shown in Fig. 4(a) and a gold wire size of 1.3 mils to connect the circuit. The 4000 series Multipurpose Bond tester machine was used for the quality control by recording the maximum shear strength, as shown in Fig. 4(b). In the experiments, the gold wire bonding capillary that is expired after manufacturing 150,000 work pieces is studied in order to investigate the effects of the factors of the shear forces in the gold wires.



(a) (b)

Fig. 3: (a) Front view of reused capillary, (b) side view of reused capillary.

(a) (b)

Fig. 4: (a) Wire bonding machine, (b) the shear testing machine.

2.2 Factors and Level of Factors in Experiments

By referring to the relevant studies, the factors of the experiments are identified. There are five factors of the strengths of the gold wire bonding, that use five factors to determine the effect of factors to create a regression model of the process. The factors include the gold bonding pressure (bond force), the gold bonding time (bond time), the ultrasonic energy used (USG current), the current released through the EFO to melt the gold wire (EFO current) and the distance between the EFO and the gold wire (EFO gap). The level of factors in the research was based on operation processes and machine

manuals to determine the appropriate level of each factor used in the experiment so as not to affect the quality of the product, as shown in Table 1.

Table 1: Factors and levels of the experiment

Factors	Level of factors			Unit
	(-1)	0	(+1)	
(A) Bond Force	70	80	90	grams
(B) Bond Time	1	2	3	ms
(C) USG Current	20	35	50	mA
(D) EFO Current	40	45	50	mA
(E) EFO Gap	25	30	35	mils

2.3 Methodology

In this research, the design of the fractional factorial experiment (FFD) was chosen because it is an experimental design that uses half the number of experiments from the design of a full number of factorial experiments [14,15]. In the case of a study using five factors, the FD experiment used a total of 32 experiments, FFD are used to halve the number of experiments, using only 16, while still maintaining efficient results analysis. There is a good method to create a regression model between the test variable and the response, resulting in a clearer understanding of the main effect and interaction effects between each factor and response.

The fractional factorial design was used to determine five factors that affect the bond shear strength. The experiment was repeated twice with 3 centre point experiments per replication to check the curvature of the regression model using a statistical significance level $\alpha = 0.05$ in the experiments. The experiment total is for 38 experiments. The experiments were applied to control other factors that may affect the research response of the three items [16]. The first principle was replication, with two repeated experiments performed; the error of the experiment was calculated and randomized to avoid factors not chosen for analysis [17]. The second was randomization, where the Minitab was used to randomize the trial order to prevent errors. The third was blocking, where all other factors were controlled factors so as not to affect the response of the experiment. The controlled factors included experimenter control, tension, and test experimenter control, which were the same for every experiment; the same machine was used throughout and the gold wire bonding material came from the same lot of material throughout the experimental process. Based on the fractional factorial analysis, results can obtain the regression model to predict bond shear strength over the experimental region. This regression model of the bond shear strength is Eq. 1.

$$y_1 = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 \quad (1)$$

where x_1, x_2, x_3, x_4, and x_5 are the factors of the experiment that represent (A), (B), (C), (D), and (E). $\beta's$ coefficients can be obtained from the effect estimates. β_0 is the average effect of all responses in the 38 experiments.

Residual analysis of the fractional factorial test results for the bond shear strength was a preliminary examination, with the results obtained from the actual experiment shown to be consistent. Three parts were considered: the requirements that the residues must have a normal distribution and, in considering the normal distribution of the residuals, the normal

probability plot and the histogram were further considered. The final test of the residuals of the trial versus the prediction value (versus fits) was considered to be the residual from the trial order (versus order) [18,19].

3. RESULTS

The results obtained from a fractional factorial experiment with 38 trials are shown in Table 2 as the ball shear strengths are recorded.

Table 2: The fractional factorial result of bond shear strength

Experiment parameter									Ball shear strength
Std Order	Run Order	Center Pt	Blocks	Ball Force	Ball Time	USG Current	EFO Current	EFO Gap	
18	1	1	1	90	1	20	40	25	55.23
5	2	1	1	70	1	50	40	25	51.55
9	3	1	1	70	1	20	50	25	52.09
7	4	1	1	70	3	50	40	35	51.51
19	5	1	1	70	3	20	40	25	52.31
10	6	1	1	90	1	20	50	35	54.04
13	7	1	1	70	1	50	50	35	52.27
8	8	1	1	90	3	50	40	25	53.94
31	9	1	1	70	3	50	50	25	51.08
30	10	1	1	90	1	50	50	25	55.35
15	11	1	1	70	3	50	50	25	51.78
37	12	0	1	80	2	35	45	30	54.02
38	13	0	1	80	2	35	45	30	51.64
6	14	1	1	90	1	50	40	35	54.79
35	15	0	1	80	2	35	45	30	52.17
33	16	0	1	80	2	35	45	30	52.95
23	17	1	1	70	3	50	40	35	53.58
2	18	1	1	90	1	20	40	25	55.25
21	19	1	1	70	1	50	40	25	52.80
22	20	1	1	90	1	50	40	35	54.50
29	21	1	1	70	1	50	50	35	51.28
12	22	1	1	90	3	20	50	25	53.69
32	23	1	1	90	3	50	50	35	52.80
16	24	1	1	90	3	50	50	35	54.99
20	25	1	1	90	3	20	40	35	54.69
1	26	1	1	70	1	20	40	35	51.88
14	27	1	1	90	1	50	50	25	53.84
3	28	1	1	70	3	20	40	25	51.80
17	29	1	1	70	1	20	40	35	52.79
28	30	1	1	90	3	20	50	25	54.77
4	31	1	1	90	3	20	40	35	53.49
34	32	0	1	80	2	35	45	30	52.51
27	33	1	1	70	3	20	50	35	51.59
36	34	0	1	80	2	35	45	30	53.39
11	35	1	1	70	3	20	50	35	50.85
26	36	1	1	90	1	20	50	35	53.28
24	37	1	1	90	3	50	40	25	53.02
25	38	1	1	70	1	20	50	25	52.05

3.1 Analysis Effect of Factors

Analysis of the results from the design of the fractional factorial experiments was performed using the Minitab program. The results were found to be a significant factor by a P-value of less than $\alpha = 0.05$, which is statistically significant [20]. The factor affected

https://doi.org/10.31436/iiumej.v22i2.1715

the bond shear strength to a statistically significant degree. The bond force is shown in Table 3, which presents the results of the sum of the square of the factors calculated from the fractional factorial test. Figure 5 shows the factors affecting the bond shear strength in the gold wire bonding process via a normal probability plot of the fractional factorial. The results were used to examine the curvature of the regression model, with the bond shear response evaluated by analysis of variance. The P-value was found to be greater than the level of statistical significance according to the linear regression model.

Table 3:. Analysis of variance (ANOVA) of bond shear strength

Source	DF	Seq SS	Adj SS	Adj MS	F Value	P Value	Significant
Main Effects	5	45.039	45.039	9.008	13.740	0.000	Significant
Bond Force	1	41.598	41.598	41.598	63.430	0.000	Significant
Bond Time	1	1.578	1.578	1.578	2.410	0.136	
USG Current	1	0.014	0.014	0.014	0.020	0.884	
EFO Current	1	1.697	1.697	1.697	2.590	0.123	
EFO Gap	1	0.152	0.152	0.152	0.230	0.635	
Two-Way Interactions	10	4.066	4.066	0.407	0.620	0.780	
Bond Force*Bond Time	1	0.228	0.228	0.228	0.350	0.562	
Bond Force*USG Current	1	0.093	0.093	0.093	0.140	0.711	
Bond Force*EFO Current	1	0.299	0.299	0.299	0.460	0.507	
Bond Force*EFO Gap	1	0.243	0.243	0.243	0.370	0.550	
Bond Time*USG Current	1	0.002	0.002	0.002	0.000	0.956	
Bond Time*EFO Current	1	0.101	0.101	0.101	0.150	0.699	
Bond Time*EFO Gap	1	0.611	0.611	0.611	0.930	0.345	
USG Current*EFO Current	1	0.241	0.241	0.241	0.370	0.551	
USG Current*EFO Gap	1	1.507	1.507	1.507	2.300	0.144	
EFO Current*EFO Gap	1	0.742	0.742	0.742	1.130	0.300	
Curvature	1	0.483	0.483	0.483	0.740	0.400	
Residual Error	21	13.771	13.771	0.656			
Pure Error	21	13.771	13.771	0.656			
Total	37	63.360					

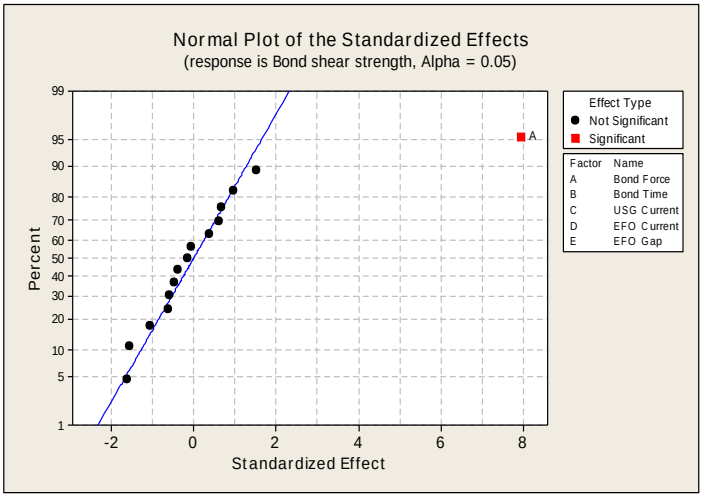


Fig. 5: Normal probability plot of factor that affect to the bond shear strength.

3.2 Residual Analysis

The results of the analysis of the graphs (normal probability plot) and histogram of the experiment showed that the residues of the bond shear strength were normal distribution, with no trend found. For versus fits, there was a random distribution, with no distribution as a trend or pattern occurring. For versus order, the nature of the residual distribution with the trial sequence was random, indicating that the error in the trial sequence was independent, as shown in Fig. 6. The bond shear strength therefore showed no abnormalities or residual tend, showing that the results of the experiment were consistent with the prerequisites of the experiment. Therefore, these experimental results were able to be used.

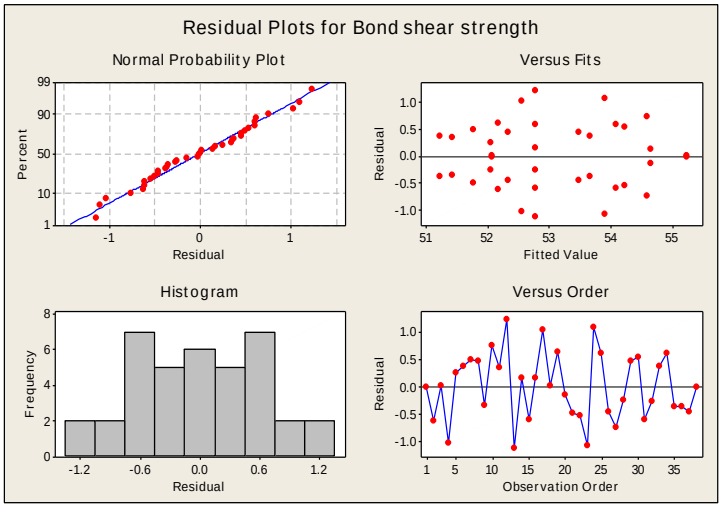


Fig. 6: Residual plot for the fractional factorial test.

3.3 Regression Model

The regression model showed that increasing the bond force factor in the gold wire bonding process would increase the bond shear force strength. For other factors, it will have the opposite effect. The production process should be carefully considered when adjustments are made in order to take advantage of the regression model to maximize performance. The regression model was concluded and is described in Eq. 2.

$$y_1 = 46.64 + 0.11x_1 - 0.86x_2 - 0.11x_3 - 0.08x_4 + 0.24x_5 \quad (2)$$

where, y_1 is the bond shear strength, x_1 is the bond force, x_2 is the bond time, x_3 is the USG current, x_4 is the EFO current, and x_5 is the EFO gap.

4. CONCLUSION

The statistical method of fractional factorial experiment is a suitable method to determine the effect of large factors in the gold wire bonding process. The experiment was used to determine the effect of five factors by two replications and 3 center point experiments per replication to check the curvature of the regression model. The statistical significance level of the experiments is $\alpha = 0.05$, a total of 38 trial experiments were performed. The experiment factors of bond force, bond time, USG current, EFO current, and EFO gap were investigated. The factor found to affect the bond shear strength in the process is the bond force factor of the gold wire bonding process. The regression model

demonstrated the correlation between coherence that would influence mutual bonding and thereby increase the bond shear strength of the gold wire and increase the strength of the bond force. In these experiments, the expired gold wire bonding capillary is studied. The factors of the shear forces of the welds are different from that of the relevant studies with many factors of the strengths of the welds. For reusing the parts, the bond force is the only factor that significantly affects the increase in the shear force. Therefore, if the manufacturers consider to re-use the expired gold wire bonding capillaries or to increase the useful lives, they can just adjust the bond force of welding to obtain the appropriate shear force.

REFERENCES

[1] Krajcsák Z. (2019) Implementing open innovation using quality management systems: the role of organizational commitment and customer loyalty. J. of Open Innovation: Technology, Market, and Complexity, 5(4): 90. https://doi.org/10.3390/joitmc5040090

[2] Prabowo R, Singgih ML, Karningsih PD, Widodo E. (2020) New product development from inactive problem perspective in Indonesian SMEs to open innovation. Journal of Open Innovation: Technology, Market, and Complexity., 6(1): 20. https://doi.org/10.3390/joitmc6010020

[3] Phimpha C, Sindhuchao S. (2020) The process of reusing machine part in gold wire bonding process of IC packaging product. The International Academic Conference on Engineering, Technology and Innovations. Bali. Indonesia, pp.13-14.

[4] Hung YH. (2007) Optimal process parameters design for a wire bonding of ultra-thin CSP package based on hybrid methods of artificial intelligence. Microelectronics International, 24(3): 3-10.

[5] Wang, Sun R. (2009) The quality test of wire bonding. Modern Applied Science, 3(12): 50-56.

[6] Bonding W. (2008) The great debate: Copper vs. gold ball bonding. Advanced Packaging, 17(7): 20.

[7] Kim BC, Ha SJ, Yang JK, Lee IC, Kang DS, Han BS, Han YJ. (2017) Process capability optimization of ball bonding using response surface analysis in light emitting diode (LED) wire bonding. Journal of the Korea Academia-Industrial Cooperation Society, 18(4): 175-182.

[8] Zhong ZW. (2009) Fine and ultra-fine pitch wire bonding: challenges and solutions. Microelectronics International, 22(2): 10-18.

[9] Sciammarella FM, Salehi Najafabadi B. (2018) Processing parameter DOE for 316L using directed energy deposition. J. of Manufacturing and Materials Processing, 2(3): 61.

[10] Wu YX, Long ZL, Lei HAN, Zhong J. (2006) Temperature effect in thermosonic wire bonding. Transactions of Nonferrous Metals Society of China, 16(30): 618-622.

[11] Satianrangsarith W, Tirakanogsathit M. (2012) Design of experiments approach for improving wire bonding quality. International Journal of Innovation, Management and Technology, 3(4): 327-331.

[12] Lu D, Wong CP. (2018) Materials for advanced packaging. Springer International Publishing, Switzerland.

[13] Antony J. (2014) Design of experiments for engineers and scientists. Elsevier.

[14] Montgomory DC, Runger GC. (2011) Applied Statistics and Probability for Engineering. John Wiley & Son Inc, U.S.A.

[15] Gong Y, Song J, Lin S, Yang J, He Y, Tan G. (2020) Design Optimization of Rubber-Basalt Fiber-Modified Concrete Mix Ratios Based on a Response Surface Method. Applied Sciences, 10(19): 6753.

[16] Mayer RM, Montgomery DC. (2011) Response surface methodology: Process and product Optimization using design experiment. John Wiley & Sons Inc, U.S.A.

[17] Sabry I, El-Kassas AM, Mourad AHI, Thekkuden DT, Abu Qudeiri J. (2019) Friction stir welding of T-Joints: Experimental and statistical analysis. J. of Manufacturing and Materials Processing, 3(2): 38.
[18] Martin J, De Adana DDR, Asuero AG. (2017) Fitting models to data: residual analysis, a primer. Chapter 7.
[19] Mongomery DC. (2009) Design and analysis of experiment. John Wiley & Sons Ltd U.S.A.
[20] Mongomery DC. (2009) Statistical Quality Control a Modern introduction. John Wiley & Sons Ltd U.S.A.

https://doi.org/10.31436/iiumej.v22i2.1747

MECHANICAL PROPERTIES, WATER ABSORPTION, AND FAILURE ANALYSES OF KENAF FIBER REINFORCED EPOXY MATRIX COMPOSITES

**IKHWAN YUSUFF, NORSHAHIDA SARIFUDDIN*
SITI NORBAHIYAH AND AFIFAH MOHD ALI**

Department of Manufacturing and Materials Engineering,
Kulliyyah of Engineering, International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia

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

(Received: 16th December 2020; Accepted: 22nd February 2021; Published on-line: 4th July 2021)

ABSTRACT: The potential of natural fibers as one of the candidate materials in the production of fiber-reinforced polymer composites have been widely investigated. In the current study, natural fiber-reinforced polymer composite was fabricated by employing woven kenaf fiber as a reinforcing agent with epoxy resin that acts as a matrix constituent. This composite sample was fabricated using the application of the vacuum infusion method in which the content of kenaf fibers was varied from 30 vol.%, 40 vol.%, and 50 vol.%. The effects of different fiber loadings toward mechanical and physical properties as well as failure properties of kenaf composite were then evaluated. Kenaf composites were subjected to mechanical tests including tensile and flexural tests. The result shows that the highest tensile strength and modulus were attained at 76.67 MPa and 2.31 GPa, respectively with kenaf composite fabricated with 40 vol.% fiber content. Meanwhile, the highest flexural strength and modulus were recorded at 61.24 MPa and 4.20 GPa, also corresponding to kenaf composite that is loaded with 40 vol.% fibers. Fiber pull-out failure was able to be detected in fabricated kenaf composites. Meanwhile, fiber breakage resulting from flexural failure could also be observed in the kenaf composite samples. Apart from that, it was found that as more kenaf fiber was loaded in the composites, the rate of water absorption tended to increase where the highest rate of water absorption was found at 43.33%, displayed by kenaf composite with 50 vol.% of fiber content.

ABSTRAK: Potensi gentian semula jadi sebagai salah satu bahan dalam penghasilan komposit polimer bertetulang gentian telah banyak dikaji. Dalam kajian terkini, komposit polimer yang diperkuat dengan gentian semula jadi dibuat dengan menggunakan serat kenaf tenunan sebagai agen penguat dan resin epoksi yang bertindak sebagai matriks. Sampel komposit ini dibuat menggunakan kaedah infusi vakum di mana kandungan serat kenaf digunakan adalah 30 vol.%, 40 vol.%, dan 50 vol.%. Kesan kandungan serat yang berbeza terhadap sifat mekanikal dan fizikal serta sifat kegagalan komposit kenaf kemudiannya dinilai. Komposit Kenaf diuji dengan ujian tegangan dan lenturan. Hasilnya menunjukkan bahawa kekuatan tegangan dan modulus tertinggi dicapai pada 76.67 MPa dan 2.31 GPa, milik komposit kenaf yang dibuat dengan kandungan serat 40 vol.%. Sementara itu, kekuatan dan modulus lenturan tertinggi dicatatkan pada 61.24 MPa dan 4.20 GPa juga milik komposit kenaf yang dimuatkan dengan serat 40% vol. Kegagalan serat terkeluar dapat dikesan pada komposit kenaf buatan. Sementara itu, kerosakan serat akibat kegagalan lenturan juga dapat dilihat pada sampel komposit kenaf. Selain itu, didapati bahawa semakin banyak serat kenaf yang dimuatkan dalam komposit, cendurung meningkatkan kadar penyerapan air di mana

kadar penyerapan air tertinggi didapati pada 43.33% yang ditunjukkan oleh komposit kenaf dengan kandungan serat 50% vol.

KEYWORDS: *natural fiber; kenaf composite; mechanical properties*

1. INTRODUCTION

The exploitation of natural fibers such as kenaf, hemp, sisal, and bamboo has been widely adopted in the production of fiber-reinforced composite materials [1,2]. The growth of natural-based composites is due to the properties of natural fibers, especially in terms of cost and biodegradability characteristics [3,4]. Moreover, the excellent biodegradability displayed by natural fibers can impart a positive impact on the environment due to the ability to form green bio-composite materials. Furthermore, the capability of natural fibers in offering outstanding attributes including non-abrasiveness, low density, as well as exhibiting a comparable specific strength, is another driving force in the utilization of this fiber as a reinforcement in fiber-reinforced composite materials [5,6].

Therefore, there are many works available in open literature that have reported on the potential of natural fiber utilization in many engineering applications such as construction, electronic, and automotive industries [7,8]. For example, Javadian et al. [9] have investigated the mechanical properties of bamboo fiber to be embedded in composites for construction uses. Moreover, Gupta [10] has studied the potential of jute composites for electronic applications. The author has found that jute composite shows acceptable dynamic mechanical properties that are suitable for application in making the casing for electronic devices such as mobile phones and laptops. Recently, Yang et al. [11] have reviewed the future of cellulosic fiber-based (kenaf, bamboo, banana, and pineapple leaf) composites for marine applications. The authors have mentioned that cellulosic fibers (natural fibers) gain great attention in marine applications due to the drawbacks of synthetic fibers, including high cost, poor recyclability, and high embodied energy.

Moreover, synthetic-based composites are often disposed of in improper ways. Incineration after end-of-service may cause serious air pollution and this is another reason that prompted the idea of developing natural fiber-based composites [12]. It is believed that the introduction of natural fibers in composite industries has created a new option, especially for the manufacturers in developing cost-effective green products with comparable mechanical properties. Recently, the potential of kenaf fibers to be used as door panels in automotive parts has been discovered [13]. Thus, it is necessary to have continuous research that explores the potentials, limitations, and barriers in growing natural fiber composites so that the properties and behaviors of this composite can be properly verified. Therefore, the current research is one of the approaches taken to ensure continuity in the development of natural fiber-based composites in order to discover the abilities of these natural resources for future use.

2. MATERIALS AND EXPERIMENTAL DESIGNS

2.1 Materials

Plain woven kenaf fiber with an average thickness of 0.8×10^{-3} m and density of 1220 kg/m^3 was utilized as a reinforcing element. Meanwhile, epoxy resin grade INF – 114 with corresponding hardener INF – 212 and density of 1140 kg/m^3 was applied as a matrix

constituent to bind the fibers where the ratio of epoxy-to-hardener was fixed at 100:24.7 (g).

2.2 Preparation of Kenaf Reinforced Epoxy Matrix Composites

Kenaf composites were fabricated based on three different fiber loadings including 30 vol.%, 40 vol.%, and 50 vol.%. Woven kenaf fibers were cut into dimensions of 300 mm × 150 mm and 150 mm × 150 mm for tensile test and flexural test, respectively. Meanwhile, pristine epoxy was also fabricated as a control sample. The fabrication process involves the application of vacuum infusion method in which epoxy resins were infused within the layer of kenaf fibers using constant pressure at -100 kPa supplied from a vacuum pump. The curing process at room temperature was performed for 24 hours once the epoxy resin completely covered the plies of kenaf fibers in order to solidify the composite panels. Fig. 1 illustrates the schematic diagram of experimental set up for fabrication of kenaf composite panels via vacuum infusion method.

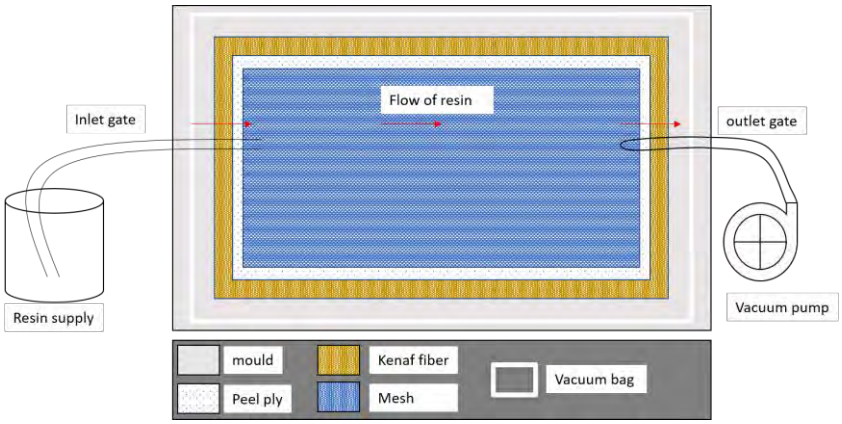


Fig. 1: Schematic diagram of the vacuum infusion layout in the fabrication of kenaf composite panel.

2.3 Characterization of Kenaf Composite Panels

Kenaf composites were subjected to three main characterizations including mechanical, physical, and morphological analyses. The mechanical properties were characterized under tensile and flexural tests. For tensile test, samples were cut into dimension of 250 mm × 25 mm in accordance with ASTM D3039. This testing was performed using a Universal Testing Machine (UTM) Instron 5582 with a load of 100 kN and crosshead speed of 2 mm/min. Meanwhile, for the flexural test, the same machine was used where samples were prepared with a dimension of 127 mm × 13 mm corresponding to ASTM D970. The same load of 100 kN and crosshead speed 2 mm/min were applied during the bending test.

The water absorption (WA) test was carried out based on ASTM D570 in which the fabricated kenaf composites were dried inside an oven at a temperature of 70 ℃ for 24 hours in order to identify the initial weight. Then, samples were immersed inside the distilled water. The weight of composites was measured every day until 30 days and the rate of water absorption was calculated using the following formula:

$$WA = \frac{M_1 - M_0}{M_0} \times 100 \quad (1)$$

where WA stands for water absorption, M_0 denotes the initial mass of the specimen, and M_1 represents the dried mass of the sample after removing from the distilled water.

The tensile fracture surfaces were characterized using a scanning electron microscope (SEM) model JOEL JSM-5600. The specimens were coated with Palladium (Pd) where the coating process was performed using a Quarom SC7620 Sputter Coater at 10 Kv voltage. Flexural failure modes were examined using an optical microscope (Nikon Measuring Microscope Trinocular Head, model MM-TRF).

3. RESULTS AND DISCUSSION

3.1 Tensile Properties of Kenaf Reinforced Epoxy Matrix Composites

Figure 2(a,b) shows the tensile properties of fabricated kenaf composites at three different fiber loadings (30 vol.%, 40 vol.%, and 50 vol.%) and also pristine epoxy as a control sample. From the figure, the tensile strength and modulus of pristine epoxy was recorded at 45 MPa and 1.52 GPa, respectively. The incorporation of kenaf fiber has improved the values of tensile strength and the modulus of the fabricated composite panel. Moreover, it can be noticed that, when the fiber content was increased from 30 vol.% to 40 vol.%, the tensile strength of the fabricated kenaf composites increased. However, when kenaf fibers were loaded up to 50 vol.% in the composite system, the value of tensile strength was prone to be decrease. Specifically, the highest tensile strength attained was 76.67 MPa when 40 vol.% fibers were loaded into the system where the performance of tensile strength decreased by 7.45% (70.96 MPa) when 30 vol.% fiber content was utilized. Surprisingly, the value of tensile strength drastically decreased by 34.08% (50.54 MPa) when the level of fibers increased to 50 vol.%. The same trend as tensile strength was observed in the value of the tensile modulus, which increased in the level of fiber content from 30 vol.% (2.09 GPa) to 40 vol.% (2.31 GPa). This improved the behavior of tensile modulus. Whereby, the increment of fiber content up to 50 vol.% (1.88 GPa) resulted in a decrement in the performance of the tensile modulus.

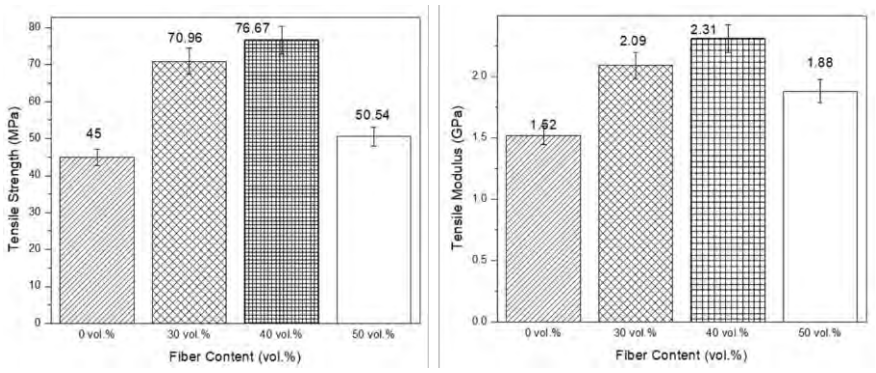


Fig. 2: (a) Tensile strength and (b) tensile modulus of fabricated kenaf composites at different fiber contents.

It is believed that the behavior of the tensile properties was influenced by the adhesion characteristic between the plies of kenaf fibers with epoxy resin. This is in agreement with Cisneros-López et al. [14] where outstanding tensile properties were probably due to the formation of a better fiber-matrix interface that resulted from excellent adhesion between fiber and matrix phases. With regards to the current work, the results indicate that the epoxy resin can provide sufficient support in binding the kenaf fibers at 40 vol.% fiber content in which the formation of high lamination quality was achieved at this condition. With regards to the current work, the results indicate that the epoxy resin can provide sufficient support in binding the kenaf fibers at 40 vol.% fiber content in which the

formation of high lamination quality was achieved at this condition. It is due to its ability to form a strong fiber-matrix interface between the kenaf fibers and the epoxy resin. Therefore, 40 vol.% kenaf composite is able to endure more tension load as compared to 30 vol.% and 30 vol.% before experiencing permanent tensile failure. Figure 3 shows the stress-strain curve of kenaf composites at different fiber loadings. From the stress-strain curve, a steep slope can be observed in the kenaf composite that was loaded with 40 vol.% fibers. In fact, the steep slope indicates the composite material tends to exhibit higher tensile modulus as compared to the composite that displays the low slope [15]. It is due to the ability to reach a high fracture point before experiencing external tensile failure. It seems that kenaf composites fabricated with 40 vol.% display higher fracture points than 30 vol.% and 50 vol.% kenaf composites. Meanwhile, the lowest fracture point was noticed in kenaf composite that produced using 50 vol.% of fibers. These findings have validated the results obtained in tensile properties discussed in the previous point.

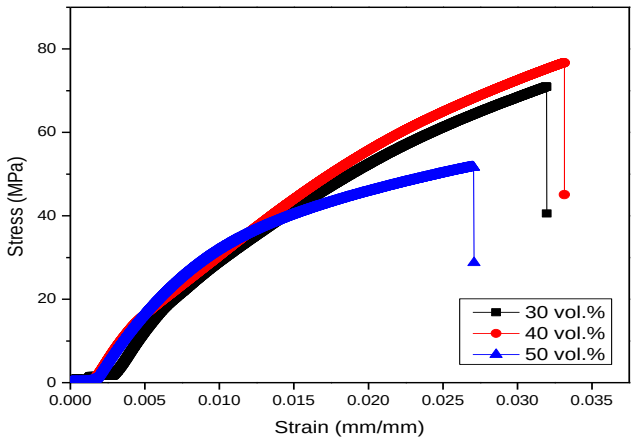


Fig. 3: The stress-strain curve of fabricated kenaf composites at different fiber loadings.

Figure 4(a–f) shows the micrographs for each of the kenaf composites. The appearance of fiber agglomeration due to poor wettability between fibers and matrix can be detected in the fabricated kenaf composites especially in kenaf composites 50 vol.% of fiber content. Severe fiber agglomeration in the 50 vol.% kenaf composite signifies this composite panel experienced poor adhesion between fibers and matrix constituent in which fibers can separate easily from the matrix during the application of tension loads. Based on Fig. 4(b), kenaf composite that was loaded with 40 vol.% fibers was able to display better fiber-matrix lamination than 30 vol.% and 50 vol.% kenaf composites. At high magnification, all fabricated kenaf composites exposed toward fiber pull-out failure as shown in Fig. 4(d–f). From Fig. 4(f), it can be noticed that the presence of the gap between matrix phase and fiber appeared in kenaf composite with 50 vol.% fiber loading. This reveals that epoxy resin was unable to bind the fibers efficiently and therefore disturbed the process of transferring tension load from matrix to fibers. As a result, kenaf composite with 50 vol.% was prone to show poor tensile strength and modulus. This is in agreement with the findings from the previous discussion.

3.2 Flexural Properties of Kenaf Reinforced Epoxy Matrix Composites

Flexural properties of fabricated pristine epoxy and kenaf composites at three different fiber loadings, namely 30 vol.%, 40 vol.%, and 50 vol.% are shown Fig. 5(a,b). Based on Fig. 5(a), the value of flexural strength and modulus of pristine epoxy stated at 48.52 MPa and 3.26 GPa, respectively.

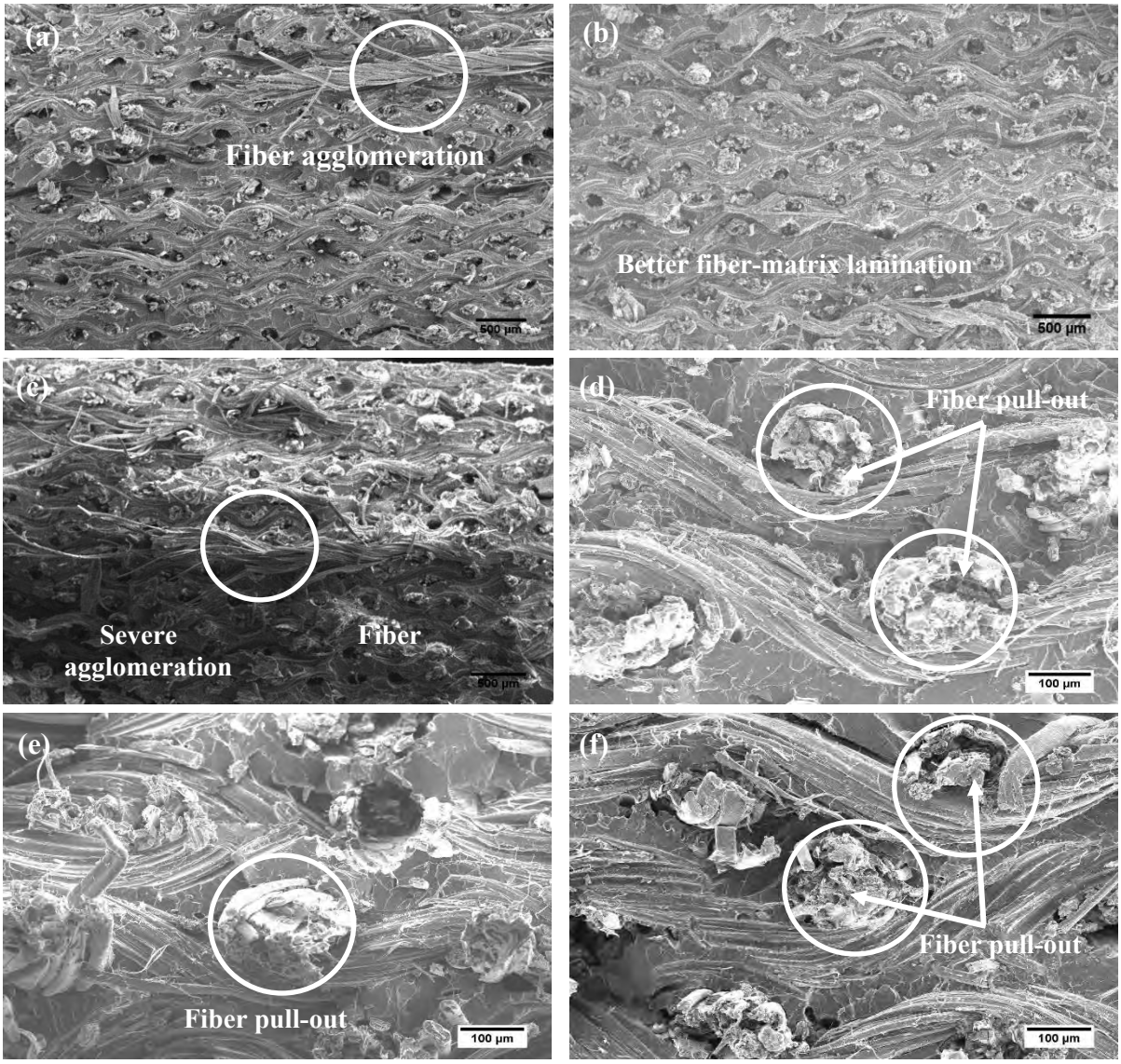
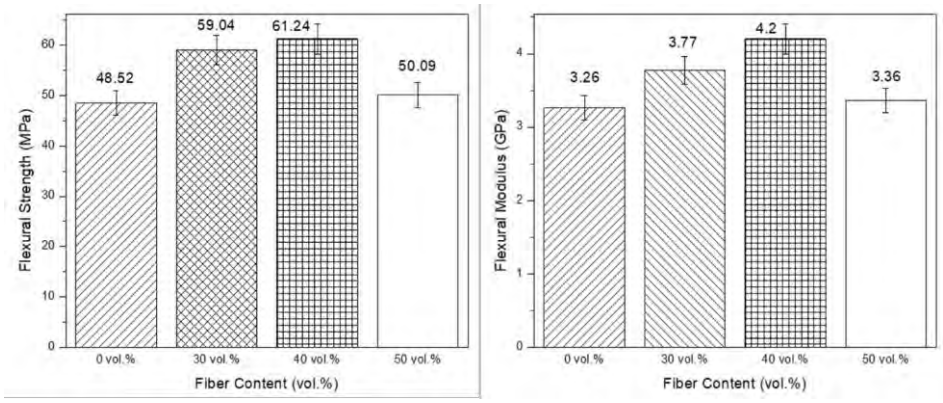


Fig. 4: Tensile fracture micrographs of fabricated kenaf composites – (a) 30 vol.%, (b) 40 vol.%, and (c) 50 vol.% fiber loadings at ×25 magnification and (d) 30 vol.%, (e) 40 vol.%, and (f) 50 vol.% fiber loadings at ×150 magnification.

The increment in the trend of flexural strength can be spotted when kenaf fibers were introduced into the composite system. Moreover, when kenaf fiber was increased from 30 vol.% to 40 vol.%, the highest flexural strength can be observed at the value recorded at 61.24 MPa. It is believed that this increment is due to the adequacy of epoxy resins to laminate the fiber efficiently [16]. As a result, once the bending load touched the 40 vol.% kenaf composites, the loads were carried by the whole system rather than the individual ply. Therefore, greater loads are needed to break the sample before suffering from flexural failure. The introduction of 50 vol.% kenaf fibers result in reducing the performance of the composite material where the value of flexural strength was reduced by 16.88% in comparison to 40 vol.% fiber content. It indicates that, the utilization of kenaf fibers up to 50 vol.% was unable to provide optimum reinforcing effect toward the composite system. Moreover, at high level of fiber content, it is difficult for natural-based composite materials to maintain their strong fiber-matrix interface due to poor compatibility with most polymer matrices [17]. Therefore, this composite material is incapable of staying longer under bending loads, hence, resulting in poor flexural properties. A similar pattern

of flexural strength was found in the values of flexural modulus in which the highest flexural modulus gained at 4.2 GPa belongs to kenaf composite loaded with 40 vol.% fibers. The values of flexural modulus decreased by 10.23% and 25% with the employment of 30 vol.% and 50 vol.% fiber content, respectively.



(a) (b)

Fig. 5: (a) Flexural strength and (b) flexural modulus of kenaf composites fabricated with different fiber contents.

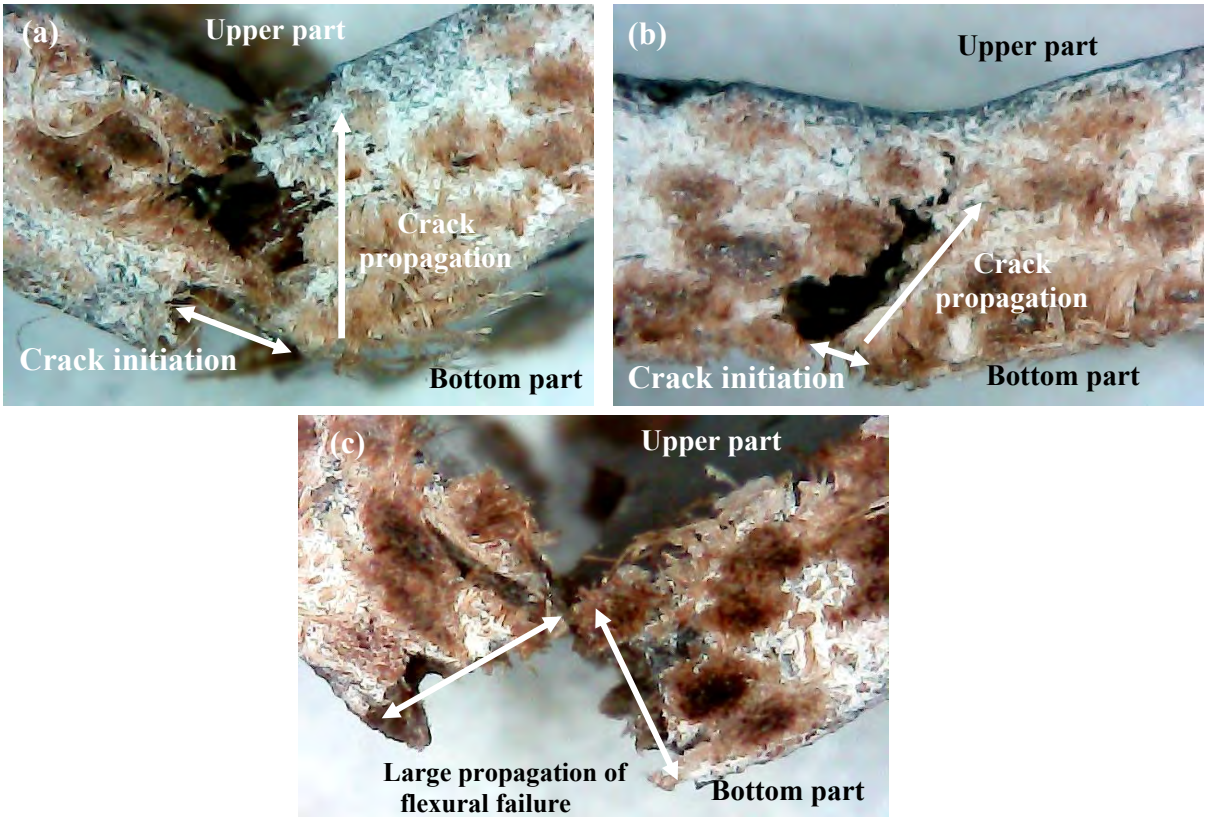


Fig. 6: Flexural failure of kenaf composites at (a) 30 vol.%, (b) 40 vol.%, and (c) 50 vol.% of fiber content.

Figure 6(a–c) visualizes failure modes of kenaf composites under a flexural condition. The behavior of crack initiation and propagation can be explained based on this figure. The cracks start to initiate at the bottom part of the kenaf composites when the upper part of composite panels are unable to resist the applied bending load due to high stress

concentration. Then, these cracks spread and propagate in the middle layer of kenaf composites until reaching the maximum fracture point and causing permanent bending failure. Based on Fig. 6(a-c), all the fabricated kenaf composites suffered from fiber breakage failure in which this failure was probably due to the presence of microcracks resulting from the incapability of composite panels to maintain their strong interfacial bonding between fiber-matrix interfaces [18]. Severe fiber breakage occurred in the kenaf composite that was fabricated using 50 vol.% fibers where it caused the composite panel to separate. This indicates that cracks can infiltrate easily within the layers of 50 vol.% kenaf composite due to poor interfacial bonding between layers of kenaf fiber and epoxy resins. Based on Fig. 6(b), less crack propagation can be noticed in the kenaf composite with 40 vol.% fiber content contrasted over 30 vol.% and 50 vol.% composite panels. The formation of a better fiber-matrix interface due to high wettability between kenaf fibers and epoxy resins has slowed down the propagation of cracks within the composite body. As a result, kenaf composite with 40 vol.% compositions of fibers capable of enduring more bending loads and consequently helping this composite to exhibit better flexural properties as compared to the other kenaf composites.

3.3 Water Absorption Behavior of Kenaf Composite Panels

Figure 7(a) shows the rate of water absorption for kenaf composites fabricated with three different fiber loadings. The rate of water absorption was measured every day up to 30 days. At the beginning of the process, the rate of water absorption for kenaf composites in all-fiber loadings increased linearly, then the trend of water uptake started to slow down until reaching saturation after a prolonged period. This trend was in accordance with the diffusion theory known as the Fickian diffusion process [19]. Based on the graph, the highest rate of water absorption was displayed by kenaf composite fabricated using 50 vol.% fibers (43.33%), followed by kenaf composite with 40 vol.% fiber content (30.46 %). Whereby, the lowest rate of water absorption recorded at 25.49% belongs to kenaf composite that loaded with 30 vol.% fibers. These values indicate that increases in the amount of kenaf fibers in the composite system tend to increase the rate of water absorption of fabricated composite panels.

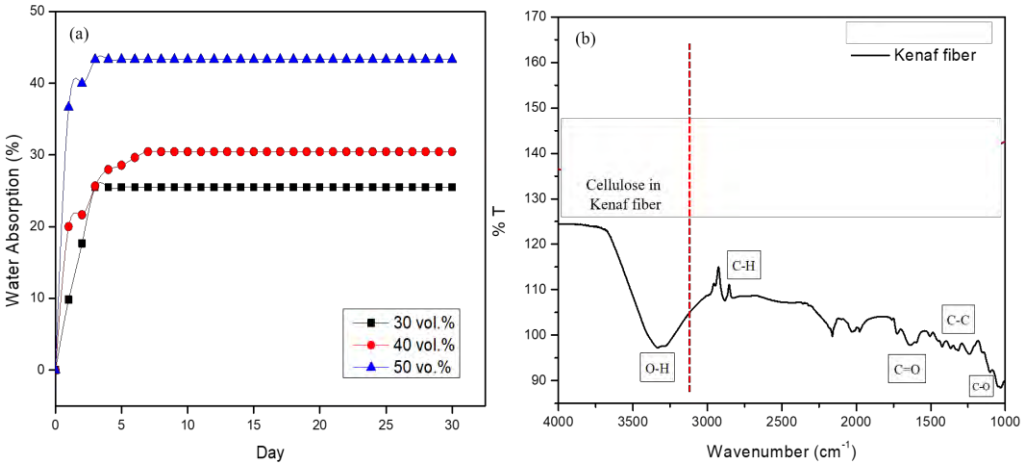


Fig. 7: The rate of water absorption of kenaf composites at 30 vol.%, 40 vol.%, and 50 vol.% fiber content for 30 days.

This phenomenon occurred due to the nature of natural fibers in which they are very responsive to the watery environment [20]. The presence of cellulose components in kenaf fibers denotes hydroxyl groups that are the principal contributors to moisture absorption

and increase the attractiveness to the watery condition [21]. The presence of OH groups from cellulose has been confirmed by FTIR spectra as shown in Fig. 7(b). Thus, increasing the level of kenaf fibers will increase the tendency of the composite panel to absorb more water as compared to the composite with fewer kenaf fibers composition. Moreover, it also believed that high water absorption, especially in kenaf composite with 50 vol.% fiber contents, is probably due to the ineffectiveness of epoxy resins to play their role as a barrier in protecting the composite from the watery environmental attack. Therefore, it may contribute to the fiber swelling phenomena in which this event can promote the appearance of microcracks [21]. Thus, the process of water uptakes becomes dominant and active in the microcrack region. As a result, kenaf composite panels with higher fiber content (50 vol.%) are prone to absorb more water in comparison to kenaf composites that are loaded with low fiber content (30 vol.% and 40 vol.%).

4. CONCLUSION

The behaviors of kenaf composites in terms of mechanical properties, water absorption, and failure analyses have been studied where the effects of fiber contents toward these properties were carefully evaluated. Thus, it can be concluded that:

a) The incorporation of kenaf fibers from 30 vol.% to 40 vol.% successfully improved the mechanical properties (tensile and flexural properties) of fabricated composite panels as compared to the pristine epoxy. Whereas the utilization of kenaf fibers up to 50 vol.% has reduced the performances of tensile and flexural properties of kenaf composites.

b) Common tensile failure, which is fiber pull-out able to be detected in all fabricated kenaf composites, and fiber breakage failures appeared in the flexural failure. These failures were severely found in the kenaf composite that was fabricated using 50 vol.% fiber content.

c) Increasing in the level of kenaf fibers is prone to increase the rate of water absorption of fabricated kenaf composite panels.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to the Ministry of Higher Education (MOHE) of Malaysia for the support through the awarded grant (FRGS 17-033-0599), Kulliyyah of Engineering, International Islamic University Malaysia (IIUM), and also to the IIUM Research Management Centre for facilitating this project.

REFERENCES

[1] Saadati Y, Lebrun G, Chatelain JF and Beauchamp Y (2020) Experimental investigation of failure mechanisms and evaluation of physical/mechanical properties of unidirectional flax–epoxy composites. J. Composite Materials, 54(20): 1-21. doi:10.1177/0021998320902243

[2] Dhaliwal GS, Dueck SM, Newaz GM. (2019) Experimental and numerical characterization of mechanical properties of hemp fiber reinforced composites using multiscale analysis approach. SN Appl. Sci., 1: 1361. https://doi.org/10.1007/s42452-019-1383-6

[3] Madhu P, Sanjay MR, Senthamaraikannan P, Pradeep S, Saravanakumar SS and Yogesha B (2018) A review on synthesis and characterization of commercially available natural fibers: Part II. Journal of Natural Fibers, 16: 25-36. doi:10.1080/15440478.2018.1453433

[4] Thyavihalli Girijappa YG, Mavinkere Rangappa S, Parameswaranpillai J and Siengchin S (2019) Natural fibers as sustainable and renewable resource for development of eco-friendly composites: a comprehensive review. Frontiers in Materials, 6: 1-14. doi:10.3389/fmats.2019.00226

[5] Mochane MJ, Mokhena TC, Mokhothu TH, Mtibe A, Sadiku ER, Ray SS, Ibrahim ID, Daramola OO (2019) Recent progress on natural fiber hybrid composites for advanced applications: A review. Express Polymer Letters, 13(2): 159-198. doi:10.3144/expresspolymlett.2019.15

[6] Siakeng R, Jawaid M, Ariffin H, Sapuan SM, Asim M and Saba N (2018) Natural fiber reinforced polylactic acid composites: A review. Polymer Composites, 40: 446-463. doi:10.1002/pc.24747

[7] Navaneethakrishnan G, Karthikeyan, T, Saravanan S, Selvam V, Parkunam N, Sathishkumar G and Jayakrishnan S (2019) Structural analysis of natural fiber reinforced polymer matrix composite. Materials Today: Proceedings, 21: 7-9. doi:10.1016/j.matpr.2019.05.295

[8] Kumar R, Ul Haq MI, Raina A and Anand A (2018) Industrial applications of natural fibre-reinforced polymer composites – challenges and opportunities. International Journal of Sustainable Engineering, 12: 212-220. doi:10.1080/19397038.2018.1538267

[9] Javadian A, Smith Ian FC, Nazanin F and Hebel DE (2019) Mechanical properties of bamboo through measurement of culm physical properties for composite fabrication of structural concrete reinforcement. Frontiers in Materials, 6: 1-15.

[10] Gupta MK. (2018) Effect of variation in frequencies on dynamic mechanical properties of jute fibre reinforced epoxy composites. Journal of Materials and Environmental Sciences, 9: 100-106. https://doi.org/10.26872/jmes.2018.9.1.12

[11] Yang MFM, Hamid H, Abdullah AM. (2018) Potential use of cellulose fibre composites in marine environment - A Review. in: Öchsner A. (eds) Engineering Applications for New Materials and Technologies. Advanced Structured Materials, Springer, Cham, 86: 22-55.

[12] Kaiser K, Schmid M, Schlummer M. (2017) Recycling of polymer-based multilayer packaging: A review. Recycling, 3: 1-26. doi:10.3390/recycling3010001

[13] Yuhazri MY, Amirhafizan MH, Abdullah A, Husin MA, Kamarul AM and Lau STW (2018) Kenaf fibre composites as promising green-composites for automotive car door map pocket application. International Journal of Mechanical & Mechatronics Engineering, 8: 15-21.

[14] Cisneros-López EO, González-López ME, Pérez-Fonseca AA, González-Núñez R, Rodrigue D and Robledo-Ortíz JR (2016) Effect of fiber content and surface treatment on the mechanical properties of natural fiber composites produced by rotomolding. Composite Interfaces, 24: 35-53. doi:10.1080/09276440.2016.1184556

[15] Alberto M. (2013) Introduction of fibre-reinforced polymers − Polymers and Composites: Concepts, Properties and Processes. Fiber Reinforced Polymers - The Technology Applied for Concrete Repair IntechOpen, pp. 3-40.

[16] Harsha AP and Arjula S. (2010) Influence of fibre orientation on three-body abrasive wear behaviour of unidirectional carbon fibre-reinforced polyetherimide composite. Journal of Materials Science, 46: 983-991. doi:10.1007/s10853-010-4856-y

[17] Adekomaya O, Jamiru T, Sadiku R and Huan Z (2017) Negative impact from the application of natural fibers. Journal of Cleaner Production, 143: 843-846. doi:10.1016/j.jclepro.2016.12.037

[18] Wang W, Chouw N, Jayaraman K. (2016) Effect of thickness on the impact resistance of flax fibre-reinforced polymer. J. Reinforced Plastics and Composites, 35: 1277-1289. doi:10.1177/0731684416648780

[19] Peret T, Clement A, Freour S, Jacquemin F. (2019) Homogenization of Fickian and non-Fickian water diffusion in composites reinforced by hydrophobic long fibers: Application to the determination of transverse diffusivity. Composite Structures, 226: 1-9. doi:10.1016/j.compstruct.2019.111191

[20] Rozali NA, Abu Bakar MB, Masri MN, Sulaiman MA, Mohamed M, Ahmad Thirmizir MZ (2017) Mechanical and water absorption properties of hybrid kenaf/glass fibre mat reinforced unsaturated polyester composites. Materials Science Forum, 888: 228-233. doi:10.4028/www.scientific.net/msf.888.228

[21] Maslinda AB, Abdul Majid MS, Ridzuan MJM, and Syayuthi ARA (2017). Water absorption behaviour of hybrid interwoven cellulosic fibre composites. IOP Conf. Series: Journal of Physics: Conf. Series, 908: 012015. doi:10.1088/1742-6596/908/1/012015

https://doi.org/10.31436/iiumej.v22i2.1768

EFFECT OF MAGNETIC FIELD ON THE PHYSICAL AND CHEMICAL PROPERTIES OF FLOWING LUBRICATING COOLING LIQUIDS USED IN THE MANUFACTURING PROCESS

ERKIN ODILOV[1], UMIDJON MARDONOV[2*], KHUSNIDDIN ABDIRAKHMONOV[2], ABDUGANI ESHKULOV[1] AND BEHZOD RAKHMATOV[3]

[1]Faculty of Mechanic,
[2]Faculty of Machine Construction,
Tashkent State Technical University, Tashkent, Uzbekistan
[3]Faculty of Chemical Technology of Inorganic Substances,
Tashkent Institute of Chemical Technologies, Tashkent, Uzbekistan

[*]Corresponding author: fff8uma@mail.ru

(Received: 5th January 2021; Accepted: 5th March 2021; Published on-line: 4th July 2021)

ABSTRACT: In this paper, the effect of magnetic field on lubricating cooling liquids, which are used in the cutting process in manufacturing, was studied. We chose three different lubricating cooling liquids that are commonly used in local manufacturing factories to conduct the experiment. Three main properties of these lubricoolants, boiling point, kinematic viscosity, and density, were analysed after magnetizing them. The magnetization processes were conducted in two conditions of liquids. At the first stage, the authors magnetized the liquids in stationary conditions; at the second stage, they magnetized the flowing liquids and analysed the difference among all the obtained results. This article shows the results of the comparisons and analyses the magnetic field influence on different types of fluids. Moreover, the paper investigates the dependence of magnetic field strength on the influence of magnetic field on liquids. It was found that the examined three parameters of liquids were changed after magnetic field treatment. The finding of this research offered a simple approach to improve the lubricating and cooling process in machining details in manufacturing.

ABSTRAK: Dalam makalah ini, pengaruh medan magnet pada cecair pendingin pelincir, yang digunakan dalam proses pemotongan dalam pembuatan, telah dikaji. Kami memilih tiga cecair penyejuk pelincir berbeza yang biasa digunakan di kilang pembuatan tempatan untuk menjalankan eksperimen. Tiga sifat utama pelincir ini seperti takat didih, kelikatan kinematik, dan ketumpatan dianalisis setelah memagnetkannya. Proses pembesaran dilakukan dalam dua keadaan cecair. Pada peringkat pertama, kami memagnetkan cecair dalam keadaan pegun dan membandingkan hasilnya; pada peringkat kedua, kami memagnetkan cecair semasa mengalir dan menganalisis perbezaan antara hasil yang diperoleh. Artikel menunjukkan hasil perbandingan ini dan menganalisis pengaruh medan magnet pada pelbagai jenis cecair. Lebih-lebih lagi, makalah ini meneliti pergantungan kekuatan medan magnet terhadap pengaruh medan magnet pada cecair. Didapati bahawa tiga parameter cecair yang diperiksa diubah setelah rawatan medan magnet. Penemuan penyelidikan ini menawarkan pendekatan mudah untuk meningkatkan proses pelinciran dan penyejukan dalam perincian mesin dalam pembuatan.

KEYWORDS: *boiling point; cutting, liquid; magnetic field; manufacture*

1. INTRODUCTION

There are many scientific works about the effect of magnetic field on the physical and chemical properties of water. When liquids are exposed to magnetic field effects, they become magnetized liquids. Nikolskiy et al. analytically and experimentally investigated the influence of magnetic and electric fields with tension gradient in the direction of the movement of the contacting gas-liquid phases [1]. Rashid at al. presented an investigation of water evaporation through a magnetic field of 0.5 T, which was located at different locations of tested water height (water-air interface, water mid-height, and bottom) [2]. Amiri and Dadkhah investigated whether or not a physical water treatment reduces the surface tension of water, as reported in some scientific literature [3]. Tritigin et al. studied the influence on the microflora of water-oiled lubricating cooling liquid of a weak electromagnetic pulse field [4]. Ageev noted that one of the hypotheses explaining the effect of a weak magnetic field on biological objects is that water properties were changed by magnetic field [5]. Hassan and Rahmon's study is a step towards gaining a better understanding of the effect of magnetism on water properties and on the biology of culture organisms, such as the brine shrimp, *Artemia salina*. Their study evaluated the effects of magnetic field exposure on water properties, which in turn affected the hatchability of A. salina [6]. Baresel et al. have taken the concept of water treatment by functionalized magnetic particles one step forward by integrating the technology into a complete proof of concept, which included the preparation of surface-modified beads, their use as highly selective adsorbents for heavy metals ions (zinc, nickel), and their performance in terms of magnetic separation [7]. Mardonov et al. studied the effect of magnetic field on the dynamic viscosity coefficient of flowing water. They found that the dynamic viscosity coefficient of water changed after magnetization. They magnetized flowing water using different magnetic field strengths. After magnetization, they measured the coefficient and found that dynamic viscosity of water was decreased [8].

The effect of magnetic field on liquids is still a controversial issue, and there is a lack of research in this field. Almost all of the research about magnetic field effect on liquids were conducted on water and the majority of that research was effectively used in the agriculture field. The authors of this paper suggest that use of magnetic field treatment in the manufacturing process would have a great advantage.

Although many scientific works about the influence of magnetic field on some parameters of water have been reported from many studies, no scientific works analysed the effect of magnetic field on lubricating cooling liquids, which are used in cutting processes. Kinematic viscosity, boiling point, and density are effectively the most important parameters of lubricating cooling liquids used in cutting processes. Analysing the influence of magnetic fields on the properties of lubricants would help to increase the efficiency of machining and increase the wear resistance of the cutting tool. The purpose of this research work is to study the effect of magnetic field on these three properties of lubricating cooling liquids such as boiling point (heating from room temperature until the liquids boil), kinematic viscosity, and density. The experiment suggests that some parameters of liquids, such as boiling point, kinematic viscosity, and density were changed by the magnetic field. It is essential to point out that in this research, three different liquids were studied, including water, and the results were applicable for other types of liquids. It has been proven, here, that magnetic field treatment decreases the boiling point of any liquid with respect to the magnetic field strength. In addition, the effect of magnetic field on kinematic viscosity and density of liquids were investigated in this study. Moreover, the influence of the magnetic field strength on magnetization effect was studied.

2. METHODS AND MATERIALS

2.1 Magnetization

The first liquid (liquid-1) was tap water and it was from a local water supply company ("Suvsoz", Uzbekistan), the second and the third liquids (liquid-2 and liquid-3) were taken from a local manufacturing factory (Meridian-A). These three liquids were used as lubricating cooling liquids in the "Meridian-A" manufacturing factory. The second liquid, liquid-2, was a compound of water and $K_2Cr_2O_7$ powder with 0.2% concentration. More accurately, m=50 grams $K_2Cr_2O_7$ powder was dissolved in m_0=25 kg of water. Lubricating cooling liquid with that concentration is used in cutting details made from mild steel. Liquid-3 is also a lubricating cooling fluid and it is mainly used for machining details where the material's hardness is higher (hard alloys) than other details. It is a 5% concentration of special cutting fluid marked BM-76M with water.

The magnetizing equipment UMD-1 was developed for magnetizing liquids, the equipment consisted of 8 magnets, and the size of each magnet was 120x80x16 (length, width, height) with a minimum strength of 40 Mt. The details of UMD-1 magnetizing equipment are given in Fig. 1.

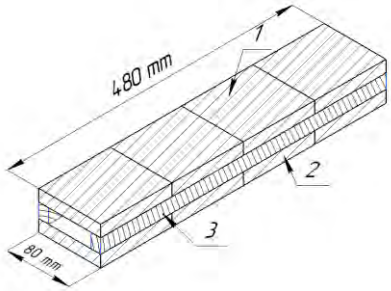


Fig. 1: UMD-1 magnetizing equipment.
1 – Magnets on the top, 2 – magnets below, 3 – wooden part for separation.

Four magnets (1) were placed on the other four (2) with a gap between them. They were separated by a non-conductor material (3) between them. Three types of magnetic field strength (40 mT, 60 mT, 80 mT) were selected to magnetize the liquids and magnetic field strength was controlled by changing the distance between magnets. More clearly, changing the non-magnetic parts with ones of different height, resulted in various magnetic field strengths. A PVC pipe was placed between the magnets, and its diameter was 10 mm. Four hundred eighty mm of the polyvinyl chloride (PVC) pipe passed through UMD-1 magnetizing equipment and this was the length of the magnetic field. When a liquid flowed through the PVC pipe, it turned into a magnetized liquid. All of the samples circulated at a flowing speed of 0.4 m/s for 30 min in UMD-1 magnetizing equipment. The scheme of the magnetizing process is given in Fig. 2.

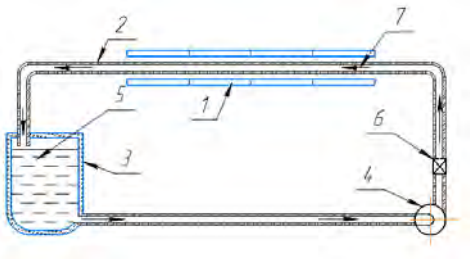


Fig. 2: 1 – UMD-1 magnetizing equipment, 2 – PVC pipe, 3 – container, 4 – pump, 5 – liquid, 6 – tap, 7 – flowing direction of the liquid.

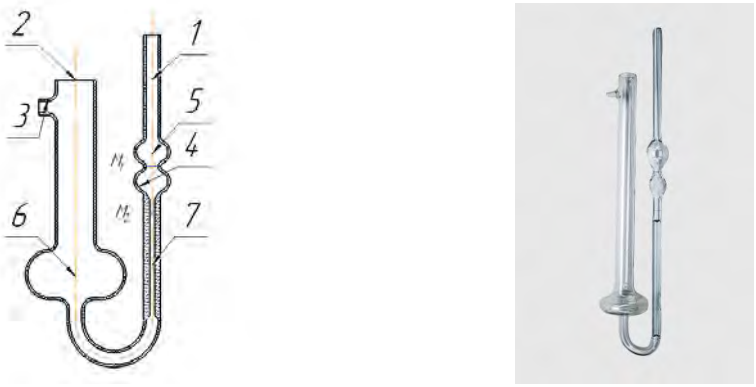
The magnetic field strength, magnetizing condition, and velocity of liquids were considered as three influential factors that affected the changes caused by the magnetic field [9,10]. Therefore, all of the experiments were conducted in the same laboratory conditions and the specification of the laboratory conditions is given in Table 1.

Table 1: Example system technical data

Item	Value
Atmospheric pressure	770 mmHg
Humidity	80%
Height above sea level	440 meter
Flowing speed of liquid	0.4 m/s
Diameter of the PVC	10 mm

2.2 Measurement of the Kinematic Viscosity

A capillary glass viscometer VPJ-4 (Fig. 3) is one of the commonly used devices to measure the viscosity coefficient of liquids and it was designed to determine the kinematic viscosity of liquids in accordance with GOST 33-66. Its range of measurement is from 0.6 mm^2/s to 10000 mm^2/s and it was very easy to measure and is also a popular method for measurement in laboratory conditions. Capillary viscometer VPJ-4 (Fig. 3) is a device in the form of a U-shaped tube, the elbow (1) is soldered with the capillary (5). Measuring viscosity by a viscometer is based on the determination of the time required for a certain volume of liquid to flow out through the capillary.



(a) (b)

Fig. 3: VPJ-4 capillary viscometer.
1, 2 – bend of the viscometer; 3 – discharge pipe; 4 – upper reservoir of the viscometer; 5 – capillary; 6 – lower reservoir of the viscometer; M_1 and M_2 – labels, limiting the measured volume of the lower reservoir viscometer.

To determine the flow time to the discharge pipe (3), a rubber tube connected to a rubber bulb was inserted. Then, holding the bend (2), the viscometer was turned over and the bend (1) was immersed in a reservoir with liquid. The liquid was sucked in (with a pear) to the M_1 mark, without letting air bubbles form in the liquid. At the moment when the liquid reached the M_2 mark, the device was removed from the vessel and immediately turned over to its original position. Excess fluid was removed from the bend (1), the rubber tube was removed from the bend (2), and it was put on the bend (1). Then, the viscometer was placed in the thermostat so that the expansion (4) was lower than the liquid level in the thermostat. It was kept in a thermostat for more than 30 minutes at a given temperature and then the

liquid was sucked into the bend to the level of one-third of the expansion (4). The rubber tube was disconnected and the time for the meniscus level to drop from the M_1 mark to the M_2 mark was measured with a stopwatch.

The diameter of the capillary was 3.55 mm. The kinematic viscosity of the liquid was determined using the following Eq. (1) [GOST 33-66],

$$V = \frac{g}{9.807} \cdot T \cdot K \quad (1)$$

where,

K – Constant of the viscometer (9.224 mm^2/s^2 according to GOST 33-66)

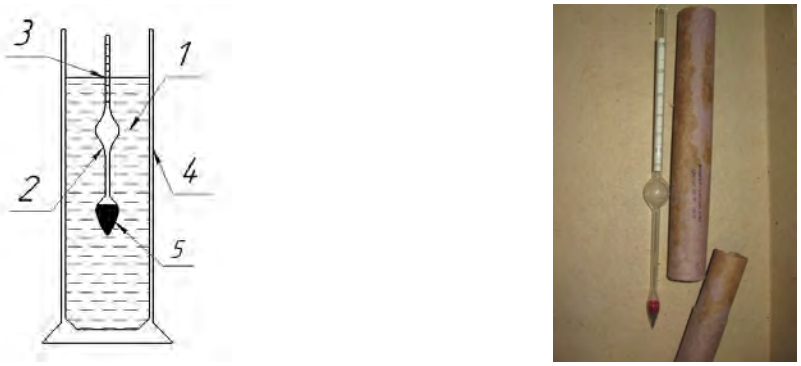
V – Kinematic viscosity of liquid [mm^2/s]

T – Liquid flow time in seconds

g – Acceleration of gravity [m/s^2]

2.3 Measurement of the Density

Density is one of the main physical quantities that characterizes the properties of a substance. When exercising control over technological processes and product quality, measuring the density of substances plays an essential role. It is known that to determine the density of a solid or liquid, it is sufficient to define the ratio of body weight to its volume. On the other hand, the ratio of the densities of a solid and a liquid determines, for example, the condition for the floating of a solid in a liquid, and the values of their density are the magnitude of the pushing forces from the liquid and the weight of the body inside the liquid. This relationship opens the ability to measure density through the interaction of a liquid and a solid immersed in that liquid. In our experiment, we used the densimeter with second class of accuracy according to GOST 1300-57 and its measurement range was from 1.000 to 1.800 (Fig. 4).



(a) (b)

Fig. 4: 1 – Liquid; 2 – Densimeter; 3 – Density measuring scale of the densimeter; 4 – tanker, 5 – Ballast.

Densimeters and hydrometers are commonly used to measure the density of different types liquids. The densimeter is a glass float of constant mass and volume, the expanded (lower) part of which is filled with ballast – clean and dry metal shot, filled with a layer of resin with a melting point of at least 80 °C. Thanks to the ballast (5), the densimeter is in a vertical position during measurements. In the upper thinned part of the float, there is a scale graduated in units of density. The use of a densimeter to determine the relative density of a liquid is based on Archimedes' law, therefore, the upper divisions of the scale correspond to

the lowest density, and the lower divisions correspond to the highest density. The readings are counted along the lower meniscus of the fluid. In the manufacturing and canning industry, sets of standard general-purpose densimeters of at least 2nd class of accuracy are used, calibrated to measure the relative density at a temperature of 20 °C with a measurement range from 0.7000 to 1.840 or from 820 to 1840, respectively.

2.4 Measurement of the Boiling Point

The TP300 (Fig. 5) digital thermometer (Xi′an Lonn M&E Equipment Co., Ltd., China) was used to measure the boiling point of magnetized and non-magnetized liquids under different conditions. The boiling point is determined by the nature of the substance and depends on external pressure. It is a characteristic of bodies located only in solid-state and for high molecular weight compounds. To determine the boiling point of fluids, we took a container and filled it with liquid (100 ml), and started to heat. While heating we put a TP300 digital thermometer into the liquid and marked the boiling point of the liquid.

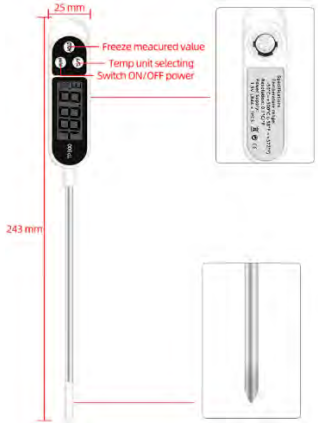


Fig. 5: TP300 digital thermometer (measurement range is from -50 ºC to +300 ºC).

3. RESULTS AND DISCUSSION

3.1 Effect of Magnetic Field on the Boiling Point of Lubricating Cooling Liquids

The digital thermometer recorded the boiling point of the three different liquids. As stated above, three different liquids were analysed under the influence of magnetic field with three different magnetic field strengths (40 mT, 60 mT, and 80 mT). To compare the results, at first, the boiling point of liquids without magnetic field effect was measured under laboratory conditions. Second, the magnetic field was affected, with a MFS of 80 mT, to liquids in a stationary condition (not flowing) and the results were compared. The boiling point results were consistent with reports in the literature that the boiling point of water increases after magnetic field treatment [11]. Figure 6 shows the difference between the obtained results.

It is noticeable from Fig. 6 that liquids under the effect of magnetic field in peace (still) condition, have lower boiling point than their natural boiling point. Every liquid's natural boiling points differed from each other, but when they were affected by the magnetic field, there was difference between boiling points of magnetized liquids in peace condition and non-magnetized liquids. Then the effect of the magnetic field with different MFS on flowing liquids was analysed. After 30 minutes of magnetic treatment under each magnetic field strength in the same laboratory condition, as specified in Table 1, the following results, given in Fig. 7, were obtained.

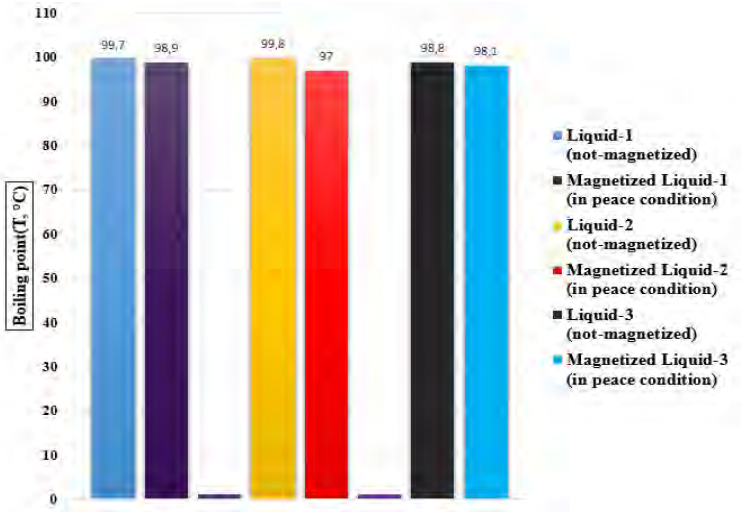


Fig. 6: The boiling point of not-magnetized and magnetized (in peace condition) liquids.

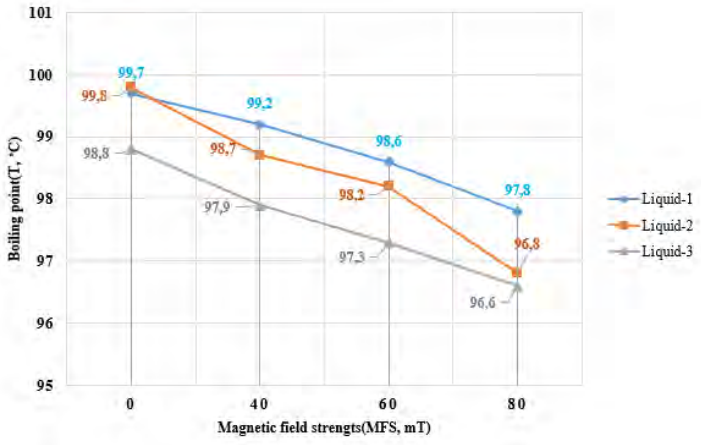


Fig. 7: The boiling point of not-magnetized and magnetized (in the flowing condition) liquids.

The boiling point is one of the main physical parameters of fluids and its value might be changed under the effect of the magnetic field. The boiling point of liquids depends on the condition and other parameters of liquids such as atmosphere and properties of fluids. As can be seen in Fig. 7, the boiling point of magnetized liquids decreases as magnetic field strength increases. The difference between boiling point of not magnetized and magnetized fluids (in the flowing condition) under the 80 mT of magnetic field strength were 1.9 ºC for liquid-1, 3 ºC for liquid-2, and 2.2 ºC for liquid-3 respectively. In addition, it is interesting to note that the higher the magnetic field strength the lower the boiling point of the liquids. The lowest boiling point was determined in 80 mT magnetic field strength. It was found that, if magnetic field strength was increased the effect of magnetic field on liquids was also increased.

3.2 Effect of Magnetic Field on the Kinematic Viscosity Coefficient of Lubricating Cooling Liquids

Viscosity coefficient is one of the important properties of lubricating cooling liquids used in the machining process in manufacturing. Analysing the effect of magnetic field on that parameter of liquids will help to increase the efficiency of the manufacturing process.

Under laboratory conditions, the VPJ-4 measuring device was used to record the results. This device commonly used in Uzbekistan.

To analyze the effect of the magnetic field on the viscosity coefficient of liquids, the three different lubricating cooling liquids in stationary condition were magnetized for 30 minutes. Eight permanent magnets with 80 mT magnetic field strength were used to make the magnetizing condition. Then, the viscosity coefficient of unmagnetized and magnetized (in peace condition) liquids (Fig. 8) were compared. It is clear from the data given in Fig. 8, that under the influence of the magnetic field, the kinematic viscosity coefficient of each liquid decreased noticeably. The difference between kinematic viscosity coefficient of unmagnetized and magnetized liquids were 0.323 mm^2/s for liquid-1, 0.057 mm^2/s for liquid-2, and 0.156 mm^2/s for liquid-3.

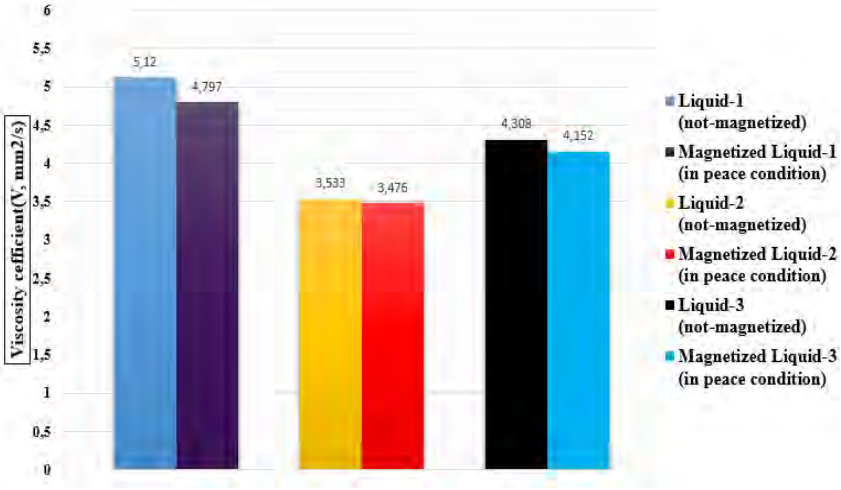


Fig. 8: The kinematic viscosity coefficient of not magnetized and magnetized (in peace condition) liquids.

Following this experiment, the influence of the magnetic field on flowing liquids was assessed. A UMD-1 magnetizing device was used to magnetize the flowing liquids. Lubricating cooling liquids passed through the UMD-1 magnetizing device at 0.4 m/s. The diameter of the pipe, which was used to pass flowing liquid through the UMD-1, was 10 mm. After 30 minutes of the magnetizing process, the kinematic viscosity of liquids was measured and the results are given in Fig. 9.

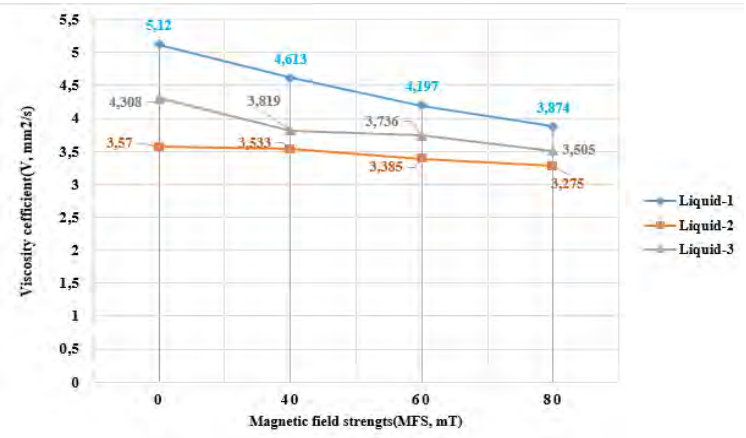


Fig. 9: Viscosity coefficient of not magnetized and magnetized (in flowing condition) liquids in different magnetic field strengths.

Results given in Fig. 9 show that a magnetic field can cause a decrease in the viscosity coefficient of flowing liquids. Moreover, the viscosity coefficient of flowing liquids depends on the magnetic field strength, because as the strength of the magnetic field increases, the kinematic viscosity of liquids decreases. The lowest viscosity coefficient was determined under the influence of the highest magnetic field strength (80 mT). Comparisons show that the value between the unmagnetized and magnetized liquids, under the highest MFS, were 1.246 mm^2/s for liquid-1, 0.803 mm^2/s for liquid-2, and 0.295 mm^2/s for liquid-3.

It was also found that the effect of magnetic field on flowing and not flowing liquids depends on the properties of the liquids. It is clear from the numbers obtained in Figs. 8 and 9 that liquid-1 was affected by the magnetic field more than the other two liquids. Liquid-3 had the least change among the three liquids after magnetization. Because the concentration of liquid-3 has more components than other liquids and its crystals are differentiated from tap water (liquid-1) more than liquid-2. It is investigated that the greater the percent concentration of liquids, the less the magnetization effect on the liquids.

3.3 Effect of Magnetic Field on the Density of Lubricating Cooling Liquids

The magnetic field effect on the density of lubricoolants is also important in machining details in the manufacturing process. For the purpose of increasing efficiency in machining details, the effect of magnetic fields on the density of liquids was given greater attention by authors. As above, at first, the difference between the density of unmagnetized and magnetized (in peace condition) liquids was compared. There were also differences among the obtained results (Fig. 10). Figure 10 shows that after magnetizing liquids, their density increased and the differences were 0.002 g/cm^3 for liquid-1, 0.001 g/cm^3 for liquid-2 and 0.002 g/cm^3 for liquid-3.

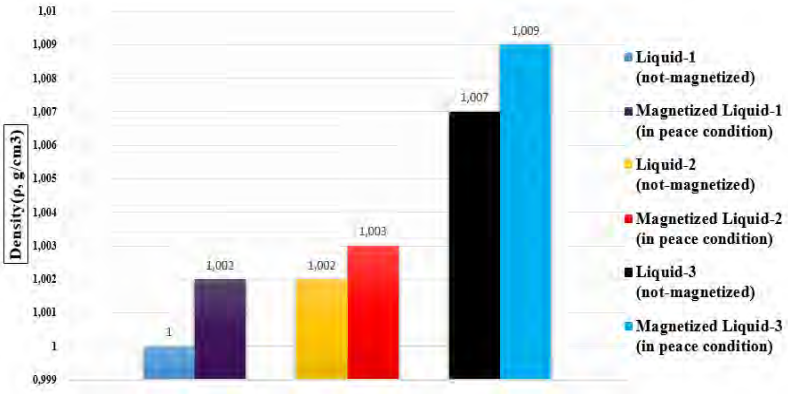


Fig. 10: The density of unmagnetized and magnetized (in peace condition) liquids.

When the flowing liquids were magnetized in different magnetic field strengths, the results changed noticeably with respect to the magnetic field strength. The results, which were taken in the same laboratory conditions, are given in Fig. 11.

Despite the fact that the density of the liquids differed from each other in the same conditions, the density of each liquid was increased after 30 minutes of the magnetization process while they were flowing (Fig. 11). The density of the first sample was 1.000 g/cm^3 before magnetization, but it reached to 1.005 g/cm^3 after magnetizing at 40 mT. When the MFS increased to 80 mT, the density reached its highest point in the experiment (1.007). Liquid-2 and liquid-3 also increased from 1.002 g/cm^3 to 1.006 g/cm^3 and from 1.009 g/cm^3

to 1.011 g/cm^3 respectively. In fact, the highest density was obtained at 80 mT of magnetic field strength.

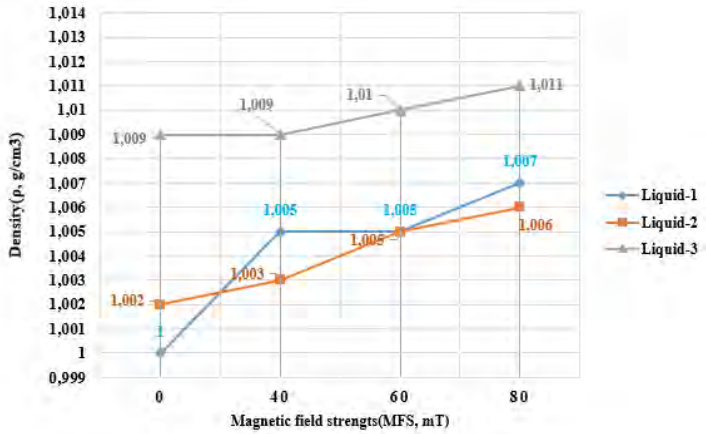


Fig. 11: The density of unmagnetized and magnetized (in flowing condition) liquids in different magnetic field strengths.

A decrease in the viscosity coefficient after magnetic field treatment was found. When magnetic field influences liquids, their crystals are regulated in the magnetic field direction by the magnetic field. This regulation of crystals causes a decrease in the viscosity coefficient after magnetic field treatment. Moreover, the regulation of crystals causes an increase in the density after magnetic field treatment. Also, it has been explored that, the effect of magnetic field on liquids depends on the properties of fluids, or more accurately, in the same laboratory condition and same magnetic field strength, magnetic field influence was different on various liquids. It is worth noting that the influence of a magnetic field on liquid depends on what kind of chemical elements are dissolved in the liquid. Moreover, the maximal change in all of the experiments was obtained at 80 mT of magnetic field strength while the lowest change was reached at 40 mT of magnetic field strength.

Obtained results are very influential to manufacturing processes because the changes in these parameters of liquids have an impact on the wear resistance of the cutting tool used in the machining process. The influence of magnetized lubricating cooling liquids on the wear resistance of cutting tools helps to increase the efficiency in the machining process and saves energy [8,15,16].

However, this issue about magnetic field treatment of liquids remains controversial. Complete understanding of the influence of magnetic fields on fluids has a great impact on agriculture, industry, and other fields [12-14]. Some scientists are working on the issue of magnetic field treatment of water, but there lack of researches on this issue, especially the magnetic field effect on different liquids.

4. CONCLUSION

The results obtained in the laboratory experiment showed that there was an increase in the density of liquids when they were affected by the magnetic field. In spite of the fact that the increase was not always noticeable on the density of fluids, it depended on the magnetic field strength. Also, the density of water was more affected by the magnetic field than the other two dissolved liquids which means that the density of dissolved liquids is less likely to be influenced by the magnetic field. However, the influence of the magnetic field on the boiling point of dissolved liquids were about 1.5 times higher than tap water. In addition,

it was found that kinematic viscosity and boiling point of experimented liquids were decreased after magnetic field treatment. However, after magnetic field treatment of flowing liquid, the kinematic viscosity of liquid-1 decreased by 24% while the kinematic viscosity of liquid-2 and liquid-3 decreased by 8% and 19% respectively. The difference between the effect of magnetic field on stationary and flowing liquids was also examined and it was explored that flowing liquids were influenced more than liquids in stationary conditions.

REFERENCES

[1] Nikolskiy VY. (2017) Primenenie elektromagnitnyx poley dlya intensifikatsii teplomassoobmena v sovmeщennyx gazojidkostnyx protsessax [Application of electromagnetic fields to intensify heat and mass transfer in combined gas-liquid processes]. Vostochno-Yevropeyskiy Jurnal Peredovyx Texnologiy, 3(8): 1729-3774. https://media.neliti.com/media/publications/307213-application-of-electromagnetic-fields-fo-6a3064eb.pdf

[2] Rashid FL, Hassan NM. (2013) Increasing water evaporation rate by magnetic field. International Science and Investigation Journal, 2: 61-68. http://isijournal.info/journals/index.php/ISIJ/article/view/12

[3] Amiri MC, Dadkhah AA. (2006) On reduction in the surface tension of water due to magnetic treatment. Colloids and Surfaces A: Physicochemical and Engineering Aspects: 278(1-3): 252-55. https://doi.org/10.1016/j.colsurfa.2005.12.046

[4] Tyrtygin VN, Sobgayda NA, Potexa VL, Shayxiev IG, Makarova YA. (2013) Vliyanie elektromagnitnogo impulsnogo polya na mikrofloru smazochno-oxlajdayuщey jidkosti [Influence of the electromagnetic pulse field on the microflora of the cutting fluid]. Vestnik Kazanskogo texnologicheskogo universiteta, 16(3):176-178. https://cyberleninka.ru/article/n/vliyanie-elektromagnitnogo-impulsnogo-polya-na-mikrofloru-smazochno-ohlazhdayuschey-zhidkosti/viewer

[5] Ageev IM, Shishkin GG, Yeskin S. (2007) Izuchenie Vliyaniya Nizkochastotnogo Magnitnogo Polya Na Vodu [Study of the Effect of Low Frequency Magnetic Field on Water]. Raditexnika i Electronica, 28: 8-16. https://www.elibrary.ru/item.asp?id=10412034

[6] Hassan SM, Ridzwan AR. (2016) Effects of exposure to magnetic field on water properties and hatchability of *Artemia salina*. ARPN Journal of Earth Sciences, 11(11): 416-23.

[7] Christian B. (2019) Functionalized magnetic particles for water treatment. Heliyon, 5(8): 1-7. https://doi.org/10.1016/j.heliyon.2019.e02325

[8] Umarov EO, Mardonov UT, Shoazimova UK. (2020) Influence of the magnetic field on the viscosity coefficient of lubricoolant that is used in the cutting proces. International Journal of Mechatronics and Applied Mechanics, 8(2): 144-149. https://www.doi.org/10.17683/ijomam/issue8.50

[9] Liu B, Gao B, Xu X. (2011) The combined use of magnetic field and iron-based complex in advanced treatment of pulp and paper wastewater. Chem Eng J, 178(1): 232-238. https://doi.org/10.1016/j.cej.2 011.10.058

[10] Manyà JJ, Jr MJA, Kinoshita CK, Stephen MM. (2011). Specific heat capacity of pure water at 4.0 MPa between 298.15 and 465.65 K. Ind Eng Chem Res, 50(10): 6470-6484. https://doi.org/10.1021/ie102462g

[11] Youkai Wang, Huinan Wei, Zhuangwen Li. (2018) Effect of magnetic field on the physical properties of water. Results in Physics, 8: 262-267. https://doi.org/10.1016/j.rinp.2017.12.022

[12] Wang XY, Zhang B, Gong Z, Gao K, Ou Y, Zhang J. (2013) The effect of a static magnetic field on the hydrogen bonding in water using frictional experiments. J Mol Struct, 1052(11): 102-104. https://doi.org/10.1016/j.molstruc.2013.08.021

[13] Cai R, Yang H, He J, Zhu W. (2009) The effects of magnetic fields on water molecular hydrogen bonds. J Mol Struct, 938(1-3): 9-15. https://doi.org/10.1016/j.molstruc.2009.08.037

[14] Toledo EJL, Ramalho TC, Magriotis ZM. (2008) Influence of magnetic field on physical–chemical properties of the liquid water: insights from experimental and theoretical models. J Mol Struct, 888(1-3): 409-415. https://doi.org/10.1016/j.molstruc.2008.01.010

[15] Umarov EO, Mardonov UT. (2020) Analysing the effect of magnetic field on Lubricoolant in machining process. International Journal of Applied Research, 6(5): 347-352. https://www.allresearchjournal.com/archives/?year=2020&vol=6&issue=5&part=F&ArticleId=6723

[16] Turakhodjaev N, Tashbulatov Sh, Zokirov R, Tursunbaev S, Baydullaev A. (2020) Studying the scientific and technological bases for the processing of dumping copper and aluminum slags. Journal of Critical Reviews, 7(11): 441-444. http://dx.doi.org/10.31838/jcr.07.11.79

https://doi.org/10.31436/iiumej.v22i2.1550

INVESTIGATION OF MULTI-ZONE MODELS FOR SPARK IGNITION ENGINE FUELED WITH ETHANOL

Olatunde Bilikis Olanrewaju[1*], Dare Adebukola Ademola[2], Ismail Olawale Saheed[2], Shote Adeola Suhud[1], Alamu Oguntola Jelili[3] and Sulaiman Adedoyin Musediq[1]

[1]Department of Mechanical Engineering, Olabisi Onabanjo University Ago Iwoye, Nigeria
[2]Department of Mechanical Engineering, University of Ibadan, Nigeria
[3]Department of Mechanical Engineering, Osun State University, Nigeria

[*]Corresponding author: olatunde.bilikis@oouagoiwoye.edu.ng

(Received: 12th June 2020; Accepted: 19th January 2021; Published on-line: 4th July 2021)

ABSTRACT: This research is aimed at investigating the effect of using ethanol (E100) in multi-zone model analysis consisting of multi-combustion chamber zoning cases. The first case considered is a three-zone model that has an unburned zone, burned zone, and transitory zone. The second case model is also three-zone, consisting of an unburned zone and two partitioned burned zones. The burned zone was imagined partitioned into burned zone-1 and burned zone-2 under uneven fuel distribution having different equivalent ratios. The third case is a four-zone model including two regions of burned zone, an unburned zone and a transitory zone, which is unburned burned zone containing a mixture of unburned and burned gases. Arbitrary constants for each of the unburned (CC1) and burned (CC2) Zone leakages in the unburned burned Zone are 0.00025, 0.0005, 0.001, 0.002, 0.005, 0.1 and 0.5. The Mass Fraction Burned (MFB) for zone-1, x1 and burned zone-2, x2 are computed using Partitioned Burnt Zones Ratios (PBZR) of 2:8, 3:7, 4:6, 5:5, 6:4, 7:3 and 8:2. Two equivalent ratios, one for each fuel MFB (ϕ1, ϕ2), (0.8, 0.6) and (0.6, 0.8) are analyzed using fuel blends of varying percentage. A comparison of values of the three zoning cases is done using peak values from the three-zone models to evaluate the four-zone model. The model was compared with a spark ignition engine (SIE) operating with a premium motor spirit (PMS) serving as baseline. The engine operating conditions were set at an engine speed of 2000 rpm, -35bTDC ignition time, and burn duration at 60 ^{o}C. The indicated mean effective pressure (IMEP), thermal efficiency (η), cylinder pressure and emission fraction from the developed models and those of two-zone analysis obtained agreed with literature values. The result showed it is undesirable to have a high volume of burned charge as infiltrate. The three-zone segmented model predicted the highest engine thermal efficiency and peak pressure at mass burn ratio of 7:3. A general reduction in N_2 emission was observed for the three-zone transitional and four-zone models.

ABSTRAK: Kajian ini menilai kesan etanol (E100) dalam analisis model zon-berbilang yang terdapat pada masalah pengezonan kebuk pembakaran-berbilang. Kes pertama yang diambil kira adalah model tiga-zon yang mempunyai zon tidak terbakar, zon terbakar dan zon peralihan. Model kedua merupakan juga tiga-zon yang terdiri daripada zon tidak-terbakar dan dua zon bahagian yang terbakar. Zon yang terbakar dibahagikan kepada zon-1 terbakar dan zon-2 terbakar di bawah kebakaran tidak sekata yang mempunyai nisbah berlainan. Kes ketiga adalah model zon-keempat termasuk dua kawasan zon terbakar, zon tidak-terbakar dan zon peralihan iaitu zon terbakar tidak-terbakar di mana ia adalah campuran gas terbakar dan tidak-terbakar. Tetapan sebarangan bagi setiap zon kebocoran tidak-terbakar (CC1) dan terbakar (CC2) dalam zon terbakar tidak-terbakar adalah 0.00025, 0.0005, 0.001, 0.002, 0.005, 0.1 dan 0.5.

Pecahan Jisim Terbakar (MFB) bagi zon-1, x1 dan zon-2 terbakar, x2 dikira menggunakan Nisbah Zon Bahagian Terbakar (PBZR) sebanyak 2:8, 3:7, 4:6, 5:5, 6:4, 7:3 dan 8:2. Nisbah dua persamaan, setiap satu bahan api MFB adalah (ϕ1, ϕ2), (0.8, 0.6) dan (0.6, 0.8) dan diuji menggunakan pelbagai peratus bahan api campuran. Nilai perbandingan bagi tiga kes zon dibuat menggunakan nilai puncak dari model tiga-zon bagi menilai model empat-zon. Model ini dibandingkan dengan enjin cucuhan bunga api (SIE) beroperasi dengan motor alkohol premium (PMS) sebagai garis asas. Keadaan operasi enjin adalah dihadkan pada 2000 rpm kelajuan enjin, masa pencucuhan -35bTDC dan tempoh pembakaran pada 60 °C. Tekanan berkesan min tertunjuk (IMEP), kecekapan haba tertunjuk (η), tekanan silinder dan pecahan pengeluaran dari model yang dibangunkan dan analisis dua-zon yang terhasil adalah sama dengan nilai literatur. Dapatan kajian menunjukkan cas terbakar pada isipadu yang banyak adalah tidak diingini sebagai penyerap. Model tiga bahagian zon menunjukkan kecekapan haba enjin tertinggi dan tekanan puncak pada jisim bakar dengan nisbah 7:3. Manakala, pengurangan umum telah diperhatikan pada pengeluaran N_2 di peralihan tiga-zon dan model empat zon.

KEYWORDS: *ethanol; multi-zone; indicated mean effective pressure; emission fraction*

1. INTRODUCTION

Modification and optimization of combustion chamber geometry has been seen as the end game in achieving stringent emission reduction and better engine performance with several researches conducted and some ongoing. Researchers have reported the importance of engine combustion chamber modification and its ability to resolve complications arising from adopting some other methods of engine optimization [1-2]. Efforts at combustion chamber modification by [3] was aimed at correcting the lean burn effect, the design modification was carried out for two compression ratios 10 and 11 which necessitated variation in piston bowl size from 16mm to 13.6mm. Increased compression ratio from 10 to 11 was reported to have resulted in leaner combustion but adversely higher combustion duration and coefficient of variance necessitating development of a new combustion chamber design. The modified combustion bowl by [3] has six leaves placed eccentrically to create the desired swirl breakup with a cutout for the inlet and exhaust valve seats. The difference in diameter between the inner and outer bowl was reduced to improve flame propagation and strengthen squish effect.

More researches on modification of the combustion chamber for the purpose of promoting faster burning of charge in a lean burn engine by using the swirl and squish principles was carried out by [2] and [4]. A chamber modification similar to the one described by [3] is the Nebula chamber, employed to retard the swirl motion close to TDC. The Nebula combustion chamber used by [5] is such that the bowl shape forces the swirl to develop into colliding air flows thereby creating the desired great turbulence and fast burn effect. [6] went a step further by implementing the Squish, Nebula, and another named Tokyo Gas, TG with similar purpose as the Nebula but different in that the TG was described as having “a cylindrical dent with semi-circular slants oppositely shaped”. The resulting effect of the TG is large turbulence at the cylinder wall and weak turbulence in the center of the cylinder to stabilize flame ignition. The Nebula, TG and squish were reported to have achieved NO_x reduction, increased thermal efficiency and stable combustion. [7-8], adopted ten different combustion chamber geometries in studying the effect of chamber geometry on flow, combustion, and emissions. Of all the geometries studied by [7-8], the cylindrical and square bowl gave the expected turbulence. However, the cylindrical bowl has desired outcome of its peak turbulence close to TDC. The

experiment further confirmed that combustion chamber shape has an effect on flow field, heat release rates, and emission rate. Use of simulation model has been proven to be a cost effective and efficient way of optimizing engine performance and emissions. In [9], three different approaches for combustion modeling in SI engines were identified, while the multi-zone approach to engine modeling seemed the cheapest, fastest and yet efficient modeling method [10]. In their study, [10] concluded the use of additional zone lead to increased heat release prediction. Furthermore, [11] observed increased engine performance using a three-zone model for predicting the effect of ethanol with diesel fuel. Several researches have been carried out on the use of a modeling approach for optimization of combustion processes. Additionally, there are still continuous environmental challenges that need to be addressed. This desire has necessitated the need for further study.

The attention received by ethanol as an alternative fuel is said to be due to its role in reducing problems such as climate change, depleting fossil oil and high oil prices in the international market. Alcohols like ethanol can be produced by fermentation of different biological feedstocks as reported by [12-13]. The most positive properties of ethanol include its ability to be produced from renewable energy sources, its high-octane number, and its high laminar flame speed [14]. The negative aspects include its low heating value due to higher oxygen content compared to gasoline, and it causes corrosion in the metal and rubber parts of an engine [15]. The engine power improves with ethanol as it has better anti-knock characteristics qualities [16]. The engine power also improves with an increase in compression ratio [17]. Ethanol has a high latent heat of vaporization, the latent heat cools the intake air and hence, increases the density and volumetric efficiency [18]. The lower heating value of ethanol enables it to have a higher compression ratio as compared to gasoline [19]. It is desirable for SI engine to have high octane and low cetane numbers. This implies higher resistance to self-ignition and lower auto-ignition possibility. Ethanol can be used as a fuel in spark ignition engines either in pure form (E100) or blended with premium motor spirit (PMS). The use of pure ethanol implies some problems during cold start due to lower vaporization compared to gasoline, which in some cases should require an electrical preheating of the engine block. The use of ethanol in conventional spark ignition engines comes with the problem of corrosion [20]. One of the problems is tackled by the invention of flexible fuel vehicles (FFVs) which implies that the fuel system is made using stainless steel, with some rubber parts replaced by nylon, fluorinated plastics, and high-density polyethylene [21,22]. The lower heating value (LHV) of ethanol requires an alteration of the injection control to increase the fuel flow [23]. Therefore, only a limited percentage of ethanol, up to 10-25%, is recommended for use in conventional gasoline engines [24]. However, FFVs designed to operate with any gasoline-ethanol mixtures are increasing. The advantage of FFV systems consists of the possibility offered by the control system of the engine to detect the concentration of ethanol in the tank and automatically optimize both injection and ignition [14,25].

According to [21], car manufacturers are capable of making vehicles compatible with hydrated ethanol, E100, at no additional cost. However, some investigation has pointed to the presence of ethanol in higher concentration as the cause of volatile organic matter (VOC) emission in the form of benzene [26]. The point line remains that there is more to be discovered for proper evaluation and optimization of ethanol and ethanol-blended fuel usage. Numerous researches [18,27-29], have been carried out concerning the use of ethanol blends in SI engines, but very few uses of 85-100% ethanol have been documented [21]. [30] investigated the effect of varying ethanol-gasoline blends up to 85% and compared the combustion of gasoline and gasoline-ethanol blends through pressure

analysis. An increase in engine efficiency was observed, coupled with CO_2 emission reduction when ethanol blends were used. This study is aimed at investigating the effect of 100% ethanol, E100 on engine performance using a thermodynamic multi-zone model approach.

2. GENERAL METHODOLOGY

To investigate the combustion process, three multi-zone models were developed, each with peculiar characteristics to study effects of additional zone(s) on spark ignition engine combustion characteristics using ethanol as fuel. The model geometry developed by the authors and adopted in this work was the same used by [31,32], where a zero-dimensional, multi-zone model was investigated. The developed models are validated making use of the model developed by [33], primarily for the two-zone model SI engine using gasoline.

In the analysis for the first multi-zone model, which is a three-zone model consisting of the primary burned and unburned zones in addition to a mixed zone described as the transitory zone, the mass infiltration from either of the primary zones, directly into the transitory zone is given in terms of constants of infiltration as shown in Eq. 1.

$$m_{ub} = CC_1 * m_u \text{ and } m_{bu} = CC_2 * m_b \quad (1)$$

where CC_1 and CC_2 are the arbitrary constants of infiltration from burned and unburned zones respectively. The second multi-zone case has its burned zone partitioned in two, the mass of the burned zone is imagined partitioned into burned zone-1 and burned zone-2 as the proportion of the mixture is uneven in the chamber and therefore, has different mass burning rates (x1, x2). The coefficients of heat transfer are predetermined to be (h_{400}, h_{450}, h_{500}), the mass fraction burned, is represented analytically as shown in Eq. 2 [34].

$$x = 0.5 * \left(1 - \cos\left(pi * \frac{(\theta - \theta_s)}{\theta_b}\right)\right), \quad (2)$$

where, x is mass fraction burned, θ, θ_s, θ_b are the crank angle, the start of energy release and duration of energy release respectively. Peak results of the two analyses of mass fraction and mass infiltration burned are employed for analysis in the third multi-zone case as a four-zone model.

The indices of performance, thermal efficiency (η), indicated mean effective pressure (IMEP) and emission fractions obtained using the developed models are compared with the standard two-zone model. The engine geometry for the model adopted for validating the developed model codes is shown in Table 1 and all are compared with the experimental data; CFD results obtained by [35] for the first multi-zone case and a four-stroke SIE (EF7) running on CNG [36] for the second multi-zone case.

2.1 Development of Mathematical Models

The first law of thermodynamics is applied to all the zones to develop the model equations. In an open system, the rate of change of mass is equal to the net flux of mass across the boundaries of the system. This is expressed by Eq. 3.

$$m = \sum_k \dot{m_k} \quad (3)$$

The mass flux of the control volume can be expressed as the addition of all unit zones as shown in Eqs. 4-6.

$$m = m_u + m_b + m_{un} \text{ (for first case)} \quad (4)$$

$$m = m_u + m_{b1} + m_{b2} \text{ (for the second case)} \quad (5)$$

https://doi.org/10.31436/iiumej.v22i2.1550

$$m = m_u + m_{b1} + m_{b2} + m_{un} \text{ (for the third case)} \quad (6)$$

From the first law of thermodynamics, the energy equation of an open system is expressed as Eq. 7.

$$\dot{E} = \sum \dot{m}h - \dot{W} + \dot{Q} \quad (7)$$

Equation 7 is therefore applied to expansion, compression, and combustion phases of the internal combustion (ICE) cycle. The model equations analysed in a computation environment are as shown in Eqs. 8 - 10 for the three modeled cases considered. Detailed mathematical derivation of these equations can be found in [31,32]. Sub-models including, air and fuel data, engine geometry inherent in the two-zone model [33] are also employed to analyze the present models.

Table 1: Engine specifications for validating modeled cases

Parameter	First case model evaluation [35]	2-zone model [33]	Second case model evaluation [36]
Number of cylinder	4	4	4
Bore (m)	0.079	0.1	0.085
Stroke (m)	0.086	0.08	0.078
Compression ratio	16	10	11
Displacement (m)	1.7	-	1.018
Inlet valve close	136BTDC	-	-
Exhaust valve open	122ATDC	-	-
Equivalence ratio	-	0.8	-
Burn duration angle	-	60	-
Start of combustion	-	-35	21
End of combustion	-	-	30
Engine speed (rpm)	1500	1500	2000

2.1.1 Model Equation for the First Case Studied

$$\frac{dP}{d\theta} = \{\frac{V_u}{T_u}\frac{dT_u}{d\theta} + \frac{V_b}{T_b}\frac{dT_b}{d\theta} + \frac{V_{un}}{T_{un}}\frac{dT_{un}}{d\theta} + (v_b - v_u)\frac{dm_{x1}}{d\theta} + (v_{un} - v_u)\frac{dm_{ub}}{d\theta} + (v_{un} - v_b)\frac{dm_{bu}}{d\theta} - v_u\frac{dm_{lu}}{d\theta} - v_b\frac{dm_{lb}}{d\theta} - v_{un}\frac{dm_{lun}}{d\theta} - \frac{dV}{d\theta}\}/\frac{V}{P} \quad (8)$$

2.1.2 Model Equation for Second Case Studied

$$\frac{dP}{d\theta} = \{\frac{V_u}{T_u}\frac{dT_u}{d\theta} + \frac{V_b}{T_b}\frac{dT_b}{d\theta} + \frac{V_{b2}}{T_{b2}}\frac{dT_{b2}}{d\theta} + (v_b - v_u)\frac{dm_{x1}}{d\theta} + (v_{b2} - v_u)\frac{dm_{x2}}{d\theta} - v_u\frac{dm_{lu}}{d\theta} - v_b\frac{dm_{lb}}{d\theta} - v_{b2}\frac{dm_{lb2}}{d\theta} - \frac{dV}{d\theta}\}/\frac{V}{P} \quad (9)$$

2.1.3 Model Equation for Third Case Studied

$$\frac{dP}{d\theta} = \{\frac{V_u}{T_u}\frac{dT_u}{d\theta} + \frac{V_b}{T_b}\frac{dT_b}{d\theta} + \frac{V_{b2}}{T_{b2}}\frac{dT_{b2}}{d\theta} + \frac{V_{un}}{T_{un}}\frac{dT_{un}}{d\theta} + (v_b - v_u)\frac{dm_{x1}}{d\theta} + (v_{b2} - v_u)\frac{dm_{x2}}{d\theta} + (v_{un} - v_u)\frac{dm_{ub}}{d\theta} + (v_{un} - v_b)\frac{dm_{bu}}{d\theta} + (v_{un} - v_{b2})\frac{dm_{b2}}{d\theta} - v_u\frac{dm_{lu}}{d\theta} - v_b\frac{dm_{lb}}{d\theta} - v_{b2}\frac{dm_{lb2}}{d\theta} - v_{un}\frac{dm_{lun}}{d\theta} - \frac{dV}{d\theta}\}/\frac{V}{P} \quad (10)$$

3. MODEL IMPLEMENTATION

3.1 Transitional Model

Relevant sub-programs such as the air and fuel were implemented for the two-zone model. Integration of the governing equations developed using ODE45 in MATLAB code is used to obtain the temperature and pressure distribution for various phases in SI engine operation using ethanol for the computation.

Implementation of the transitional three-zone model is done using selected values of infiltration in the developed code following the steps below;

- Implementing input parameters (available in engine.m)
- Cylinder volume and mass are evaluated
- Determination of relevant thermophysical properties such as u, h, cp, etc.
- Mass fraction burned is evaluated
- Heat transfer coefficient is selected using either constant estimated values or Woschni Equation
- Inputing selected constants for CC_1 and CC_2 using additional program
- Computing the required governing equations.

The model employs the parameters of an ICE with combustion and operating parameters details as stated in Table 1.

3.2 Segmented Model

The mass burned zone is imagined partitioned into burned zone-1 and burned zone-2 since there is a proportion that is unevenly mixed in the chamber and therefore they have different mass burning rates (x1, x2) and heat transfer coefficients, (h_{400}, h_{450}, h_{500}). Mass fraction burned, denoted by x, while, θ, θ_s, θ_b are the crank angle, start of energy release and duration of energy release respectively. These parameters are represented analytically and the equation solved is also presented in Eq. 2.

3.3 Transitional-Segmented Model

In this model, special consideration is given to mass infiltration into the transitory zone and mass fraction ratio such that predetermined optimal results are analyzed for the third case (four-zone model).

4. RESULTS AND DISCUSSION

4.1 Transitional Model (First Case Model)

Figure 1 shows the thermal efficiency (Ƞ), temperature, pressure, and indicated mean effective pressure (MPa) at varying CC_1 while the value of CC_2 (unburnt) is kept constant at 0.001 and at varying CC_2 while the value of CC_1 (burnt) is kept constant at 0.001 are shown in Fig. 1. Highest Ƞ, pressure and IMEP are observed at the least infiltration constant of 0.00025, values of all properties decreased as the infiltration constant increased, an evidenced in the sharp decrease observed from infiltration constant 0.005 to 0.1 for pressure, IMEP and Ƞ. It can be deduced that for the varying CC_1, the least infiltration constant from burnt region to the mixed region of unburnt burned is favorable for all performance parameters and that a greater infiltration has adverse effect on the engine efficiency. Generally, infiltration from the burned zone has greater effect on the

engine performance compared to infiltration from the unburned zone as steeper slopes of Fig. 1 suggest for CC_1 as compared to CC_2. It is observed that the difference in variation witnessed for varying CC_2 is not as steep as it was for varying CC_1 as shown in Fig.1 due to the positive implication of unburnt infiltration into the mixed region until optimum values of engine performance are observed and then a decrease in performance beyond this point is evident. It is also evident that slight backflow of burned charge will have noticeable adverse effect on engine performance while these effects are less noticeable and require a lot more unburned charge to infiltrate into the mixed zone to reduce engine performance. Practically, infiltration in the combustion chamber could be inform of crevice flow and blowby, it is essential to have a limited amount of burned gas as backflow to avoid reduction in power and efficiency as mentioned by [37].

The observed peak pressure (in Table 2) using E100 is 15.7% closer to the experimental data. However, when compared to the 2-zone model it is 16.4% closer to the experimental value. High IMEP and η are also observed for the E100 3-zone model.

In Fig. 2, concerning three-zone and two-zone models, there is no significant difference in the general emission patterns of CO_2, NO, and O_2. A reduction in values of N_2 is observed for the three-zone model which could be attributed to higher oxygen content present in ethanol to react with the N_2.

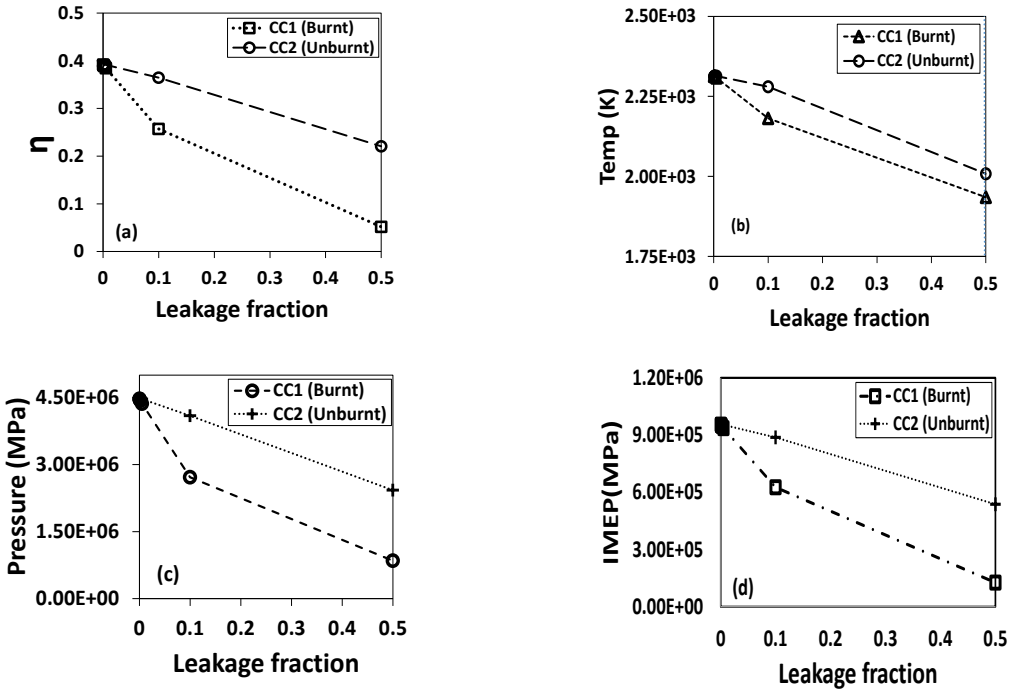


Fig. 1: 3-zone model for the fractional burnt and unburnt regions,
(a) efficiency, (b) temperature, (c) pressure and (d) Indicated Mean Effective Pressure.

Table 2: Comparison of 2-zone model, 3-zone model and experimental results

	2-zone model	Experi-mental values	3-zone model with C_1= 0.00025 and C_2= 0.001	3-zone model at 0.005	3-zone model at optimal value of C_1 = 0.00025 and C_2 = 0.005
Pressure (MPa)	4.4358e+06	5.30E+06	4.4680e+06	4.4668e+06	4.4651e+06
IMEP	9.1635e+05	-	9.5427e+05	9.5382e+05	9.5344e+05
η	0.3741	-	0.3917	0.3915	0.3914

https://doi.org/10.31436/iiumej.v22i2.1550

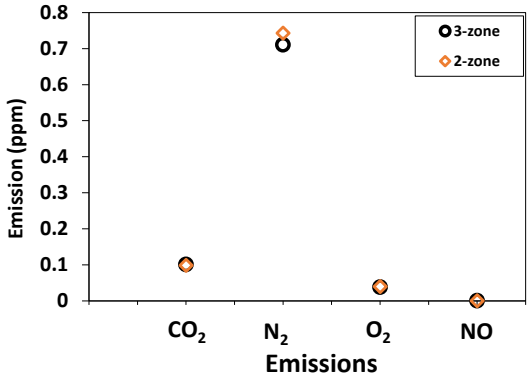


Fig. 2: Emissions for the 2-zone and case-1 models using ethanol.

4.2 Segmented Model

The η, temperature and IMEP history for varying segmented burned zones for E100 investigated are as shown in Fig. 3 using varying heat transfer coefficient as seen in Fig. 3 (a, b, c and d) for heat transfer coefficients, h = 400, 450, and 500, respectively.

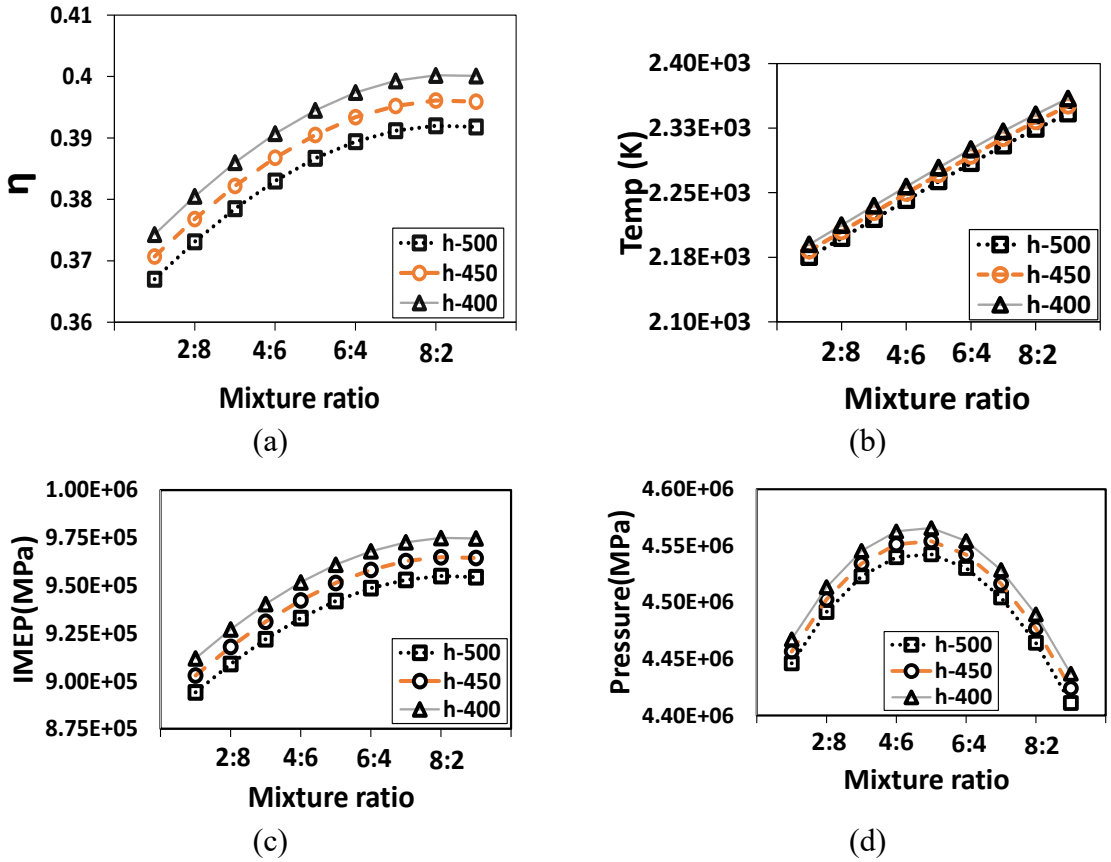


Fig. 3: Three-zone model values for the burned zone-1 and burned zone-2 using heat transfer coefficients 500 and 450 (a) efficiency, (b) temperature, (c) IMEP and (d) pressure.

The maximum peak pressures in Fig. 3(d) for h_{400}, h_{450}, h_{500} are 4.57 MPa, 4.55 MPa, 4.54 MPa, respectively are observed using 4:6, although the highest thermal efficiency and IMEP are observed at a mixture ratio of >7:3, h_{400} in Fig. 3(a and c). Expectedly, the engine temperature history increased as mass flow rate increased in Fig. 3(b), however, the

peak temperature was observed towards the end of the burn flow. Evidently, a partition in the burned mixture had an effect on the engine performance as an increase in engine performance is observed as the mixture ratio increased. It can be argued that a proper and sequential fair variation in mixture ratio to an optimal ratio can increase the engine performance while a decrease in engine performance is observed beyond the optimal value. It is also worthy of note that the sequence of the segment is important, as a ratio of 8:2 is greater compared to a ratio of 2:8. Therefore, the flow and turbulence in the engine chamber requires uniformity to a certain degree and later a little variation or reduction in flow towards the end of the combustion process to stabilize flame ignition as described by [6].

4.3 Transitional-Segmented Model

The readings from both analyses of mass fraction and mass infiltration burned from cases-1 and 2 respectively are used for the third case. This work has reported findings from combining the optimal results. Table 3 presents the four-zone model results for pressure, IMEP, η and emission fractions analyzed from the optimal mass infiltration and mass fraction burned values. It was observed that there was a reduction in CO_2 and increased N_2 emission fraction for the 4-zone model as compared to the other 3-zone models.

Experimental values, 2-zone, 3-zone, and 4-zone models are shown in Table 4. From the table, higher pressure values observed for the 4-zone model is 3.9% closer to experimental results while that of the 2-zone model is 4.4% closer to the experimental results.

Table 3: 4-zone (third case) modelled using ethanol

CC_1	CC_2	P_e (MPa)	Temp (K)	IMEP	η	CO_2	H_2O	N_2	O_2	NO
0.00025	0.005	4.4651e+06	2.3134e+03	9.5344e+05	0.3914	0.1007	0.1511	0.7104	0.0377	0.0001
X1	X2									
0.7	0.3	4.4893e+06	2.341e+03	9.7492e_05	0.4002	0.1007	0.1511	0.7104	0.0377	0.0001
Case-3 Values		4.4565e+06	2.3344e+03	9.4735e+05	0.3867	0.0983	0.1194	0.7428	0.0395	0.0001

Table 4: Showing comparison of results from 2-zone model, three cases modeled and experimental results

	2-zone model	Experi-mental values	3-zone model using X1 = 0.7 and X2 = 0.3	3-zone model using CC_1 = 0.00025 and CC_2 = 0.005	4-zone Model using X1 = 0.7 and X2 = 0.3 (h_{400}), CC_1 = 0.00025 and CC_2 = 0.005
Pe	4.4358e+06	4.640e+06	4.4893e+06	4.4651e+06	4.4565e+06
T_{be}	2305K	2800K	2.3410e+03	2.3134e+03	2.3344e+03
IMEP	9.1635e+05		9.7492e+05	9.5344e+05	9.4735e+05
η	0.3741	-	0.4002	0.3914	0.3867

Table 5 shows comparative analysis using ethanol E100 of the three-zone model involving mass infiltration at optimal values of CC1= 0.00025 and CC2 = 0.005, three-zone model involving mass burn rate of x1 = 0.7 and x2 = 0.3 and a four-zone model labeled as case 3. Emission history for the four-zone model, like the two-zone model, remains the lowest for all emission types. The highest peak pressure, IMEP, and η are observed at the three-zone case 2 model, and this is clearly shown in Fig. 4. It could be concluded that the segmented three-zone model (3zc2) is better for peak in-cylinder pressure and engine efficiency analysis using ethanol, E100. In Fig. 4(a-d), the three-zone

segmented model (3zc2) produces the best results in terms of efficiency, pressure, temperature and IMEP.

Table 5: Comparison of 2-zone, 3-zones (case-1 and case-2) and 4-zone models (case-3) using ethanol, E100

	Pressure (MPa)	Temp (K)	IMEP	η	Emission characteristics				
					CO_2	H_2O	N_2	O_2	NO
2-zone model, 2z	4.4358e+06	2305K	9.1635e+05	0.3741	0.0983	0.1194	0.7428	0.0395	0.0001
First case, 3zc1 using CC1 = 0.00025 and CC2 = 0.005	4.4651e+06	2.3134e+03	9.5344e+05	0.3914	0.1007	0.1511	0.7104	0.0377	0.0001
Second case, 3zc2 0.7 0.3	4.4893e+06	2.3410e+03	9.7492e+05	0.4002	0.1007	0.1511	0.7104	0.0377	0.0001
Third case, 4z	4.4565+06	2.3344e+03	9.4735e+05	0.3867	0.0983	0.1194	0.7428	0.0395	0.0001

2z-2zone model, 3zc1-3zone case 1 model, 3zc2- 3zone case 2 model, 4z- 4zone model

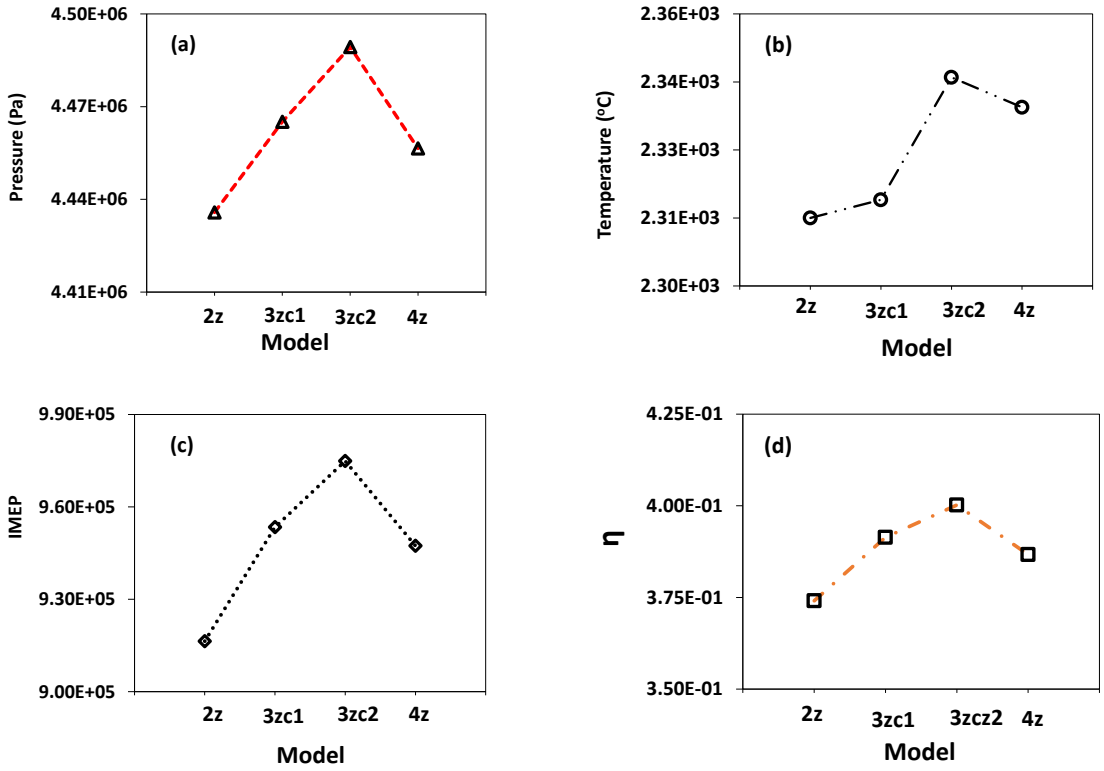


Fig. 4: Temperature, Indicated Mean Effective Pressure (IMEP) and thermal efficiency (η) for 2-zone model (2z), 3-zone transitory model (3zc1), 3-zone segmented model (3zc2) and 4-zone model (4z) using 100% ethanol fuel.

5. CONCLUSIONS

Multi-zone models have been developed for analysis of combustion in spark ignition engines using ethanol as fuel. It is evident that inclusion of additional zones into the combustion chamber analysis and use of ethanol has increased the engine performance in terms of indicated mean effective pressure, thermal efficiency and emission rate. From the developed models and analysis, the following conclusions can be drawn:

1. The infiltration of charges in the combustion chamber should be limited for burned gases as found in the range of infiltration (0.0025, 0.005) that gave optimum values of engine performance for burned and unburned gases respectively.
2. Homogenous flow; creating uniformity of charge flow of ratio 7:3 is encouraged for burned gases.
3. The three-zone segmented model was observed to give a better result of engine performance.
4. Generally, engine zoning has significant effect on engine performance and every zone modelled is user-specific. Further investigation is still required to properly mark out the effect of modification on specific engine parameters.
5. The use of ethanol in the internal combustion engine has given better results of engine performance with promising encouraging effect on emission fraction as compared to that of the gasoline standard.

REFERENCES

[1] Yan B, Wang H, Zheng Z, Qin Y, Yao M. (2018) The effect of combustion chamber geometry on in-cylinder flow and combustion process in a stoichiometric operation natural gas engine with EGR. Applied Thermal Engineering, 129: 199-211.
[2] Evans RL. (1992) Combustion chamber design for a lean-burn SI engine. SAE Transactions, 1611-1616.
[3] Wohlgemuth S, Roesler S, Wachtmeister G. (2014) Piston design optimization for a two-cylinder lean-burn natural gas engine-3D-CFD-simulation and test bed measurements (No. 2014-01-1326). SAE Technical Paper.
[4] Evans RL, Tippett EC. (1990) The effects of squish motion on the burn-rate and performance of a spark-ignition engine (No. 901533). SAE Technical Paper.
[5] Jones MK, Heaton DM. (1989) Nebula combustion system for lean burn spark ignited gas engines (No. 890211). SAE Technical Paper.
[6] Sakurai T, Iko M, Okamoto K, Shoji F. (1993) Basic research on combustion chambers for lean burn gas engines. SAE Transactions, 2240-2250.
[7] Johansson B, Olsson K. (1995) Combustion chambers for natural gas SI engines part I: Fluid flow and combustion. SAE Transactions, 374-385.
[8] Olsson K, Johansson B. (1995) Combustion chambers for natural gas SI engines part 2: Combustion and emissions. SAE transactions, 499-511.
[9] Masoudi R, Azad NL, McPhee J. (2014) Parameter identification of a quasi-dimensional spark-ignition engine combustion model (No. 2014-01-0385). SAE Technical Paper.
[10] Kaprielian L, Demoulin M, Cinnella P, Daru V. (2013) Multi-zone quasi-dimensional combustion models for Spark-Ignition engines (No. 2013-24-0025). SAE Technical Paper.
[11] Juntarakod P, Soontornchainacksaeng T. (2014) A quasi-dimensional three-zone combustion model of the diesel engine to calculate performances and emission using the diesel-ethanol dual fuel. Eng. Sci, *7*(1): 19-37.
[12] Sánche ÓJ, Cardona CA. (2012) Conceptual design of cost-effective and environmentally-friendly configurations for fuel ethanol production from sugarcane by knowledge-based process synthesis. Bioresource Technology, 104: 305-314.
[13] Alamu OJ, Waheed MA, Jekayinfa SO. (2007) Biodiesel production from Nigerian palm kernel oil: effect of KOH concentration on yield. Energy for Sustainable Development, 11(3): 77-82.
[14] Gnansounou E, Dauriat A. (2005) Ethanol fuel from biomass: A review. CSIR, 64 (11): 809-821. http://hdl.handle.net/123456789/5372

[15] Yücesu HS, Sozen A, Topgül T, Arcaklioğlu E. (2007) Comparative study of mathematical and experimental analysis of spark ignition engine performance used ethanol–gasoline blend fuel. Applied Thermal Engineering, 27(2-3): 358-368.
[16] Ceviz MA, Yüksel F. (2005) Effects of ethanol–unleaded gasoline blends on cyclic variability and emissions in an SI engine. Applied Thermal Engineering, 25(5-6): 917-925.
[17] Verma AP, Choube A. (2012) Ethanol as alternative fuel for SI engine - A review. 4(14): 90-95.
[18] Srinivasan CA, Saravanan CG. (2010) Study of combustion characteristics of an SI engine fuelled with ethanol and oxygenated fuel additives. Journal of Sustainable Energy & Environment, 1: 85-91.
[19] Taraba JL, Turner GM, Razor R. (1981) Energy in agriculture: The use of ethanol as an unmixed fuel for lnternal combustion engines. AEES, 14(8): 1-20
[20] Park IJ, Yoo YH, Kim JG, Kwak DH, Ji, WS. (2011) Corrosion characteristics of aluminum alloy in bio-ethanol blended gasoline fuel: Part 2. The effects of dissolved oxygen in the fuel. Fuel, 90(2): 633-639.
[21] Larsen U, Johansen T, Schramm J. (2009) Ethanol as a future fuel for road transportation: Main research report. DTU Mekanik,
[22] Yoo YH, Park IJ, Kim JG, Kwak DH, Ji WS. (2011) Corrosion characteristics of aluminum alloy in bio-ethanol blended gasoline fuel: Part 1. The corrosion properties of aluminum alloy in high temperature fuels. Fuel, 90(3): 1208-1214.
[23] Strong, RM. (1911). Gasoline and alcohol tests on internal-combustion engines (No. BM-BULL-32). Bureau of Mines, Washington, DC (USA).
[24] Bhetalu AD, Patil SS, Ingole NW. An overview ethanol as a motor fuel. Journal of Engineering Research and Studies, 3(2): 50-53.
[25] Bokhary AYF, Alhazmy M, Ahmad N, Albahkali A. (2014) Investigations on the utilization of ethanol-unleaded gasoline blends on SI engine performance and exhaust gas emission. International Journal of Engineering & Technology, 14(2): 88-96.
[26] Karavalakis G, Durbin TD, Shrivastava M, Zheng Z, Villela M, Jung H. (2012) Impacts of ethanol fuel level on emissions of regulated and unregulated pollutants from a fleet of gasoline light-duty vehicles. Fuel, 93: 549-558.
[27] Yusaf T, Buttsworth D, Najafi G. (2009) Theoretical and experimental investigation of SI engine performance and exhaust emissions using ethanol-gasoline blended fuels. In 2009 3rd International Conference on Energy and Environment (ICEE) (pp. 195-201). IEEE.
[28] Kumar, J, Trivedi, D, Mahara, P, Butola, R. (2013). Performance study of ethanol blended gasoline fuel in spark ignition engine. Journal of Mechanical and Civil Engineering, 7(3): 71-78.
[29] Pai S, Tasneem HA, Rao A, Shivaraju N, Sreeprakash B. (2013) Study of impact of ethanol blends on SI engine performance and emission. National Conference on Challenges in Research & Technology in the Coming Decades CP648, pp 1-7.
[30] De Simio L, Gambino M, Iannaccone S. (2012) Effect of ethanol content on thermal efficiency of a spark-ignition light-duty engine. International Scholarly Research Network Renewable Energy, 2012(219703): 1-8.
[31] Dare AA, Ismail OS, Olatunde OB. (2017) Development of three-zone transitional model for reciprocating internal combustion engine analysis using gasoline. Current Journal of Applied Science and Technology, 25(6): 1-11.
[32] Dare AA, Olatunde OB. (2018) Development of four-zone segmented transitional model for reciprocating internal combustion engine analysis using gasoline. American Journal of Science, Engineering and Technology, *3*(2): 46-52.
[33] Buttsworth DR. (2002) Spark ignition internal combustion engine modelling using Matlab. Faculty of engineering & surveying technical reports TR-2002-02: 1-41.
[34] Ferguson CR, Kirkpatrick AT. (2016) Internal combustion Engine. Applied Thermosciences. Third Edition. John Wiley & Sons, Ltd.
[35] Kodavasal J, Keum S, Babajimopoulos A. (2011) An extended multi-zone combustion model for PCI simulation. Combustion Theory and Modelling, 15(6): 893-910.

[36] Asgari O, Hannani SK, Ebrahimi R. (2012) Improvement and experimental validation of a multi-zone model for combustion and NO emissions in CNG fueled spark ignition engine. Journal of Mechanical Science and Technology, *26*(4): 1205-1212.
[37] Heywood JB. (1988) *Internal Combustion Engine Fundamentals*. McGraw Hill, Inc. New York series (11) pp.1

https://doi.org/10.31436/iiumej.v22i2.1707

ANALYTICAL AND NUMERICAL THERMAL ANALYSIS ON FRICTION STIR WELDING USING POLYGONAL TOOL PIN

STEPHEN LEON JOSEPH LEON[1*], ALFRED FRANKLIN VARGHESE[2], JOSEPH MICHEL[2] AND GOPINATH GUNASEKARAN[2]

[1]Department of Mechanical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamilnadu, India.

[2]Engineering Department, University of Technology and Applied Sciences-Ibri, Oman.

[*]Corresponding author: stephenleonj@gmail.com

(Received: 26th November 2020; Accepted: 25th January 2021; Published on-line: 4th July 2021)

ABSTRACT: Frictional heat generation in the tool/matrix interface followed by the stirring of material along the weld line causes plasticized solid state joining in friction stir welding. In this paper, the existing torque based thermo-mechanical model for the tools with cylindrical pins is remodified for the polygonal tool pin profile by introducing novel multiplication factors with respect to the number of sides in the tool pin geometry. The variation in the effective heat supply with respect to the chosen pin geometry was analyzed. A comparative analysis of the proposed analytical model with the existing model was also carried out to understand the accuracy of the proposed model. Furthermore, a transient thermal numerical modelling was carried out in the view of understanding the change in process peak temperature in the stir zone and change in temperature gradient along the heat affected zone with respect to the change in pin geometry for the opted set of process input parameters. An analytically estimated heat-input-based numerical model was adopted in the present study. It was observed that the process peak temperature was directly proportional to the number of sides in the tool pin.

ABSTRAK: Penjanaan haba geseran antara muka pada alat/matrik diikuti dengan pengacauan material sepanjang garis kimpalan menyebabkan keadaan plastik pepejal melekat bersama geseran kimpalan pengacau. Kajian ini berkaitan tork sedia ada berdasarkan model mekanikal-terma bagi alat pin silinder yang terubah suai bagi profil pin alat poligon dengan memperkenalkan faktor gandaan berdasarkan bilangan sisi geometri alat pin. Perubahan pada bekalan haba efektif berdasarkan geometri pin pilihan telah dikaji. Analisis bandingan pada model analitik yang dicadang bersama model sedia ada, telah dilakukan bagi memahami ketepatan model cadangan. Tambahan, model transien numerikal terma telah dibuat bagi memahami proses perubahan suhu puncak ketika zon pengacauan dan perubahan gradien suhu sepanjang zon terkena haba perubahan geometri pin pada set proses parameter input terpilih. Kajian ini mengaplikasi model numerik berdasarkan input anggaran haba secara analitik. Dapatan kajian menunjukkan suhu puncak proses adalah berkadar langsung dengan bilangan sisi pin alat.

KEYWORDS: *thermal mode; friction stir welding; polygonal tool pin; transient model*

1. INTRODUCTION

The process peak temperature rise in friction stir welding during the welding stage is always maintained lower than the solidus temperature of the material to be joined so that this joining process can be considered as a solid-state metal joining technique. The major heat input for this solid-state welding technique is generated along the tool/matrix contact due to friction and it is assisted by the heat generation through the plastic deformation of material during the joining process under the tool shoulder [1]. The tool rotates and slides over the stationary workpiece, resulting in a relative velocity along the tool/matrix interface that is the root cause of the friction heat. This solid-state joining process is divided into different stages based on the position of the tool (Fig.1). Temperature gradient developed in the process at various distances from the axis of rotation of the tool is the key factor that decides the post weld mechanical properties of the joint [2].

Various software assisted simulations are used to estimate the influence of various internal and external factors of the heat generation. Shi et al. [4] studied temperature results through finite element analysis conducted using ABAQUS. Their thermo-mechanical analysis was validated through experimental results. A two-dimensional finite element model developed by Lockwood and Reynolds [5] for friction stir welding in aluminum alloy 2024-T351. The obtained temperature distribution from the FE model was validated by the experimental results.

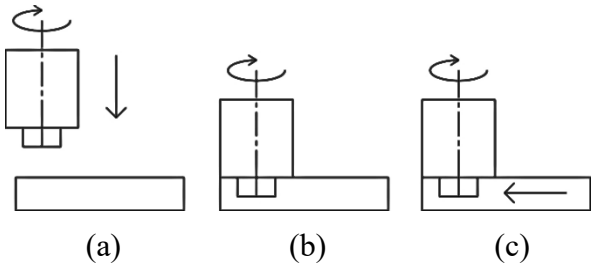


Fig. 1: Stages in FSW: (a) plunging, (b) dwelling, (c) welding.

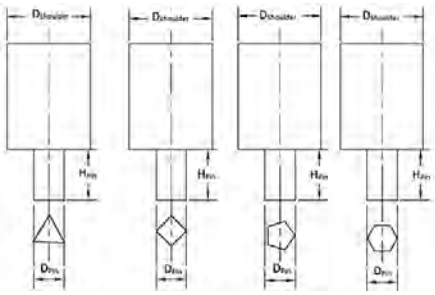


Fig. 2: Tools with polygonal tool pin.

Ullessy [2] evaluated the impact of variation in the tool speed on the thermal field by developing a three-dimensional viscoplastic model using computational fluid dynamics. A better simulation was obtained by Li et al. [6] using ABAQUS software for the fully sticking condition of tool-workpiece interface. Sanjeev et al. [7] analyzed the influence of the friction coefficient on the simulation output using ABAQUS and found that the coefficient of friction was equal to 1.0 during a sticking condition. Apart from similar material, little research was done to optimize process parameters on the joining of two dissimilar materials [8]. Labesh et al. [9] compared the effects of dwell time, tool speed, and plunge depth and found dwell period had higher influence on post weld joint strength over other factors.

In this solid-state welding, major amounts of effective heat supply are obtained along the interface of shoulder/matrix and a part of heat generated along the contact surface of the tool pin and material assists to maintain a constant temperature gradient along the stir zone. Analysis on the joint efficiency of the various welds developed on the usage of a tool with different non-circular straight pin profiles concluded that the tool with a square pin delivered a superior weld quality [10]. Better weld quality [11] was observed on a taper cylindrical pin profile compared to a square pin profile and it was concluded that the

increase in number of sides of a noncircular pin profile, from square profile to hexagon profile, increased the heat generation rate during welding. An experimental study [12] was carried out to understand the effects of tool pin geometries on the joint strength and it was found that tool pin geometry with square cross-sectional shape delivered a better weld quality than other polygonal shapes. Apart from these, various research [13-17] concluded that geometrical change in the tool pin geometry resulted in considerable changes in heat generation that in turn affected the temperature gradient during the joining process. So, tool pin geometry can also be a key factor with other parameters like tool feed rate and tool rotational speed. In this paper, for a given constant working condition, analytical expressions are derived and validated with the published results to analyze the quantity of heat supplied by the different polygonal pin profiles with the number of sides varying from three to six. Furthermore, heat generation during dwell, not only plasticizes the metal under the tool but also softens the neighboring material, which reduces the opposing forces to the tool movement [18]. This underlines the importance of analyzing the transient temperature increase during the dwell period. So, the effect of variation in tool pin geometry on the transient temperature gradient during the dwell period in the stir zone were compared through the results obtained from combined analytical and numerical approaches.

2. ANALYTICAL MODELLING

In the view of simplifying the analytical solution, the following assumptions were made: (i) Estimation of total torque developed during the process is the combined effect of friction along the contact surface and plastic deformation under the tool shoulder; (ii) The small quantity of heat developed during the initial stage (plunging) is neglected [19].

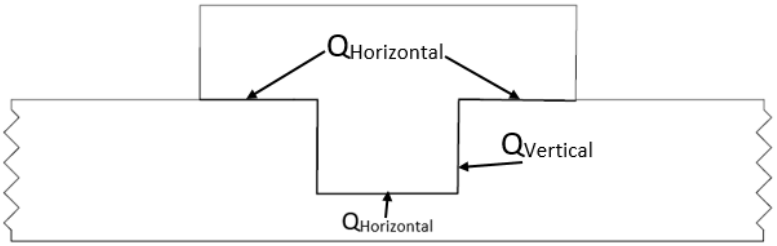


Fig 3: Tool/matrix frictional contact surfaces.

Total heat-generating surfaces are divided into horizontal and vertical surfaces as shown in Fig. 3. Frigaard et al. [20] estimated the total heat generation along the horizontal contact surface as,

$$Q_{Horizontal} = Q_{Shoulder} + Q_{pin\text{-}tip} = \frac{2}{3}\pi\sigma_{contact}\omega R_s^3$$

In the tool shoulder contact surface,

$$Q_{Shoulder} = \left[\frac{2}{3}\pi\sigma_{contact}\omega R_s^3\right] - Q_{pin\text{-}tip}$$

$$Q_{Vertical} = Q_{Pin\text{-}side}$$

For every tool pin geometry, the geometrical shape of the shoulder does not vary considerably. The uniform shoulder/matrix contact surface indicates that there will not be any change in heat generated by the tool shoulder irrespective of its pin geometrical shape. But there is a considerable change in tool pin contact surface area when the pin geometry varies.

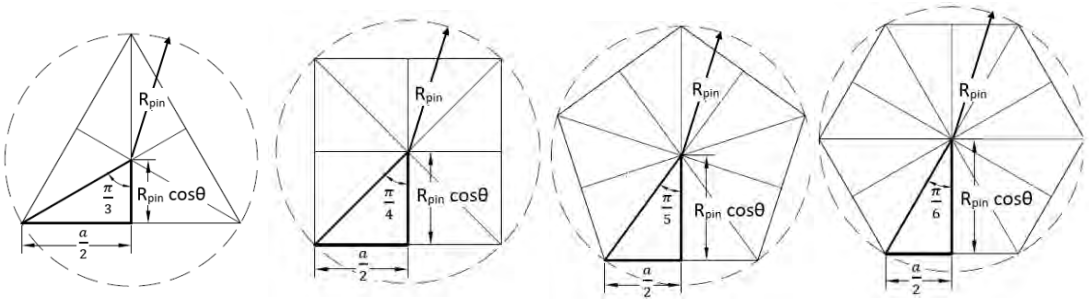


Fig. 4: Cross section of pin profiles and segments.

Polygonal tool pins shown for the current analysis were divided into small triangles as denoted in Fig. 4. The total amount of triangular divisions in every polygonal pin geometry was equal to two times the number of sides in the cross-sectional shape of the tool pin.

Schmidt et al. [21] estimated total torque developed in FSW as,

$$\text{Rotating torque} = \sigma_c\, \omega\, x^2\, d\theta\, dr \quad (1)$$

where σ_c denotes contact stress, ω represents tool rotational speed, x is the distance of any point form the axis of rotation.

Considering one segment in Fig. 4,

$$Q_{pin-tip} = \iint \sigma_c\, \omega x^2 d\theta dx$$

$$\text{D} = \text{f(x,}\theta) \begin{cases} 0 \le \theta \le \frac{\pi}{z} \\ 0 \le x \le R_{pin} cos\theta \end{cases}$$

$$Q_{\text{Pin-tip}} = \sigma_c \omega \int_0^{\frac{\pi}{z}} d\theta \int_0^{R_{pin}cos\theta} f(x) dx$$

Torque developed for the entire pin can be obtained by,

$$\text{Total torque developed} = \int_0^{\pi/z} \int_0^{R_{pin}cos\theta} \tau_c\, \omega\, x^2\, d\theta\, dr$$

$$= \sigma_c\, \omega\, R_{pin}^3 \left[\frac{1}{3} sin\theta - \frac{sin^3\theta}{9}\right]_0^{\pi/z} \quad (2)$$

For the tool pin/matrix horizontal contact surface,

$$Q_{\text{pin-tip}} = z\, \sigma_c\, \omega R_{pin}^3 \left[\frac{2}{3} \sin\left(\frac{\pi}{z}\right) - \frac{2sin^3\left(\frac{\pi}{z}\right)}{9}\right] \quad (3)$$

where z represents the sides in the chosen polygonal pin geometry.

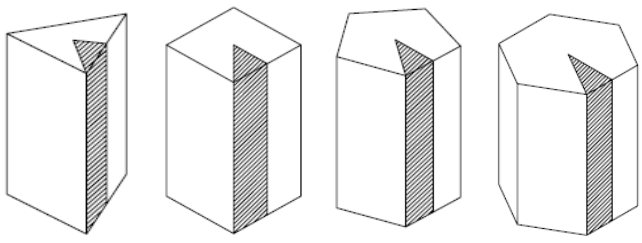


Fig. 5: Vertical surface segments.

https://doi.org/10.31436/iiumej.v22i2.1707

Like the horizontal contact surface, the vertical contact surface can also be divided into small segments as denoted in Fig. 5. The quantity of rectangular divisions in each shape is equal to two times of the number of sides in the tool pin. For every rectangular vertical surface,

$$Q_{pin-vertical} = \iint \sigma_c \, \omega x^2 d\theta dx$$

$$D = f(x,\theta) \begin{cases} 0 \le \theta \le \frac{\pi}{z} \\ 0 \le x \le h \end{cases}$$

$$Q_{Pin\text{-}tip} = \sigma_c \omega \int_0^{\frac{\pi}{z}} d\theta \int_0^h f(x) dx$$

$$\text{Torque developed} = \int_0^{H_{pin}} \int_0^{(\frac{\pi}{z})} \sigma_c \, \omega R_{pin}^2 \, cos^2\theta d\theta \, dh$$

$$= \sigma_c \, \omega \, R_{pin}^2 \; H_{pin} \left[\frac{\pi}{2n} + \frac{1}{4}\sin(2\pi/z)\right] \quad (4)$$

For the pin vertical/matrix contact surface,

$$Q_{pin\text{-}vertical} = \frac{z}{2} \sigma_c \, \omega \, R_{pin}^2 \, H_{pin} \left[\frac{2\pi}{z} + \sin(2\pi/z)\right] \quad (5)$$

where, contact stress $\sigma_c = \mu P$, μ is the friction coefficient, ω is the angular velocity of the tool, pin height (H_{pin}) and shoulder radius (R_s) represent the dimensions of the pin. P is the vertical force given by the tool.

A standard method of prescribing tool pin dimension is through its side length. So the above derived equations have to be rearranged by side length (a). The relationship between pin circumferential radius and pin side length can be represented as,

$$R_{pin} = \frac{a}{2Sin\left(\frac{\pi}{z}\right)} \quad (6)$$

Applying this, equations (3 – 5) can be remodified as,

$$Q_{pin\text{-}tip} = n \, \sigma_c \, \omega \, \frac{a^3}{8Sin^3\left(\frac{\pi}{n}\right)} \left[\frac{2}{3} \sin\left(\frac{\pi}{z}\right) - \frac{2sin^3\left(\frac{\pi}{z}\right)}{9}\right] \quad (7)$$

$$Q_{pin\text{-}tip} = \frac{1}{12} z \, \sigma_c \, \omega \, a^3 \left[\frac{1}{Sin^2\left(\frac{\pi}{z}\right)} - \frac{1}{3}\right] \quad (8)$$

$$Q_{pin\text{-}vertical} = z \, \sigma_c \, \omega \frac{a^2}{8Sin^2\left(\frac{\pi}{z}\right)} \; H_{pin} \left[\frac{2\pi}{z} + \sin\left(\frac{2\pi}{z}\right)\right] \quad (9)$$

$$Q_{pin\text{-}vertical} = z \, \sigma_c \, \omega \, a^2 H_{pin} \left[\frac{\pi}{4Sin^2\left(\frac{\pi}{z}\right)} - \frac{1}{4}\frac{\cos\left(\frac{\pi}{z}\right)}{\sin\left(\frac{\pi}{z}\right)}\right] \quad (10)$$

$$Q_{pin\text{-}vertical} = z \, \sigma_c \, \omega \, a^2 H_{pin} \left[\frac{\pi}{4Sin^2\left(\frac{\pi}{z}\right)} - \frac{1}{4\tan\left(\frac{\pi}{z}\right)}\right] \quad (11)$$

These equations can be simplified by introducing multiplication factors as,

$$Q_{pin\text{-}tip} = X_R \, \tau_c \, \omega \, R_{pin}^3 \quad (12)$$

$$Q_{pin\text{-}vertical} = Y_R \; \tau_c \, \omega \, R_{pin}^2 \, H_{pin} \quad (13)$$

$$Q_{pin\text{-}tip} = X_S \, \tau_c \, \omega \, a^3 \quad (14)$$

$$Q_{pin\text{-}vertical} = Y_S \, \tau_c \, \omega \, a^2 H_{pin} \quad (15)$$

The values of multiplication factors for various number of pin sides are given in Table 1.

Table 1: Multiplication factor values

Pin shape	X_R Value	Y_R Value	X_S Value	Y_S Value
Triangular (n=3)	1.33	4.50	0.27	3.40
Square (n=4)	1.54	5.01	0.64	7.15
Pentagonal (n=5)	1.72	5.44	1.25	12.72
Hexagonal (n=6)	1.89	5.82	2.16	20.35

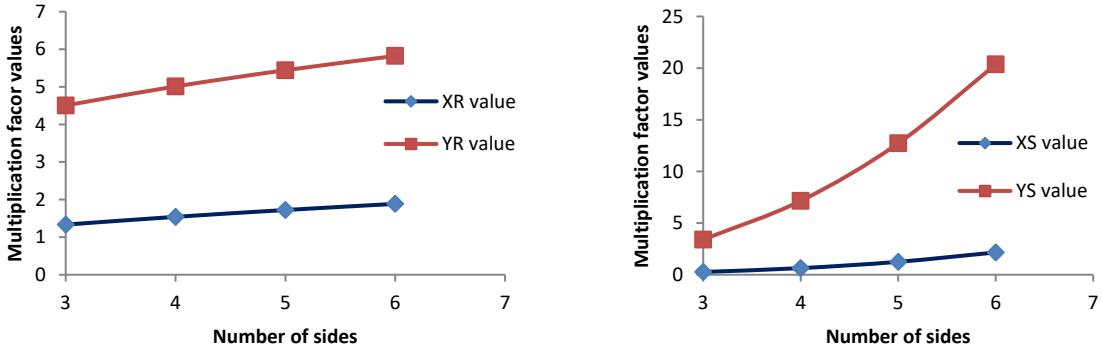
Based on the obtained values for the given number of sides, these multiplication factors can be written as a function of number of sides (z),

$$X_R = 0.77z^{0.5} \quad (16)$$

$$Y_R = 3z^{0.37} \quad (17)$$

$$X_s = 0.01z^3 \quad (18)$$

$$Y_s = 0.2z^{2.58} \quad (19)$$



(a) (b)

Fig. 6: Change in multiplication values (X_R and Y_R) (a) when radius of pin is considered, (b) when pin side length is considered.

3. NUMERICAL MODELLING

In ANSYS workbench, transient temperature rises for different pin profiles, Triangular(T), Square (S), Pentagonal (P) and Hexagonal (H)) were analyzed using analytically estimated heat input for different pin profiles to analyze variation in dwell period to reach the required peak temperature in order to switch over from dwell period to weld period. Applied heat flux does not change in the X and Z directions and it varies only with respect to the Y direction as the thermal conductivity reduces heat flow towards the depth of the workpiece. So, a two-dimensional cross section geometry model was adopted for the analysis. A AA2024-T3 plate was considered as the base material for the current analysis. Other properties considered for the analysis are denoted in Tables 2 and 3.

Table 2: Temperature dependent thermal property of base metal [15]

Temperature (K)	290	373	473	573	673
Specific heat capacity (J kg^{-1} K^{-1})	864	921	1047	1130	1172
Thermal conductivity (W m^{-1} K^{-1})	120	134.4	151.2	172.2	176.4

Table 3: process parameters and tool material properties

Property/parameter	Value
Diameter of tool shoulder (mm)	15
Diameter of the pin (mm)	5
Height of the tool pin (mm)	5.7
Base metal thickness (mm)	6
Material density (tool) (kg/m^3)	7930
Density of base metal (kg/m^3)	2780
Tool specific heat capacity (J kg^{-1} K^{-1})	502
Thermal conductivity of tool (W m^{-1} K^{-1})	21.4
Base metal bottom side heat transfer coefficient at the bottom of the workpiece (W m^{-2} K^{-1})	300
Base metal top side Heat transfer coefficient (W m^{-2} K^{-1})	50
Base metal other side heat transfer coefficient except top and bottom surfaces (W m^{-2} K^{-1})	200

3.1 Boundary and Initial Conditions

At tool shoulder/matrix interface

$$k\left(\left.\frac{\partial T}{\partial n}\right|_{sholder}\right) = \eta(Q_{\text{horizontal}} - Q_{\text{pin-tip}}) \Big|^{R_{pin} \le X \le R_s}_{t > 0}$$

At Pin tip/matrix interface

$$k\left(\left.\frac{\partial T}{\partial n}\right|_{Pin-tip}\right) = \eta\, Q_{\text{pin-tip}} \Big|^{0 \le X \le R_{pin}}_{t > 0}$$

At Pin vertical/matrix interface

$$k\left(\left.\frac{\partial T}{\partial n}\right|_{Piin-vertical}\right) = \eta\, Q_{\text{pin-vertical}} \Big|^{0 \le Y \le H_{pin}}_{t > 0}$$

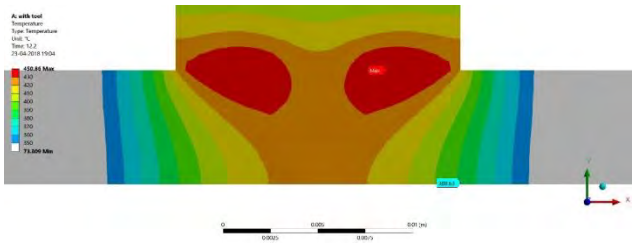
Convective heat transfer from other surfaces to atmosphere

$$k\left(\left.\frac{\partial T}{\partial n}\right|_{Top-bottom-side}\right) = h_x\,(T_x - T_{amb}),$$

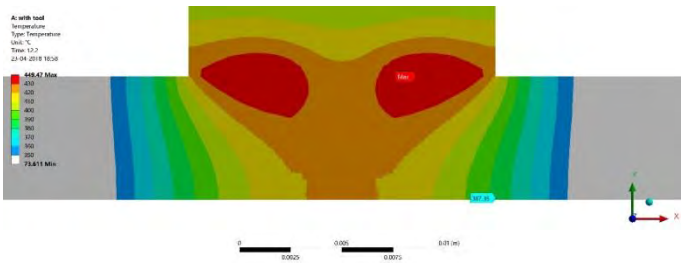
where h_x represents convective heat transfer coefficient between base metal and atmosphere, T_x is surface temperature of the workpiece during the joining process at different points and T_{amb} refers to ambient temperature. For the model simplicity, the bottom surface is also considered to be directly in contact with ambient air.

Here, η refers to the percentage of heat transferred to the workpiece. For the current analysis it had been considered as 95%, with the remaining 5% transferred through the tool.

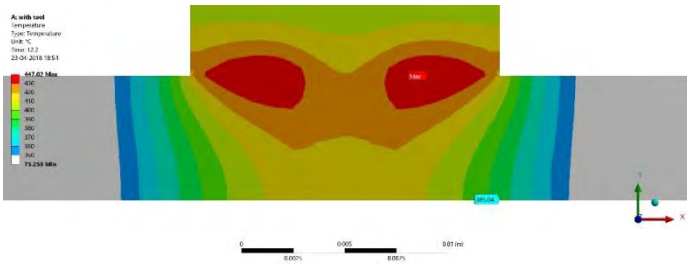
$$T(x,y) = T_{initial} = 22\ ^{o}C \text{ at any point when time } t > 0$$



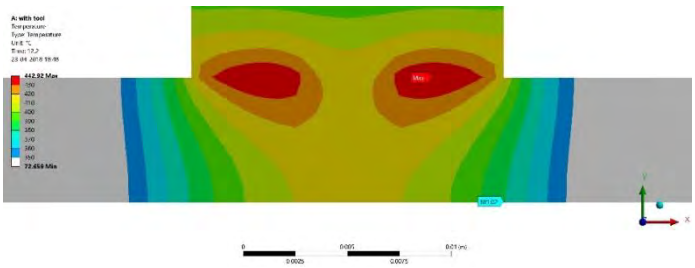
(a) Thermal distribution when hexagonal shaped pin is used.



(b) Thermal distribution when pentagonal shaped pin is used.



(c) Thermal distribution when square shaped pin is used.



(d) Thermal distribution when triangular shaped pin is used.

Fig. 7: Temperature distribution with different pin shapes used at 12.2 s of dwell period.

4. RESULTS AND DISCUSSION

For the current analysis dwell period is considered as the time required to reach a peak temperature of 90% of the melting temperature (775 K) of AA2024-T3. Length of the pin side for different pin geometries (C, S, P and H) were selected in such a way that their pin radius remains constant. In the view of validation of the obtained transient model, the estimated amount of effective heat supply during different process conditions is compared with an analytical model developed by Gadakh et al. [15],

$$Q_{Total} = \frac{2}{3}\pi\tau_{contact}\omega(R_s^3 + X.R_{pin}^2 H_{pin}) \quad (20)$$

where multiplication factor X varies from 0.72 to 3 for different pin profiles.

https://doi.org/10.31436/iiumej.v22i2.1707

Comparative analysis of heat generated along the different contact surfaces of the tool with the base metal are denoted in the Table.4.

Table 4: Total heat generation at by the different pin profiles

Pin profile	$Q_{Shoulder}$ (kW)	$Q_{Pin\text{-}tip}$ (kW)	$Q_{Pin\text{-}Vertical}$ (kW)	Q_{total} (kW)	Q_{total} (kW) Gadhakh et al. [15]
Triangular	4335.55	94.2	397.33	4827.09	4699.08
Square	4306.51	123.24	458.18	4887.94	4785.11
Pentagon	4293.69	136.06	492.19	4921.95	4874.89
Hexagon	4288.45	141.3	510.09	4939.84	4964.67

Gahakh et al. [15] obtained a heat generation model [Eq. 22] based on the assumption that any tool pin shape occupies a circular contact area when it rotates. While obtaining the effective heat supply along the tool horizontal surface ($Q_{shoulder}$ + $Q_{pin\text{-}tip}$), this assumption was justified as the combined geometry of the horizontal surface of tool shoulder and pin occupies a circular contact shape. But while obtaining the heat generation values separately for shoulder/matrix, tool pin tip/matrix, tool side/matrix interfaces, it did not show any variation in the effective heat supply by the horizontal tool contact surfaces of all pin shapes and variations were absorbed along the vertical contact surface of the tool pin.

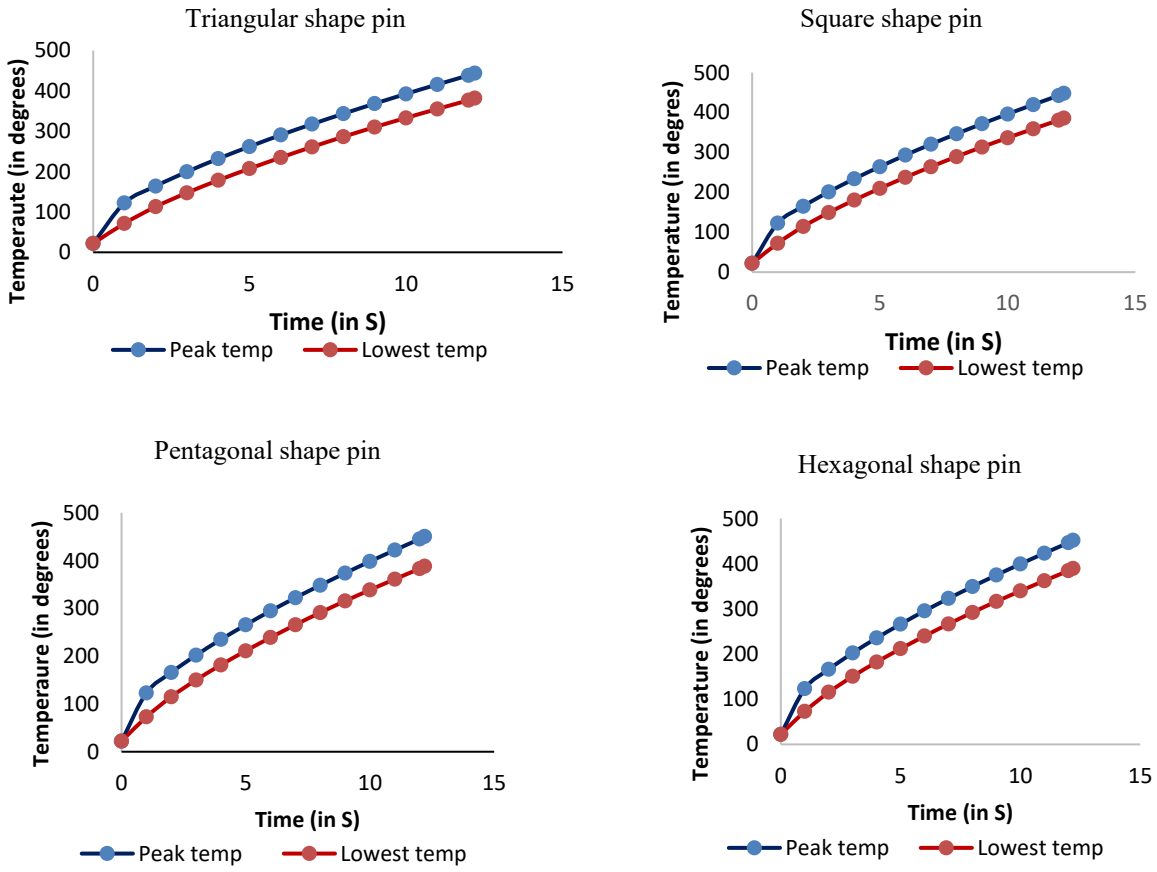


Fig. 8: Transient temperature rises for different pin profiles.

Although the tool pin occupies a circular area irrespective of pin shape while rotating, at any instance of time, contact surface cannot be circular and it varies with

respect to the number of sides considered. As in the derived equations (Eq. 14, 15, 16 & 17) the exact geometry of the tool pin was considered, the accuracy of the obtained result was improved. It can be noticed from Table 4. that the estimated heat generation by the horizontal surface varies with the increase in tool pin sides.

From Fig. 9 it can be observed that the time taken to attain process peak temperature by the hexagonal pin profile is low compared to other shapes and time required increases when the number of sides decreases from pentagonal shape to triangular shape as the heat generated by the contact surface decreases.

Transient temperature rise in a point where the minimum temperature was absorbed in the stir zone is also denoted in Fig. 8. For a constant input rotational speed, maximum and minimum temperatures were observed at different points along the stir zone and shown in Fig. 8 to analyze the temperature gradient at an instance of time. Rapid peak temperature achievement of the hexagonal pin profiled tool reveals the high intensity of effective heat input by the vertical pin surface which assists in quick plasticization of the nearby layer. Isolines obtained from the numerical modelling (Fig. 7) explain the relationship between the intensity of heat supplied and the variation of thermal field along the stir zone for an instance of time. The increase in the number of tool pin sides increases the temperature gradient in the material flow zone at the given instance of time.

The purpose of the dwell period is to increase the temperature of the material under the tool shoulder until it plasticizes in order to reduce flow stress to obtain a high rate of strain around the tool pin. Plasticized material not only has less flow stress but also has lower strength to resist the movement of the tool during the weld, which increases tool life. Increase in the heat intensity reduces the ideal dwell period and increases the welding speed [22]. Increase in welding speed in turn reduces the heat affected zone. As a result, it is understood that an increase in the tool pin sides results in an increase in energy input which drastically decreases temperature rise in the heat affected zone. Failure in FSW joints often occurs in the heat affected zone due to the existence of fewer needle-shaped precipitates. When the effective heat supply intensity is increased, tool feed rate can be increased. When the tool feed rate increased, the heat affected zone is reduced and the temperature gradient in the stir zone also becomes uniform around the tool pin, which in turn reduces the residual stress in the joints.

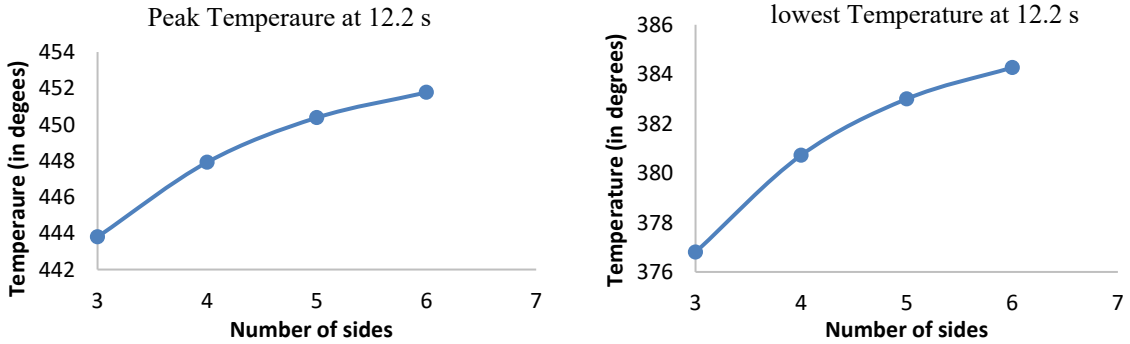


Fig. 9: Change in peak and low temperatures in the stir zone.

Obtained results indicate that although the vertical surface has a limited contribution on total heat supply, it has a major influence on the distribution of heat along the stir zone during the joining process in friction stir welding. From the graph (Fig. 9) it is evident that

increase in the number of tool pin sides, increases the percentage contribution of the vertical surface in the total heat generation.

5. CONCLUSIONS

A novel heat generation model was derived and validated with the goal of analyzing the change in the effective heat supply according to the geometrical changes in tool pin. From the analytical and numerical study on the thermal field developed during the joining process, it can be concluded that:

- Numerical analysis on transient temperature increase during the process suggests that the dwell time is directly proportional to the tool pin sides in the polygonal shape. The increase in tool pin sides has considerable effect on the temperature rise along the heat affected zone.
- Process peak temperature increases with the increase in tool pin sides. A maximum of 452 ºC is absorbed for the hexagonal pin geometry and a minimum of 443.6 ºC is observed for the triangular tool pin geometry.
- For hexagonal tool pin geometry, the effective heat supply intensity is high, which facilities increase in tool feed rate. When the tool feed rate increased, the heat affected zone is reduced and the temperature gradient in the stir zone is also maintained uniform around the tool pin, which in turn reduces the residual stress in the joints.
- From the analytical heat generation estimation, it was understood that the increase in the number of sides in the tool pin increases the percentage contribution of the tool pin on the effective heat supply during the process. For the given process input conditions, usage of the hexagonal tool pin increases the total quantity of heat input up to 330 J/s compared to the heat supply by the tool with the triangular shaped tool pin.

REFERENCES

[1] Leon JS, Jayakumar V. (2019) An investigation of analytical modelling of friction stir welding. Int. J. Mech. Prod. Eng. Res. Dev., 9(1): 179-189. doi: 10.24247/ijmperdfeb201918.

[2] Ulysse P. (2022) Three-dimensional modeling of the friction stir welding process. Int. J. Mach. Tools. Manuf., 42: 1549–1557.

[3] Shi Q, Dickerson T, Shercliff HR. (2003) Thermo-mechanical FE modelling of friction stir welding of Al-2024 including tool loads, 4th International Symposium on Friction Stir Welding, Park City, Utah, USA, 14-16 May 2003.

[4] Lockwood WD, Reynolds AP. (2003) Simulation of the global response of a friction stir weld using local constitutive behavior. Materials Science and Engineering, A339: 35-42.

[5] Li W, Shi S, Wang F, Zhang Z, Ma T, Li J. (2012) Numerical simulation of friction welding processes based on ABAQUS environment. J. Eng. Sci. Tech. Rev., 5: 10-19.

[6] Sanjeev NK, Vinayak M, Suresh Hebbar H. (2014) Effect of coefficient of friction in finite element modeling of friction stir welding and its importance in manufacturing process modelling applications. Int. J. Appl. Sci. Eng. Res., 3: 755-762.

[7] Palanivel R, Mathews PK, Murugan N, Dinaharan I. (2012) Effect of tool rotational speed and pin profile on microstructure and tensile strength of dissimilar friction stir welded AA5083-H111 and AA6351-T6 aluminium alloys. Mater. Des., 40: 7-16.

[8] Labesh Kumar C, Jayakumar V, Bharathiraja G. (2020) Optimization of welding parameters for friction stir spot welding of AA6062with similar and dissimilar thicknesses. Materials Today: Proceedings. https://doi.org/10.1016/j.matpr.2019.07.204.

[9] Labesh Kumar C, Jayakumar V, Bharathiraja G. (2019) Optimisation of friction stir spot welding parameters of AA6062-ST1020 lap joint. Int. J. Mech. Prod. Eng. Res. Dev., 3: 10727-10732.

[10] Suresha CN, Rajaprakash BM, Upadhya S. (2011) A study of the effect of tool pin profiles on tensile strength of welded joints produced using friction stir welding process. Mater. Manuf., 26(9): 1111-1116.

[11] Fujii H, Cui L, Maeda M, Nogi K. (2006) Effect of tool shape on mechanical properties and microstructure of friction stir welded aluminium alloys. Mater. Sci. Eng., 419: 25-31.

[12] Stephen Leon J, Bharathiraja G, Jayakumar V. (2021) Experimental analysis on friction stir welding using flat-faced Pins in AA2024-T3 Plate. FME Transactions, 49: 78-86. doi: 10.5937/fme2101078S.

[13] Vijay SJ, Murugan N. (2010) Influence of tool pin profile on the metallurgical and mechanical properties of friction stir welded Al–10 wt.% TiB2 metal matrix composite. Mater. Des., 31: 3585-3589.

[14] Leon SJ, Bharathiraja G, Jayakumar V. (2020) Durability map for the friction stir welding tools with flat faced Pins. Int J of Integrated Engineering, 2(8): 83-96 doi: 10.30880/ijie.2020.12.08.008.

[15] Gadakh VS, Kumar A. (2014) Friction stir welding window for AA6061-T6 aluminium alloy. Proc IMechE, Part B: J Engineering Manufacture, 228(9): 1172-1181.

[16] Stephen Leon J, Bharathiraja G, Jayakumar V. (2020) Analytical and experimental investigations of optimum thermomechanical conditions to use tools with non-circular pin in friction stir welding. Int. J. Adv. Manuf. Technol., 107: 11-12. doi: 10.1007/s00170-020-05341-7.

[17] Shi L, Wu CS. (2017) Transient model of heat transfer and material flow at different stages of friction stir process. J. Manuf. Proc., 25: 323-339.

[18] Hamilton C, Dymek S, Sommers A. (2008) A thermal model of friction stir welding in aluminium alloys. Int J Mach Tools & Manuf., 48(10): 1120-1130.

[19] Schmidt H, Hattel J. (2005) Modelling heat flow around tool probe in friction stir welding. Sci Tech Weld Join, 10: 176-186.

[20] Frigaard O, Grong O, Midling OT. (2001) A process model for friction stir welding of age hardening aluminum alloys. Metall Mater Trans, 32(5): 1189-1200.

[21] Nandan R, DebRoy T, Bhadeshia HKDH. (2008) Recent advances in friction-stir welding – Process, weldment structure and properties. Prog Mater Sci, 53(6): 980-1023.

[22] Leon JS, Jayakumar V. (2020) Transient heat input model for friction stir welding using non-circular tool pin. FME Transactions, 48(1): 137-142. doi: 10.5937/fmet2001137L.

https://doi.org/10.31436/iiumej.v22i2.1691

MULTI-OBJECTIVE OPTIMIZATION OF SNAKE ROBOT IN SERPENTINE LOCOMOTION

MARWAN BADRAN, MD RAISUDDIN KHAN*, SITI FAUZIAH TOHA AND ZULKIFLI ZAINAL ABIDIN

Department of Mechatronics Engineering, International Islamic University Malaysia, Jalan Gombak, 53100 Kuala Lumpur, Malaysia

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

(Received: 8th November 2020; Accepted: 14th February 2021; Published on-line: 4th July 2021)

ABSTRACT: This paper presents multi-objective optimization for a snake robot with serpentine locomotion. Genetic algorithm (GA) is used to achieve two objectives: minimizing the total *travelling time* and minimizing the total *energy consumption*. The effect of initial values of *winding angle* and *acceleration* on *energy consumption* and *average speed* is depicted. The simulation results show a periodic pattern of the joint torques when the robot maintains a serpenoid curve during travel. Moreover, a Pareto-optimal front was generated for optimal solutions of both of the objectives, while the weighted sum method was used for selecting the best solution. Finally, the simulation results were verified experimentally on an eight-link snake robot considering the limitations of the servomotors used in the experiment. The experimental results with the *winding angle of* 30° was found as the optimum *winding angle* that can achieve both objectives of minimizing the *energy consumption* and the *travelling time*.

ABSTRAK: Kajian ini berkenaan pelbagai-objektif optimum bagi robot ular dengan gerakan serpentin. Algoritma genetik (GA) diguna bagi mencapai dua objektif ini iaitu mengurangkan jumlah masa gerakan dan guna tenaga. Gambaran kesan awal nilai sudut belitan dan pecutan pada guna tenaga dan purata kelajuan dihasilkan. Dapatan simulasi menunjukkan corak berkala tork sendi yang tetap terhasil semasa robot ini berkeadaan lengkung serpenoid ketika bergerak. Tambahan, Pareto-optimal berdepan terhasil bagi solusi optimum pada kedua-dua objektif, sementara kaedah berat campuran digunakan bagi menentukan solusi terbaik. Akhirnya, dapatan simulasi disahkan secara eksperimen pada robot ular lapan-bahagian dengan menimbangkan kekurangan servomotor yang digunakan dalam eksperimen. Dapatan eksperimen menunjukkan sudut belitan 30° adalah sudut belitan optimum bagi kedua-dua objektif iaitu mengurangkan tenaga dan masa gerakan.

KEYWORDS: *snake robot; serpentine locomotion; genetic algorithm; multi-objective optimization; Pareto optimality*

1. INTRODUCTION

Snake robots are robotic mechanisms inspired by the motion of biological snakes. Snake robots were studied during the last five decades, although the early work on real snakes was conducted by the zoologist, Gray [1]. The first snake robot was developed by Hirose in 1972 [2]. A considerable number of research projects on snake robots focused on their design and modelling; thus, various models of snake-like robots were proposed in the literature [3-7]. On the other hand, several studies focused on the control of the snake

robot locomotion [8,9], while others focused on applications for snake robots such as pipe inspection [10], and rescue tasks [11,12].

The pioneer model of a snake robot was presented by Hirose in his book [3], where he presented a thorough study of the locomotion patterns of the natural snake and formulated what he called a *serpenoid curve*, which closely mimics the locomotion of living snakes. Hirose used a snake robot with passive wheels in his experiments to achieve normal friction much higher than tangential friction, which was an essential principle in snake movement. On contrary, Saito et al. [4] designed a wheel-less snake robot with direct contact to the ground. The snake has the potential of adapting to the environment at the cost of increased power consumption. Saito developed his model based on Hirose's principle and obtained equations of motion for a multi-link snake robot travelling with serpentine locomotion. Their model, which was based on the directional friction coefficients, had the feature of decoupling the inertial locomotion from the internal shape motion. Another detailed work on snake robot was introduced in [13]. They presented a systematic complete description of a snake robot in travelling wave locomotion based on the snake robot model introduced by Hirose. The effects of the initial *winding angle* and friction coefficient on the joint input torques were also discussed. Simulation and experiment in that study were carried out using locomotion along a vertical plane. However, this type of locomotion usually needs two periods of serpenoid curve to guarantee two supporting points in contact with the ground to keep the snake robot balanced.

Unlike traditional robots, snake robots are unique in terms of their shape variant, degree of freedom, and locomotion types. Hence, a snake robot has a controllable shape that can achieve a specific locomotion that helps the snake to follow a certain path in a certain direction. More precisely, the snake robot's amplitude (which is controlled by the initial *winding angle*) and the number of cycles can be controlled so that the shape and the size of the snake robot are both variable. Besides, the acceleration of the snake robot plays a role that affects the *travelling speed* and *energy consumption*, and that leads to the need for optimization of locomotion parameters. Therefore, optimization was another area of interest in the research on snake robots, as noticed in [14]. The authors in this study introduced the snake robot model in such a way that optimization could be used to find the optimal parameters for power and speed efficiency. In their study, they used Pareto optimality to obtain the optimal values of snake parameters. For example, they found that one period of the serpenoid curve could achieve the optimal power efficiency. Similar work was proposed in [15] for the locomotion efficiency of snake robots. Optimization was also used by [16], who proposed a path planning technique based on a serpenoid curve and real time dual genetic algorithm to control the snake-like robot. They used multi-objective optimization to find the shortest path and the minimum curvature deviation, as well as decrease the influence of motion error. Another study in [17] used genetic algorithm (GA) to find the optimal central pattern generator (CPG) parameters in terms of moving speed of a snake robot considering the effect of friction coefficients. The study included the design of a fuzzy tuner for CPG-network parameters during robot motion to achieve maximum speed. Genetic algorithm (GA) was also used to optimize the control parameters of a snake robot in [18]. The proposed PI/Backstepping control in that study aimed to reduce the tracking error and control energy in snake robot motion.

The focus of this paper is on multi-objective optimization of a snake robot's parameters to enhance its efficiency in terms of *travelling speed* and *energy consumption*. This study considers the serpentine locomotion of the snake robot while travelling along a straight line.

2. SYSTEM DESCRIPTION

In this section, a brief description is introduced for the snake robot model used in this study including the dynamics of the snake robot. Furthermore, the parameters that play role in achieving the objective functions are also discussed.

2.1 Model of Snake Robot

The model of snake robot used in this study is the one introduced in [3]. In that model, the serpenoid curve is used to control the body shape of the snake robot during travel in serpentine locomotion. The serpenoid curve can be achieved by the curvature function as follows:

$$\rho(s) = \frac{-2\pi\alpha}{L} sin\left(\frac{2\pi}{L}s\right) \tag{1}$$

where L is the length of the snake body, α is the initial *winding angle* of the curve, and s is the body length along the body curve, as shown in Fig. 1.

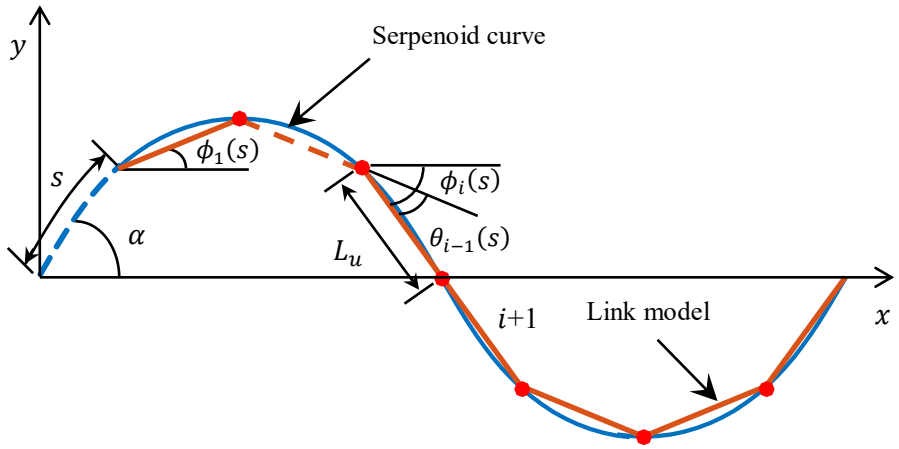


Fig. 1: Snake robot scheme with serpenoid curve.

Here, all the links are assumed to have the same length L_u. Hence, by integration of the curvature function in Eq. (1), the approximate values of relative angle $\theta_i(s)$ can be derived as follows

$$\theta_i(s) = \int_{s+(i-1)L_u}^{s+iL_u} \rho(u)du = -2\alpha\, sin\left(\frac{\pi}{n}\right).sin\left(\frac{2\pi s}{L} + \frac{2\pi i}{n} - \frac{\pi}{n}\right) \tag{2}$$

Consequently, angular velocity $\dot{\theta}_i(s)$ and angular acceleration $\ddot{\theta}_i(s)$ can be obtained as follows:

$$\dot{\theta}_i(s) = -\frac{4\pi\alpha}{L}.sin\left(\frac{\pi}{n}\right).cos\left(\frac{2\pi s}{L} + \frac{2\pi i}{n} - \frac{\pi}{n}\right)\dot{s} \tag{3}$$

$$\ddot{\theta}_i(s) = -\frac{4\pi\alpha}{L}.sin\left(\frac{\pi}{n}\right).cos\left(\frac{2\pi s}{L} + \frac{2\pi i}{n} - \frac{\pi}{n}\right)\ddot{s} + \frac{8\pi^2\alpha}{L^2}.sin\left(\frac{\pi}{n}\right).sin\left(\frac{2\pi s}{L} + \frac{2\pi i}{n} - \frac{\pi}{n}\right)\dot{s}^2 \tag{4}$$

where

$$i = 1,2,\dots,n-1$$

n = number of links
L_u = length of single link
$\dot{s}$ is the velocity along the body curve
$\ddot{s}$ is the acceleration along the body curve given as

$$\ddot{s} = \begin{cases} a & 0 \le t < R \\ 0 & R \le t < T - R \\ -a & T - R \le t < T \end{cases} \quad (5)$$

where, a is the initial acceleration, R is the acceleration/deceleration time and T is the total travelling time.

Since $\ddot{s}$ is given, $\dot{s}$ and s can be found by finding the first integral and second integral of $\ddot{s}$, respectively. Based on the above equations, $\theta_i(s)$ can be varied to control the body shape and *travelling speed* of the snake robot.

2.2 Dynamics of the Snake Robot

As mentioned above, the locomotion of the snake robot relies on the difference between tangential friction and normal friction. Thus, the dynamics of the snake robot describes the relationship between the joint torques, frictional force and the robot locomotion as illustrated in Fig. 2.

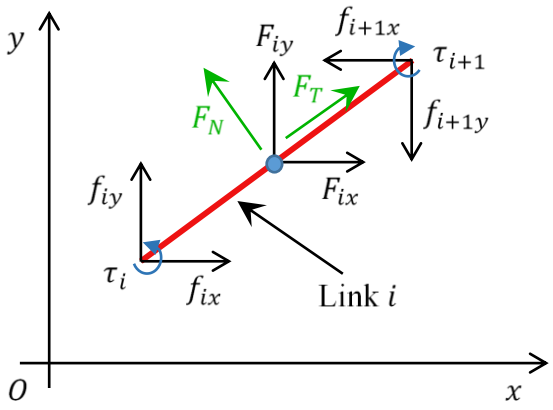


Fig. 2: Scheme of forces acting on link i.

To represent the environment dynamics, *Coulomb friction* is used so that the direction of motion is taken in consideration in terms of the sign of the velocity. Hence, tangential friction force (F_T) and normal friction force (F_N) for link i can be calculated as follows:

$$F_{iT} = \mu_t . \text{sgn}(v_{iT}) . mg \quad (6)$$

$$F_{iN} = \mu_n . \text{sgn}(v_{iN}) . mg \quad (7)$$

where μ_t and μ_n are tangential and normal friction coefficients of link i, while v_{iT} and v_{iN} are tangential and normal velocity at the center of gravity of link i. The latter two velocities can be calculated as follows:

$$v_{iT} = \dot{x}_{iG} \cos\phi_i + \dot{y}_{iG} \sin\phi_i \quad (8)$$

$$v_{iN} = \dot{x}_{iG} \sin\phi_i + \dot{y}_{iG} \cos\phi_i \quad (9)$$

https://doi.org/10.31436/iiumej.v22i2.1691

where $\dot{x}$ and $\dot{y}$ are the velocity components at the centre of gravity of link i. Using Eqs. (6) and (7), the x and y components of the friction force are expressed as:

$$F_{ix} = F_{iT}.\cos\phi_i - F_{iN}.\sin\phi_i \tag{10}$$

$$F_{iy} = F_{iT}.\sin\phi_i - F_{iN}.\cos\phi_i \tag{11}$$

Referring to Fig. 2 and by applying Newton's second law of motion, the motion for the link i with respect to the link i+1 can be described as follows

$$f_{ix} - f_{i+1x} + F_{ix} = m\ddot{x}_{iG} \tag{12}$$

$$f_{iy} - f_{i+1y} + F_{iy} = m\ddot{y}_{iG} \tag{13}$$

$$\tau_i - \tau_{i+1} + (f_{ix} + f_{i+1x} + F_{ix})\,L_u \sin\phi_i/2 - (f_{iy} + f_{i+1y} + F_{iy})\,L_u \cos\phi_i/2 = I_{i\ddot{\phi}_i} \tag{14}$$

However, joint 1 and joint n+1 represent the tail and the head of the snake robot where no actuators are attached. Therefore, torque and force at both ends will be zeros as expressed in the following two equations:

$$f_{1x} = f_{1y} = f_{n+1x} = f_{n+1y} = 0 \tag{15}$$

$$\tau_1 = \tau_{n+1} = 0 \tag{16}$$

By taking summation of both sides of the recursive formulas Eqs. (12) and (13) and then satisfying the conditions in Eqs. (15) and (16), the force equations Eqs. (17) and (18) of link i are obtained.

$$f_{ix} = \sum_{j=i}^{n} m\ddot{x}_{jG} - \sum_{j=i}^{n} F_{jx} \tag{17}$$

$$f_{iy} = \sum_{j=i}^{n} m\ddot{y}_{jG} - \sum_{j=i}^{n} F_{jy} \tag{18}$$

Substituting Eq. (17) and (18) in the torque Eq. (14), we get:

$$\begin{aligned}\tau_i - \tau_{i+1} &+ \left(2\sum_{j=i}^{n} m\ddot{x}_{jG} - 2\sum_{j=i}^{n} F_{jx} - m\ddot{x}_{iG} + F_{ix}\right) L_u \sin\phi_i \Big/ 2 \\ &- \left(2\sum_{j=i}^{n} m\ddot{y}_{jG} - 2\sum_{j=i}^{n} F_{jy} - m\ddot{y}_{iG} + F_{iy}\right) L_u \cos\phi_i \Big/ 2 = I_i\ddot{\phi}_i\end{aligned} \tag{19}$$

Consequently, summation of both sides of Eq. (19) gives:

$$\begin{aligned}\sum_{i=1}^{n}\Bigg\{\left(2\sum_{j=i}^{n} m\ddot{x}_{jG} - 2\sum_{j=i}^{n} F_{jx} - m\ddot{x}_{iG} + F_{ix}\right) L_u \sin\phi_i \Big/ 2 \\ - \left(2\sum_{j=i}^{n} m\ddot{y}_{jG} - 2\sum_{j=i}^{n} F_{jy} - m\ddot{y}_{iG} + F_{iy}\right) L_u \cos\phi_i \Big/ 2\Bigg\} = \sum_{i=1}^{n} I_i\ddot{\phi}_i\end{aligned} \tag{20}$$

By solving Eq. (20) we obtain the only unknown variable ($\ddot{\phi}_1$) which is the rotation acceleration of the first joint. However, from kinematics of link i at the center of gravity we have:

$$\ddot{x}_{iG} = \ddot{x}_i - \frac{L_u}{2}\cos\phi_i\dot{\phi}_i - \frac{L_u}{2}\sin\phi_i\ddot{\phi}_i \tag{21}$$

$$\ddot{y}_{iG} = \ddot{y}_i - \frac{L_u}{2}\sin\phi_i\dot{\phi}_i + \frac{L_u}{2}\cos\phi_i\ddot{\phi}_i \tag{22}$$

Hence, substituting Eqs. (21) and (22) in Eqs. (17) and (18) gives:

$$n\ddot{x}_1 - \sum_{i=1}^{n}\left[\sum_{k=1}^{i-1}\left(L_u\cos\phi\dot{\phi}_k^2 + L_u\sin\phi_k\ddot{\phi}_k\right) + \frac{L_u}{2}\cos\phi_i\dot{\phi}_i + \frac{L_u}{2}\sin\phi_i\ddot{\phi}_i\right] = \sum_{i=1}^{n+1}F_{ix}/m \tag{23}$$

$$n\ddot{y}_1 - \sum_{i=1}^{n}\left[\sum_{k=1}^{i-1}\left(L_u\sin\phi\dot{\phi}_k^2 - L_u\cos\phi_k\ddot{\phi}_k\right) + \frac{L_u}{2}\sin\phi_i\dot{\phi}_i - \frac{L_u}{2}\cos\phi_i\ddot{\phi}_i\right] = \sum_{i=1}^{n+1}F_{iy}/m \tag{24}$$

Substituting $(\ddot{\phi}_1)$ obtained from Eq. (20) into Eqs. (23) and (24), we get the linear acceleration of the first link $(\ddot{x}_1)$ and $(\ddot{y}_1)$.

Moreover, substituting the obtained variables from Eqs. (20), (23) and (24) into Eq. (19), all the joints' torques required to generate the robot motion are obtained. Consequently, the power consumption of each joint can be calculated as

$$P_i = \left|\tau_i\dot{\theta}_i\right| \tag{25}$$

Hence, total *energy consumption* of all joints for the travelling time, *T* can be calculated as

$$E_T = \sum_{i=1}^{n}\int_0^T \left|\tau_i\dot{\theta}_i\right| dt \tag{26}$$

Since time segment used in this case is discrete, *trapezoidal numerical integration* can be used to find an approximation of the total *energy consumption*.

2.3 Energy Consumption and Average Speed

As mentioned earlier in this paper, the target of this study is to achieve the maximum *average speed* (or minimum *travelling time*) of the snake robot using minimum *energy consumption*. For this purpose, we need to study the factors that may affect the *average speed* and the *energy consumption* of the snake robot. Namely, we focus on the *winding angle* and the *acceleration* of the snake robot, as illustrated in the following subsections.

2.3.1 Effect of Winding Angle on Energy Consumption and Average Speed

In order to study the effect of *acceleration* and *winding angle* of the snake robot on *energy consumption* and *average speed* of the forward motion, several simulations were performed. In these simulations, the snake robot travelled forward using serpentine locomotion for a distance of 3 meters. In the next step, *winding angle* was varied within the range 0.3 to 1.5 radians, while the *acceleration* was kept constant at 0.1 m/s^2. The results of this simulation are shown in Fig. 3.

As seen in Fig. 3, increasing the *winding angle* of the snake robot has an undesired effect on both *energy consumption* and *average speed*. In other words, if the *winding angle* is increased, the *average speed* of the snake robot is decreased, and the *energy consumption* is increased. Since the objective is to minimize the *energy consumption* as well as the *travelling time* of the snake robot, it is obviously clear that selecting a minimum *winding angle* would achieve both objectives. The detailed results of this simulation are listed in Table 1.

https://doi.org/10.31436/iiumej.v22i2.1691

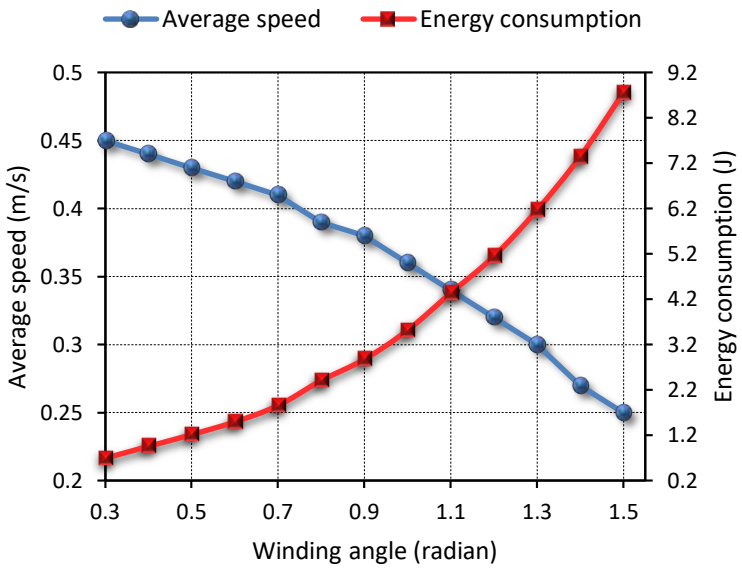


Fig. 3: Effect of *winding angle* on *energy consumption* and *average speed.*

Table 1: Effect of *winding angle* on *energy consumption* and *average speed*

Winding Angle		Travelling Time	Average Speed	Energy Consumption
radian	**(°)**	**(s)**	**(m/s)**	**(J)**
0.3	**17.2**	06.70	**0.45**	**0.70**
0.4	22.9	06.80	0.44	0.96
0.5	28.6	06.90	0.43	1.22
0.6	34.4	07.10	0.42	1.50
0.7	40.1	07.30	0.41	1.86
0.8	45.8	07.70	0.39	2.41
0.9	51.6	08.00	0.38	2.88
0.1	05.7	08.30	0.36	3.52
1.1	63.0	08.90	0.34	4.34
1.2	68.8	09.40	0.32	5.15
1.3	74.5	10.20	0.30	6.17
1.4	80.2	11.10	0.27	7.35
1.5	85.9	12.20	**0.25**	**8.76**

As illustrated in Table 1, while the *winding angle* is gradually increased from 0.3 to 1.5 radians, the *travelling time* of the snake robot is also increased. Accordingly, the *average speed* (travelled distance divided by *travelling time*) is decreased. Table 1 shows that applying minimum *winding angle* (0.3 radian) can achieve maximum *average speed* (0.45 m/s) and minimum *energy consumption* (0.70 J). On the contrary, applying maximum *winding angle* (1.5 radian) leads to undesired minimum *average speed* (0.25 m/s) and undesired maximum *energy consumption* (8.76 J).

2.3.2 Effect of Acceleration on Energy Consumption and Average Speed

To study the effect of *acceleration* along the body of the snake robot on *energy consumption* and *average speed*, simulation tests were carried out using different values of

https://doi.org/10.31436/iiumej.v22i2.1691

acceleration ranging between 0.1 and 1.3 m/s^2, while the *winding angle* was kept constant at $\pi/3$ radian. Simulation results are shown in Fig. 4.

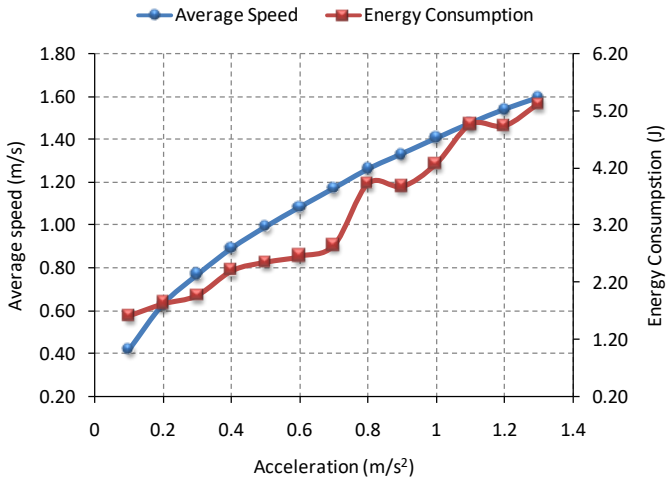


Fig. 4: Effect of *acceleration* on *energy consumption* and *average speed.*

Figure 4 demonstrates that the *acceleration* along the body of the snake robot has a contradictory effect on *average speed* and *energy consumption*. Hence, increasing the *acceleration* will increase *average speed*, and that matches the first objective. On the other hand, increasing the *acceleration* will increase *energy consumption* of the snake robot, which mismatches the second objective. Therefore, there is no single value of *acceleration* that can achieve both objectives.

For more details, Table 2 illustrates the effect of the initial *acceleration* of the snake robot on *average speed* and *energy consumption*. For example, applying the maximum *acceleration* (1.3 m/s^2) achieves the maximum *average speed* (1.60 m/s), however, that leads to undesired maximum *energy consumption* (5.32 J). While applying the minimum *acceleration* (0.1 m/s^2) achieves the minimum *energy consumption* (1.61 J), but leads to undesired minimum *average speed* (0.42 m/s).

Table 2: Effect of *acceleration* on *energy consumption* and *average speed*

Acceleration (m/s^2)	**Travelling Time (s)**	**Average Speed (m/s)**	***Energy Consumption* (J)**
0.1	7.2	**0.42**	**1.61**
0.2	4.9	0.63	1.83
0.3	4.0	0.77	1.96
0.4	3.5	0.89	2.40
0.5	3.1	0.99	2.55
0.6	2.8	1.08	2.65
0.7	2.6	1.17	2.85
0.8	2.5	1.27	3.93
0.9	2.3	1.33	3.88
0.1	2.2	1.41	4.27
1.1	2.1	1.48	4.97
1.2	2.0	1.54	4.94
1.3	1.9	**1.60**	**5.32**

In conclusion, both objective functions (minimizing of *energy consumption* and maximizing of *average speed*) are dependent on *winding angle* and *acceleration* of the

snake robot. Consequently, the latter two parameters need to be optimized in order to achieve the mentioned objective functions. Once the optimization parameters are defined, the simulation setup can be started as described in the following section.

3. SIMULATION SETUP

Multi-objective optimization described in the previous section was implemented using the *gamultiobj* function in MATLAB that applies a genetic algorithm to find *Pareto-optimal front* of multiple objective functions. Pareto-optimal front contains optimal solutions that fit the objective functions. The general form of the *gamultiobj* function is expressed as follows:

[X, FVAL] = gamultiobj (FitnessFcn, nVars, A, b, , , LB, UB, options)

where X: is a matrix of all solutions of the design variables (*winding angle* and *acceleration*)

FVAL: is a matrix of all values of objective functions defined in *FitnessFcn* (*average speed* and *energy consumption*)

A fundamental step in optimization techniques is that initial values need to be set before moving forward in the optimization process. The initial values of the function's parameters that have been used in this study are described in Table 3.

Table 3: Initial values for parameters of *gamultiobj* function

Parameter	Value	Description
FitnessFcn	multiObjFcn	Fitness Function
nVars	2	Number of design variables
LB	[π/6, 0.1]	Lower bounds
UB	[π/3, 2.0]	Upper bounds
PopulationSize	50	Number of individuals
Generations	15	Total number of generations

The fitness function in this simulation uses the snake locomotion model as described in Section 2. The fitness function has two parameters as input: *winding angle* and *acceleration*, the values of which are generated by the genetic algorithm. On the other hand, the fitness function returns the values of the two parameters: *energy consumption* and *average speed*, which will be used by genetic algorithm to determine the fitness value of the input variables. Moreover, the other variables related to snake robot are initialized in the fitness function as described in Table 4.

Table 4: Initial values for snake robot variables

Variable	Value	Description
n	8	Number of links
L	1	Length of snake body (m)
m	0.1	Mass of single link (kg)
I	0.0001	Moment of inertia (kg m^2)
Mu_t	0.07	Tangential friction coefficient
Mu_n	0.7	Normal friction coefficient

https://doi.org/10.31436/iiumej.v22i2.1691

The travelling distance was set to 3 meters, which is enough to get more than one full cycle of the snake shape. Meanwhile, the speed of the snake $\dot{s}$ along the serpenoid curve is determined by rise time R as described in $\ddot{s}$ function in Eq. (5). Arbitrarily, the value of R is set to 5 seconds. The process of multi-objective optimization is illustrated in the flowchart shown in Fig. 5.

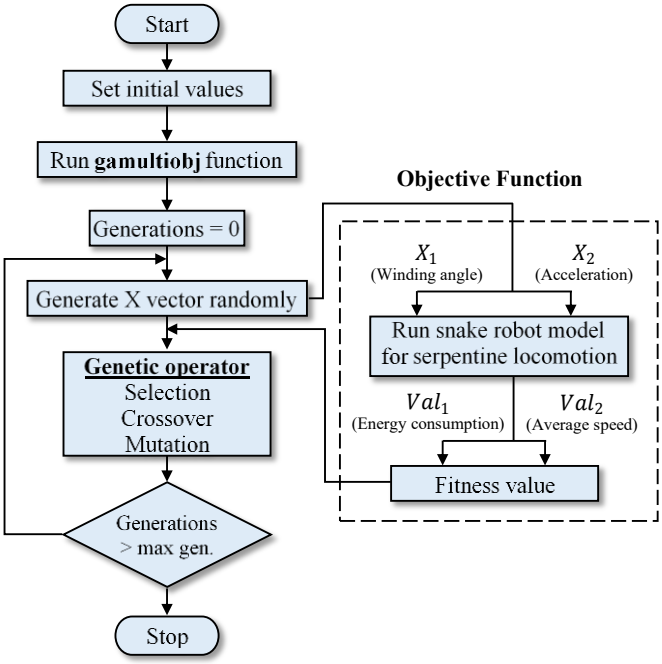


Fig. 5: Flowchart of multi-objective optimization process.

4. SIMULATION RESULTS

At first, the fitness function was simulated using settings in Section III. The fitness function uses kinematics and dynamics of the snake robot to achieve the serpentine locomotion. The results show a variety of input torques of the snake robot joints as seen in Fig. 6.

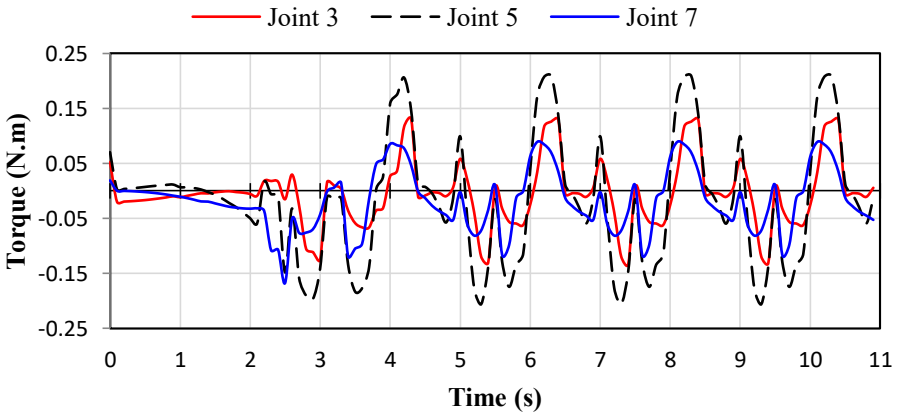


Fig. 6: Input torque for snake robot joints.

The pattern of the joints' torques shows a periodic shape because the snake is following a repeated sinusoidal path. However, it can be noticed that the middle joint

(joint 5) has the highest torque compared to the other joints. Consequently, the torque is lower when joints are farther from the centre of the snake robot (joints 3 and 7), which makes sense, as each joint is affected by the forces of the preceding and succeeding links. Several simulations were carried out using *gamultiobj* tool in order to optimize the *winding angle* and *acceleration* parameters. Therefore, different values of these two parameters were used for multi-objective optimization. On the other hand, *Population Size* was selected as follows: 20, 30, 40 and 50, along with different values of the number of *Generations* and *Pareto Fraction*. The travelling distance was selected as 3 meters. Eventually, the configuration that showed the least gaps between solution points was selected, which was as follows: *PopulationSize*: 50, *Generations*: 15, *Pareto Fraction*: 0.7. Using the previous configuration, simulation was repeated for three trials giving the optimal solutions using Pareto-optimal front as shown in Fig. 7.

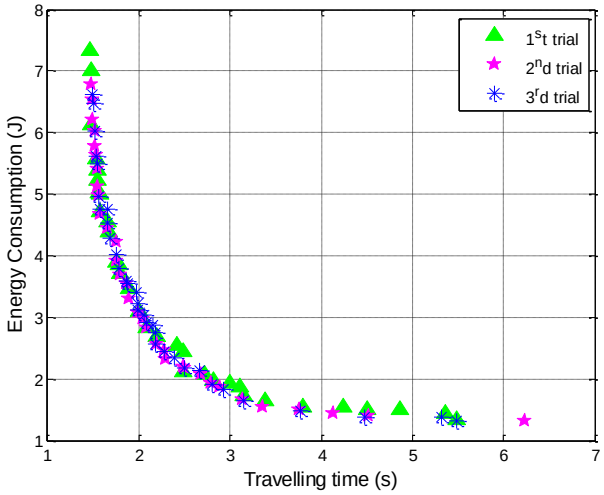


Fig. 7: Pareto-optimal front using (*PopulationSize* = 50, *Generations* = 15).

Despite using the same parameters in the three trials, the possible solutions are different from one trial to another as seen in Fig. 7. That is referred to the algorithm of the optimization technique, which uses random numbers for the initialization process. Nevertheless, by looking at the Pareto-optimal front in Fig. 7, all the possible solutions of the three trials lie within the same precision, which denotes the stability of the optimization process. The three trials have almost the same range of solutions that is located between 1.3 and 6.8 Joules for *energy consumption*, and between 1.5 and 6.2 seconds for *travelling time*. It is also noticed that concentration of solutions decreases as *travelling time* increases.

The solutions in the Pareto-optimal front are used to determine the relative angles for each joint of the snake robot to achieve the objective functions. Table 5 shows the relative joints' angles based on the optimized parameters. By looking at Table 5, it is noticed that all values of *winding angle* (α) are approaching $\pi/6$ which is the minimum value within the angle bounds as initialized in *gamultiobj* function. These values match with *winding angle* curve behaviour in Fig. 3. On the other side, values of *acceleration* have wider range that almost covers the whole range of *acceleration* (0.1 to 2.0) as initialized in the *gamultiobj* function. However, all Pareto optimal solutions in Table 5 are slightly different from each other, which denotes the fact that all solutions can be accepted with a little difference in fitting the objective functions

https://doi.org/10.31436/iiumej.v22i2.1691

Table 5: Relative joints' angles for optimized parameters Pareto-optimal front

Pareto Solution	Winding Angle (°)	Acceleration (m/s^2)	Joints' Angles (°)						
			θ_1	θ_2	θ_3	θ_4	θ_5	θ_6	θ_7
1	31.0	0.22	-09.2	-22.0	-21.8	-8.9	09.2	22.0	21.8
2	31.0	0.62	-09.5	-22.1	-21.7	-8.6	09.5	22.1	21.7
3	31.0	1.80	-10.3	-22.4	-21.4	-7.8	10.3	22.4	21.4
4	31.0	1.57	-10.1	-22.3	-21.5	-8.0	10.1	22.4	21.5
5	31.1	0.51	-09.5	-22.1	-21.8	-8.7	09.5	22.1	21.8
6	31.1	1.12	-09.9	-22.3	-21.7	-8.3	09.9	22.3	21.7
7	31.2	0.72	-09.6	-22.2	-21.8	-8.6	09.6	22.2	21.8
8	31.2	0.93	-09.8	-22.3	-21.8	-8.5	09.8	22.3	21.8
9	31.3	0.57	-09.6	-22.3	-21.9	-8.8	09.6	22.3	21.9
10	31.3	0.83	-09.7	-22.3	-21.9	-8.6	09.7	22.3	21.9
11	31.3	1.27	-10.0	-22.5	-21.7	-8.3	10.0	22.5	21.7
12	31.3	0.31	-09.4	-22.2	-22.0	-8.9	09.4	22.2	22.0
13	31.3	1.87	-10.5	-22.6	-21.6	-7.8	10.5	22.6	21.6
14	31.3	1.61	-10.3	-22.6	-21.6	-8.0	10.3	22.6	21.6
15	31.3	1.41	-10.1	-22.5	-21.7	-8.2	10.1	22.5	21.7
16	31.3	0.78	-09.7	-22.4	-21.9	-8.6	09.7	22.4	21.9
17	31.3	1.78	-10.4	-22.6	-21.6	-7.9	10.4	22.6	21.6
18	31.4	0.95	-09.9	-22.5	-21.9	-8.5	09.9	22.5	21.9
19	31.5	1.26	-10.1	-22.6	-21.9	-8.3	10.1	22.6	21.9
20	31.5	1.13	-10.0	-22.6	-22.0	-8.4	10.0	22.6	22.0
21	31.6	1.03	-10.0	-22.6	-22.0	-8.5	10.0	22.6	22.0
22	31.6	0.16	-09.4	-22.4	-22.3	-9.1	09.4	22.4	22.3
23	31.8	1.84	-10.6	-23.0	-21.9	-8.0	10.6	23.0	21.9
24	31.8	1.93	-10.7	-23.0	-21.9	-7.9	10.7	23.0	21.9
25	31.8	0.15	-09.4	-22.5	-22.5	-9.2	09.4	22.6	22.5

4.1 Selection of the Best Solution

After obtaining the Pareto optimal solutions, which can be considered as equally accepted solutions, only one solution is needed to achieve the objective functions. For such a case, several methods are available and can be used to select a single solution from all Pareto optimal solutions. One of these methods *the weighted sum method* (also called *utility function*), which is one of the simplest and most meaningful methods for solving multi-objective optimization problems [19]. Moreover, *the weighted sum method* is more suitable for a convex Pareto frontier (which is the case in this study). In this method, a weighting factor (w_i) is given for each of the objective functions (f_i) based on its importance compared to the other objective functions. Then a total or overall utility function U can be defined as follows [20]:

$$U = \sum_{i=1}^{k} w_i f_i(X) \tag{27}$$

where

w_i is the weighting factor
k is the number of objective functions
X is a vector of design variables.

The solution can be found by maximizing (or minimizing) the total utility U based on the objective function. The result of this method is a single value, but it may denote more than one solution when they have the same utility value.

Let us say there are two objective functions: f_1 and f_2, then there should be two preselected weighting factors: w_1 and w_2. The weights used in this method usually satisfy the normalization condition:

$$\sum_{i=1}^{k} w_i = 1 \tag{28}$$

Using the solution set of trial 1 in Fig. 7, the *weighted sum method* was applied several times using various weights for each of the objective functions. However, all values of the objective functions need to be normalized (between 0 and 1) to eliminate the effect of the big difference among values of different units. Hence, the following formula was used for normalization purposes:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{29}$$

Each time the *weighted sum method* was applied, only one solution was selected out of all solutions in the Pareto curve. The results of using different combination of weights are shown in Table 6, where f_1 and f_2 represent the objective functions, while w_1 and w_2 represent their weights.

As mentioned earlier, the *weighted sum method* needs a predefined weight for each of the objective functions. These weights are user-defined parameters, and vary from one case to another. Therefore, Table 6 has various values to include a wide range of weights for each objective function to show the effect of this variety on the selection of the best answer. The last two columns in Table 6 clearly show that the values match with the behaviour shown on the curves in Fig. 3 and Fig. 4. For example, when high priority is given to minimize *energy consumption*, as in the first case in Table 6, the minimum *energy consumption* is achieved (1.323 J) with the cost of minimum *average speed* (0.119 m/s). On the other hand, when high priority is given to maximize the *average speed*, as in the last case in Table 6, the maximum *average speed* is achieved (2.039 m/s) with the cost of maximum *energy consumption* (6.784 J). Nevertheless, in some of the cases, different pairs of weights may lead to the same result, like cases 6 and 7 in the same table. For more clarification, Fig. 8 illustrates the optimization results, including all possible solutions, the Pareto-optimal front, and the best solution (using w_1= 0.4 and w_2= 0.6).

Table 6: Applying *weighted sum method* using variant weights for objectives functions

w_1	w_2	$w_1 f_1$	$w_2 f_2$	$\sum_{i=1}^{2} w_i f_i$	α (radian)	Acceleration (m/s^2)	Avg. Speed (m/s)	Energy cons. (J)
0.10	0.90	0.100	0.000	0.100	0.541	0.119	**0.481**	**1.323**
0.20	0.80	0.147	0.034	0.181	0.541	0.393	0.894	1.552
0.34	0.66	0.236	0.041	0.277	0.545	0.449	0.956	1.662
0.40	0.60	0.188	0.110	0.298	0.541	0.848	1.309	2.326
0.54	0.46	0.158	0.167	0.325	0.543	1.244	1.584	2.326
0.60	0.40	0.051	0.246	0.297	0.541	1.785	1.907	4.679
0.70	0.30	0.059	0.184	0.244	0.541	1.785	1.907	4.679
0.82	0.18	0.054	0.126	0.180	0.541	1.823	1.937	5.131
0.90	0.10	0.000	0.100	0.100	0.551	1.975	**2.039**	**6.784**

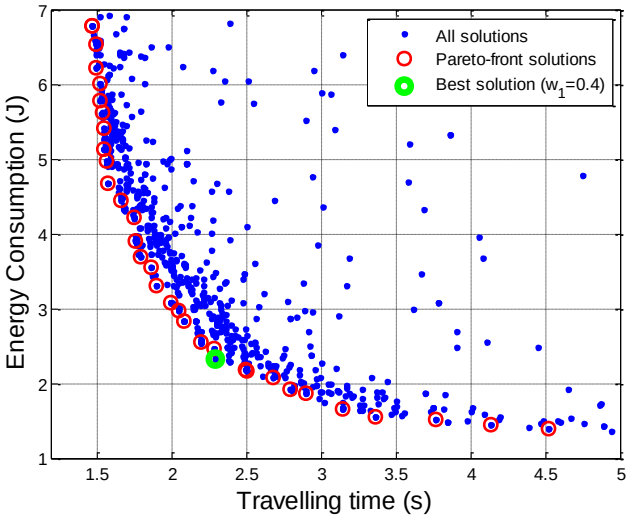


Fig. 8: Pareto-optimal front showing dominated points.

4.2 Validation of the Obtained Results

In order to validate the obtained results of multi-objective optimization, arbitrarily variant values of *winding angle* and *acceleration* were tested and compared with the optimized parameters. The tested values were selected from the same range of *winding angle* ($\pi/6$ to $\pi/3$) and *acceleration* (0.1 to 2.0) as initialized in *gamultiobj* function. The comparison of results is shown in Table 7.

Table 7: Comparison of selected and optimized parameters

Winding angle radian	**(°)**	**Acceleration (m/s^2)**	**Average Speed (m/s)**	**Energy Cons. (J)**	**Remarks**
1.047	60.0	0.100	0.349	03.978	Not optimized
1.047	60.0	0.500	0.860	07.907	Not optimized
1.047	60.0	1.100	1.285	17.501	Not optimized
1.047	60.0	1.400	1.440	19.691	Not optimized
0.785	45.0	0.200	0.599	02.743	Not optimized
0.785	45.0	0.700	1.118	04.749	Not optimized
0.785	45.0	1.000	1.345	07.136	Not optimized
0.785	45.0	1.500	1.659	11.385	Not optimized
0.628	36.0	0.300	0.770	01.962	Not optimized
0.628	36.0	0.600	1.085	02.653	Not optimized
0.628	36.0	0.120	0.473	01.644	Not optimized
0.542	31.1	0.314	0.797	01.511	Optimized
0.541	31.0	0.848	1.309	02.326	Optimized
0.541	31.0	1.272	1.611	03.549	Optimized
0.543	31.1	1.427	1.709	03.909	Optimized
0.541	31.0	1.823	1.937	05.131	Optimized

The results in Table 7 clearly show that using any combination of values for *winding angle* and *acceleration* other than the values obtained by Pareto-optimal front do not achieve better results in terms of fitting the objective functions. Therefore, these results show the validity of using multi-objective optimization technique for snake robot locomotion.

https://doi.org/10.31436/iiumej.v22i2.1691

5. EXPERIMENTAL RESULTS

To verify the simulation results, several experiments were conducted to check the effect of changing the related factors on power consumption, namely, acceleration, *winding angle* and surface friction.

5.1 Effect of Acceleration on Energy Consumption

Acceleration here means the change of speed in the forward motion of the snake robot along the serpentine path after it starts moving. Consequently, this acceleration will determine the forward speed of the snake robot, which is theoretically assumed to be constant. However, direct control of acceleration is not an easy job; instead, controlling the rotation speed of the servomotors is much easier. Accordingly, the effect of the acceleration on power consumption can be simply found by varying the rotational speed of the servomotors and getting the final reading of *energy consumption*.

To conduct the experiment, the value of the *winding angle* needs to be fixed, while the rotation speed of the servomotor needs to be changed within a specific range. For more analysis, the experiment was repeated several times using a range of *winding angle* from 20° to 45° with a 5-step increment, while the range of the servomotor *rotation speed* was varied for each *winding angle* between the *rotation speed code* '30' which represents 0.77 rad/s, and the *rotation speed code* '60' which represents 1.43 rad/s.

In this experiment, the snake robot was run 42 times for 6-meter travel. Meanwhile, the *winding angle* and *rotation speed* were varied for each experiment. Digital wattmeter was used to measure the total *energy consumption* for each travel of the snake robot. The results are illustrated in Fig. 9.

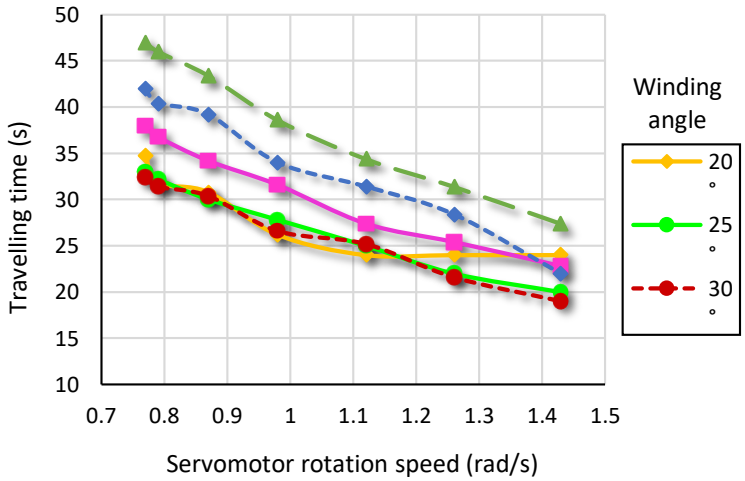


Fig. 9: Travelling time for 6-meter distance.

Figure 9 illustrates the relationship between the servomotor *rotation speed* and the travelling time of the snake robot. The experiments were conducted using different values of servomotor *rotation speed*, starting from *rotation speed code* '30' (0.77 rad/s) until *rotation speed code* '60' (1.43 rad/s) with a 5-step increment. Meanwhile, the *winding angle* was varied from 20° to 45° with a 5-step increment also. As seen in the above figure, all the curves demonstrate a similar trend, which shows that the travelling time is inversely proportional to the servomotor *rotation speed*. In other words, increasing the servomotor *rotation speed* can decrease the travelling time of the snake robot. It is also noticed that the travelling time increased when the *winding angle* was above 30° as the body of snake

robot became more twisted. Therefore, the objective of decreasing the travelling time of the snake robot when it is used for exploration task can be achieved by simply increasing the rotation speed of the servomotors.

The speed range used in these experiments considered the limitations of the servomotors. Therefore, the minimum rotation speed selected was 0.77 rad/s, since the speed below this value is relatively slow, while the maximum speed was 1.43 rad/s, as controlling the snake robot became difficult beyond that speed.

5.2 Effect of Winding Angle on Energy Consumption

The following experiment aims to find the relationship between the *winding angle* of a snake robot and the *energy consumption* after travelling a specific distance. This experiment was actually a part of the previous experiment, where the actual time needed by the snake robot to travel for 3 meters and 6 meters was measured and recorded. As mentioned in the previous experiment, the range of *winding angle* was 20° to 45° while the range of servomotor *rotation speed* was from 0.77 rad/s to 1.43 rad/s.

In this experiment, the *acceleration* of the snake robot or, in other words, the servomotor *rotation speed* was kept constant (e.g. 1.43 rad/s) while the *winding angle* was varied within a range of 30° to 60°. Later on, the experiment was repeated using different values of *winding angles* while the servomotor *rotational speed* was kept constant. In this case, increasing the *winding angle* will obviously increase the undulation of the snake robot and consequently reduce its length (along the body axis) as shown in Fig. 10, which demonstrates snapshots of a snake robot with three different *winding angle*s.

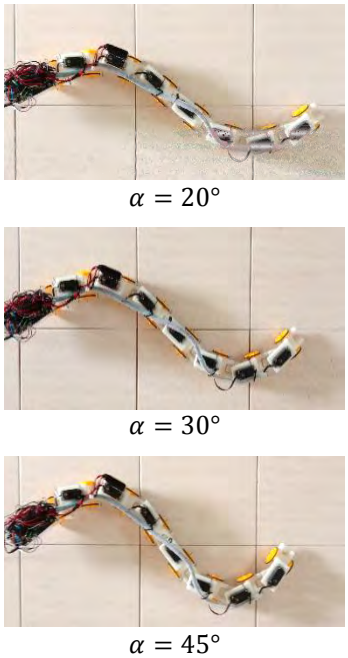


Fig. 10: Undulation of the snake robot with different *winding angles*.

Based on the above figure, increasing the undulation of the snake robot increases the travelling distance (because of the body shrinking), thus, the snake robot will take a longer time to reach its destination. The *energy* consumptions were measured starting from the beginning of each travel until the end of each travel of the snake robot using a digital wattmeter. The results of these experiments for 3 meters and 6 meters travelling distances are presented in Tables 8 and 9, and plotted in Figs. 11 and 12 respectively.

https://doi.org/10.31436/iiumej.v22i2.1691

Table 8: *Energy consumption* for 3-meter distance (Wh)

Winding Angle (°)	Servomotor rotation speed (rad/s)						
	0.77	**0.79**	**0.87**	**0.98**	**1.12**	**1.26**	**1.43**
20	72	72	86	65	65	65	58
25	72	72	65	65	65	72	58
30	79	79	72	72	72	58	65
35	79	79	86	72	72	72	65
40	86	94	94	94	79	79	72
45	101	101	94	94	101	94	86

Table 9: *Energy consumption* for 6-meter distance (Wh)

Winding Angle (°)	Servomotor rotation speed (rad/s)						
	0.77	**0.79**	**0.87**	**0.98**	**1.12**	**1.26**	**1.43**
20	137	122	115	122	115	108	108
25	137	122	122	122	122	115	108
30	130	122	115	122	115	115	101
35	137	144	137	130	130	130	115
40	158	166	158	151	144	144	137
45	173	166	166	166	173	158	144

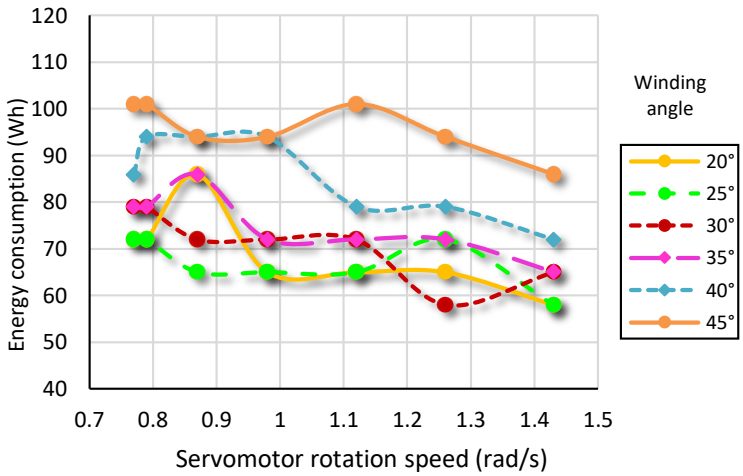


Fig. 11: *Energy consumption* for 3-meter distance.

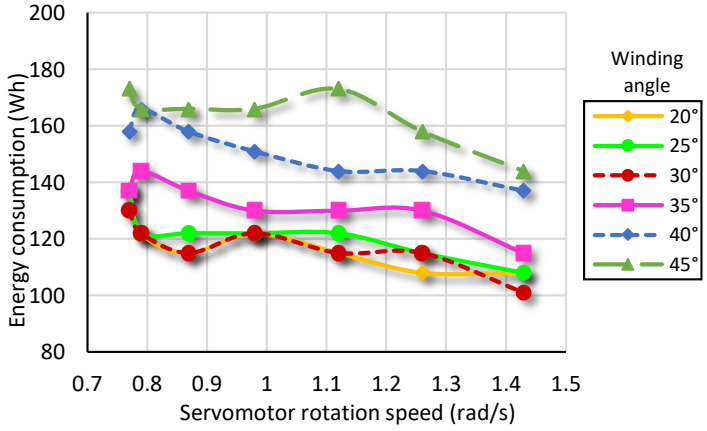


Fig. 12: *Energy consumption* for 6-meter distance.

The experiments were conducted using different values of servomotor *rotation speed*, starting from *rotation speed code* '30' (0.77 rad/s) until *rotation speed code* '60' (1.43 rad/s) with a 5-step increment. Meanwhile, the *winding angle* was varied from 20º to 45º with a 5-step increment. The curves in both the figures demonstrated a general trend showing that *energy consumption* is inversely proportional to the servomotor *rotation speed*. However, that is not the case in some ranges such as in the range between the rotation speed of 0.87 rad/s and 1.12 rad/s for the *winding angles* 25° and 30°.

Furthermore, it can be noticed from Figs. 11 and 12 that doubling the travelling distance did not necessarily double the *energy consumption*. Instead, the *energy consumption* for the 6 meters travelling distance increased in the range of 1.3 to 2 times the energy consumption of the 3 meters travelling distance for different winding angles and servomotor speeds.

On the other hand, the direct proportional relationship between the *winding angle* and the *energy consumption* can be seen clearly in both figures, which means that increasing the *winding angle* increases the *energy consumption*. However, the curve of 30° *winding angle* demonstrates an exception as it shows less *energy consumption* than the 25° curve, and even less than the 20° curve in some points.

It can be deduced from the above analysis that minimizing *energy consumption* can be achieved by selecting a low value of *winding angle*. However, the practical experiment proved that the optimal *winding angle* is not necessarily the one with minimum value as seen in Fig. 11, where the 30° curve shows the minimum *energy consumption* despite it not being the minimum *winding angle*. Therefore, from an optimization perspective, selection of the *winding angle* of 30° could achieve the minimum *energy consumption* in serpentine locomotion of the snake robot.

6. COMPARISON OF SIMULATION AND EXPERIMENTAL RESULTS

A comparison is conducted between simulation results as shown in Fig. 13 and experimental results shown in Fig. 12, with respect to *energy consumption* and *travelling speed* of the snake robot. In both simulation and experiment, the snake robot was tested for travelling for a specific distance using different values of *winding angle* and *rotation speed* at each time. The trend of energy consumption is divided into three regions as shown in Fig. 13. In the first region (*Region 1*) and the last region (*Region 3*), the *energy consumption* is directly proportional to *acceleration*, while the middle region (*Region 2)* demonstrates an inversely proportional relationship.

By comparing both figures, it can be seen that *energy consumption* in Fig. 12 is inversely proportional to servomotor *rotation speed*, which matches the trend of the middle region (marked with a red rectangle) in Fig. 13, where the *energy consumption* is inversely proportional to *acceleration.* Although the x-axes are not the same in both figures, they both lead to the average speed of the snake robot.

It can be concluded from the above figures that the rotation speed of the servomotors used in this experiment is limited. Therefore, the experimental *energy consumption* shown in Fig. 12 cannot demonstrate the trend of *Region 3* of the simulation results as seen in Fig. 13.

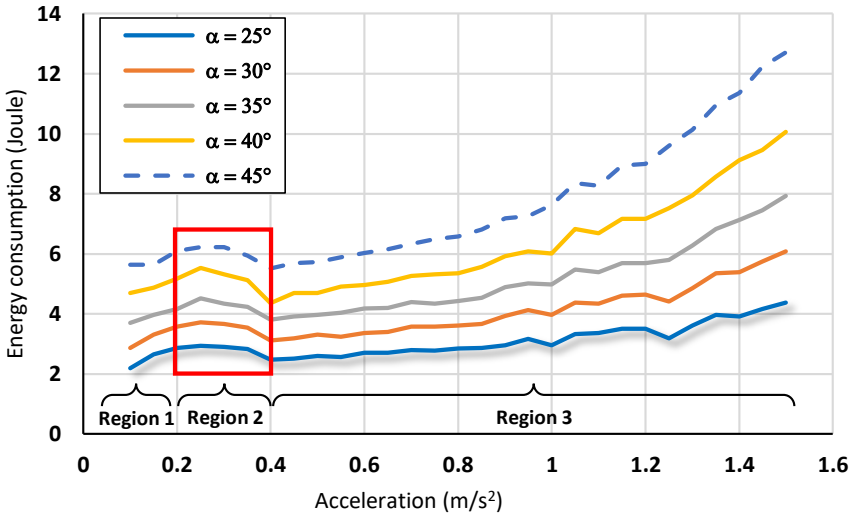


Fig. 13: *Energy consumption* in simulation results.

Moreover, simulation results show that for achieving the objectives of minimizing *travelling time* and *energy consumption* the minimum *winding angle* and fastest *rotation speed* should be selected. On the other hand, the experimental results demonstrated different behaviour when servomotor *rotation speed* was set above 1.2 rad/s as seen in Fig. 9. In the mentioned figures, *winding angle*s 25° and 30° showed less *travelling time* than the *winding angle* 20°. Moreover, the experimental results in Fig. 12 show that the minimum *energy consumption* can be achieved using the *winding angle* of 30°.

Finally, it was experimentally found that the *winding angle* of 30° is the optimum *winding angle* that can achieve the objectives of minimizing the *energy consumption* and the *travelling time* when the snake robot travels in serpentine locomotion.

7. CONCLUSION

In this paper, a multi-objective optimization for snake robot locomotion is presented. The model of the snake robot is introduced including the dynamic effects of the environment including friction. The equations listed in this paper describe the model of the snake robot travelling in serpentine locomotion besides necessary calculation of *energy consumption* and *average speed* of the snake robot. Simulation results show that *acceleration* has a contradictory effect on *average speed* and *energy consumption*, while the initial value of *winding angle* affects both *energy consumption* and *average speed* analogously. MATLAB *gamultiobj* function was used to achieve two objectives: minimizing the total *travelling time* and minimizing the total *energy consumption* by obtaining the Pareto-optimal front. The comparison of selected and optimized parameters shows the validity of using a multi-objective optimization technique for snake robot locomotion to achieve both objectives. On the other hand, the experimental results conformed with the simulation results regarding the effect of acceleration on *energy consumption* except for the *winding angles* of 25°and 30°, which demonstrated less *energy consumption* than the *winding angle* of 20° when the servomotor *rotation speed* was above 1.2 rad/s. Furthermore, the experimental results showed that the *winding angle* 30° was the optimum *winding angle* that can achieve both objectives of minimizing the *energy consumption* and the *travelling time*.

ACKNOWLEDGEMENT

The authors profoundly acknowledge the Ministry of Higher Education (MOHE), Malaysia for funding this research through the Fundamental Research Grant Scheme (grant no. FRGS130720313).

REFERENCES

[1] Gray J. (1946) The mechanism of locomotion in snakes. Journal of Experimental Biology, 23(2): 101-120.
[2] Hirose S, Mori M. (2004) Biologically inspired snake-like robots. Robotics and Biomimetics, 2004. ROBIO 2004. IEEE International Conference on (pp. 1-7). IEEE.
[3] Hirose S. (1993) Biologically inspired robots: snake-like locomotors and manipulators, vol. 1093. Oxford: Oxford university press.
[4] Saito M, Fukaya M, Iwasaki T. (2002) Modeling, analysis, and synthesis of serpentine locomotion with a multilink robotic snake. IEEE Control Systems Magazine, 22(1): 64-81.
[5] Chen L, Ma S, Wang Y, Li B, Duan D. (2007) Design and modelling of a snake robot in traveling wave locomotion. Mechanism and Machine Theory, 42(12): 1632-1642.
[6] Chernousko FL. (2005) Modelling of snake-like locomotion. Applied mathematics and computation, 164(2): 415-434.
[7] Ariizumi R, Matsuno F. (2017) Dynamic analysis of three snake robot gaits. IEEE Transactions on Robotics, 33(5): 1075-1087.
[8] Sanfilippo F, Azpiazu J, Marafioti G, Transeth AA, Stavdahl Ø, Liljebäck P. (2017) Perception-driven obstacle-aided locomotion for snake robots: the state of the art, challenges and possibilities. Applied Sciences, 7(4): 336.
[9] Wang Z, Gao Q, Zhao H. (2017) CPG-inspired locomotion control for a snake robot basing on nonlinear oscillators. Journal of Intelligent & Robotic Systems, 85(2): 209-227.
[10] Trebuňa F, Virgala I, Pástor M, Lipták T, Miková Ľ. (2016) An inspection of pipe by snake robot. International Journal of Advanced Robotic Systems, September-October, 1-12.
[11] Bando Y, Suhara H, Tanaka M, Kamegawa T, Itoyama K, Yoshii K, ... & Okuno HG. (2016) Sound-based online localization for an in-pipe snake robot. in 2016 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR) (pp. 207-213). IEEE.
[12] Whitman J, Zevallos N, Travers M, Choset H. (2018) Snake robot urban search after the 2017 Mexico City earthquake. 2018 IEEE international symposium on safety, security, and rescue robotics (SSRR) (pp. 1-6). IEEE.
[13] Chen L, Ma S, Wang Y, Li B, Duan D. (2007) Design and modelling of a snake robot in traveling wave locomotion. Mechanism and Machine Theory, 42(12): 1632-1642.
[14] Saito M, Fukaya M, Iwasaki T. (2002) Modeling, analysis, and synthesis of serpentine locomotion with a multilink robotic snake. IEEE Control Systems Magazine, 22(1): 64-81.
[15] Kelasidi E, Jesmani M, Pettersen KY, Gravdahl JT. (2018) Locomotion efficiency optimization of biologically inspired snake robots. Applied Sciences, 8(1): 80.
[16] Liu J, Wang Y, Ii B, Ma S. (2004) Path planning of a snake-like robot based on serpenoid curve and genetic algorithms. in Intelligent Control and Automation, 2004. WCICA 2004. Fifth World Congress on (6: 4860-4864). IEEE.
[17] Hasanzadeh S, Tootoonchi AA. (2008) Adaptive optimal locomotion of snake robot based on CPG-network using fuzzy logic tuner. in Robotics, Automation and Mechatronics, 2008 IEEE Conference on (pp. 187-192). IEEE.
[18] Jafari M, Shahmansoorian A. (2014). PI/backstepping control of snake robot optimized by genetic algorithm. In Electrical and Computer Engineering (CCECE), 2014 IEEE 27th Canadian Conference on (pp. 1-6). IEEE.
[19] Rao SS, Rao SS. (2009) Engineering optimization: theory and practice. John Wiley & Sons.
[20] Marler RT, Arora JS. (2010). The weighted sum method for multi-objective optimization: new insights. Structural and multidisciplinary optimization, 41(6): 853-862.

https://doi.org/10.31436/iiumej.v22i2.1706

MODELLING A 1-DOF FINGER EXTENSOR MACHINE FOR HAND REHABILITATION

IFRAH SHAHDAD, NORSINNIRA ZAINUL ABIDIN* AND AHMAD JAZLAN

Department of Mechatronics Engineering,
International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia

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

(Received: 24th November 2020; Accepted: 22nd December 2020; Published on-line: 4th July 2021)

ABSTRACT: It is essential to have an accurate representation of a robotic rehabilitation device in the form of a system model in order to design a robust controller for it. This paper presents mathematical modelling and validation through simulation and experimentation of the 1-DOF Finger Extensor rehabilitation machine. The machine's design is based on an iris mechanism, built specifically for training open and close movements of the hand. The goal of this research is to provide an accurate model for the Finger Extensor by taking into consideration various factors affecting its dynamics and to present an experimental validation of the devised model. Dynamic system modelling of the machine is performed using Lagrangian formulation and the involved physical parameters are obtained experimentally. To validate the developed model and demonstrate its effectiveness, hardware-in-the-loop experiments are conducted in the Simulink-MATLAB environment. Mean absolute error between the simulated and experimental response is 1.38° and the relative error is 1.13%. The results obtained are found to be within the human motion resolution limits of 5 mm or 5º and exhibit suitability of the model for application in robotic rehabilitation systems. The model accurately replicates the actual behavior of the machine and is suitable for use in controller design.

ABSTRAK: Gambaran tepat mengenai model sistem peranti rehabilitasi robotik adalah sangat penting bagi pembangunan sesebuah reka bentuk alat kawalan tahan lasak. Kajian mengenai model matematik dan pengesahan melalui simulasi dan eksperimentasi mesin pemulihan 1-DOF 'Finger Extensor'. Mesin ini direka bentuk berdasarkan mekanisme iris, dibangunkan khusus bagi melatih gerakan buka dan tutup tangan. Tujuan kajian ini adalah bagi menyediakan model Finger Extensor yang tepat dengan mengambil kira faktor mempengaruhi dinamik dan pengesahan model eksperimen yang dirancang. Model sistem dinamik mesin ini diuji menggunakan formula Lagrangian dan parameter fizikal yang terlibat diperoleh melalui eksperimen. Model ini disahkan dan diuji keberkesanannya menggunakan eksperimen Perkakasan-dalam-gelung melalui MATLAB-Simulink. Purata ralat mutlak antara dapatan simulasi dan respon eksperimen adalah 1.38° dan ralat relatif 1.13%. Dapatan kajian adalah dalam had resolusi gerakan tangan manusia iaitu 5 mm atau 5º dan didapati model ini sesuai bagi aplikasi sistem rehabilitasi robotik. Model ini tepat dalam mereplikasi kelakuan sebenar mesin dan sesuai digunakan bagi reka bentuk kawalan.

KEYWORDS: *modelling; simulation; experimental validation; hand rehabilitation; hardware-in-the-loop*

1. INTRODUCTION

Every year, up to 15 million people suffer from stroke, making it one of the leading causes of severe disability in the world [1]. After a neurological injury such as stroke, patients suffer from several impairments, of which the most common and impeding is the impairment to the hand sensorimotor function. The finger extensor muscles develop weakness during voluntary movements as a result of activity deficit. This leads to spasticity which is a velocity-dependent increase in muscle tone. Spasticity causes intense muscle contractions and stiffness that manifest in the form of a clenched fist, tensed fingers, a curled wrist and other deformities. These impairments lead to a decreased ability to open and close the hand, control individual finger movements and perform force coordination, thereby severely affecting the quality of life of stroke patients [2].

Patients undergo intense physical therapy under the guidance of an occupational therapist to regain functionality in the affected limb. Activity-based exercises are known to be effective in helping overcome the motor deficits induced by post-stroke spasticity [3]. However, occupational therapy is expensive and as a result, a vast majority of the patients do not have continuous access to it. Besides, conventional therapy is monotonous, repetitive and difficult to be sustained for prolonged periods of time, leading to high drop-out rates. Furthermore, with conventional therapy, it is not possible to measure the exact progress of patients over time and provide them with well-regulated and a wide range of rehabilitative forces. Due to these factors, robot-assisted rehabilitation has seen a growing interest in the last few decades. Robotic devices, due to their features of high repeatability, interactivity, ability to provide a wide range of accurate forces and automatic measurement of patient progress, are becoming increasingly popular. The interactive nature of robot-aided therapy not only increases physical engagement but also encourages cognitive engagement of the patient through immersive and challenging exercises. Patient engagement in rehabilitation is vital because encouraging human users to perform self-initiated movements is an essential requirement to achieve motor learning and neuroplasticity [4].

Furthermore, robot-aided rehabilitation is capable of producing *assistive* as well as resistive forces that can be tailored to the specific impairment of the patient. When patients gain some amount of motor function in the impaired part, they are subjected to active-resistive therapy whereby, forces are applied on the patient in a direction opposite to the one they are tracing. Resistive strategies enhance patient performance as they require additional effort from the patient to resist opposite forces. Such efforts induce neuroplasticity and help in development of neural pathways from the impaired muscles to the brain which aids in regaining of motor function [5].

However, all these features of robotic rehabilitation devices are possible only when efficient controllers are developed [6]. A variety of control strategies have been proposed for rehabilitation robots over the past few decades. These control schemes are designed around the central idea of modulating assistance/resistance provided to the patient based on various factors such as force exerted by the patient, limb velocity, tracking error, interaction force/torque, EEG/EMG activity [7].

In recent literature, approaches such as force/stiffness control [8] Assist-as-Needed control (AAN) [9-11] and performance-based impedance control [7,12,13] have been implemented on rehabilitation robots. AAN control aims at providing minimal intervention in patient recovery. Performance-based impedance control modulates the assistance provided, based on real-time measurements of the patient's biomechanical characteristics.

Before the implementation of these control strategies, dynamic and kinematic modelling of the rehabilitation machine is necessary. Modelling is the mathematical representation of a physical system [14]. A system model is developed from mathematical equations and then represented graphically in software such as MATLAB, Simulink and StateFlow, etc. The developed model is then validated through simulations, reducing the cost associated with experimental validation. During the process of model development, trade-offs have to be made between the fidelity and simplicity of the model. Fidelity implies the extent to which the system model mimics the behavior of the real system and simplicity ensures that the model remains simple enough for smooth controller development in later stages. Thus, there are some deviations between the behavior of the actual system and its model. After balancing simplicity and fidelity, the model is validated using a Hardware-in-the-Loop (HiL) approach. In this approach, one or more real subsystems interact in a closed loop with the virtual subsystems of the model. A comparison between the response of such a closed-loop interaction with the simulation output reveals performance of the developed model [15].

In [16], dynamic and kinematic modelling of the multipurpose rehabilitation robot, the Universal Haptic Pantograph, was performed. Validation of the model was carried out through experiments with healthy subjects reflecting its performance. Dynamic modelling and experimental validation of a parallel robot for upper limb rehabilitation was presented in [17]. Approximately zero steady-state errors points to a high degree of accuracy and suitability of the developed model. Dynamic modelling of a haptic finger actuated by a McKibben artificial muscle built for tele-operated rehabilitation was performed in [18]. Upon experimental validation, the model demonstrated a satisfactory performance.

In this work, mathematical modelling and system simulation of a 1 degree-of-freedom (DOF) Finger Extensor rehabilitation machine was performed. A Model-Based Development (MBD) [15] approach to controller design was taken. Such an approach to design enabled verification and validation at each stage. Based on simulation results, modifications to the system could be made easily without incurring high costs. As a first step, mathematical modelling was done, followed by simulation of the system behavior. Simulation results were then compared to hardware measurements which demonstrated the accuracy of the realized model. Hardware-in-the-Loop testing of the model carried out in this study ensures minimization of risks and costs associated with experimental validation. The developed model aids in the design and implementation of a control algorithm for the Finger Extensor machine.

2. THE FINGER EXTENSOR MACHINE

The Finger Extensor rehabilitation machine is a 1-DOF device with an iris mechanism lying at the heart of its dynamics. The design of the iris or diaphragm is inspired by the constriction and dilation of the iris of the human eye, which varies size of the pupil, adjusting the aperture of the eye. The iris mechanism built into the Finger Extensor constricts and dilates the poles at the top of its surface, to allow openings of variable sizes. These poles, as shown in Fig. 1 and Fig. 2 are grasped by the patient during therapy and enable him/her to practice finger flexion and extension [19].

There are six poles in the current Finger Extensor design and each pole has a blade associated with it. On the upper end, each blade is connected to a slider and at the bottom, each blade goes into a slot in a disc using a bearing, as illustrated in Fig. 2. The disc with slots is in turn connected to the motor actuation system through a chain and sprocket mechanism. As the motor turns, the sprocket induces motion into the slotted disc. This

causes the blades to slide and move the poles in their own respective slots at the top of the machine.

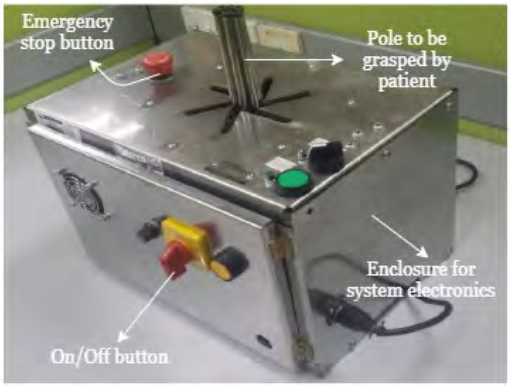


Fig. 1: 1-DOF Finger Extensor machine.

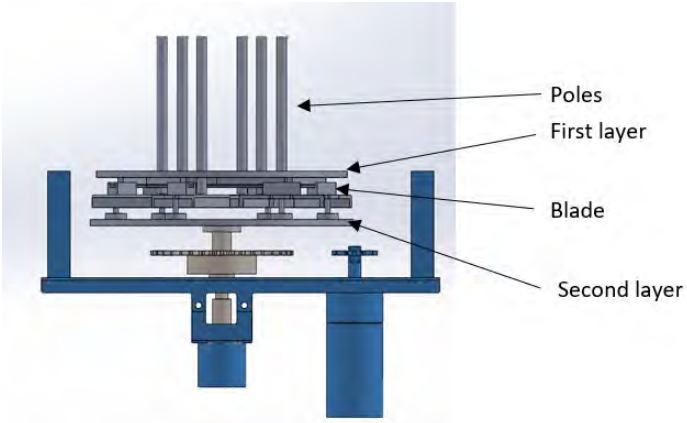


Fig. 2: CAD drawing of the side view of the iris mechanism.

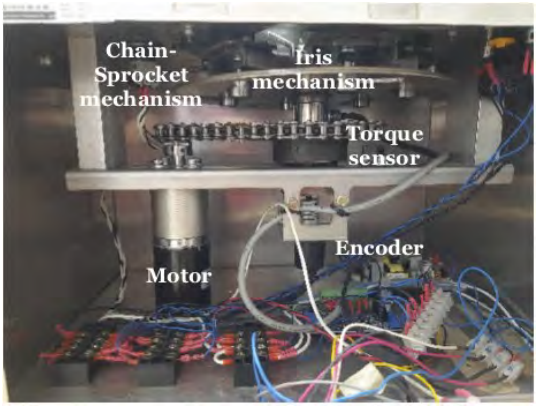


Fig. 3: Inner view of the machine.

Figure 3 shows the actuation system of the machine comprising of a brushed planetary geared DC motor (Gear Ratio- 32.5:1). Rotation of the motor is recorded by a rotary encoder by ESB electronics, Japan (Counts per Revolution- 500) connected to the sprocket using a flexible shaft coupling. A torque sensor, connected close to the chain-sprocket mechanism provides the torque applied by the machine on to the patient. Actuation system and sensors are controlled by Arduino microcontroller board which is programmed in the Simulink environment.

https://doi.org/10.31436/iiumej.v22i2.1706

3. SYSTEM MODELLING

3.1 Mechanical Model

Newton-Euler and the Lagrange-Euler methods are the most commonly used approaches for deriving dynamic equations of mechanical parts of robots [20]. The former is ideal for on-line control applications due to reduced computational time. However, it falls short in providing an adequate insight into control design due to recursive computation. The Lagrange-Euler method is simple and systematic, providing dynamic equations in a matrix form and was chosen for deriving the mechanical model for the Finger Extensor. Dynamic equation describing motion of the Finger Extensor was written as

$$M(\theta(t),\xi)\ddot{\theta}(t) + D\big(\theta(t),\dot{\theta}(t),\xi\big) + G(\theta(t),\xi) = T(t) \quad (1)$$

where,

$M(\theta(t),\xi)$ is the N x N inertia matrix.

$D\big(\theta(t),\dot{\theta}(t),\xi\big)$ is the N x 1 vector of coriolis and centrifugal forces.

$G(\theta(t),\xi)$ is the N x 1 vector of gravitational forces.

$T(t)$ is the N x 1 vector of torques applied by the actuators.

$\theta(t),\dot{\theta}(t),\ \ddot{\theta}(t)$ are the N x 1 vectors of joint displacements, velocities and accelerations respectively.

ξ is a vector (with appropriate dimension) of parameters of the mechanism such as payload.

For the Finger Extensor, $M(\theta(t),\xi)$ is a 1x1 matrix, with one element equal to the moment of inertia of a solid cylinder with iris radius of 0.055m. $D\big(\theta(t),\dot{\theta}(t),\xi\big)$ and $G(\theta(t),\xi)$ for a 1-DOF machine were null matrices. $T(t)$ for the Finger Extensor is calculated in the subsequent sections.

For developing the mechanical model of the Finger Extensor, the next step was to derive equations of motion of the moving parts of the mechanism. These were represented by flywheel and spring mechanical elements as shown in Fig. 4. J_1, J_2 represent inertia elements of the motor and iris mechanisms respectively, k_1, b_1 represent the spring stiffness effect and viscous and coulomb friction.

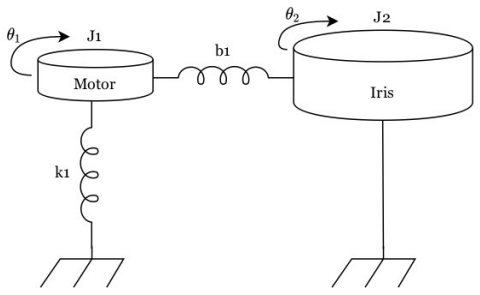


Fig. 4: Block diagram representation.

Forces acting on the motor and iris sprockets are shown in Fig. 5, where θ_1 is the angle of rotation of the motor sprocket, θ_2 is the angle of rotation of the iris mechanism, T_g is the torque on the driver sprocket, T_e is the torque exerted by the user, $f_c r_1$ and $f_c r_2$

are the contact force torques. Figures 4 and 5, hence represent Eq. (1) in terms of sum of forces (inertial, gravitational, frictional) acting on the machine.

As per D'Alembert's law, sum of torques on both flywheels was given by the equations

$$T_g = J_1\ddot{\theta_1} + k_1\theta_1 + b_1(\dot{\theta_1} - \dot{\theta_2}) - f_c r_1 \tag{2}$$

$$-f_c r_2 - J_2\ddot{\theta_2} - b_1(\dot{\theta_2} - \dot{\theta_1}) - T_e = 0 \tag{3}$$

From Eq. (3), value of f_c is calculated as

$$f_c = \frac{1}{r_2}\left[-J_2\ddot{\theta_2} - b_1(\dot{\theta_2} - \dot{\theta_1}) - T_e\right] \tag{4}$$

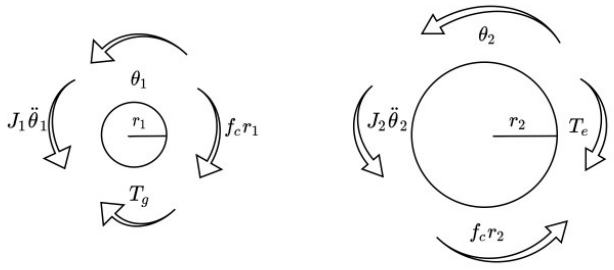


Fig. 5: Forces acting on mechanism.

Substituting value of f_c in Eq. (2)

$$T_g = k_1\theta_1 + b_1(\dot{\theta_1} - \dot{\theta_2}) - \frac{r_1}{r_2}\left[-J_2\ddot{\theta_2} - b_1(\dot{\theta_2} - \dot{\theta_1}) - T_e\right] \tag{5}$$

Using the inverse gear ratio, $\frac{r_2}{r_1}$ = N and the equality of arc lengths, $r_1\theta_1 = r_2\theta_2$, it follows that $\theta_1 = \frac{r_2}{r_1}\theta_2 = N\theta_2$. Substituting in Eq. (5),

$$T_g = \frac{1}{N}J_2\ddot{\theta_2} + \frac{1}{N}T_e + Nk_1\theta_2 + V_c(\dot{\theta_2}) + F_c Sgn(\dot{\theta_2}) \tag{6}$$

where, $V_c(\dot{\theta_2})$ and $F_c Sgn(\dot{\theta_2})$ are the viscous and coulomb friction terms. Therefore, dynamic equation of the 1-DOF Finger Extensor was represented as

$$T_g(t) = \frac{1}{N}J_2(\theta_2)\ddot{\theta_2}(t) + \frac{1}{N}T_e(t) + Nk_1\theta_2(t) + V_c(\dot{\theta_2})(t) + F_c Sgn(\dot{\theta_2}) \tag{7}$$

3.2 Actuator Model

The Finger Extensor is powered by a planetary-geared, 24 V brushed DC motor. A schematic of the motor is shown in Fig. 6. Using Kirchhoff's voltage law and Newton's law of motion, dynamic equations of the actuator were written as

$$J_m\ddot{\theta_m}(t) = -B_v\dot{\theta_m}(t) + k_t i_a(t) - T_L(t) \tag{8}$$

$$Li_a(t) = -k_v\dot{\theta_m}(t) - Ri_a(t) + V(t) \tag{9}$$

where,

J_m is the moment of inertia of the motor (kg. m^2)

https://doi.org/10.31436/iiumej.v22i2.1706

$\theta_m(t)$ is the angular displacement of the motor (rad)

B_v is the viscous friction coefficient (Nm/rad/s)

k_t, k_v are the torque and back emf constants (Nm/A, V/rad/s)

L, R are the armature inductance and resistance (H, Ω)

$i_a(t)$ is the armature current for the motor (A)

$T_L(t)$ is the load torque for the motor (Nm)

$V(t)$ is the voltage input (V)

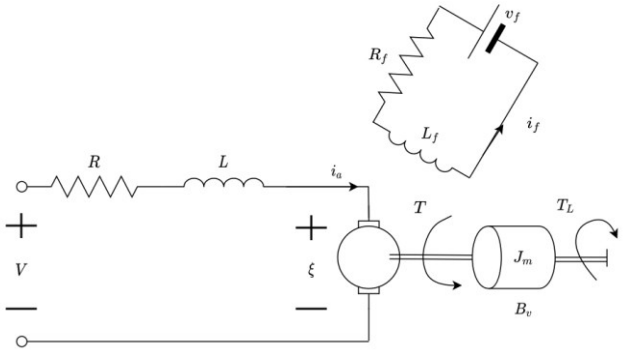


Fig. 6: Motor schematic.

Utilizing the relationship between loading torque $T(t)$ (from the motor) and load torque $T_L(t)$ (on the motor shaft) and the relationship between the motor's angular position and the machine's angular position, we formulated

$$T_L(t) = \frac{T(t)}{N}, N \geq 1 \tag{10}$$

$$\theta_m(t) = N\theta(t) \tag{11}$$

where *N* is the inverse gear ratio.

Substituting Eq. (10) and Eq. (11) in Eq. (8) and Eq. (9), the otherwise linear, time-invariant actuator dynamic was converted into a non-linear time-variant one. The actuator model in the form of a third-order differential equation was represented as:

$$\dddot{\theta}_2(t) + \left\{\frac{B_v L + J_m R}{J_m L}\right\}\ddot{\theta}_2(t) + \left\{\frac{k_v k_t + B_v R}{J_m L}\right\}\dot{\theta}_2(t) + \left\{\frac{R}{N^2 J_m L}\right\}T_g(t) + \left\{\frac{1}{N^2 J_m}\right\}\dot{T}_g = \left\{\frac{k_t}{N J_m L}\right\}V(t) \tag{12}$$

To represent the actuator dynamic model in the state variable form, θ, $\dot{\theta}, \ddot{\theta}$ were chosen as state variables and the state vector for the actuator was given by

$$X_B(t) = \begin{bmatrix}\theta(t) & \dot{\theta}(t) & \ddot{\theta}(t)\end{bmatrix}^T \tag{13}$$

The state equation was formulated as

$$\dot{X}_B(t) = A_B X_B(t) + B_B U(t) + F_B T_g(t) + Q_B \dot{T}_e(t) \tag{14}$$

where

$$A_B(X_B, t) = \begin{bmatrix}0 & 0 & 1\\ 0 & 1 & 0\\ 0 & a_{3_2} & a_{3_3}\end{bmatrix}, B_B(X_B, t) = \begin{bmatrix}0\\ 0\\ b_{3_1}\end{bmatrix}, F_B(X_B, t) = \begin{bmatrix}0\\ 0\\ f_{3_1}\end{bmatrix}, Q_B(X_B, t) = \begin{bmatrix}0\\ 0\\ q_{3_1}\end{bmatrix}$$

https://doi.org/10.31436/iiumej.v22i2.1706

Here, A_B, B_B, F_B, Q_B are the system, input, load distribution and rate of load distribution respectively. $U(t)$, $T_g(t)$, $T_e(t)$ are the input vector, mechanical link torque and toque exerted by the patient. The values of elements a_{3_2}, a_{3_3}, b_{3_1}, f_{3_1}, and q_{3_1} are

$$a_{3_2} = \left\{\frac{-k_v k_t + B_v R}{J_m L}\right\}; a_{3_3} = \left\{\frac{-B_v L + J_m R}{J_m L}\right\}; b_{3_1} = \left\{\frac{k_t}{N J_m L}\right\}; f_{3_1} = -\left\{\frac{R}{N^2 J_m L}\right\}; q_{3_1} = -\left\{\frac{1}{N^2 J_m}\right\}$$

3.3 Integrated Model

For developing the integrated model of the machine, a time derivative of the dynamic equation of the machine was found. Differentiating Eq. (7) gave

$$\dot{T}_g(t) = \left\{\frac{J_2(\theta_2)}{N}\right\}\ddot{\theta}_2(t) + \left\{\frac{\dot{T}_e(t)}{N}\right\} + Nk\dot{\theta}_2(t) + F_c\{2\delta(\dot{\theta}_2)\ddot{\theta}_2\} + V_c\ddot{\theta}_2(t) \tag{15}$$

Substituting Eq. (15) in Eq. (12),

$$\begin{aligned}\ddot{\theta}_2(t)\left\{\frac{J_m N^3 + J_2(\theta_2)}{J_m N^3}\right\} + \dot{\theta}_2(t)\left\{\frac{N^2(B_v L + J_m R) + 2F_c L\delta(\dot{\theta}_2) + V_c L}{J_m N^2 L}\right\} \\ + \dot{\theta}_2(t)\left\{\frac{N(k_v k_t + B_v R) + Lk_1}{J_m N L}\right\} + T_g(t)\left\{\frac{R}{N^2 J_m L}\right\} + \dot{T}_e(t)\left\{\frac{1}{N^3 J_m}\right\} \\ = \left\{\frac{k_t}{N J_m L}\right\}V(t)\end{aligned} \tag{16}$$

Equation (16) was solved for θ and put in state-space form, the integrated model of the Finger Extensor was obtained as

$$\dot{X}_B(t) = A_B(X_B, t)X_B(t) + B_B(X_B, t)U(t) + F_B T_g(t) + Q_B\dot{T}_e(t) \tag{17}$$

where $X_B = \begin{bmatrix}\theta_2 & \dot{\theta}_2 & \ddot{\theta}_2\end{bmatrix}^T$

A_B, B_B, F_B, Q_B are the system, input, load distribution and rate of load distribution respectively. $U(t)$, $T_g(t)$, $T_e(t)$ are the input vector, mechanical link torque and toque exerted by the patient.

$$A_B(X_B, t) = \begin{bmatrix}0 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & a_{3_2} & a_{3_3}\end{bmatrix}, B_B(X_B, t) = \begin{bmatrix}0 \\ 0 \\ b_{3_1}\end{bmatrix}, F_B(X_B, t) = \begin{bmatrix}0 \\ 0 \\ f_{3_1}\end{bmatrix}, Q_B(X_B, t) = \begin{bmatrix}0 \\ 0 \\ q_{3_1}\end{bmatrix}$$

Equation (17) was put in the state-space form and solved for $\dot{X}_B$.

$$a_{3_2} = -\left\{\frac{N^2(Nk_v k_t + NB_v R + Lk_1)}{L(J_m N^3 + J_2(\theta_2))}\right\}, a_{3_3} = -\left\{\frac{N(N^2 B_v L + N^2 J_m R + 2F_c L\delta(\theta_2) + V_c L}{L(J_m N^3 + J_2(\theta_2))}\right\}$$

$$b_{3_1} = \left\{\frac{k_t N^2}{L(J_m N^3 + J_2(\theta_2))}\right\}, f_{3_1} = -\left\{\frac{RN}{L(J_m N^3 + J_2(\theta_2))}\right\}, q_{3_1} = -\left\{\frac{1}{(J_m N^3 + J_2(\theta_2))}\right\}$$

The state-space model derived in Eq. (17) represents the integrated dynamics of the Finger Extensor and was utilized in the simulation of the developed system.

3.4 Mechanical and Electrical Parameter Determination

The development of the integrated model for the Finger Extensor was followed by determination of mechanical and electrical parameters involved. This step aids in enhancing the fidelity of the developed model. Mechanical parameters such as radius of

the iris and moment of inertia of the inner mechanism were adapted from [20]. Electrical parameters associated with the actuator model were experimentally found from the motor characteristic curve as shown in Fig. 7. Torque and back EMF constants k_t, k_v were found from the torque-current and torque-speed curves with similar values of 0.791 and 0.886, respectively. Other parameters such as the viscous friction constant B_v, armature resistance R and inductance L were adapted from [20].

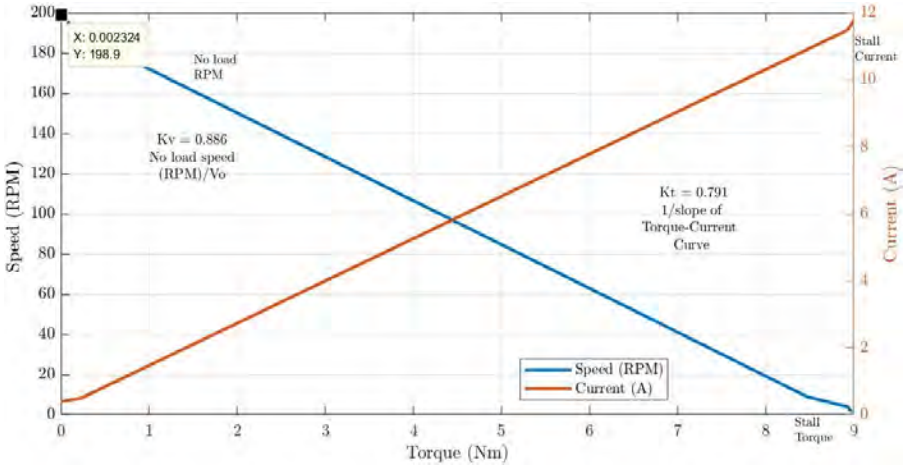


Fig. 7: Motor characteristic curves.

4. MODEL VALIDATION

Using the MBD approach for designing a controller for the Finger Extensor, dynamic system modelling was followed by graphical representation of the mathematical formulation in the MATLAB-Simulink environment. Simulation results were validated through hardware-in-the-loop experimentation. The same input was provided to both the developed model as well as the machine and comparisons were drawn between resultant outputs.

4.1 Simulation

For carrying out simulation of the mathematical model of the Finger Extensor, the voltage and torques from the motor and user acted as inputs to the Finger Extensor plant while the angle of rotation of the iris was the output. A PID controller was used for optimal position control. Rotational position output and its derivatives were used to calculate velocity and acceleration. These values helped in calculating the force acting on the patient. A desired angle of $\theta = 121°$ (2.11 radians) was fed into the plant and the simulated position of the iris was plotted as shown in Fig. 8.

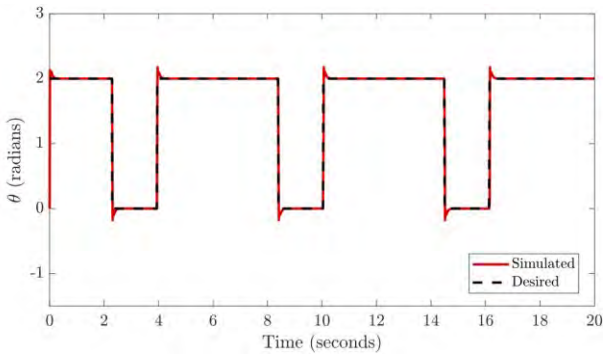


Fig. 8: Comparison between target and simulated angular position of the iris.

4.2 Hardware-in-the-Loop Experiments

To validate the developed dynamic model and demonstrate its effectiveness, experiments were conducted making the Finger Extensor a part of the simulation loop. These experiments were conducted with a healthy subject, a female of 29 years of age. A simple open-close movement of the machine was performed. Results from the simulations run in MATLAB were compared with the experimental tests with the Finger Extensor. The experimental set-up as shown in Fig. 9 consisted of the Finger Extensor connected to an Intel Core i5 PC by means of an Arduino Uno microprocessor board. This board received sensory data from the encoder as well as the torque sensor. Communication between the Arduino and MATLAB 2018b software, running on the PC was established using a MATLAB s-function. This function read data from the encoder i.e. the position sensor of the machine and converted the encoder counts to angular displacement of the iris.

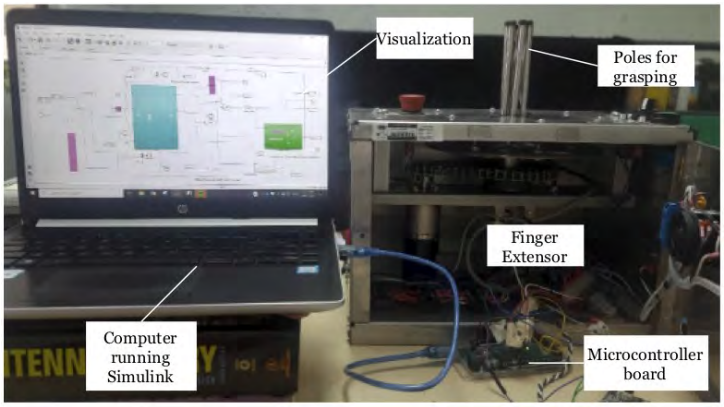


Fig. 9: Experimental setup for hardware-in-the-loop simulation.

The movement commands to the motor and consequently to the encoder were provided using Simulink's support package for Arduino which allowed Simulink blocks to control hardware connected to the Arduino using PWM signals. In this case, hardware controlled was the motor driver connected to the motor. The desired input angle was fed into the hardware and the Finger Extensor was run for about 30 seconds, performing open-close movements. The comparison between the simulated and the experimental angular displacement of the machine is shown in Fig. 10.

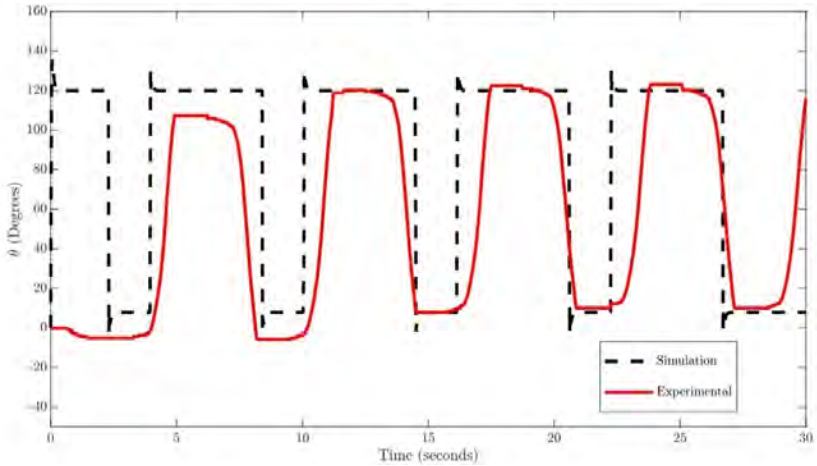


Fig. 10: Simulated and experimental responses during open-close movements.

The figure shows a complete open and close cycle of the Finger Extensor for 30 seconds. The poles of the Extensor open to their maximum rotation angle of 121° followed by closing motion to 7.89°. From the figure, it is clear that the hardware had a slower response than the simulation and took about 11.20 seconds to sync with the simulation results. However, beyond that time, error was negligible and actual behavior of the Finger Extensor overlapped with the simulated, modeled behavior. From experimental data, it was observed that the mean absolute error was 1.38° and the relative error was 1.13%. However, for rehabilitation applications, movement accuracy is less critical as compared to surgical applications. Also, position resolution of human arm movement is 5 mm or 5º [21], which indicates that a mean absolute error of 1.38° is likely to be acceptable for rehabilitation purposes.

Figure 11 shows the comparison between the simulated result and the hardware-in-the-loop response, when the Finger Extensor was programmed to open to an angle of 59.58° (1.04 radians). Simulation and experimental curves took 0.3 seconds to reach the desired angle. The Root Mean Squared Error (RMSE) was calculated as 0.9875. There was a slight overshoot of 2.86º (0.05 radians) which caused the poles of the Finger Extensor to open to an angle 2.86º greater than the target. However, this negligible overshoot did not inflict any discomfort on the subject. Also, the response settled down to its desired value within 0.08 seconds.

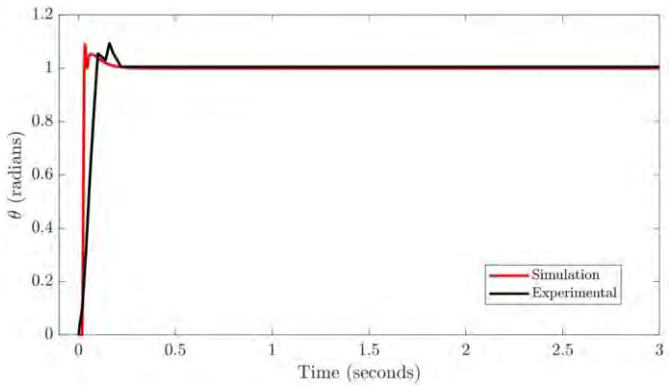


Fig. 11: Simulated and experimental responses for target angle experiment.

5. CONCLUSION

This paper presented a mathematical model for the 1-DOF Finger Extensor rehabilitation machine, taking into consideration the dynamics of the mechanism as well as that of the actuation system. Simulation results show that the developed model tracks the desired position efficiently with minimal error. Hardware-in-the-Loop experiments carried out with the machine demonstrate the validity of the model. System response using the developed model was compared with that obtained through hardware experimentation. The results illustrate the accuracy of the model to the actual behavior of the rehabilitation system and make it suitable for use in the design of a control scheme for it. Obtained results with mean absolute error of only 1.38° are within the human motion resolution limits and exhibit suitability of the model for application in robotic rehabilitation systems.

Future research with the Finger Extensor and the developed model will be directed towards development of a control scheme for the machine. This will be aimed at modulating assistance/resistance provided to the patient based on his/her performance using various metrics such as limb velocity, applied force, position error etc. Considering patient's strength and residual ability, primary goal of the controller would be to increase

self-initiated movements and patient engagement in training exercises which would lead to an increase in neuroplasticity. As part of future study, clinical testing of the Finger Extensor will also be carried out with stroke patients. This will include a detailed study of the level of comfort of the device for use with such patients.

ACKNOWLEDGEMENTS

The authors would like to acknowledge International Islamic University Malaysia (IIUM) for supporting this research under the grant number: P-RIGS18-019-0019.

REFERENCES

[1] Khor KX, Chin PJH, Rahman HA, Yeong CF, Su ELM, Narayanan ALT. (2014) A novel haptic interface and control algorithm for robotic rehabilitation of stoke patients. In 2014 IEEE Haptics Symposium (HAPTICS), Houston, Texas-USA; pp. 421-426.

[2] Plantin J, Pennati GV, Roca P, Baron JC, Laurencikas E, Weber K, Godbolt AK, Borg J, Lindberg PG. (2019) Quantitative assessment of hand spasticity after stroke: imaging correlates and impact on motor recovery. Frontiers in neurology, 10:1-11.

[3] Cramer SC. (2019) Intense rehabilitation therapy produces very large gains in chronic stroke. Journal of Neurology, Neurosurgery & Psychiatry, 90: 497.

[4] Lambercy O, Ranzani R, Gassert R. (2018) Robot-assisted rehabilitation of hand function. In Rehabilitation Robotics. Volume 1. 1st edition. Academic Press; pp 205–225.

[5] Popovic MB. (2019) Biomechatronics. Academic Press.

[6] Miao Q, Zhang M, Cao J, Xie SQ. (2018) Reviewing high-level control techniques on robot-assisted upper-limb rehabilitation. Advanced Robotics, 32(24): 1253-1268.

[7] Marini F, Hughes CM, Squeri V, Doglio L, Moretti P, Morasso P, Masia L. (2017) Robotic wrist training after stroke: Adaptive modulation of assistance in pediatric rehabilitation. Robotics and Autonomous Systems, 91: 169-178.

[8] Su YY, Yu YL, Lin CH, Lan CC. (2019) A compact wrist rehabilitation robot with accurate force/stiffness control and misalignment adaptation. International Journal of Intelligent Robotics and Applications, 3(1): 45-58.

[9] Luo L, Peng L, Wang C, Hou ZG. (2019) A greedy assist-as-needed controller for upper limb rehabilitation. IEEE Transactions on Neural Networks and Learning Systems, 30(11): 3433-3443

[10] Pehlivan AU, Losey DP, O'Malley MK. (2015) Minimal assist-as needed controller for upper limb robotic rehabilitation. IEEE Transactions on Robotics; 32(1): 113-124.

[11] Pehlivan AU, Losey DP, Rose CG, O'Malley MK. (2017) Maintaining subject engagement during robotic rehabilitation with a minimal assist-as-needed (mAAN) controller. In 2017 International Conference on Rehabilitation Robotics (ICORR): 2017; London; pp. 62–67.

[12] Fricke SS, Bayon C, Rocon E, van der Kooij H, van Asseldonk EH. (2018) Pilot study of a performance based adaptive assistance controller for stroke survivors. In International Conference on Neuro Rehabilitation. Springer, pp. 302–306.

[13] Stroppa F, Loconsole C, Marcheschi S, Mastronicola N, Frisoli A. (2018) An improved adaptive robotic assistance methodology for upper-limb rehabilitation. In International Conference on Human Haptic Sensing and Touch Enabled Computer Applications. Springer; pp. 513–525.

[14] Bringmann E, Krämer A. (2008). Model-based testing of automotive systems. in 1st international conference on software testing, verification, and validation. IEEE; pp. 485-493.

[15] Soltani A, Assadian F. (2016) A hardware-in-the-loop facility for integrated vehicle dynamics control system design and validation. IFAC-Papers Online, 49(21): 32-38.

[16] Mancisidor A, Zubizarreta A, Cabanes I, Bengoa P, Brull A, Jung JH. (2019) Inclusive and seamless control framework for safe robot-mediated therapy for upper limbs rehabilitation. Mechatronics, 58: 70-79.

[17] Guang H, Ji L, Shi Y, Misgeld BJ. (2018) Dynamic modeling and interactive performanceof PARM: A parallel upper-limb rehabilitation robot using impedance control for patients after stroke. Journal of Healthcare Engineering, 2018: 1-11.

[18] Franco W, Maffiodo D, De Benedictis C, Ferraresi C. (2018) Dynamic modeling and experimental validation of a haptic finger based on a mckibben muscle. in IFToMM Symposium on Mechanism Design for Robotics. Springer; pp. 251–259.

[19] Ali MAA, Azlan NZ. (2006) Design of iris mechanism for flexion and extension training in hand rehabilitation. ARPN Journal of Engineering and Applied Sciences; 11: 4115-4122.

[20] Osman JHS. (1992) Integrated Model of Industrial Robot for Control Applications. Jurnal Teknologi, 19(1): 27-41.

[21] Germanotta M, Vasco G, Petrarca M, Rossi S, Carniel S, Bertini E, Cappa P, Castelli E. (2015) Robotic and clinical evaluation of upper limb motor performance in patients with Friedreich's ataxia: an observational study. Journal of Neuroengineering and Rehabilitation, 12(1): 1-3.

.